

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

STRATÉGIE DE MINIMISATION LOCALE
DU RISQUE QUADRATIQUE
D'UN FOURNISSEUR D'ÉLECTRICITÉ

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN MATHÉMATIQUES

PAR
PATRICIA BORDELEAU-PICHÉ

NOVEMBRE 2020

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens en premier lieu à remercier mon directeur de recherche, Clarence Simard, du département de mathématiques de l'Université du Québec à Montréal. Je ne le remercierai jamais assez de m'avoir orienté, aidé et conseillé. Sa patience a été essentielle à la réussite de ce mémoire.

Je remercie mes parents pour leur éternel soutien, peu importe l'envergure de mes projets. Je remercie aussi mes amis qui ont toujours été là pour moi. Leur soutien inconditionnel et leurs encouragements ont été d'une grande aide.

Enfin, je tiens à remercier mon employeur qui a su comprendre l'importance de l'achèvement de ma maîtrise et qui était concilient lorsque je devais prendre des congés.

À tous ces gens, je présente mes remerciements et ma gratitude.

TABLE DES MATIÈRES

LISTE DES FIGURES	iv
LISTE DES TABLEAUX	v
RÉSUMÉ	vi
INTRODUCTION	1
1 GESTION DE RISQUE	5
1.1 Notation, terminologie et rappels	6
1.2 Minimisation locale du risque quadratique	8
1.3 Minimisation globale du risque quadratique	15
1.4 Comparaison des stratégies de couverture du risque quadratique	17
2 MARCHÉS D'ÉLECTRICITÉ ET MODÈLES	22
2.1 Demande d'électricité pondérée par les prix	25
2.2 Contrats à terme sur l'électricité	28
2.3 Corrélation entre la demande et les contrats à terme	34
3 ALGORITHME DE MINIMISATION LOCALE DU RISQUE	38
3.1 Méthode Monte Carlo	38
3.2 Réduction du temps de calculs	39
3.3 Algorithme	42

4 RÉSULTATS NUMÉRIQUES	45
4.1 Données	45
4.2 Erreur de couverture du risque	46
4.3 Résultats	47
CONCLUSION	53
ANNEXE A : PROGRAMMATION DYNAMIQUE	56
ANNEXE B : MARCHÉ D'ÉLECTRICITÉ PJM	58
BIBLIOGRAPHIE	63

LISTE DES FIGURES

2.1	Marché d'électricité simpliste	23
2.2	Demande hebdomadaire et tendance estimée g	26
2.3	Comparaison entre les résidus réalisés absolus $\sqrt{\hat{v}(t)} \hat{\epsilon}_t $ et la tendance estimée $\sqrt{\hat{v}}$	27
2.4	Rendements des contrats à terme pour les différentes échéances . .	32
3.1	Histogramme des différences absolues entre les coefficients de corrélation calculés avec MODULO 365.25/7 et MODULO 52	41
4.1	Erreur de couverture du risque $\Delta_T^* + \mathcal{C}_T$ pour la stratégie locale .	50
4.2	Comparaison entre les résidus réalisés absolus $ \hat{\sigma}_{i,T} \hat{z}_{i,T} $ et l'écart-type estimé $\sigma_{i,T}$	52
B1	Rendements réalisés vs estimés d'un modèle ARIMA(2,0,2)(2,0,0)[52] pour les rendements des contrats à terme d'échéance 3 semaines .	61
B2	Étude des résidus du modèle ARIMA(2,0,2)(2,0,0)[52] pour les rendements des contrats à terme d'échéance 3 semaines	61
B3	Rendements réalisés vs estimés d'un modèle ARIMA(3,0,3)(2,0,1)[52] pour les rendements des contrats à terme d'échéance 1 semaine . .	62
B4	Étude des résidus du modèle ARIMA(3,0,3)(2,0,1)[52] pour les rendements des contrats à terme d'échéance 1 semaine	62

LISTE DES TABLEAUX

2.1	Tableau des informations connues au temps t	34
4.1	Résultats des tests « <i>in-sample</i> » de 2007 à 2012	49
4.2	Résultats des tests « <i>in-sample</i> » de 2007 à 2009 pour la stratégie locale et 2007 à 2012 pour les autres stratégies	51
B1	Tests de stationnarité et d'autocorrélation	60

RÉSUMÉ

La gestion des risques auxquels font face les fournisseurs d'électricité est impérative. Un fournisseur d'électricité doit acheter son électricité au prix du système et la revendre à ses clients à un prix prédéterminé dans leurs contrats. Il fait donc face au risque provenant de la volatilité du prix du marché. De plus, comme la demande totale des clients fluctue en temps réel, il fait aussi face à un risque de volume. En termes d'erreur de couverture terminale, la stratégie de minimisation du risque quadratique global produit des résultats supérieurs à la stratégie de minimisation locale des risques. Par contre, l'algorithme de sa solution est plus complexe que celui de sa contrepartie locale. Dans un horizon court terme, les deux méthodes devraient donner approximativement les mêmes résultats, tel que démontré par Augustyniak et al. (2017) et Heath et al. (2001a). Dupuis et al. (2016) ont développé une stratégie court terme de minimisation du risque quadratique global pour un fournisseur d'électricité de Nord Pool. Nous développons un algorithme pour une stratégie de minimisation locale des risques. Grâce à la formule de Taylor-Young et la fonction génératrice des moments de la distribution NIG, nous parvenons à réduire le temps de calculs. Nous appliquons l'algorithme sur les mêmes données et modèles que Dupuis et al. (2016) et comparons la performance de notre stratégie avec leurs résultats.

Mots clés : gestion du risque, marché de l'électricité, minimisation locale du risque quadratique, couverture du risque

INTRODUCTION

La libéralisation des marchés d'électricité a commencé dans les années 1980. Cette déréglementation a modifié la structure des marchés, auparavant des monopoles réglementés, en ouvrant ceux-ci à la concurrence. L'augmentation de la compétition dans les marchés a aidé à réduire les prix élevés créés par la privatisation.

Plusieurs marchés déréglementés ont vu le jour depuis, dont le marché de la Norvège en 1991. En 1996, la Suède se joint à la Norvège pour former Nord Pool. Au cours des années 1998-2000, la Finlande et le Danemark se joint aussi à Nord Pool. Ces quatre pays forment le marché nordique de Nord Pool. D'après le rapport de NordREG (2007), un peu plus de la moitié de la capacité de production d'électricité de 2006 provenait du système hydroélectrique, 30% des centrales d'énergie thermique, 13% des centrales nucléaires et le reste de diverses autres sources, incluant l'énergie éolienne.

Bien qu'il soit parfois possible de mettre en réserve la ressource nécessaire à la production d'électricité, par exemple le gaz, l'huile ou le charbon pour l'énergie thermique ou à l'aide de réservoirs d'eau pour certaines centrales hydroélectriques, il n'est pas possible de stocker l'électricité une fois produite. La difficulté de garder de l'électricité en stock pour plus tard se traduit par l'impossibilité d'utiliser la relation traditionnelle de « *cost-of-carry* » pour la tarification d'un contrat à terme puisque cette technique suppose aucun arbitrage entre les prix de ce dernier et les prix de l'actif sous-jacent. À titre de rappel, l'arbitrage se définit comme une opération financière assurée de résulter en un gain positif ou nul en profitant d'écart temporels de prix entre plusieurs actifs.

Les prix d'électricité des marchés déréglementés présentent des caractéristiques qui sont rarement observées dans d'autres commodités. Les prix ont une saisonnalité fortement basée sur la demande qui varie selon les saisons et les températures. On observe aussi beaucoup de volatilité due à la difficulté de stocker la production, la demande non-élastique ainsi que la production non-élastique des énergies renouvelables. Une panne imprévue dans une centrale aura pour conséquence qu'une autre source de production, habituellement plus chère, devra remplacer l'électri-

ité manquante. Les centrales pouvant se mettre en marche très rapidement ont des coûts de production très élevés, ainsi le prix de l'électricité subira cette hausse.

Un fournisseur d'électricité doit acheter son électricité au prix du système et la revendre à ses clients à un prix prédéterminé dans leurs contrats. Ce flux monétaire peut être séparé en deux parties : la première étant une marge de profit de référence déterminée en grande partie par le prix dans les contrats et la deuxième étant les coûts d'approvisionnement dépendants du prix du système et de la demande. Le fournisseur fait donc face au risque provenant de la volatilité du prix du marché ainsi qu'au risque de volume puisque la demande totale des clients fluctue en temps réel.

Pour couvrir ces risques, le fournisseur peut faire un investissement dans les produits dérivés sur le marché de l'électricité. Une stratégie de couverture est un investissement supplémentaire servant à couvrir partiellement ou totalement le risque financier. Ainsi, l'objectif de la stratégie dans le contexte d'un fournisseur d'électricité sera de répliquer les coûts d'approvisionnement à l'aide d'un portefeuille de couverture comprenant des actifs sans risque ainsi que des produits dérivés.

La gestion des risques est très importante pour contrôler la probabilité d'événements défavorables et réduire l'impact de ceux-ci. Black, Scholes et Merton ont été des pionniers dans la gestion des risques d'un marché complet. Dans la réalité, les marchés sont généralement incomplets dû, entre autres, aux rebalancements en temps discret, à la volatilité stochastique des prix du marché et à la complexité des facteurs qui influencent les prix.

En mathématiques financières, la problématique de couverture de risque dans les marchés incomplets a souvent été étudiée. Entre autres, la relation entre les mesures martingales et le principe d'absence d'arbitrage a été étudiée par Harrison and Kreps (1979) et les techniques de super-réplication par El Karoui and Quenez (1995) et Karatzas and Shreve (1998). Plusieurs autres méthodes ont été analysées dans la littérature mais les chercheurs ne peuvent s'entendre sur celle qui simultanément donne les meilleurs résultats et nécessite les calculs les moins complexes. Dans les marchés incomplets, il est très rare, voire impossible, d'avoir une

méthode qui satisfait ces deux propriétés.

Föllmer and Sondermann (1986) ont été les premiers à présenter l'idée de minimiser l'erreur moyenne quadratique sous une mesure martingale. Cette idée a été généralisée à une mesure arbitraire par Schweizer (1988) et est connue sous le nom de minimisation locale des risques (*local risk-minimization*). Föllmer and Schweizer (1988) ont présenté une solution simple pour déterminer une stratégie de couverture de risques locale en temps discret. De façon générale, l'idée est d'investir une somme initiale et de réduire l'erreur moyenne quadratique, c'est-à-dire les coûts incrémentaux, à chaque période d'investissement. Un dernier investissement se fait à la fin de la stratégie de telle sorte que la valeur du portefeuille est égale à la valeur terminale du droit contingent, c'est-à-dire $V_T = H_T$. Ceci équivaut à une séquence de régression des moindres carrés sur une période, implémentée de manière récursive dans le temps.

Une autre méthode est de couvrir son risque terminal et de ne pas gérer ses profits et ses pertes à chaque incrément dans le temps. Cette méthode est mieux connue sous le nom de minimisation globale des risques (*global risk-minimization*). De façon générale, l'idée est d'investir une somme initiale qui fluctuera dans le temps en fonction des positions prises sur le produit dérivé alors qu'aucun montant additionnel n'est investi en cours de stratégie.

Si le critère d'optimisation repose sur le risque quadratique, on appelle cette méthode la minimisation du risque quadratique global (*global quadratic hedging, mean-variance hedging, variance-optimal hedging*). Bouleau and Lamberton (1989) définissent la solution de cette stratégie dans le cas de marchés markoviens. Duffie and Richardson (1991) et Schweizer (1992), entre autres, démontrent les solutions dans des cas plus généraux. La solution en temps discret est présentée par Schweizer (1995).

Ces deux approches se basent sur un critère quadratique qui est critiqué dû à la pénalité symétrique des pertes et des gains. Bien que ce critère ne reflète pas comment les participants de marché évaluent leur risque, son grand avantage est sa traçabilité qui permet des solutions explicites dans plusieurs cas.

De par sa définition, la minimisation du risque quadratique global produit des ré-

sultats supérieurs à sa contrepartie locale en termes d'erreur quadratique moyenne (*Root-mean-square deviation/RMSE*). Par contre, l'algorithme permettant de trouver la stratégie globale optimale est généralement plus complexe que celui de la minimisation du risque local.

Il a toutefois été démontré par Černý and Kallsen (2009) qu'en temps discret, l'algorithme peut être aussi simple. De plus, il a été observé par Heath et al. (2001a) et Augustyniak et al. (2017) que les deux méthodes donnent approximativement les mêmes résultats pour des stratégies court terme.

Dupuis et al. (2016) ont développé une stratégie de minimisation du risque quadratique global pour un fournisseur d'électricité de Nord Pool sur une courte période de 3 semaines. L'objectif du mémoire sera d'appliquer la théorie de minimisation locale des risques sur les mêmes données afin d'évaluer si les résultats obtenus démontrent aussi que l'efficacité des deux méthodes est semblable à court terme.

Le mémoire se divise comme suit. Dans le premier chapitre, les stratégies de couverture du risque quadratique locale et globale sont définies et comparées théoriquement. Le deuxième chapitre se concentre sur les marchés d'électricité et détaille les modèles utilisés pour simuler les intrants nécessaires à la couverture de risque pour les fournisseurs d'électricité. Le chapitre 3 présente l'algorithme développé pour résoudre le problème de minimisation locale des risques spécifiques aux fournisseurs d'électricité ainsi que certaines techniques employées pour améliorer le temps d'exécution du code. Grâce à la formule de Taylor-Young et la fonction génératrice des moments de la distribution NIG, nous parvenons à réduire le temps de calculs. Le dernier chapitre présente les données sur lesquelles l'algorithme est appliqué, la mesure permettant de comparer la performance des stratégies de couverture du risque quadratique ainsi que l'analyse des résultats. Finalement, nous proposons des développements additionnels suite aux résultats obtenus.

1 GESTION DE RISQUE

Dans un marché incomplet, un droit contingent n'admet généralement pas de protection parfaite contre les risques. Les stratégies de gestion de risque peuvent être séparées en deux catégories : locale ou globale.

Le principe de super-réplication, qui consiste en une stratégie autofinancée couvrant la valeur du droit contingent H , une variable aléatoire à valeurs non-négatives \mathcal{F}_T -mesurable, a un coût généralement trop élevé pour être appliqué. À titre de rappel, une stratégie est dite autofinancée s'il y a un investissement initial et qu'il n'y a aucun ajout ou retrait de fonds pendant la durée de la stratégie de couverture.

La technique de delta-hedging est une technique locale par rapport au mouvement du sous-jacent. Elle consiste en un portefeuille pour lequel les variations imitent les variations du droit contingent lorsque son actif sous-jacent change de valeur. Par contre, dans la réalité, il est impossible de transiger en temps continu. De plus, le delta-hedging est sujet à une mauvaise spécification du modèle du droit contingent. Il est néanmoins un atout comme point de référence à la comparaison de d'autres stratégies.

D'autres stratégies locales se caractérisent par une minimisation locale du risque dans le temps. Ces stratégies sont dites myopes car elles ne tiennent pas compte des résultats sur toute la période de couverture. Parmi celles-ci, on retrouve la stratégie ayant pour objectif de réduire le risque quadratique, défini comme l'espérance du carré de l'erreur de couverture.

De leurs côtés, les stratégies globales consistent à minimiser l'erreur de couverture terminale. Ces stratégies sont autofinancées et appliquent une pénalité sur la différence entre la valeur du portefeuille et la valeur du droit contingent à la fin de la stratégie. Plusieurs fonctions de pénalité ont été étudiées dans la littérature : l'espérance du risque, le quantile du risque (VaR), la TVaR et l'espérance de l'erreur quadratique, entre autres.

Dans ce chapitre, nous présentons la théorie de la minimisation locale des risques

et résumons les concepts importants de la stratégie de minimisation du risque quadratique global. Nous terminons par une comparaison de ces deux méthodes. Or, dans un premier temps, nous devons définir certains termes.

Remarque. Les concepts de ce chapitre sont dérivés de la théorie présentée dans le chapitre 10 de Föllmer and Schied (2002). Les preuves des propositions présentées se retrouvent dans cet ouvrage.

1.1 Notation, terminologie et rappels

Nous considérons l'espace probabilisé filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \{0,1,\dots,T\}}, \mathbb{P})$, où \mathbb{P} est une mesure de probabilité sur \mathcal{F} et correspond à la mesure physique/objective/réelle, et travaillons en temps discret puisque la réalité ne nous permet pas d'investir en temps continu. Pour l'ensemble du mémoire, nous considérons seulement le cas où il n'y a qu'un seul actif risqué, bien que certains des résultats présentés fonctionnent même s'il y en a plus qu'un.

Définition 1.1. *Stratégie d'investissement*

Une stratégie d'investissement $\theta = (\theta^0, \theta^1)$ est une paire de processus stochastiques tels que

- $\theta^0 = (\theta_t^0)_{t=0,1,\dots,T}$ est un processus stochastique adapté par rapport à la filtration $(\mathcal{F}_t)_{t=0,1,\dots,T}$, c'est-à-dire θ_t^0 est \mathcal{F}_t -mesurable. Il représente le nombre d'unités détenus de l'actif sans risque ;
- $\theta^1 = (\theta_t^1)_{t=1,\dots,T}$ est un processus stochastique prévisible par rapport à la filtration $(\mathcal{F}_t)_{t=0,1,\dots,T}$, c'est-à-dire θ_t^1 est \mathcal{F}_{t-1} -mesurable. Il représente le nombre d'unités détenus de l'actif risqué et on définit $\theta_0^1 = 0$.

Définition 1.2. *Valeur du portefeuille*

La valeur du portefeuille de couverture avant le rebalancement au temps $t \in \{0, 1, \dots, T\}$ pour la stratégie θ est

$$\bar{V}_t(\theta) = \theta_t^0 B_t + \theta_t^1 \bar{X}_t$$

où

- $B = (B_t)_{t=0,1,\dots,T}$ est un processus déterministe représentant le prix de l'actif sans risque. Nous supposons que $B_t = \exp(rt)$ où r est le taux sans risque.
- $\bar{X} = (\bar{X}_t)_{t=0,1,\dots,T}$ est un processus stochastique de carré intégrable représentant le prix de l'actif risqué.

Nous rappelons qu'un processus $Y = (Y_t)_{t \geq 0}$ est dit de carré intégrable par rapport à \mathbb{P} si $\mathbb{E}[Y_t^2] < +\infty \forall t \geq 0$. On note $Y_t \in \mathcal{L}^2(\Omega, \mathcal{F}_t, \mathbb{P}) \forall t$.

Dans le contexte qui nous intéresse, nous supposons aussi que le droit contingent H est un processus de carré intégrable.

Remarque. La valeur actualisée du portefeuille au temps 0 pour la stratégie θ est

$$B_t^{-1}\bar{V}_t(\theta) = \theta_t^0 + \theta_t^1 B_t^{-1}\bar{X}_t.$$

Nous utiliserons les processus actualisés $V_t(\theta) = B_t^{-1}\bar{V}_t(\theta)$ et $X_t = B_t^{-1}\bar{X}_t$ pour la suite.

De plus, nous dénoterons par

$$\sigma_t^2 = \text{Var}[X_t - X_{t-1} | \mathcal{F}_{t-1}]$$

la variance conditionnelle et par

$$\alpha_t = \mathbb{E}[X_t - X_{t-1} | \mathcal{F}_{t-1}]$$

l'espérance conditionnelle de l'incrément des prix actualisés X pour $t \in \{1, \dots, T\}$.

Définition 1.3. Processus de gains

Le processus de gains au temps $t \in \{1, \dots, T\}$ (actualisé au temps 0) pour la stratégie θ est

$$G_t(\theta) = \sum_{k=1}^t \theta_k^1 (X_k - X_{k-1}).$$

Il représente les gains cumulatifs actualisés des rendements de l'actif risqué. On définit $G_0(\theta) = 0$.

Définition 1.4. *Processus de coûts*

Le processus de coûts au temps $t \in \{0, 1, \dots, T\}$ (actualisé au temps 0) pour la stratégie θ est

$$C_t(\theta) = V_t(\theta) - G_t(\theta).$$

Notons que $C_0(\theta) = V_0(\theta) - G_0(\theta) = V_0(\theta)$ représente l'investissement initial pour la stratégie d'investissement θ .

Définition 1.5. *Processus de coûts incrémentaux*

Le processus de coûts incrémentaux actualisé représente le flux monétaire nécessaire au temps $t \in \{1, \dots, T\}$ dans l'objectif de rebalancer le portefeuille selon la stratégie θ . Celui-ci est dénoté par $C_t(\theta) - C_{t-1}(\theta)$, ce qui est équivalent à $V_t(\theta) - V_{t-1}(\theta) - \theta_t^1(X_t - X_{t-1})$.

Remarque. Dans le cas d'une stratégie autofinancée, nous avons $C_t(\theta) = V_0(\theta) \forall t$ puisque nous n'injectons aucun argent supplémentaire au temps $t \in \{1, \dots, T\}$. Ainsi, le processus de coûts incrémentaux actualisé est nul pour $t \in \{1, \dots, T\}$, tel que démontré par

$$C_t(\theta) - C_{t-1}(\theta) = V_0(\theta) - V_0(\theta) = 0.$$

1.2 Minimisation locale du risque quadratique

L'objectif d'une stratégie de couverture locale quadratique en temps discret est de trouver la paire (θ_t^0, θ_t^1) qui minimise le carré des coûts incrémentaux au temps $t \in \{1, \dots, T\}$, c'est-à-dire

$$\operatorname{argmin}_{(\theta_t^0, \theta_t^1)} \mathbb{E}[(C_t(\theta) - C_{t-1}(\theta))^2 | \mathcal{F}_{t-1}].$$

La stratégie doit aussi satisfaire $V_T = H$ \mathbb{P} -p.s., où H représente la valeur actualisée du droit contingent que nous souhaitons répliquer avec la stratégie d'investissement. Comme la stratégie n'est pas contrainte d'être autofinancée, la valeur H peut toujours être atteinte en posant $\theta_T^0 = H - \theta_T^1 X_T$.

Dans cette section, nous dérivons la solution optimale explicite de la stratégie de minimisation locale du risque. De plus, nous donnons une condition essentielle pour l'existence de celle-ci.

Définition 1.6. *Stratégie admissible (locale)*

On dit qu'une stratégie θ est admissible (locale) pour H si $V_T = H$ et si ses processus $V(\theta)$ et $G(\theta)$ sont de carré intégrable par rapport à \mathbb{P} .

Définition 1.7. *Processus de risque quadratique local*

Le processus de risque quadratique local pour la stratégie θ est défini par

$$R_{t-1}^L(\theta) := \mathbb{E}[(C_t(\theta) - C_{t-1}(\theta))^2 | \mathcal{F}_{t-1}] \text{ pour } t \in \{1, \dots, T\}.$$

Définition 1.8. *Stratégie optimale de minimisation du risque quadratique local*

Une stratégie $\hat{\theta}$ est appelée la stratégie optimale de minimisation du risque quadratique local si, pour tout t et toute stratégie admissible θ ,

$$R_{t-1}^L(\hat{\theta}) \leq R_{t-1}^L(\theta) \text{ } \mathbb{P}\text{-p.s.}$$

Notons qu'une stratégie autofinancée minimise le carré des coûts incrémentaux mais ne satisfait généralement pas la contrainte $V_T = H$.

Lemme 1.9. *Soit X le processus stochastique de carré intégrable représentant le prix de l'actif risqué. Il existe une constante $c \in \mathbb{R}$ pour laquelle*

$$\begin{aligned} \mathbb{E}[(X_t - X_{t-1})^2 | \mathcal{F}_{t-1}] &= \text{Var}[X_t - X_{t-1} | \mathcal{F}_{t-1}] + \mathbb{E}[X_t - X_{t-1} | \mathcal{F}_{t-1}]^2 \\ &\leq c \text{Var}[X_t - X_{t-1} | \mathcal{F}_{t-1}] \\ &= c \times \sigma_t^2 \text{ } \mathbb{P}\text{-p.s. } \forall t \end{aligned}$$

Théorème 1.10. *Solution explicite de la stratégie optimale (local)*

Dans le cas où le modèle de marché ne contient qu'un actif risqué, la solution explicite à la stratégie de minimisation locale des risques est donnée par la récursion

suivante :

$$\begin{aligned}
V_T(\hat{\theta}) &:= H; \\
\hat{\theta}_t^1 &:= \frac{\text{Cov}[V_t(\hat{\theta}), X_t - X_{t-1} | \mathcal{F}_{t-1}]}{\sigma_t^2} \mathbb{1}_{\{\sigma_t^2 \neq 0\}}; \\
\hat{\theta}_t^0 &:= V_t(\hat{\theta}) - \hat{\theta}_t^1 X_t; \\
V_{t-1}(\hat{\theta}) &:= \mathbb{E}[V_t(\hat{\theta}) | \mathcal{F}_{t-1}] - \hat{\theta}_t^1 \mathbb{E}[X_t - X_{t-1} | \mathcal{F}_{t-1}]
\end{aligned}$$

pour $t \in \{1, \dots, T\}$.

Démonstration. Nous fixons t et $V_t(\theta)$ pour $t \in \{1, \dots, T\}$.

Nous souhaitons trouver $V_{t-1}(\theta)$ et θ_t^1 qui minimisent

$$R_{t-1}(\theta) = \mathbb{E}[(C_t(\theta) - C_{t-1}(\theta))^2 | \mathcal{F}_{t-1}] = \mathbb{E}[(V_t(\theta) - V_{t-1}(\theta) - \theta_t^1(X_t - X_{t-1}))^2 | \mathcal{F}_{t-1}].$$

Notons que

$$\begin{aligned}
\mathbb{E}[(C_t(\theta) - C_{t-1}(\theta))^2 | \mathcal{F}_{t-1}] &= \text{Var}[C_t(\theta) - C_{t-1}(\theta) | \mathcal{F}_{t-1}] \\
&\quad + (\mathbb{E}[C_t(\theta) - C_{t-1}(\theta) | \mathcal{F}_{t-1}])^2
\end{aligned}$$

où

$$\begin{aligned}
\text{Var}[C_t(\theta) - C_{t-1}(\theta) | \mathcal{F}_{t-1}] &= \text{Var}[V_t(\theta) - V_{t-1}(\theta) - \theta_t^1(X_t - X_{t-1}) | \mathcal{F}_{t-1}] \\
&= \text{Var}[V_t(\theta) - \theta_t^1(X_t - X_{t-1}) | \mathcal{F}_{t-1}]
\end{aligned}$$

et

$$\begin{aligned}
(\mathbb{E}[C_t(\theta) - C_{t-1}(\theta) | \mathcal{F}_{t-1}])^2 &= (\mathbb{E}[V_t(\theta) - V_{t-1}(\theta) - \theta_t^1(X_t - X_{t-1}) | \mathcal{F}_{t-1}])^2 \\
&= (\mathbb{E}[V_t(\theta) - \theta_t^1(X_t - X_{t-1}) | \mathcal{F}_{t-1}] - V_{t-1}(\theta))^2
\end{aligned}$$

car $V_{t-1}(\theta)$ est \mathcal{F}_{t-1} -mesurable.

Notons que faire varier la valeur de $V_{t-1}(\theta)$ n'influence pas $V_t(\theta)$ ni θ_t^1 puisque

que nous pouvons fixer θ_t^0 de sorte que ceux-ci ne varient pas. Ainsi, faire varier $V_{t-1}(\theta)$ influence seulement l'équation $(\mathbb{E}[C_t(\theta) - C_{t-1}(\theta)|\mathcal{F}_{t-1}])^2$.

On doit donc trouver la valeur de $V_{t-1}(\theta)$ qui minimise $(\mathbb{E}[C_t(\theta) - C_{t-1}(\theta)|\mathcal{F}_{t-1}])^2$, soit $\hat{V}_{t-1}(\theta)$:

$$\begin{aligned} 0 &= \frac{\partial}{\partial V_{t-1}} \left(\mathbb{E}[V_t(\theta) - \theta_t^1(X_t - X_{t-1})|\mathcal{F}_{t-1}] - V_{t-1}(\theta) \right)^2 \\ &= -2 (\mathbb{E}[V_t(\theta) - \theta_t^1(X_t - X_{t-1})|\mathcal{F}_{t-1}] - V_{t-1}(\theta)) \end{aligned}$$

$$\begin{aligned} \Rightarrow \hat{V}_{t-1}(\theta) &= \mathbb{E}[V_t(\theta) - \theta_t^1(X_t - X_{t-1})|\mathcal{F}_{t-1}] \\ &= \mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}] - \theta_t^1 \mathbb{E}[X_t - X_{t-1}|\mathcal{F}_{t-1}] \end{aligned}$$

Posons $U = V_t(\theta)$ et $Y = X_t - X_{t-1}$.

Nous avons donc que

$$\begin{aligned} \text{Var}[C_t(\theta) - C_{t-1}(\theta)|\mathcal{F}_{t-1}] &= \text{Var}[V_t(\theta) - \theta_t^1(X_t - X_{t-1})|\mathcal{F}_{t-1}] \\ &= \text{Var}[U - \theta_t^1 Y|\mathcal{F}_{t-1}] \\ &= \text{Var}[U|\mathcal{F}_{t-1}] + \text{Var}[-\theta_t^1 Y|\mathcal{F}_{t-1}] + 2 \text{Cov}[U, -\theta_t^1 Y|\mathcal{F}_{t-1}] \\ &= (\theta_t^1)^2 \text{Var}[Y|\mathcal{F}_{t-1}] - 2(\theta_t^1) \text{Cov}[U, Y|\mathcal{F}_{t-1}] + \text{Var}[U|\mathcal{F}_{t-1}] \end{aligned}$$

est une équation de second degré en fonction de θ_t^1 de la forme $a(\theta_t^1)^2 + b(\theta_t^1) + c$ où

- $a = \text{Var}[Y|\mathcal{F}_{t-1}]$;
- $b = -2 \text{Cov}[U, Y|\mathcal{F}_{t-1}]$; et
- $c = \text{Var}[U|\mathcal{F}_{t-1}]$.

Puisque $a > 0$; on trouve le minimum de l'équation de second degré à

$$\begin{aligned}\theta_t^1 &= \frac{-b}{2a} \\ &= \frac{\text{Cov}[U, Y | \mathcal{F}_{t-1}]}{\text{Var}[Y | \mathcal{F}_{t-1}]} \\ &= \frac{\text{Cov}[V_t(\theta), X_t - X_{t-1} | \mathcal{F}_{t-1}]}{\text{Var}[X_t - X_{t-1} | \mathcal{F}_{t-1}]} \\ &= \frac{\text{Cov}[V_t(\theta), X_t - X_{t-1} | \mathcal{F}_{t-1}]}{\sigma_t^2}.\end{aligned}$$

On pose ensuite $\theta_t^0 = V_t(\theta) - \theta_t^1 X_t$ et on obtient la paire (θ_t^0, θ_t^1) .

Nous devons maintenant prouver que θ est bien une stratégie admissible. Par le lemme précédent, nous avons $\mathbb{E}[(X_t - X_{t-1})^2 | \mathcal{F}_{t-1}] \leq c \text{Var}[X_t - X_{t-1} | \mathcal{F}_{t-1}]$, où $c \in \mathbb{R}$.

Soit $t \in \{1, \dots, T\}$.

$$\begin{aligned}\mathbb{E}[(\theta_t^1 (X_t - X_{t-1}))^2] &= \mathbb{E} \left[\frac{(\text{Cov}[V_t(\theta), X_t - X_{t-1} | \mathcal{F}_{t-1}])^2}{(\sigma_t^2)^2} \mathbb{E}[(X_t - X_{t-1})^2 | \mathcal{F}_{t-1}] \mathbb{1}_{\{\sigma_t^2 \neq 0\}} \right] \\ &\leq c \mathbb{E} \left[\frac{(\text{Cov}[V_t(\theta), X_t - X_{t-1} | \mathcal{F}_{t-1}])^2}{\sigma_t^2} \mathbb{1}_{\{\sigma_t^2 \neq 0\}} \right] \\ &\leq c \mathbb{E} [\text{Var}[V_t(\theta) | \mathcal{F}_{t-1}]] \\ &= c (\mathbb{E}[V_t(\theta)^2 | \mathcal{F}_{t-1}] - \mathbb{E}[V_t(\theta) | \mathcal{F}_{t-1}]^2) \\ &\leq c \mathbb{E}[V_t(\theta)^2 | \mathcal{F}_{t-1}].\end{aligned}$$

La dernière espérance est finie si V_t est de carré intégrable. Dans ce cas, $\theta_t^1 (X_t - X_{t-1})$ est aussi de carré intégrable.

Nous utilisons le raisonnement par récurrence pour compléter la preuve puisque

V_{t-1} est aussi de carré intégrable :

$$\begin{aligned}
& \mathbb{E}[V_{t-1}(\theta)^2] \\
&= \mathbb{E}[(\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}] - \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}])^2] \\
&= \mathbb{E}[\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}]^2 + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}]^2 - 2\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}]\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}]] \\
&\leq \mathbb{E}[\mathbb{E}[V_t(\theta)^2|\mathcal{F}_{t-1}] + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2|\mathcal{F}_{t-1}] - 2\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}]\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}]] \\
&= \mathbb{E}[V_t(\theta)^2] + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2] - 2\mathbb{E}[\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}]\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}]] \\
&\leq \mathbb{E}[V_t(\theta)^2] + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2] + 2\mathbb{E}[|\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}]\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}]|] \\
&\leq \mathbb{E}[V_t(\theta)^2] + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2] + 2\mathbb{E}[|\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}]|^2]\mathbb{E}[|\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}]|^2] \\
&= \mathbb{E}[V_t(\theta)^2] + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2] + 2\mathbb{E}[\mathbb{E}[V_t(\theta)|\mathcal{F}_{t-1}]^2]\mathbb{E}[\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))|\mathcal{F}_{t-1}]^2] \\
&\leq \mathbb{E}[V_t(\theta)^2] + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2] + 2\mathbb{E}[\mathbb{E}[V_t(\theta)^2|\mathcal{F}_{t-1}]]\mathbb{E}[\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2|\mathcal{F}_{t-1}]] \\
&= \mathbb{E}[V_t(\theta)^2] + \mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2] + 2\mathbb{E}[V_t(\theta)^2]\mathbb{E}[(\theta_t^1(X_t - X_{t-1}))^2] \\
&< \infty
\end{aligned}$$

□

Remarque. Toute autre stratégie optimale de minimisation du risque quadratique local coïncide avec la stratégie définie dans le théorème 1.10 à l'exception de modifications de θ_t sur l'ensemble $\{\sigma_t^2 = 0\}$.

Une stratégie de minimisation du risque quadratique local n'est pas autofinancée comme l'est sa contrepartie globale. Par contre, celle-ci a une propriété similaire. Pour la définir, nous devons d'abord rappeler le concept de martingale.

Définition 1.11. Martingale

Un processus stochastique $M = (M_t)_{t=0,1,\dots,T}$ sur l'espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \{0,1,\dots,T\}}, \mathbb{P})$ est une \mathbb{P} -martingale si

1. M est adapté ;
2. $\mathbb{E}[|M_t|] < \infty$ pour tout t ; et
3. $M_s = \mathbb{E}[M_t|\mathcal{F}_s]$ pour $0 \leq s \leq t \leq T$.

Lemme 1.12. *Pour deux martingales de carré intégrable M et N , les deux énoncés suivants sont équivalents :*

1. M et N sont fortement orthogonales.
2. Le produit MN est une martingale.

Démonstration. Puisque M et N sont des martingales,

$$\begin{aligned}
\text{Cov}[M_t - M_{t-1}, N_t - N_{t-1} | \mathcal{F}_{t-1}] &= \mathbb{E}[(M_t - M_{t-1})(N_t - N_{t-1}) | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[M_t N_t | \mathcal{F}_{t-1}] + \mathbb{E}[M_{t-1} N_{t-1} | \mathcal{F}_{t-1}] - \mathbb{E}[M_t N_{t-1} | \mathcal{F}_{t-1}] \\
&\quad - \mathbb{E}[M_{t-1} N_t | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[M_t N_t | \mathcal{F}_{t-1}] + M_{t-1} N_{t-1} - N_{t-1} \mathbb{E}[M_t | \mathcal{F}_{t-1}] \\
&\quad - M_{t-1} \mathbb{E}[N_t | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[M_t N_t | \mathcal{F}_{t-1}] + M_{t-1} N_{t-1} - N_{t-1} M_{t-1} - M_{t-1} N_{t-1} \\
&= \mathbb{E}[M_t N_t | \mathcal{F}_{t-1}] - M_{t-1} N_{t-1}.
\end{aligned}$$

Si M et N sont fortement orthogonales, alors la covariance ci-dessus est égale à 0. Cette expression est égale à 0 si et seulement si le produit MN est une martingale. \square

Définition 1.13. *Stratégie autofinancée en moyenne*

Une stratégie θ est dite autofinancée en moyenne si son processus de coûts est une martingale, c'est-à-dire si $\forall t$

$$\mathbb{E}[C_t(\theta) - C_{t-1}(\theta) | \mathcal{F}_{t-1}] = 0.$$

De par la solution explicite donnée dans le théorème 1.10, nous avons que la stratégie optimale de minimisation du risque quadratique local est autofinancée en moyenne car

$$\begin{aligned}
\mathbb{E}[C_t(\theta) - C_{t-1}(\theta) | \mathcal{F}_{t-1}] &= \mathbb{E}[V_t(\theta) - V_{t-1}(\theta) - \theta_t^1 (X_t - X_{t-1}) | \mathcal{F}_{t-1}] \\
&= (\mathbb{E}[V_t(\theta) | \mathcal{F}_{t-1}] - \theta_t^1 \mathbb{E}[X_t - X_{t-1} | \mathcal{F}_{t-1}]) - V_{t-1}(\theta) \\
&= 0.
\end{aligned}$$

Dernièrement, nous donnons une condition essentielle à l'existence d'une stratégie

optimale.

Proposition 1.14. *Il existe une stratégie optimale de minimisation du risque quadratique local si et seulement si H admet une décomposition orthogonale par rapport à X de la forme*

$$H = c + \sum_{t=1}^T \theta_t^1 (X_t - X_{t-1}) + L_T \text{ } \mathbb{P}\text{-p.s.},$$

où c est une constante, θ^1 est un processus prévisible tel que $\theta_t^1 (X_t - X_{t-1})$ est de carré intégrable et où L est une \mathbb{P} -martingale de carré intégrable fortement orthogonale à X et qui satisfait $L_0 = 0$.

Dans ce cas, la stratégie optimale de minimisation du risque quadratique local $\hat{\theta} = (\hat{\theta}^0, \hat{\theta}^1)$ est donnée par $\hat{\theta}^1 = \theta^1$ et par le processus adapté $\hat{\theta}^0$ défini par $\hat{\theta}_0^0 = c$ et

$$\hat{\theta}_t^0 = c + \sum_{k=1}^t \theta_k^1 (X_k - X_{k-1}) + L_t - \theta_t^1 X_t, \text{ pour } t \in \{1, \dots, T\}.$$

De plus, cette décomposition est unique dans le sens où L est uniquement déterminée.

1.3 Minimisation globale du risque quadratique

Contrairement à la couverture locale quadratique qui minimise les risques à chaque période d'investissement, la couverture globale quadratique minimise l'erreur de couverture $\Lambda_T = H - V_T$. De plus, uniquement les stratégies autofinancées sont considérées.

Dans cette section, nous définissons quelques termes et donnons une solution explicite dans le cas où le ratio σ_t^2/α_t^2 est déterministe pour chaque t .

Définition 1.15. *Processus de risque quadratique global*

Un processus de risque quadratique global pour la stratégie admissible (locale) $\theta = (\theta^0, \theta^1)$ est défini par

$$R^G(\theta) := \mathbb{E}[(C_T(\theta) - C_0(\theta))^2] := \mathbb{E}[(H - V_0(\theta) - G_T(\theta))^2].$$

Remarque. Puisque $V_0(\theta) = \theta_0^0$ et $G_T(\theta)$ dépend seulement de θ^1 , nous avons

$$R^G(\theta^0, \theta^1) \equiv R^G(V_0, \theta^1).$$

Définition 1.16. *Stratégie admissible (globale)*

La paire (V_0^*, θ^*) est une stratégie admissible (globale) pour le droit contingent actualisé H si θ^* est un processus prévisible représentant le nombre d'unités détenus de l'actif risqué et si $G_t(\theta^*)$ est de carré intégrable. On note \mathcal{S} l'ensemble des processus prévisibles θ^* satisfaisant ceci.

L'objectif d'une stratégie de couverture globale quadratique en temps discret est de trouver la stratégie autofinancée (V_0^*, θ^*) qui minimise le carré de l'erreur de réplication au temps T , c'est-à-dire

$$\operatorname{argmin}_{(V_0^*, \theta^*)} \mathbb{E}[(H - V_T(\theta^*))^2] = \operatorname{argmin}_{(V_0^*, \theta^*)} \mathbb{E}[(H - V_0^* - G_T(\theta^*))^2].$$

Définition 1.17. *Stratégie optimale de minimisation du risque quadratique global*

La paire (V_0^*, θ^*) est une stratégie optimale de minimisation du risque quadratique global pour le droit contingent actualisé H si

$$R^G(V_0^*, \theta^*) \leq R^G(V_0, \theta)$$

pour toute stratégie admissible (V_0, θ) où $\theta \in \mathcal{S}$.

Dans le cas où le ratio σ_t^2/α_t^2 est déterministe pour chaque t , la solution de la stratégie optimale de minimisation du risque quadratique global peut être définie en terme de la solution explicite de la stratégie optimale de minimisation du risque quadratique local tel que démontré par le théorème suivant.

Proposition 1.18. *Solution explicite de la stratégie optimale (global)*

Supposons que σ_t^2/α_t^2 est déterministe pour chaque t et que $\theta = (\theta^0, \theta^1)$ est la solution optimale de minimisation du risque quadratique local définie dans le théorème 1.10. Alors $V_0^* = V_0$ et

$$\theta_t^* = \theta_t^1 + \frac{\alpha_t}{\sigma_t^2 + \alpha_t^2} (V_{t-1} - V_0 - G_{t-1}(\theta^*)), \text{ pour } t = 1, \dots, T \quad (1.1)$$

définissent une stratégie optimale de minimisation du risque quadratique global (V_0^*, θ^*) .

De plus, le risque quadratique global se quantifie par

$$\mathbb{E}[(H - V_0^* - G_T(\theta^*))^2] = \sum_{t=1}^T \gamma_t \mathbb{E}[(C_t - C_{t-1})^2] \quad (1.2)$$

où C est le processus de coûts de θ et

$$\gamma_t = \prod_{k=t+1}^T \frac{\sigma_k^2}{\sigma_k^2 + \alpha_k^2}. \quad (1.3)$$

1.4 Comparaison des stratégies de couverture du risque quadratique

Dans cette section, nous comparons les stratégies de couverture du risque quadratique local et global et présentons une troisième façon de quantifier le risque quadratique. Nous devons d'abord définir la notion de mesure martingale.

Définition 1.19. *Mesure martingale*

Une mesure de probabilité \mathbb{P} sur (Ω, \mathcal{F}) est une mesure martingale si le processus de prix actualisé X est une \mathbb{P} -martingale.

Définition 1.20. *Mesure martingale équivalente*

Une mesure martingale \mathbb{P}^* est appelée mesure martingale équivalente, ou mesure de risque-neutre, si elle est équivalente à \mathbb{P} sur \mathcal{F} , c'est-à-dire si les deux mesures

définissent les mêmes ensembles négligeables. L'ensemble de toutes les mesures martingales équivalentes est noté \mathcal{P} .

Puisqu'un marché est sans arbitrage si et seulement si l'ensemble \mathcal{P} est non-nul, nous savons qu'il existe une mesure de probabilité équivalente $\mathbb{P}^* \approx \mathbb{P}$ telle que X est une martingale sous \mathbb{P}^* et une semimartingale sous \mathbb{P} . Un modèle de marché sans arbitrage est dit complet si et seulement si tous les droits contingents sont accessibles, c'est-à-dire si $\forall H$, il existe une stratégie autofinancée θ pour laquelle $\mathbb{P}(V_T(\theta) = H) = 1$. Ainsi, dans un marché complet sans arbitrage, un droit contingent H peut être parfaitement répliqué par une stratégie d'investissement dynamique autofinancée consistant à investir dans l'actif risqué X et un actif sans risque. Dans ce cas, le risque à priori peut être réduit à 0.

Dans un marché incomplet, un droit contingent n'admet généralement pas de protection parfaite contre les risques, il y aura donc un risque intrinsèque à réduire. Dans le cas où \mathbb{P} est une mesure martingale, Föllmer and Schied (2002) démontrent que si une stratégie optimale locale $\theta = (\theta^0, \theta^1)$ existe alors $(V_0^*, \theta^*) = (\theta_0^0, \theta^1)$ est la stratégie optimale de minimisation du risque quadratique global. Dans le cas où \mathbb{P} n'est pas une mesure martingale, il pourrait être intéressant d'utiliser une transformation de Girsanov pour travailler avec le cas où X est une martingale. Par contre, comme le marché est incomplet, la mesure martingale équivalente n'est pas unique et il est délicat de choisir la mesure martingale adéquate $\mathbb{P}^* \approx \mathbb{P}$ pour déterminer la stratégie optimale en terme de \mathbb{P}^* .

La propriété martingale du processus de coûts C d'une stratégie de minimisation du risque quadratique local permet de définir un cas spécifique où la stratégie globale est supérieure à la stratégie locale en terme d'erreur terminale.

Lemme 1.21. *La propriété martingale du processus de coûts C d'une stratégie de minimisation du risque quadratique local implique*

$$\mathbb{E}[(C_T(\theta) - C_0(\theta))^2] = \sum_{t=1}^T \mathbb{E}[(C_t(\theta) - C_{t-1}(\theta))^2].$$

Démonstration. À l'aide du lemme 1.12, nous obtenons

$$\begin{aligned}
\mathbb{E}[(C_T - C_0)^2] &= \mathbb{E} \left[\left(\sum_{t=1}^T (C_t - C_{t-1}) \right)^2 \right] \\
&= \mathbb{E} \left[\sum_{t=1}^T (C_t - C_{t-1})^2 + 2 \sum_{s < t} (C_s - C_{s-1}) (C_t - C_{t-1}) \right] \\
&= \sum_{t=1}^T \mathbb{E}[(C_t - C_{t-1})^2] + 2 \sum_{s < t} \mathbb{E}[(C_s - C_{s-1}) (C_t - C_{t-1})] \\
&= \sum_{t=1}^T \mathbb{E}[(C_t - C_{t-1})^2] + 2 \sum_{s < t} (\mathbb{E}[C_s C_t] - \mathbb{E}[(C_s C_{t-1})] - \mathbb{E}[C_{s-1} C_t] + \mathbb{E}[C_{s-1} C_{t-1}]) \\
&= \sum_{t=1}^T \mathbb{E}[(C_t - C_{t-1})^2] \\
&\quad + 2 \sum_{s < t} (\mathbb{E}[\mathbb{E}[C_s C_t | \mathcal{F}_s]] - \mathbb{E}[\mathbb{E}[C_s C_{t-1} | \mathcal{F}_s]] - \mathbb{E}[\mathbb{E}[C_{s-1} C_t | \mathcal{F}_{s-1}]] + \mathbb{E}[\mathbb{E}[C_{s-1} C_{t-1} | \mathcal{F}_{s-1}]]) \\
&= \sum_{t=1}^T \mathbb{E}[(C_t - C_{t-1})^2] + 2 \sum_{s < t} (\mathbb{E}[C_s^2] - \mathbb{E}[C_s^2] - \mathbb{E}[C_{s-1}^2] + \mathbb{E}[C_{s-1}^2]) \\
&= \sum_{t=1}^T \mathbb{E}[(C_t - C_{t-1})^2].
\end{aligned}$$

□

Dans le cas où le ratio σ_t^2/α_t^2 est déterministe pour chaque t , si le processus de coûts de la stratégie optimale de minimisation du risque quadratique local $C(\theta)$ n'est pas équivalent à C_0 et si, tel que défini par l'équation 1.3, γ_1 est plus petit que 1, alors par le lemme précédent et par l'équation 1.2, le risque quadratique local terminal de la stratégie (θ^0, θ^1) doit être supérieur au risque quadratique

global de la stratégie (V_0^*, θ^*) .

$$\begin{aligned} \mathbb{E}[(C_T(\theta) - C_0(\theta))^2] &= \sum_{t=1}^T \mathbb{E}[(C_t(\theta) - C_{t-1}(\theta))^2] \\ &> \sum_{t=1}^T \gamma_t \mathbb{E}[(C_t(\theta) - C_{t-1}(\theta))^2] \\ &= \mathbb{E}[(H - V_0^* - G_T(\theta^*))^2]. \end{aligned}$$

Si \mathbb{P} n'est pas une mesure martingale, alors $\alpha_t = \mathbb{E}[X_t - X_{t-1} | \mathcal{F}_{t-1}] \neq 0$ pour au moins un t et $\gamma_1 > 1$. Donc la stratégie optimale de minimisation du risque quadratique global (V_0^*, θ^*) , donnée par l'équation 1.1, n'est pas équivalente à la stratégie optimale de minimisation du risque quadratique local (θ^0, θ^1) .

On peut minimiser le risque quadratique d'une stratégie de manière locale ou globale. Une troisième manière est de calculer le risque conditionnel restant.

Définition 1.22. *Risque conditionnel restant*

Le risque conditionnel restant d'une stratégie admissible (locale) θ est défini par le processus

$$R_t^{rcr}(\theta) = \mathbb{E}[(C_T(\theta) - C_t(\theta))^2 | \mathcal{F}_t] \text{ pour } t \in \{0, 1, \dots, T\}.$$

Définition 1.23. *Minimisation du risque conditionnel restant*

Une stratégie admissible θ minimise le risque conditionnel restant si

$$R_t^{rcr}(\theta) \leq R_t^{rcr}(\hat{\theta}) \text{ } \mathbb{P}\text{-p.s.}$$

pour tout t et pour toute stratégie $\hat{\theta} = (\hat{\theta}^0, \hat{\theta}^1)$ qui a les mêmes positions que la stratégie θ du temps 0 au temps t .

Dans le cas d'une mesure martingale, minimiser le risque conditionnel restant est équivalent à minimiser le risque quadratique local. Par contre, il n'existe pas toujours de stratégie $\hat{\theta}$ minimisant le risque conditionnel restant puisque les deux contraintes suivantes ne sont généralement pas compatibles :

- La minimalité de $R_0^{rcr}(\hat{\theta}) = \mathbb{E}[(C_T - C_0)^2] = R^G(\hat{\theta})$ requiert $\hat{\theta}^1 = \theta^*$ par la définition de stratégie globale optimale ;
- La minimalité de $R_{T-1}^{rcr}(\hat{\theta}) = \mathbb{E}[(C_T - C_{T-1})^2 | \mathcal{F}_{T-1}] = R_{T-1}^L(\hat{\theta})$ requiert $\hat{\theta}_T^1 = \theta_T^1$ par la définition de stratégie locale optimale.

La technique de réduction du risque quadratique global se fait sous une mesure risque-neutre \mathbb{P}^* qui ne coïncide pas nécessairement avec la mesure objective \mathbb{P} . Or, ceci implique qu'une stratégie optimale globale est toujours définie alors que ce n'est pas le cas pour une stratégie optimale locale. Il a d'ailleurs été démontré par Schweizer (2001) que dans le cas où X n'est pas une martingale, un droit contingent H n'admet généralement pas de stratégie θ optimale de minimisation du risque quadratique local telle que $V_T(\theta) = H$ \mathbb{P} -p.s.

Dans le cas d'options modélisées sous GARCH, il a été démontré par Augustyniak et al. (2017) que la stratégie globale présente une amélioration par rapport à la stratégie locale plus la période de couverture est grande :

1. Pas d'améliorations notables pour les contrats court terme de 3 mois ;
2. Progrès modestes pour les contrats moyen terme de 1 an ; et
3. Plus-value important pour les contrats long terme de 3 ans.

Sur de très longues périodes de couverture, il a été démontré par Heath et al. (2001b) que, pour certains modèles avec volatilité stochastique spécifiques, l'erreur quadratique du risque global tend vers 0 alors que l'erreur quadratique du risque local tend vers une valeur fixe plus grande que 0. Ils mentionnent aussi que le temps de calculs est environ 20% plus élevé dans le cas de la stratégie globale versus sa contrepartie locale.

2 MARCHÉS D'ÉLECTRICITÉ ET MODÈLES

Dans ce chapitre, nous présentons certains concepts des marchés d'électricité déréglementés et nous discutons ensuite de la modélisation des intrants à l'algorithme de couverture locale de risque qui sera présenté dans le prochain chapitre. Plus particulièrement, nous présentons le modèle pour la demande d'électricité ainsi que le modèle illustrant les log-rendements des contrats à terme. Nous terminons avec la modélisation de la dépendance entre la demande et les log-rendements.

Dans la majorité des marchés d'électricité déréglementés, il existe un prix « *day-ahead* » et un prix « *real-time* ». Le premier est défini pour chaque heure de la journée d'opération comme l'intersection des courbes d'offre et de demande reçues par le marché le jour précédent. Le second est défini à chaque 5 minutes pendant la journée d'opération selon les conditions réelles du marché.

Les prix d'électricité comportent habituellement un facteur de congestion variant selon la demande et la capacité des lignes électriques. Prenons, par exemple, un système électrique très simple avec seulement deux zones tel qu'illustré dans la Figure 2.1. Supposons une ligne électrique ayant une capacité de 100MW du point A au point B, un générateur au point A pouvant produire 150MW à un prix moins élevé qu'un générateur au point B pouvant produire 200MW et que la demande, située au point B, est de 200MW. Comme la demande dépasse la capacité de la ligne électrique, le fournisseur d'électricité achètera 100MW au générateur A et 100MW au générateur B, le tout au prix du générateur B. Ainsi, la composante congestion du prix payé par le fournisseur aura la valeur de la différence entre le prix de la zone B et celui de la zone A.

Le fournisseur de marché prévoit la demande totale de ses clients dans le marché « *day-ahead* » mais devra acheter la différence de volume provenant des fluctuations en temps réel dans le marché « *real-time* ».

Tout comme Dupuis et al. (2016), nous supposons que la demande totale est achetée dans le marché « *day-ahead* » au prix du système sans congestion et que la stratégie de couverture de risque a pour objectif de couvrir le risque hebdo-

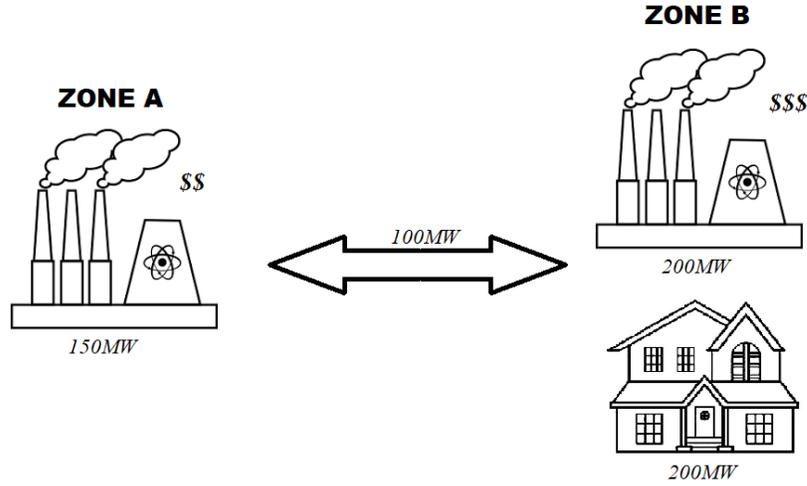


FIGURE 2.1 – Marché d’électricité simpliste

madaire d’un fournisseur d’électricité en utilisant des contrats à terme jusqu’à trois semaines d’échéance. Nous référons à Dupuis et al. (2016) pour les détails techniques des données mais certaines notations sont tout de même nécessaires.

La demande d’électricité hebdomadaire pour la semaine t est donnée par

$$L_t = \sum_{d=1}^7 \sum_{h=1}^{24} L_{t,d,h}, \quad (2.1)$$

où $L_{t,d,h}$ représente la demande à l’heure h du jour d pour la semaine t et où la semaine est définie du lundi au dimanche.

Le prix du système de Nord Pool sans congestion sera le prix spot. Ce prix horaire pour l’heure h du jour d pour la semaine t est dénoté par $S_{t,d,h}$. Le prix hebdomadaire moyen payé par le fournisseur d’électricité pour la semaine t est donné par

$$S_t^* = \frac{\sum_{d=1}^7 \sum_{h=1}^{24} L_{t,d,h} S_{t,d,h}}{L_t}. \quad (2.2)$$

Soit Π_t le prix que les clients du fournisseur paient pour l’électricité, prédéterminé

dans leurs contrats. Alors le flux monétaire du fournisseur pour la semaine t sera $L_t(\Pi_t - S_t^*)$ si aucune couverture de risque n'est implémentée. Ce flux monétaire admet un risque par rapport à la variabilité de la demande L_t mais aussi par rapport à celle du prix S_t^* .

Notons que le prix hebdomadaire moyen payé par le fournisseur diffère du prix hebdomadaire moyen

$$S_t = \frac{\sum_{d=1}^7 \sum_{h=1}^{24} L_{t,d,h} S_{t,d,h}}{7 \times 24}. \quad (2.3)$$

Or, les contrats à terme hebdomadaires reposent sur le prix S_t plutôt que sur le prix S_t^* . Notons $F_{t_0,t}$ le prix au temps t_0 du contrat à terme d'échéance t .

On peut donc séparer le flux monétaire en deux composantes :

$$\begin{aligned} L_t(\Pi_t - S_t^*) &= L_t(\Pi_t - S_t^* - \frac{S_t^*}{S_t} F_{t_0,t} + \frac{S_t^*}{S_t} F_{t_0,t}) \\ &= L_t(\Pi_t - \frac{S_t^*}{S_t} F_{t_0,t}) - L_t(S_t^* - \frac{S_t^*}{S_t} F_{t_0,t}) \\ &= L_t(\Pi_t - \frac{S_t^*}{S_t} F_{t_0,t}) - L_t \frac{S_t^*}{S_t} (S_t - F_{t_0,t}). \end{aligned}$$

Comme la première partie peut être gérée par la détermination adéquate du prix du contrat Π_t , la stratégie de couverture s'appliquera seulement à la deuxième partie où nous définissons

$$\mathcal{L}_t = L_t \frac{S_t^*}{S_t} \quad (2.4)$$

la demande d'électricité pondérée par les prix.

Remarque. Étant donné que nous avons présenté la théorie sous la forme de processus actualisés, nous utiliserons l'équivalence $X_t = B_t^{-1} F_{t,T}$ pour le prix actualisé du contrat à terme d'échéance T .

L'objectif sera de construire un portefeuille d'investissement composé de l'actif sans risque B et de contrats à terme X , dont la valeur actualisée sera donnée par V . Cette stratégie d'investissement tâchera de répliquer le risque $\mathcal{L}_T(X_T - X_{t_0})$

pour la semaine T .

2.1 Demande d'électricité pondérée par les prix

La demande d'électricité pondérée par les prix montre une tendance saisonnière dans la moyenne. Ceci n'est pas inhabituel puisque la demande est grandement corrélée à la météo. La température dans les pays nordiques est sous la barre des 0°C pendant l'hiver alors le chauffage est grandement requis. Par contre, la température douce en été ne nécessite pas la climatisation selon Vehvilainen and Pyykkonen (2005). Ces particularités entraînent une plus grande demande d'électricité dans les mois hivernaux seulement.

Plusieurs outils de prévision sont disponibles pour modéliser des séries de données qui fluctuent. Ceux-ci incluent les méthodes classiques, telles que les régressions et les modèles ARIMA, et les méthodes d'intelligence artificielle, telles que les réseaux de neurones. Lorsqu'appliquées sur les données de demande d'électricité du marché de l'Espagne, les séries de Fourier performant mieux comparativement aux réseaux de neurones ou à un modèle ARIMA, tel qu'observé par Gonzalez-Romera et al. (2008). Comme les données de ce marché sont aussi périodiques sur une base annuelle, une série de Fourier a été choisie par Dupuis et al. (2016) comme modèle pour prévoir la tendance saisonnière de la demande d'électricité dans le marché Nord Pool.

À partir du modèle et des paramètres estimés par Dupuis et al. (2016), nous avons pu répliquer les graphiques représentant la demande et sa tendance saisonnière dans la Figure 2.2 ainsi que la volatilité de la demande et sa tendance saisonnière dans la Figure 2.3. Ceci nous a permis d'illustrer l'adéquation du modèle de demande et de s'assurer que nous travaillions bien avec les mêmes données.

Dans la Figure 2.2, on note de l'autocorrélation lorsque la demande d'électricité s'éloigne de sa tendance habituelle, d'où le choix d'un modèle autorégressif AR(1) pour modéliser la demande d'électricité désaisonnalisée. Par exemple, autour de janvier 2008 et janvier 2010, on observe que quelques semaines se sont écoulées avant que la demande retourne au niveau de sa tendance.

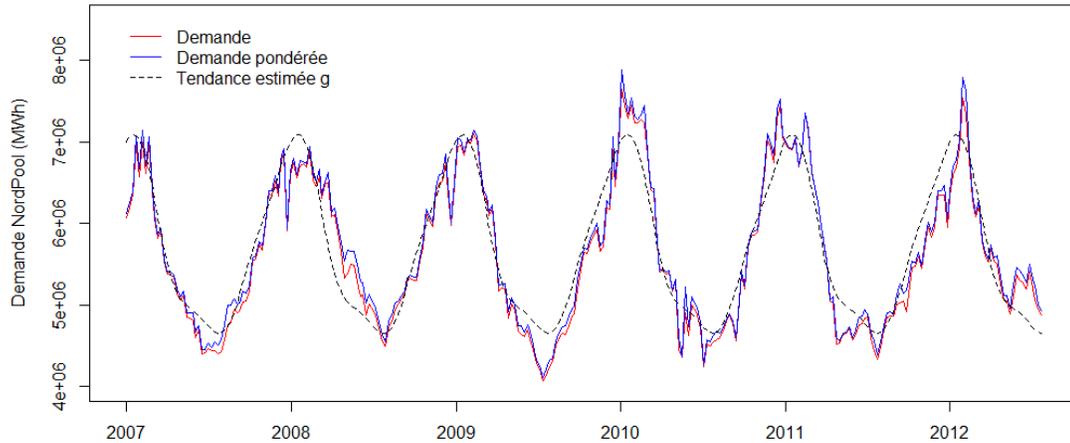


FIGURE 2.2 – Demande hebdomadaire et tendance estimée g

Définition 2.1. *Modèle autorégressif d'ordre 1*

Le modèle autorégressif d'ordre 1, dénoté AR(1), suppose que la série chronologique $(y_t)_{\{t=0,1,\dots\}}$ est stationnaire. Il implique que la valeur de y au temps t est une fonction linéaire de la valeur y au temps $t - 1$. On a donc

$$y_t = \mu + \gamma y_{t-1} + \epsilon_t$$

où μ est la valeur de la moyenne stationnaire long-terme de y_t et les ϵ_t sont indépendants de y_t et suivent une loi normale de moyenne 0 et de variance σ_ϵ^2 .

On peut observer dans la Figure 2.3 qu'une tendance saisonnière existe aussi dans la variance, alors qu'elle est plus grande en hiver, particulièrement en début de saison, et plus basse à la fin de l'été. Ainsi, une autre série de Fourier a été utilisée pour modéliser la variance. Il n'est pas mentionné par Dupuis et al. (2016) si un autre modèle a été envisagé et testé. Notez que la série de Fourier modélise le logarithme de la variance, évitant ainsi de traiter les cas où une valeur négative serait obtenue.

Le modèle ainsi proposé par Dupuis et al. (2016) pour \mathcal{L}_t , la demande d'électricité

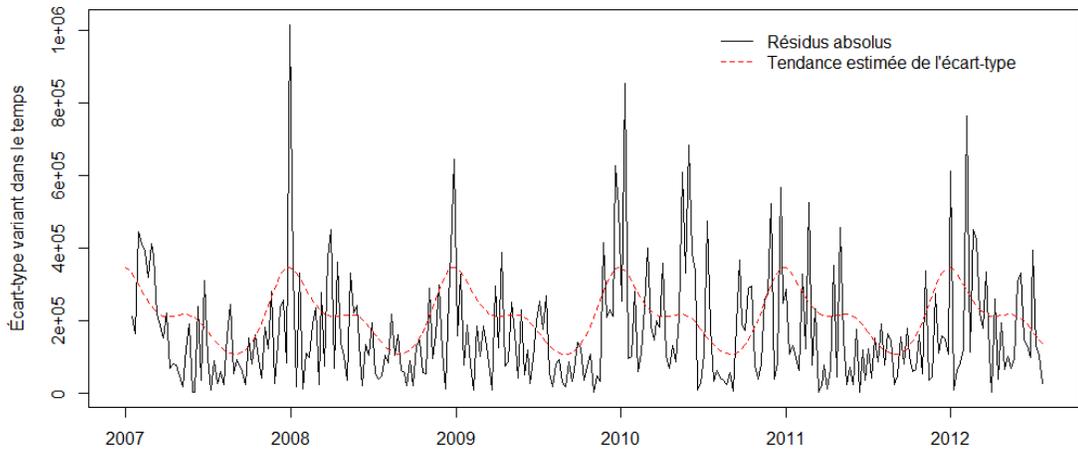


FIGURE 2.3 – Comparaison entre les résidus réalisés absolus $\sqrt{\hat{v}(t)}|\hat{\epsilon}_t|$ et la tendance estimée \sqrt{v}

hebdomadaire pondérée sur les prix, est le suivant :

$$\begin{aligned}\mathcal{L}_t - g(t) &= \gamma(\mathcal{L}_{t-1} - g(t-1)) + \sqrt{v(t)}\epsilon_t, \\ g(t) &= \beta_0 + \sum_{j=1}^P \beta_j \sin\left(\frac{3\pi}{2} + \frac{2\pi jt}{365.25/7}\right) + \sum_{j=1}^P \beta_{j+P} \cos\left(\frac{3\pi}{2} + \frac{2\pi jt}{365.25/7}\right), \\ \log v(t) &= \alpha_0 + \sum_{j=1}^Q \alpha_j \sin\left(\frac{3\pi}{2} + \frac{2\pi jt}{365.25/7}\right) + \sum_{j=1}^Q \alpha_{j+Q} \cos\left(\frac{3\pi}{2} + \frac{2\pi jt}{365.25/7}\right),\end{aligned}$$

où $g(t)$ représente la fonction de tendance saisonnière et donc $\mathcal{L}_t - g(t)$ la demande désaisonnalisée, $v(t)$ représente la variance de cette dernière et ϵ est un bruit blanc Gaussien standard.

Notons que par les équivalences

$$\sin(A + B) = \sin(A) \cos(B) + \cos(A) \sin(B),$$

$$\cos(A + B) = \cos(A) \cos(B) - \sin(A) \sin(B),$$

nous obtenons

$$\begin{aligned}\sin\left(\frac{3\pi}{2} + \frac{2\pi jt}{365.25/7}\right) &= -\cos\left(\frac{2\pi jt}{365.25/7}\right), \\ \cos\left(\frac{3\pi}{2} + \frac{2\pi jt}{365.25/7}\right) &= \sin\left(\frac{2\pi jt}{365.25/7}\right).\end{aligned}$$

Ainsi, le modèle équivaut à

$$\begin{aligned}\mathcal{L}_t - g(t) &= \gamma(\mathcal{L}_{t-1} - g(t-1)) + \sqrt{v(t)}\epsilon_t, \\ g(t) &= \beta_0 - \sum_{j=1}^P \beta_j \cos\left(\frac{2\pi jt}{365.25/7}\right) + \sum_{j=1}^P \beta_{j+P} \sin\left(\frac{2\pi jt}{365.25/7}\right), \\ \log v(t) &= \alpha_0 - \sum_{j=1}^Q \alpha_j \cos\left(\frac{2\pi jt}{365.25/7}\right) + \sum_{j=1}^Q \alpha_{j+Q} \sin\left(\frac{2\pi jt}{365.25/7}\right).\end{aligned}$$

2.2 Contrats à terme sur l'électricité

Il y a deux approches principales à la modélisation des prix de produits dérivés. La première consiste à modéliser les prix de l'actif sous-jacent pour ensuite dériver le prix du contrat à terme à partir des flux monétaires actualisés. Cette technique est connue sous le nom d'approche risque-neutre et l'existence d'un prix dépend de l'hypothèse de non-arbitrage. Puisque l'électricité n'est pas un actif facilement stockable et que le marché n'est pas ouvert au grand public, donc illiquide, le marché n'est pas complet. On peut donc se retrouver avec plusieurs mesures de probabilité risque-neutre équivalentes et ainsi obtenir des prix différents pour le même contrat à terme si on se base uniquement sur l'hypothèse de non-arbitrage. On peut spécifier davantage, sans pour autant identifier avec exactitude, la mesure équivalente en connaissant les caractéristiques de la prime de risque. Il est ainsi débatable que cette approche ne soit pas la plus appropriée. Néanmoins, Benth et al. (2008) l'ont tout de même étudiée avec un modèle de forme réduite répliquant la dynamique des prix « *day-ahead* » du marché Nord Pool. Leur modèle est basé sur une série de processus Ornstein-Uhlenbeck et ils utilisent la transformée d'Esscher qui préserve les propriétés de la distribution du processus aléatoire. Ils

argumentent que dans le cas d'un modèle géométrique, il n'y a habituellement pas de formule de prix explicite pour un contrat à terme si on souhaite que la volatilité soit bien représentée et que dans le cas d'un modèle arithmétique, il faut ajouter une contrainte empêchant le prix d'être négatif.

La deuxième approche consiste à directement modéliser les prix des contrats à terme, sans passer par le prix de l'actif sous-jacent. La méthode Heath-Jarrow-Morton utilisée dans les modèles de taux d'intérêt est étudiée par Benth et al. (2008) mais comme les contrats à terme sur le marché de l'électricité couvrent une période et non un temps fixe, l'adaptation de cette méthode à ce contexte n'est pas simple.

Dupuis et al. (2016) retiennent donc de Benth et al. (2008) qu'un modèle statistique représentant la dynamique des prix des contrats à terme serait le plus approprié. Contrairement à Benth et al. (2008), les auteurs choisissent d'utiliser un modèle avec volatilité stochastique GARCH et une distribution NIG pour les résidus.

Définition 2.2. *Modèle d'hétéroscédasticité conditionnelle auto-régressive généralisée*

Une variable y_t peut être illustrée par la somme de son espérance μ et d'un terme aléatoire ϵ_t . Les modèles d'hétéroscédasticité conditionnelle auto-régressive généralisée, nommés GARCH, permettent de modéliser le comportement de la variance conditionnelle du terme aléatoire ϵ . Le modèle de base GARCH(1,1) est souvent utilisé dans la littérature pour modéliser les log-rendements $y_t = \log(S_t/S_{t-1})$ conditionnellement aux informations connues à $t - 1$.

On suppose $\epsilon_t = z_t\sigma_t$ où z_t est une innovation de moyenne 0 et variance 1 indépendante de \mathcal{F}_{t-1} . La variance conditionnelle de ϵ_t sachant \mathcal{F}_{t-1} est donnée

par

$$\begin{aligned}
 h_t &= \text{Var}[\epsilon_t | \mathcal{F}_{t-1}] \\
 &= \text{Var}[z_t \sigma_t | \mathcal{F}_{t-1}] \\
 &= \text{Var}[z_t] \text{Var}[\sigma_t | \mathcal{F}_{t-1}] \\
 &= \text{Var}[\sigma_t | \mathcal{F}_{t-1}] \\
 &= \sigma_t^2.
 \end{aligned}$$

Dans le cas du modèle GARCH(1,1), la variance conditionnelle du terme d'erreur ϵ_t est modélisée par la somme de la moyenne long-terme de la volatilité, du premier décalage de la variance conditionnelle et du premier décalage de l'innovation au carré :

$$h_t = \kappa + \gamma h_{t-1} + \xi \epsilon_{t-1}^2.$$

On obtient donc le modèle

$$\begin{aligned}
 y_t &= \mu + z_t \sigma_t, \\
 \sigma_t^2 &= \kappa + \gamma \sigma_{t-1}^2 + \xi z_{t-1}^2 \sigma_{t-1}^2.
 \end{aligned}$$

Définition 2.3. Loi Normale-Inverse Gaussienne

La loi Normale-Inverse Gaussienne (NIG) est une sous-classe des distributions hyperboliques généralisées. Elle est utilisée dans le domaine des mathématiques financières puisque ses paramètres permettent de représenter les queues épaisses souvent observées dans les rendements d'actifs financiers. Ses paramètres sont :

- μ : paramètre de position ;
- δ : paramètre d'échelle ;
- α : paramètre de kurtosis (épaisseur des queues) ;
- β : paramètre de symétrie.

À des fins de simplification des formules, on écrit $\gamma = \sqrt{\alpha^2 - \beta^2}$. De plus, on parle de NIG standard lorsque $\mu = 0$ et $\delta = 1$. L'attrait de la NIG est, entre autres, dû

à sa fonction génératrice des moments qui offre des expressions simples pour tous ses moments.

Pour $t \in \mathbb{R}$, la fonction génératrice des moments de $X \sim \text{NIG}(\mu, \delta, \alpha, \beta)$ est

$$\mathcal{M}_X(t) = \mathbb{E}[e^{tX}] = e^{\mu t + \delta(\gamma - \sqrt{\alpha^2 - (\beta+t)^2})}.$$

Pour la NIG standard, cette même fonction est encore plus simple

$$\mathcal{M}_X(t) = \mathbb{E}[e^{tX}] = e^{\gamma - \sqrt{\alpha^2 - (\beta+t)^2}}.$$

Les stratégies de couverture de risque proposées par Dupuis et al. (2016) ont une période de couverture de trois semaines. On s'intéresse donc seulement aux contrats à terme hebdomadaires jusqu'à trois semaines avant l'échéance. Comme le prix du contrat à terme d'échéance trois semaines $F_{T-3,T}$ est connu au début de la période de couverture, on s'intéresse donc à la modélisation de

$$F_{T-2,T}, F_{T-1,T} \text{ et } F_{T,T} = S_T,$$

où T représente la semaine pour laquelle nous souhaitons gérer le risque. L'actif sous-jacent à ces contrats à terme est le prix « *day-ahead* » dans le système d'électricité de Nord Pool.

Les log-rendements des contrats à terme, $\omega_{T-i,T}$ pour $i = 0, \dots, T$, sont modélisés puisque la série de données a davantage de probabilité d'être stationnaire, comme c'est le cas avec une majorité d'actifs financiers. Dans le contexte actuel, il y a donc trois log-rendements à prévoir :

$$\begin{aligned} \omega_{T-2,T} &= \log(F_{T-2,T}/F_{T-3,T}), \\ \omega_{T-1,T} &= \log(F_{T-1,T}/F_{T-2,T}), \\ \omega_{T,T} &= \log(F_{T,T}/F_{T-1,T}). \end{aligned}$$

À partir du modèle et des paramètres estimés par Dupuis et al. (2016), nous avons

pu répliquer les graphiques représentant les log-rendements des contrats à terme pour les différentes échéances dans la Figure 2.4. Ceci nous a permis de s'assurer que nous travaillions bien avec les mêmes données puisque certaines manipulations sont nécessaires. De plus, les graphiques originaux n'étaient pas aussi clairs et ne présentaient pas les rendements de 2007.

En observant la série de rendements dans la Figure 2.4, on note de l'autocorrélation, des regroupements de volatilité qui ne sont pas saisonniers ainsi qu'une probabilité d'événements extrêmes plus élevée qu'une distribution normale. De plus, on peut observer une certaine corrélation entre chaque échéance.

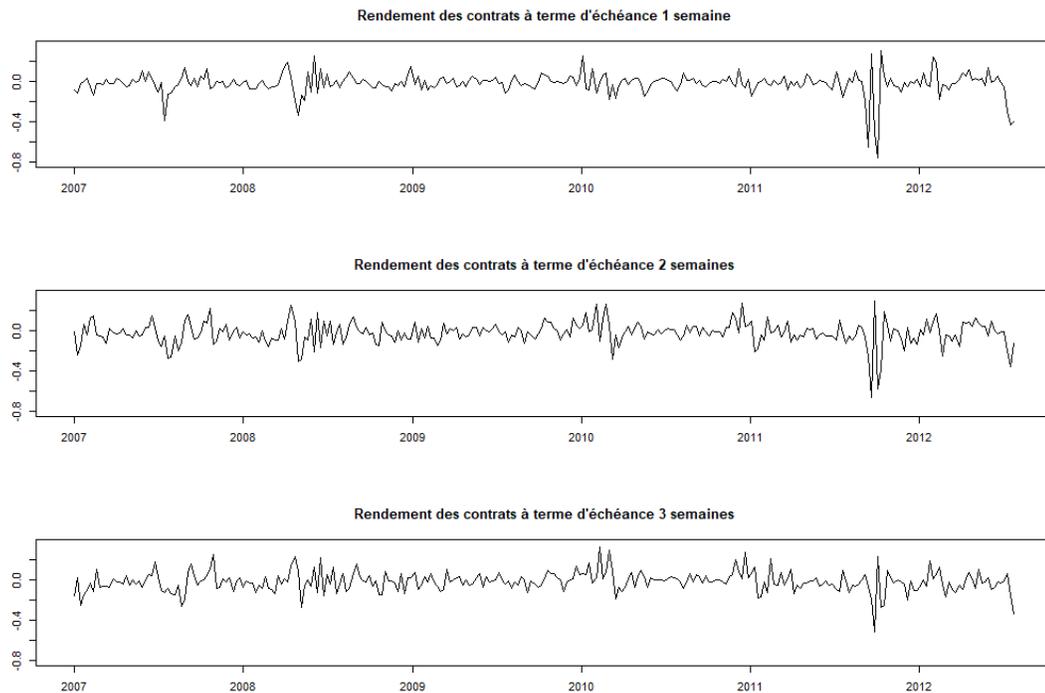


FIGURE 2.4 – Rendements des contrats à terme pour les différentes échéances

Le modèle pour les log-rendements des contrats à terme est un $AR(1)$ -GARCH(1,1) avec des résidus suivant une distribution NIG standard. La partie $AR(1)$ est retenue pour représenter l'autocorrélation précédemment observée, le GARCH(1,1) représente les regroupements de volatilité non saisonniers et les queues épaisses de la distribution sont illustrées par une loi NIG standard qui contrôle à la fois

l'asymétrie et le kurtosis. De plus, la corrélation entre les différentes échéances est modélisée par une copule dont nous discuterons dans la prochaine section. On a donc

$$\begin{aligned}\omega_{t,t} &= \mu_0 + a_0\omega_{t-1,t-1} + \sigma_{0,t}z_{0,t}, \\ \omega_{t-1,t} &= \mu_1 + a_1\omega_{t-2,t-1} + \sigma_{1,t-1}z_{1,t-1}, \\ \omega_{t-2,t} &= \mu_2 + a_2\omega_{t-3,t-1} + \sigma_{2,t-2}z_{2,t-2},\end{aligned}$$

où $\sigma_{i,t+1}^2 = \min\{\zeta^2, \kappa_i + \gamma_i\sigma_{i,t}^2 + \xi_i\sigma_{i,t}^2z_{i,t}^2\}$ pour $i \in \{0, 1, 2\}$ et où

$z_t = (z_{0,t}, z_{1,t}, z_{2,t})$ satisfait les propriétés suivantes :

- si $s \neq t$, z_s et z_t sont indépendants ;
- $z_{i,t}$ suit une loi NIG(α_i, β_i) standard ;
- conditionnellement aux informations connues au temps $t - 1$, $z_{0,t}$, $z_{1,t}$ et $z_{2,t}$ sont liés par une copule Gaussienne.

La limite de la volatilité, représentée par la constante ζ , est définie de telle sorte que le processus de prix X soit un processus de carré intégrable.

On note en observant le modèle que seulement les rendements de même temps restant avant l'échéance sont nécessaires pour chaque équation. Par exemple, supposons que $t = T - 1$ et que nous voulons prévoir $\omega_{T,T}$, c'est-à-dire le rendement de la prochaine semaine du contrat échéant au terme de cette même semaine. Nous utiliserons la première équation et nous aurons seulement besoin de connaître $\omega_{T-1,T-1}$, c'est-à-dire le rendement du contrat d'échéance 1 semaine venant d'expirer. Nous n'aurons, entre autres, pas besoin de connaître $\omega_{T-1,T}$, le rendement du contrat d'échéance 2 semaines venant à terme à $t = T$, qui sera plutôt utile pour prévoir $\omega_{T,T+1}$ à l'aide de la deuxième équation.

Comme $z_{i,t}$ est indépendant de $z_{i,t-1}$ pour $i \in \{0, 1, 2\}$ et que $z_{0,t}$, $z_{1,t}$ et $z_{2,t}$ sont liés par une copule Gaussienne, on a donc que les rendements des contrats à terme à différents temps restants avant l'échéance sont liés que par les résidus.

Au temps initial de la stratégie de couverture, les informations pour $t \leq T - 3$ sont connues, tel qu'illustré dans le tableau 2.1.

Temps t	T-3	T-2	T-1	T
Prix hebdomadaires	S_{T-3}	S_{T-2}	S_{T-1}	S_T
Prix des contrats à terme	$F_{T-3,T}$ $F_{T-3,T-1}$ $F_{T-3,T-2}$ $F_{T-3,T-3}$	$F_{T-2,T}$ $F_{T-2,T-1}$ $F_{T-2,T-2}$	$F_{T-1,T}$ $F_{T-1,T-1}$	$F_{T,T}$
Rendements des contrats à terme	$\omega_{T-3,T}$ $\omega_{T-3,T-1}$ $\omega_{T-3,T-2}$ $\omega_{T-3,T-3}$	$\omega_{T-2,T}$ $\omega_{T-2,T-1}$ $\omega_{T-2,T-2}$	$\omega_{T-1,T}$ $\omega_{T-1,T-1}$	$\omega_{T,T}$

TABLE 2.1 – Tableau des informations connues au temps t

Le premier rendement que l'on souhaite prévoir est $\omega_{T-2,T}$ qui est donné par le modèle

$$\omega_{T-2,T} = \mu_2 + a_2\omega_{T-3,T-1} + \sigma_{2,T-2}z_{2,T-2}$$

où

- $\omega_{T-3,T-1} = \log(F_{T-3,T-1}/F_{T-4,T-1})$ est connu ;
- $z_{2,T-2}$ est simulé à partir d'une distribution $\text{NIG}(\alpha_2, \beta_2)$;
- $\sigma_{2,T-2}^2 = \min\{\zeta^2, \kappa_2 + \gamma_2\sigma_{2,T-3}^2 + \xi_2\sigma_{2,T-3}^2z_{2,T-3}^2\}$ est connu car toutes les composantes sont connues à $t = T - 3$.

Comme $\omega_{T-2,T} = \log(F_{T-2,T}/F_{T-3,T})$ et $F_{T-3,T}$ est connu, nous pouvons ensuite déduire $F_{T-2,T}$. Similairement, le deuxième rendement que l'on souhaite prévoir est $\omega_{T-1,T} = \log(F_{T-1,T}/F_{T-2,T})$, duquel nous pourrions déduire $F_{T-1,T}$. Le troisième et dernier rendement que nous obtiendrons avec la même technique est $\omega_{T,T} = \log(F_{T,T}/F_{T-1,T})$, duquel nous pourrions déduire $F_{T,T} = S_T$.

2.3 Corrélation entre la demande et les contrats à terme

La théorie des copules permet de modéliser la structure de dépendance entre plusieurs variables aléatoires indépendamment de leurs lois marginales. Nous référons le lecteur à Haugh (2016) pour une introduction mathématique aux copules et à Wiecki (2018) pour une excellente démonstration visuelle des copules.

Une fonction C définie sur $[0, 1]^d$ et à valeurs dans $[0, 1]$ est une copule si et seulement si elle admet les propriétés suivantes :

- $C(u_1, \dots, u_d) = 0$ si $u_i = 0$ pour $i \in \{1, \dots, d\}$;
- $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ pour $i \in \{1, \dots, d\}$;
- Croissance : pour tous $u^1 = (u_1^1, \dots, u_d^1)$ et $u^2 = (u_1^2, \dots, u_d^2)$ tels que $u_i^1 \leq u_i^2$, alors

$$\sum_{k_1=1}^2 \dots \sum_{k_d=1}^2 (-1)^{k_1+\dots+k_d} C(u_1^{k_1}, \dots, u_d^{k_d}) \geq 0.$$

Étant donné sa définition, une copule est une fonction de répartition d'un vecteur de variables aléatoires uniformes sur l'intervalle $[0, 1]$.

À titre de rappel, si F est une fonction de répartition d'une variable aléatoire X , alors

1. $F(x) = P(X \leq x)$;
2. F est continue à droite avec des limites à gauche (càdlàg) ;
3. $F^{-1}(y) = \inf\{x | F(x) \geq y\}$ est la fonction inverse de F .

De plus, la variable $U = F(X)$ est distribuée selon une loi uniforme sur $[0, 1]$, puisque $P(U \leq u) = P(F(X) \leq u) = P(X \leq F^{-1}(u)) = u$.

Le théorème de Sklar déclare que toute fonction de répartition multivariée $F(x_1, \dots, x_d)$ d'un vecteur de variables aléatoires dans \mathbb{R}^d peut être exprimée en termes des fonctions de répartition marginales $F_i(x_i)$ et d'une copule C sous la forme

$$F(X) = C(F_1(x_1), \dots, F_d(x_d)).$$

De plus, si les marginales sont continues, la copule est unique. La copule associe donc les distributions marginales pour définir la fonction de distribution cumulative jointe, permettant de définir la structure de dépendance entre les variables aléatoires.

Il existe plusieurs familles de copules, chacune définissant à sa façon la relation entre les variables aléatoires. Une de celle-ci est la famille des copules archimédiennes qui sont intéressantes puisqu'elles admettent généralement une formule explicite.

À titre d'exemple simple, la copule d'indépendance bivariée fait partie de cette famille. Soient X et Y deux variables aléatoires réelles indépendantes, F leur fonction de répartition jointe, ainsi que G et H leurs fonctions de répartition marginales respectives telles que $U = G(X)$ et $V = H(Y)$ sont uniformes, alors

$$\begin{aligned}
C(u, v) &= C(G(x), H(y)) \\
&= F(x, y) \\
&= P(X \leq x, Y \leq y) \\
&= P(X \leq x) P(Y \leq y) \\
&= G(x) H(y) \\
&= u \times v.
\end{aligned}$$

Une autre famille de copules couramment utilisée est la copule gaussienne. Elle peut être écrite sous la forme $C_R(u) = \Phi_R(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$, où Φ^{-1} est la fonction de distribution cumulative inverse d'une loi normale standard et où Φ_R est la fonction de distribution cumulative jointe d'une distribution normale multivariée de vecteur de moyenne 0 et matrice de corrélation R .

La copule gaussienne n'admet pas de formule explicite mais sa valeur peut être estimée à l'aide de l'intégration numérique de la fonction de densité

$$c_R(u) = \frac{1}{\sqrt{\det(R)}} \exp \left(-\frac{1}{2} \begin{pmatrix} \Phi^{-1}(u_1) \\ \dots \\ \Phi^{-1}(u_d) \end{pmatrix}^T (R^{-1} - I) \begin{pmatrix} \Phi^{-1}(u_1) \\ \dots \\ \Phi^{-1}(u_d) \end{pmatrix} \right).$$

Notons que la copule gaussienne avec une matrice de corrélation R égale à la matrice identité I est aussi une copule d'indépendance et admet une formule explicite.

Dans l'article de Dupuis et al. (2016), la dépendance entre les log-rendements des contrats à terme de diverses échéances est modélisée par une copule gaussienne où les coefficients de corrélation $(\sigma_{0,1}, \sigma_{0,2}, \sigma_{1,2})$ sont fixes et ne varient pas dans le temps. La dépendance entre la demande et les log-rendements des contrats à

terme variant pendant l'année, ils ont introduit un coefficient de corrélation $\sigma_{LF,t}$ représenté par une fonction trigonométrique pour illustrer la saisonnalité cyclique, soit

$$\sigma_{LF,t} = \tau_1 0.5 \left(1 + \sin \left(\frac{3\pi}{2} + \frac{2\pi \min \left((t - \tau_3) \bmod \frac{365.25}{7}, \tau_2 \right)}{\tau_2} \right) \right).$$

La matrice de corrélation de $(z_{0,t}, z_{1,t}, z_{2,t}, \epsilon_t)$ varie donc dans le temps et est donnée par

$$R_t = \begin{bmatrix} 1 & \sigma_{0,1} & \sigma_{0,2} & \sigma_{LF,t} \\ \sigma_{0,1} & 1 & \sigma_{1,2} & \sigma_{LF,t} \\ \sigma_{0,2} & \sigma_{1,2} & 1 & \sigma_{LF,t} \\ \sigma_{LF,t} & \sigma_{LF,t} & \sigma_{LF,t} & 1 \end{bmatrix}.$$

3 ALGORITHME DE MINIMISATION LOCALE DU RISQUE

Nous avons testé l'implémentation de l'algorithme de programmation dynamique présenté par Augustyniak et al. (2017) pour répliquer la solution obtenue dans le théorème 1.10. Par contre, dû à la complexité du modèle, cette méthode n'était pas fonctionnelle. Les détails sont disponibles dans l'Annexe A.

Comme la stratégie de couverture de risque se fait avec seulement 3 rebalancements, nous avons opté pour la méthode Monte Carlo. L'algorithme se fait vers l'avant contrairement à l'algorithme de programmation dynamique qui se fait vers l'arrière dans le temps, c'est-à-dire les valeurs calculées à $t = T - 1$ servent au calcul des valeurs à $t = T - 2$.

Dans ce chapitre, nous présentons la méthode Monte Carlo et les modifications faites dans le but de réduire le temps de calculs. De plus, nous décrivons l'algorithme à l'aide d'une partie du pseudo-code.

3.1 Méthode Monte Carlo

La précision des estimations de Monte Carlo varie selon le nombre de trajectoires. Par la loi des grands nombres, l'espérance d'une variable aléatoire peut être estimée par la moyenne empirique d'observations indépendantes. De plus, le théorème central limite nous permet de déterminer l'intervalle de confiance de notre estimation.

Déterminer la précision optimale devient donc un choix personnel en considérant les temps de calculs. Nous nous sommes d'abord fiés aux paramètres choisis par Dupuis et al. (2016) pour simuler leur arbre stochastique. Ceux-ci utilisaient entre 500 et 1000 scénarios pour chacune des 15 positions possibles de la stratégie.

À partir de la solution optimale de minimisation locale des risques du théorème 1.10, la première étape est de dériver les formules nécessaires.

Remarque. Pour des fins de simplification, la notation de l'espérance et de la variance conditionnelle s'écrit maintenant sous la forme $\mathbb{E}_{T-3}[X_{T-2}]$ au lieu de $\mathbb{E}[X_{T-2}|\mathcal{F}_{T-3}]$. De plus, définissons $\Delta X_t = X_t - X_{t-1}$.

Par exemple,

$$\begin{aligned}\theta_T^1 &= \frac{\text{Cov}[V_T, X_T - X_{T-1}|\mathcal{F}_{T-1}]}{\text{Var}[X_T - X_{T-1}|\mathcal{F}_{T-1}]} \\ &= \frac{\text{Cov}_{T-1}[V_T, \Delta X_T]}{\text{Var}_{T-1}[\Delta X_T]} \\ &= \frac{\mathbb{E}_{T-1}[H\Delta X_T] - \mathbb{E}_{T-1}[H] \times \mathbb{E}_{T-1}[\Delta X_T]}{\mathbb{E}_{T-1}[\Delta X_T^2] - \mathbb{E}_{T-1}[\Delta X_T]^2}.\end{aligned}$$

3.2 Réduction du temps de calculs

Rappelons que les résidus $z_{i,t}$ du modèle des log-rendements suivent une loi $\text{NIG}(\alpha_i, \beta_i)$ standard, pour $i \in \{0, 1, 2\}$. Une simplification notable des calculs est due à la fonction génératrice des moments de la NIG puisque nous pouvons calculer les espérances conditionnelles $\mathbb{E}_{T-i-1}[\exp(\sigma_{i,T-i} z_{i,T-i})]$ et $\mathbb{E}_{T-i-1}[\exp(2\sigma_{i,T-i} z_{i,T-i})]$ à l'aide de celle-ci :

$$\mathcal{M}_{z_{i,T-i}}(\sigma_{i,T-i}) = \mathbb{E}_{T-i-1}[\exp(\sigma_{i,T-i} z_{i,T-i})] = e^{\sqrt{\alpha_i^2 - \beta_i^2} - \sqrt{\alpha_i^2 - (\beta_i + \sigma_{i,T-i})^2}},$$

$$\mathcal{M}_{z_{i,T-i}}(2\sigma_{i,T-i}) = \mathbb{E}_{T-i-1}[\exp(2\sigma_{i,T-i} z_{i,T-i})] = e^{\sqrt{\alpha_i^2 - \beta_i^2} - \sqrt{\alpha_i^2 - (\beta_i + 2\sigma_{i,T-i})^2}}.$$

Ainsi, conditionnellement aux informations connues au temps $t = T - 1$, les seules espérances conditionnelles à calculer sont

$$\mathbb{E}_{T-1}[\epsilon_T \exp(\sigma_{0,T} z_{0,T})]$$

et

$$\mathbb{E}_{T-1}[\epsilon_T \exp(2\sigma_{0,T} z_{0,T})].$$

Pour réduire le temps de calculs, nous avons décidé d'utiliser la formule de Taylor-Young pour pouvoir sortir $\sigma_{0,T}$ du calcul. Ceci réduit grandement le nombre d'es-

pérances à calculer puisque nous n'avons pas à calculer une espérance pour chaque valeur possible de $\sigma_{0,T}$.

Nous avons donc approximé ces espérances avec la formule de Taylor-Young :

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + \dots$$

Ainsi, avec $f(x) = e^{cx}$ et $a = 0$, où $c \in \mathbb{R}$, la formule d'ordre 2 est

$$e^{cx} \cong 1 + cx + \frac{c^2 x^2}{2}.$$

Nous obtenons donc

$$\mathbb{E}_{T-1}[\epsilon_T \exp(\sigma_{0,T} z_{0,T})] \cong \sigma_{0,T} \mathbb{E}_{T-1}[\epsilon_T z_{0,T}] + \frac{\sigma_{0,T}^2}{2} \mathbb{E}_{T-1}[\epsilon_T z_{0,T}^2]$$

et

$$\mathbb{E}_{T-1}[\epsilon_T \exp(2\sigma_{0,T} z_{0,T})] \cong 2\sigma_{0,T} \mathbb{E}_{T-1}[\epsilon_T z_{0,T}] + \frac{\sigma_{0,T}^2}{2} \mathbb{E}_{T-1}[\epsilon_T z_{0,T}^2].$$

En conséquence, il suffit simplement de calculer de manière préemptive les espérances $\mathbb{E}_{T-1}[\epsilon_T z_{0,T}]$ et $\mathbb{E}_{T-1}[\epsilon_T z_{0,T}^2]$ pour les 291 périodes et d'ensuite les intégrer en tant que valeur connue dans l'algorithme. Nous utilisons la méthode Monte Carlo pour estimer la valeur de ces espérances puisque le calcul direct nécessiterait une intégrale de quatre dimensions.

De plus, en modifiant le diviseur du modulo de $\frac{365.25}{7}$ à 52, nous observons dans la Figure 3.1 que les coefficients de corrélation entre la demande et les log-rendements divergent peu, en moyenne de seulement 0.0066 avec un écart-type de 0.0083.

Nous avons pu réduire le nombre de périodes de 291 à 52. Parmi les 52 périodes, le coefficient de corrélation $\sigma_{LF,t}$ était nul pour 18 d'entre elles. Nous avons donc à calculer seulement 35 fois les espérances $\mathbb{E}_{T-1}[\epsilon_T z_{0,T}]$ et $\mathbb{E}_{T-1}[\epsilon_T z_{0,T}^2]$.

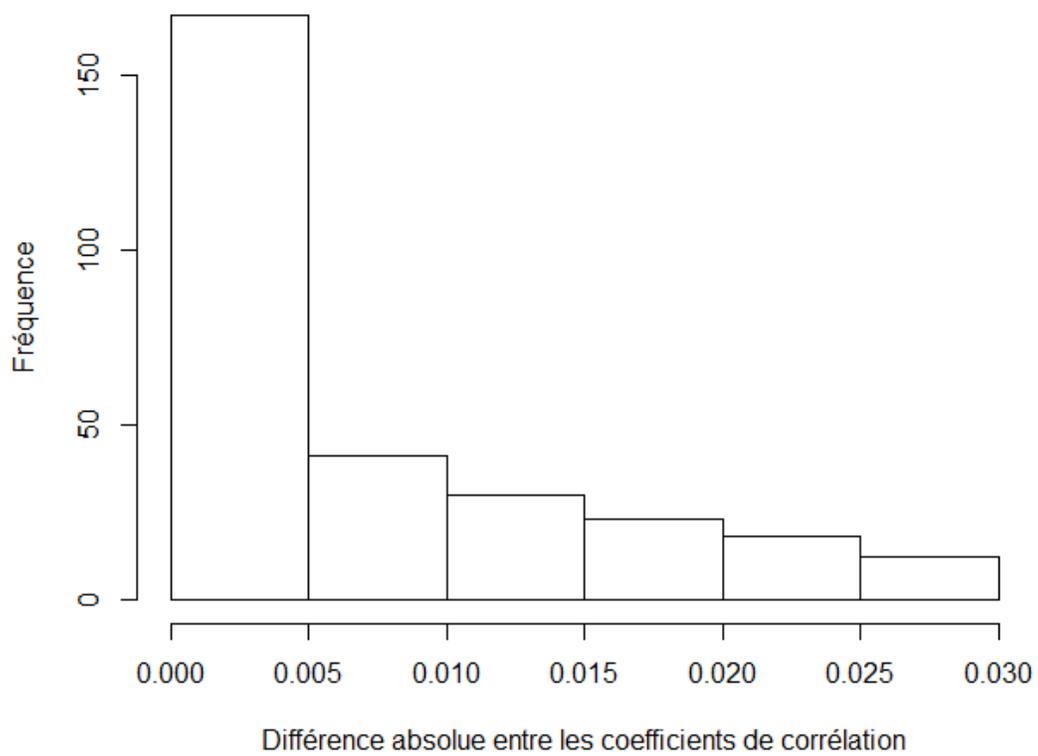


FIGURE 3.1 – Histogramme des différences absolues entre les coefficients de corrélation calculés avec MODULO 365.25/7 et MODULO 52

3.3 Algorithme

On rappelle que la solution de la stratégie optimale θ de la minimisation locale du risque du fournisseur d'électricité pour la semaine T est donnée par la récursion suivante :

$$\begin{aligned} V_T(\theta) &:= H; \\ \theta_t^1 &:= \frac{\text{Cov}_{t-1}[V_t(\theta), \Delta X_t]}{\text{Var}_{t-1}[\Delta X_t]}; \\ \theta_t^0 &:= V_t(\theta) - \theta_t^1 X_t; \\ V_{t-1}(\theta) &:= \mathbb{E}_{t-1}[V_t(\theta)] - \theta_t^1 \mathbb{E}_{t-1}[\Delta X_t] \end{aligned}$$

La première étape est de calculer θ_{T-2}^1 à l'aide des informations connues au temps $T - 3$:

$$\begin{aligned} \theta_{T-2}^1 &= \frac{\text{Cov}_{T-3}[V_{T-2}, \Delta X_{T-2}]}{\text{Var}_{T-3}[\Delta X_{T-2}]} \\ &= \frac{(\mathbb{E}_{T-3}[(\mathbb{E}_{T-2}[H] - \mathbb{E}_{T-2}[\theta_{T-1}^1 \Delta X_{T-1}] - \mathbb{E}_{T-2}[\theta_T^1 \Delta X_T]) \Delta X_{T-2}])}{\mathbb{E}_{T-3}[(\Delta X_{T-2})^2] - \mathbb{E}_{T-3}[\Delta X_{T-2}]^2} \\ &\quad - \frac{(\mathbb{E}_{T-3}[\mathbb{E}_{T-2}[H] - \mathbb{E}_{T-2}[\theta_{T-1}^1 \Delta X_{T-1}] - \mathbb{E}_{T-2}[\theta_T^1 \Delta X_T]] \mathbb{E}_{T-3}[\Delta X_{T-2}])}{\mathbb{E}_{T-3}[(\Delta X_{T-2})^2] - \mathbb{E}_{T-3}[\Delta X_{T-2}]^2} \end{aligned}$$

Ceci est effectué en estimant chaque espérance conditionnelle au temps $T - 3$ avec la méthode de Monte Carlo :

- $\mathbb{E}_{T-3}[\mathbb{E}_{T-2}[H] \Delta X_{T-2}]$;
- $\mathbb{E}_{T-3}[\mathbb{E}_{T-2}[\theta_{T-1}^1 \Delta X_{T-1}] \Delta X_{T-2}]$;
- $\mathbb{E}_{T-3}[\mathbb{E}_{T-2}[\theta_T^1 \Delta X_T] \Delta X_{T-2}]$;
- $\mathbb{E}_{T-3}[H]$;
- $\mathbb{E}_{T-3}[\theta_{T-1}^1 \Delta X_{T-1}]$;
- $\mathbb{E}_{T-3}[\theta_T^1 \Delta X_T]$;
- $\mathbb{E}_{T-3}[\Delta X_{T-2}]$;
- $\mathbb{E}_{T-3}[(\Delta X_{T-2})^2]$.

Prenons par exemple $\mathbb{E}_{T-3}[\Delta X_{T-2}] = \mathbb{E}_{T-3}[X_{T-2}] - X_{T-3}$.

Pour calculer l'espérance de X_{T-2} au temps $T - 3$, nous utilisons la définition de X ainsi que la fonction génératrice des moments de la NIG.

$$\begin{aligned}\mathbb{E}_{T-3}[X_{T-2}] &= \mathbb{E}_{T-3} \left[\frac{X_{T-3}}{\Delta B} \exp(\omega_{T-2,T}) \right] \\ &= \frac{X_{T-3}}{\Delta B} \exp(\mu_2 + a_2 \omega_{T-3,T-1}) \mathbb{E}_{T-3} [\exp(\sigma_{2,T-2} z_{2,T-2})] \\ &= \frac{X_{T-3}}{\Delta B} \exp(\mu_2 + a_2 \omega_{T-3,T-1}) \exp \left(\sqrt{\alpha_2^2 - \beta_2^2} - \sqrt{\alpha_2^2 - (\beta_2 + \sigma_{2,T-2})^2} \right)\end{aligned}$$

Dans le cas de $\mathbb{E}_{T-3}[H]$, nous devons utiliser les copules puisque $H = \mathcal{L}_T(X_T - X_{T-3})$ dépend à la fois de la demande d'électricité et des prix des contrats à terme.

On génère M_2 échantillons de $(z_{0,T-2}, z_{1,T-2}, z_{2,T-2}, \epsilon_{T-2})$ à l'aide de la copule. Pour chaque échantillon m_2 , on calcule \mathcal{L}_{T-2} , $\omega_{T-2,T}$, $\omega_{T-2,T-1}$, $\omega_{T-2,T-2}$, $\sigma_{0,T-1}$ et $\sigma_{1,T-1}$. On génère M_1 échantillons de $(z_{0,T-1}, z_{1,T-1}, z_{2,T-1}, \epsilon_{T-1})$ avec lesquels on calcule \mathcal{L}_{T-1} , $\omega_{T-1,T}$, $\omega_{T-1,T-1}$ et $\sigma_{0,T}$.

On est en mesure de calculer

$$\begin{aligned}\mathbb{E}_{T-3}[H] &= \frac{1}{M_2 \times M_1} \sum_{m_2=1}^{M_2} \sum_{m_1=1}^{M_1} \mathbb{E}_{T-1}[H] \\ &= \frac{1}{M_2 \times M_1} \sum_{m_2=1}^{M_2} \sum_{m_1=1}^{M_1} (\mathbb{E}_{T-1}[\mathcal{L}_T X_T] - X_{T-3} \mathbb{E}_{T-1}[\mathcal{L}_T])\end{aligned}$$

L'espérance $\mathbb{E}_{T-1}[\mathcal{L}_T]$ est directement calculée par sa définition, c'est-à-dire $\mathcal{L}_T = g(T) + \gamma(\mathcal{L}_{T-1} - g(T-1))$.

L'espérance $\mathbb{E}_{T-1}[\mathcal{L}_T X_T]$ est calculée à l'aide de la fonction génératrice des moments de la NIG ainsi que les espérances préalablement calculées, tel que discuté dans la section précédente.

La prochaine étape est de calculer θ_{T-3}^0 . On sait que θ_{T-3}^0 représente l'investissement

ment initial V_{T-3} puisque $\theta_{T-3}^1 = 0$ par définition :

$$\begin{aligned}
\theta_{T-3}^0 &= V_{T-3} \\
&= \mathbb{E}_{T-3}[V_{T-2}(\theta)] - \theta_{T-2}^1 \mathbb{E}_{T-3}[\Delta X_{T-2}] \\
&= \mathbb{E}_{T-3}[\mathbb{E}_{T-2}[V_{T-1}(\theta)] - \theta_{T-1}^1 \mathbb{E}_{T-2}[\Delta X_{T-1}]] - \theta_{T-2}^1 \mathbb{E}_{T-3}[\Delta X_{T-2}] \\
&= \mathbb{E}_{T-3}[\mathbb{E}_{T-2}[\mathbb{E}_{T-1}[H] - \theta_T^1 \mathbb{E}_{T-1}[\Delta X_T]] - \theta_{T-1}^1 \mathbb{E}_{T-2}[\Delta X_{T-1}]] - \theta_{T-2}^1 \mathbb{E}_{T-3}[\Delta X_{T-2}] \\
&= \mathbb{E}_{T-3}[H] - \mathbb{E}_{T-3}[\theta_T^1 \Delta X_T] - \mathbb{E}_{T-3}[\theta_{T-1}^1 \Delta X_{T-1}] - \theta_{T-2}^1 \mathbb{E}_{T-3}[\Delta X_{T-2}]
\end{aligned}$$

On remarque que toutes les composantes de la dernière équation ont été calculées à l'étape précédente et nous utilisons donc ces valeurs.

Au temps $T-3$, le fournisseur d'électricité achètera θ_{T-2}^1 contrats à terme d'échéance T au prix X_{T-3} . Une semaine plus tard, au temps $T-2$, il vendra ses θ_{T-2}^1 contrats au prix X_{T-2} et achètera θ_{T-1}^1 contrats à ce prix.

Il faut donc déterminer la valeur de θ_{T-1}^1 de la même manière qui a été utilisée pour θ_{T-2}^1 :

$$\begin{aligned}
\theta_{T-1}^1 &= \frac{\text{Cov}_{T-2}[V_{T-1}, \Delta X_{T-1}]}{\text{Var}_{T-2}[\Delta X_{T-1}]} \\
&= \frac{\mathbb{E}_{T-2}[(\mathbb{E}_{T-1}[H] - \mathbb{E}_{T-1}[\theta_T^1 \Delta X_T]) \Delta X_{T-1}] - (\mathbb{E}_{T-2}[H] - \mathbb{E}_{T-2}[\theta_T^1 \Delta X_T]) \mathbb{E}_{T-2}[\Delta X_{T-1}]}{\mathbb{E}_{T-2}[(\Delta X_{T-1})^2] - \mathbb{E}_{T-2}[\Delta X_{T-1}]^2}
\end{aligned}$$

On fait la même chose pour les prochaines positions jusqu'à l'échéance.

4 RÉSULTATS NUMÉRIQUES

Dans cette section, nous allons dans un premier temps définir les données sur lesquelles l'algorithme présenté précédemment a été appliqué. Par la suite, nous définissons le critère de comparaison de la performance des stratégies de couverture de risque quadratique. Dans la dernière section, nous analysons les résultats obtenus.

4.1 Données

Nous référons à Dupuis et al. (2016) pour la description des données et leur provenance. Les auteurs nous ont fourni cinq séries de données afin que nous puissions générer les résultats d'une stratégie de couverture locale comparative à leur travail :

- Prix horaire « *day-ahead* » de Nord Pool de Janvier 1999 à Juillet 2012 ;
- Demande horaire de Nord Pool de Janvier 1997 à Juillet 2012 ;
- Prix de fermeture quotidien des contrats à terme d'échéance 1 semaine de Août 2006 à Octobre 2012 ;
- Prix de fermeture quotidien des contrats à terme d'échéance 2 semaines de Août 2006 à Octobre 2012 ; et
- Prix de fermeture quotidien des contrats à terme d'échéance 3 semaines de Août 2006 à Octobre 2012.

Ces données horaires ont dû être manipulées pour être utilisées dans notre stratégie de couverture de risque locale hebdomadaire. Tout comme Dupuis et al. (2016), nous avons gardé les données pour la période de janvier 2007 à juillet 2012.

À l'aide des équations 2.1 à 2.4, les prix hebdomadaires S_t et S_t^* ont été calculés pour ensuite calculer L_t et \mathcal{L}_t . Nous obtenons ainsi 291 observations hebdomadaires.

Puisque le rebalancement de notre stratégie se fait à chaque vendredi, nous avons uniquement gardé les prix de fermeture des contrats à terme du vendredi. De plus,

le prix terminal $F_{T,T} = S_T$ est donné par la série de données de prix d'électricité et est calculé en date du dimanche.

Notons que le rebalancement du vendredi se fait sur des données incomplètes puisque la semaine commence le lundi et se termine le dimanche. Ainsi, à 16h le vendredi de la semaine t , l'heure de la fermeture des marchés financiers, nous connaissons la demande d'électricité que pour $24 \times 4 + 16 = 112$ des 168 heures de la semaine. Dupuis et al. (2016) ont déterminé que le modèle

$$\hat{\mathcal{L}}_t = \tilde{L}_t \times \left(0.716 + 0.171 \sum_{j=1}^3 \frac{\mathcal{L}_{t-j}}{\tilde{L}_{t-j}} \right),$$

où $\tilde{L}_t = \sum_{d=1}^4 \sum_{h=1}^{24} L_{t,d,h} + \sum_{h=1}^{16} L_{t,5,h}$, était approprié pour estimer la demande hebdomadaire pondérée par les prix en date du vendredi.

4.2 Erreur de couverture du risque

On rappelle que la stratégie de minimisation locale du risque est une stratégie qui n'est pas contrainte d'être autofinancée, ce qui permet donc la réplication parfaite à l'échéance en posant $\theta_T^0 = H - \theta_T^1 X_T$. Il est alors incohérent d'utiliser l'erreur terminal $H - V_T$ pour comparer la performance de cette dernière avec celle de la stratégie de minimisation du risque quadratique global.

L'erreur de couverture Δ_T d'une stratégie θ est définie comme la différence entre la valeur du droit contingent à son échéance T et la valeur terminale du portefeuille de couverture net des coûts incrémentaux,

$$\Delta_T := H - \left(V_T(\theta) - \sum_{t=1}^T (C_t(\theta) - C_{t-1}(\theta)) \right).$$

Cette valeur est aussi équivalente au flux monétaire supplémentaire à la suite de l'investissement initial.

Pour une minimisation du risque global où la stratégie est autofinancée, l'erreur

de couverture correspond au flux monétaire nécessaire à la fin de la période pour couvrir la différence entre la valeur du droit contingent et la valeur finale du portefeuille de couverture,

$$\Delta_T^G = H - V_T(\theta).$$

Pour une stratégie de minimisation locale du risque, l'erreur correspond au flux monétaire supplémentaire investi tout au long de la période de couverture pour que la valeur finale du portefeuille atteigne la valeur du droit contingent. Ceci représente donc la somme des coûts incrémentaux,

$$\Delta_T^L := \sum_{t=1}^T (C_t(\theta) - C_{t-1}(\theta)) = C_T(\theta) - C_0(\theta).$$

Pour comparer deux stratégies sur un même pied d'égalité, il faut ajouter la contrainte que l'investissement initial des deux stratégies est identique. On pourrait aussi définir l'erreur de couverture comme la différence entre la valeur du droit contingent à son échéance T et la valeur terminale du portefeuille de couverture, net des coûts incrémentaux et de l'investissement initial.

$$\Delta_T^* := H - \left(V_T(\theta) - \sum_{t=1}^T (C_t(\theta) - C_{t-1}(\theta)) - V_0(\theta) \right) = \Delta_T + V_0(\theta).$$

Dupuis et al. (2016) choisissent $V_0 = 0$ pour toutes les stratégies. Pour notre stratégie de minimisation locale du risque, nous aurons donc à calculer

$$\Delta_T^* = \Delta_T^L + V_0(\theta) = C_T(\theta) - C_0(\theta) + V_0(\theta) = C_T(\theta).$$

4.3 Résultats

Le théorème 1.18 définit la solution de la stratégie de couverture de risque quadratique global en fonction de la solution de minimisation locale du risque dans le cas où le ratio $\frac{\sigma_t^2}{\alpha_t^2}$ est déterministe pour chaque t . En conséquence de la fonction génératrice des moments de la NIG, c'est le cas dans notre modèle. En particulier,

pour $i \in \{0, 1, 2\}$,

$$\frac{\sigma_{T-i}^2}{\alpha_{T-i}^2} = \frac{\exp(2\mu_i + 2a_i\omega_{T-i-1,T-1})(\mathcal{M}_{z_i}(2\sigma_{i,T-i}) - \mathcal{M}_{z_i}(\sigma_{i,T-i})^2)}{(\exp(\mu_i + a_i\omega_{T-i-1,T-1})\mathcal{M}_{z_i}(\sigma_{i,T-i}) - \exp(r))^2}.$$

Par contre, par la proposition 1.14, une solution à la stratégie de minimisation locale du risque existe si et seulement si le droit contingent H admet une décomposition orthogonale. Dans notre cas,

$$H = \mathcal{L}_T(X_T - X_{T-3}) = \sum_{t=T-2}^T \mathcal{L}_T(X_t - X_{t-1}).$$

Sous cette forme, H n'est pas une décomposition orthogonale adéquate puisque \mathcal{L}_T n'est pas un processus prévisible. Par contre, ceci n'implique pas qu'une décomposition orthogonale n'existe pas puisque l'existence de celle-ci dépend de la structure de l'espace de probabilité filtré.

Nous allons vérifier ces résultats théoriques avec les résultats numériques de l'algorithme présenté dans le dernier chapitre.

Notez que l'algorithme de Dupuis et al. (2016) inclut les frais de transactions dans l'optimisation. Nous avons décidé de ne pas le faire et de seulement les ajouter à la fin. Le total des frais est donné par

$$\mathcal{C}_T = 0.04 \times (|\theta_{T-2}^1| + |\theta_{T-1}^1 - \theta_{T-2}^1| + |\theta_T^1 - \theta_{T-1}^1| + |\theta_T^1|)$$

et équivaut en moyenne à 6,000€ avec un écart-type de 1,220€.

Puisque nous n'estimons pas les paramètres du modèle, nous nous contentons de comparer les résultats « *in-sample* » sur la même période que Dupuis et al. (2016), soit les 287 semaines du 29 janvier 2007 au 23 juillet 2012. Nous utiliserons les mêmes mesures, soit l'erreur quadratique moyenne (RMSE), la semi erreur quadratique moyenne (semi-RMSE) ainsi que la VaR et la TVaR pour des quantiles de 95% et 99% :

$$RMSE = \sqrt{\frac{1}{287} \sum_{T=1}^{287} (\Delta_T^* + \mathcal{C}_T)^2},$$

$$semi - RMSE = \sqrt{\frac{1}{287} \sum_{T=1}^{287} (\Delta_T^* + \mathcal{C}_T)^2 \mathbb{1}_{\{(\Delta_T^* + \mathcal{C}_T) > 0\}}},$$

$$TVaR_\alpha = \frac{\sum_{T=1}^{287} (\Delta_T^* + \mathcal{C}_T) \mathbb{1}_{\{(\Delta_T^* + \mathcal{C}_T) \geq q(1-\alpha)\}}}{\sum_{T=1}^{287} \mathbb{1}_{\{(\Delta_T^* + \mathcal{C}_T) \geq q(1-\alpha)\}}},$$

où $VaR_\alpha = q_{1-\alpha}$ est le quantile de niveau $1 - \alpha$ de l'erreur de couverture incluant les frais $\Delta_T^* + \mathcal{C}_T$.

Les résultats sont présentés dans la Table 4.1 en comparaison avec la stratégie de minimisation du risque quadratique global, la stratégie du delta-hedging naïf (DHN) ainsi que les résultats si aucune stratégie de couverture est implémentée. La stratégie DHN est une stratégie autofinancée où la position dans l'actif risqué θ_t^* équivaut à la demande pondérée par les prix de la semaine t .

Modèle	Global	Local	NDH	Aucun
Moyenne	7.37	16.66	7.06	-70.7
RMSE	26.04	112.9	35.03	474.1
Semi-RMSE	24.55	80.18	32.15	319.5
$VaR_{5\%}$	53.46	143.1	73.85	688.7
$VaR_{1\%}$	118.3	376.9	153.5	1471
$TVaR_{5\%}$	92.79	257.9	121.0	1201
$TVaR_{1\%}$	134.5	431.4	180.9	1831

TABLE 4.1 – Résultats des tests « *in-sample* » de 2007 à 2012

On observe que le portefeuille de couverture du delta-hedging naïf semble plus centré sur la valeur du droit contingent puisque la méthode performe mieux en moyenne.

D'une autre part, on note que la stratégie de couverture globale des risques performe mieux par rapport aux autres mesures. Le portefeuille de couverture est légèrement plus biaisé que celui du delta-hedging mais il mitige mieux les risques extrêmes, tel que démontré par les mesures de VaR et de TVaR.

Dernièrement, on trouve que la stratégie de minimisation locale des risques est la méthode qui performe le moins bien. On sait que théoriquement cette méthode performe moins bien que la stratégie de couverture globale. Par contre, on s'atten-

dait à ce qu'elle performe mieux que le delta-hedging naïf qui ne requière aucun calcul complexe.

Puisque nous avons apporté quelques modifications au modèle pour améliorer les temps de calculs, entre autres la modification du diviseur du modulo, nous étudions $\Delta_T + \mathcal{C}_T$ dans le temps. La série temporelle est illustrée dans la Figure 4.1.

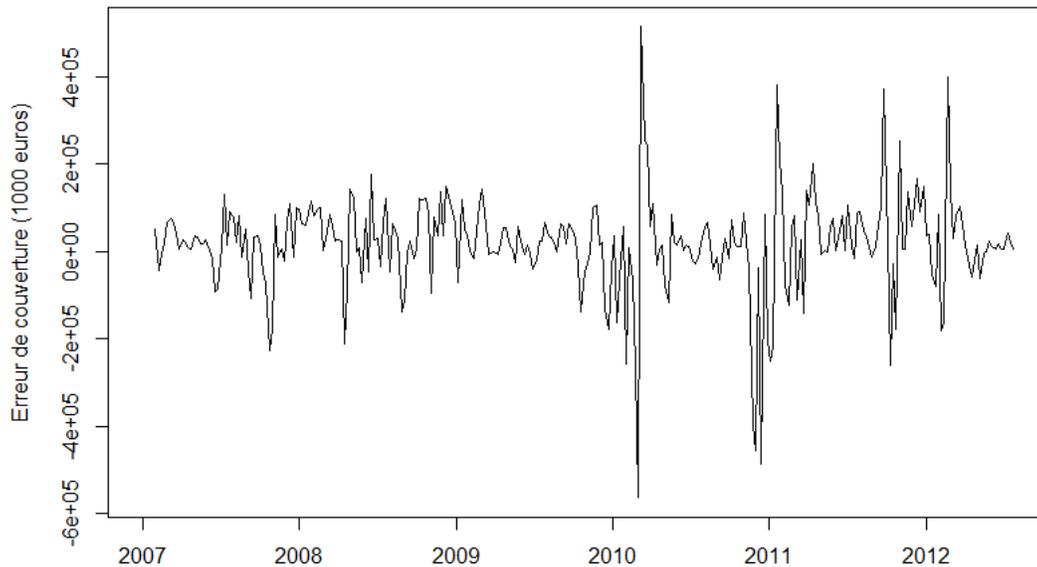


FIGURE 4.1 – Erreur de couverture du risque $\Delta_T^* + \mathcal{C}_T$ pour la stratégie locale

Nous observons que l'erreur est beaucoup plus volatile après 3 ans, ce qui pourrait être conséquence d'un décalage causé par la modification du diviseur du modulo. Nous recalculons les mêmes mesures sur les données des 3 premières années, soit 156 observations et présentons les résultats dans la Table 4.2 en comparaison avec les mêmes résultats de Dupuis et al. (2016) calculés sur les 287 périodes.

On observe que notre stratégie de minimisation locale du risque performe mieux que le delta-hedging naïf dans la gestion des risques extrêmes, illustré par la $VaR_{1\%}$ et la $TVaR_{1\%}$. Par contre, il faut garder en tête que les mesures ne sont pas calculées sur les mêmes périodes puisque nous n'avons pas accès à la série chronologique des résultats de Dupuis et al. (2016).

En comparant la Figure 4.1 représentant la série temporelle de $\Delta_T^* + \mathcal{C}_T$ et la

Modèle	Global	Local	NDH	Aucun
Moyenne	7.37	20.77	7.06	-70.7
RMSE	26.04	74.14	35.03	474.1
Semi-RMSE	24.55	57.76	32.15	319.5
$Var_{5\%}$	53.46	121.6	73.85	688.7
$Var_{1\%}$	118.3	149.3	153.5	1471
$TVaR_{5\%}$	92.79	140.9	121.0	1201
$TVaR_{1\%}$	134.5	163.3	180.9	1831

TABLE 4.2 – Résultats des tests « *in-sample* » de 2007 à 2009 pour la stratégie locale et 2007 à 2012 pour les autres stratégies

Figure 2.4 illustrant les rendements des contrats à terme pour les différentes échéances, nous observons que l’erreur de couverture du risque semble subir les mêmes sauts que les rendements. Il pourrait être possible que le modèle des log-rendements ne soit pas adéquat et ainsi qu’une partie de l’erreur de couverture de risque soit causée par le risque de modèle.

Contrairement au modèle de demande pondérée par les prix, l’adéquation du modèle de variance n’a pas été illustré par Dupuis et al. (2016). Dans la Figure 4.2, on compare les résidus réalisés absolus $|\sigma_{i,T} z_{i,T}|$ à l’écart-type $\sigma_{i,T}$ estimé par le modèle. On trouve que l’adéquation diminue avec l’augmentation du temps restant avec l’échéance. De plus, on observe que le modèle tend à surestimer la variance lorsque la volatilité est faible.

Contrairement à ce qui a été mentionné dans la littérature, les résultats de la stratégie de minimisation locale des risques ne s’apparentent pas aux résultats de la stratégie de couverture globale.

Le nombre de simulations dans les estimations de Monte Carlo concordait avec le nombre de scénarios utilisés par Dupuis et al. (2016). Pour le calcul de θ_{T-2}^1 sachant \mathcal{F}_{T-3} , $M_2 = 1000$ et $M_1 = 500$ simulations ont été utilisées, où M_2 représente le nombre de scénarios simulés pour $(z_{0,T-2}, z_{1,T-2}, z_{2,T-2}, \epsilon_{T-2})$ et M_1 le nombre de scénarios simulés pour $(z_{0,T-1}, z_{1,T-1}, z_{2,T-1}, \epsilon_{T-1})$. Pour le calcul de θ_{T-1}^1 sachant \mathcal{F}_{T-2} , $M_1 = 1000$ scénarios ont été simulés. Nous n’avons pas besoin de simuler de scénarios pour $(z_{0,T}, z_{1,T}, z_{2,T}, \epsilon_T)$ puisque nous avons réduit

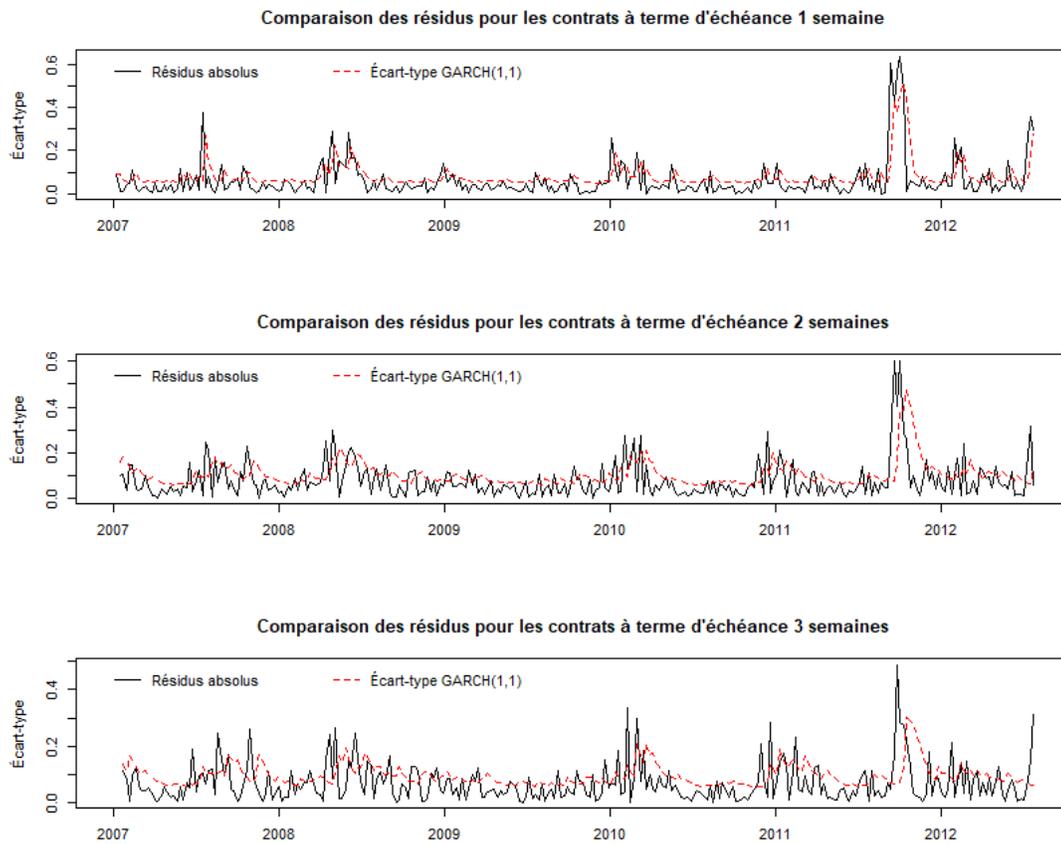


FIGURE 4.2 – Comparaison entre les résidus réalisés absolus $|\hat{\sigma}_{i,T} \hat{z}_{i,T}|$ et l'écart-type estimé $\sigma_{i,T}$

le problème à des fonctions explicites grâce à la fonction génératrice des moments de la NIG et la formule de Taylor-Young.

Le temps de calcul de la stratégie pour une semaine t avec les paramètres $M_2 = 1000$ et $M_1 = 500$ pour la première étape et $M_1 = 1000$ pour la deuxième était d'environ 7h. Doubler la valeur de M_1 pour la première étape doublait le temps de calculs.

CONCLUSION

La demande d'électricité du marché de Nord Pool est marquée par une forte saisonnalité dans la moyenne, due à la température, ainsi que dans la variance. La série chronologique présente aussi de l'autocorrélation dans les différences consécutives de la demande et de sa tendance. Les rendements des contrats à terme présentent de l'autocorrélation, de la volatilité non-saisonnale ainsi que des queues épaisses.

Toutes ces caractéristiques ne se ressemblent pas dans chaque marché. Par exemple, un marché nordique n'aura pas nécessairement les mêmes variations de température qu'un marché américain. Les besoins hivernaux pourraient être semblables mais les besoins estivaux impliquant la climatisation pourraient différer.

Les réalités gouvernementales auront aussi un impact sur le niveau des prix moyens dû aux diverses réglementations propres à chaque marché ainsi qu'à une répartition de génération différente de marché en marché. Par exemple, plusieurs marchés se tournent de plus en plus vers l'énergie renouvelable. Un autre bon exemple est dans le marché de l'Ontario où des contrats long termes ont été signés pour certaines centrales nucléaires et hydroélectriques dans le passé. La différence entre ces taux régulés et le prix déduit de la courbe de l'offre et de la demande est chargée aux consommateurs.

Nous avons procédé à l'analyse des données du marché de PJM aux États-Unis dans le but de transposer les modèles utilisés dans l'article de Dupuis et al. (2016). Par contre, certaines caractéristiques différant du marché de Nord Pool, tel que l'autocorrélation des rendements, nous empêchait d'utiliser les mêmes modèles. Les résultats de cette analyse sont présentés dans l'Annexe B.

Dans le cadre de ce mémoire, nous avons présenté la solution optimale d'une stratégie de minimisation locale du risque afin d'appliquer cette théorie sur les données de Nord Pool et de comparer les résultats avec ceux observés par Dupuis et al. (2016) qui appliquent une stratégie de minimisation globale du risque quadratique.

On trouve que contrairement à la littérature, notre stratégie sur une courte période de 3 semaines ne se compare pas avec les résultats de la stratégie de minimisation globale du risque quadratique. Le plus étonnant a été de trouver que même le delta-hedging naïf surpassait notre stratégie. Par contre, Augustyniak et al. (2017) ont démontré que la stratégie de minimisation globale du risque quadratique performe mieux que sa contrepartie locale en cas d'erreur de spécification du modèle.

En comparant avec le delta-hedging naïf qui se base uniquement sur le modèle de la demande, on déduit que l'erreur de spécification du modèle provient probablement des rendements des contrats à terme.

Puisque la littérature est sans équivoque par rapport au fait que les deux stratégies de minimisation du risque quadratique donnent approximativement les mêmes résultats pour les stratégies court terme, il serait intéressant de poursuivre la comparaison en définissant un meilleur modèle pour les rendements.

Puisqu'il a été démontré qu'il y a une dépendance entre les contrats à terme de différentes échéances, une nouvelle approche à considérer serait un modèle multivarié tel que

$$\begin{pmatrix} \omega_{t,t} \\ \omega_{t,t+1} \\ \omega_{t,t+2} \end{pmatrix} = \begin{pmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \end{pmatrix} + A \begin{pmatrix} \omega_{t-1,t-1} \\ \omega_{t-1,t} \\ \omega_{t-1,t+1} \end{pmatrix} + \begin{pmatrix} \sigma_{0,t} & 0 & 0 \\ 0 & \sigma_{1,t} & 0 \\ 0 & 0 & \sigma_{2,t} \end{pmatrix} \begin{pmatrix} z_{0,t} \\ z_{1,t} \\ z_{2,t} \end{pmatrix},$$

où A serait de la forme

$$\begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} \\ a_{1,0} & a_{1,1} & a_{1,2} \\ a_{2,0} & a_{2,1} & a_{2,2} \end{pmatrix}$$

au lieu de la forme originale

$$\begin{pmatrix} a_0 & 0 & 0 \\ 0 & a_1 & 0 \\ 0 & 0 & a_2 \end{pmatrix}.$$

De plus, le modèle proposé pour la demande d'électricité inclut un seul décalage dans le but de préserver la parcimonie et ainsi limiter la complexité numérique de

l'algorithme. Une nouvelle approche serait d'inclure une variable reliée à la météo, telle que la température ou les degrés-jours de chauffage (« *heating-degree days* »). Puisque l'erreur de prévision de la météo entraîne d'importantes erreurs de prévision dans la demande d'électricité, l'ajout d'une telle variable augmenterait la précision du modèle en indiquant lorsque la météo ne suit pas sa valeur moyenne. Étant donné que la stratégie couvre une courte période de trois semaines, il pourrait être vérifié si l'addition d'une variable reliée à la météo augmente la précision du modèle et si cette augmentation de précision vaut l'augmentation du temps de calculs.

ANNEXE A : PROGRAMMATION DYNAMIQUE

Nous avons premièrement utilisé l'algorithme de programmation dynamique présenté par Augustyniak et al. (2017) pour répliquer la solution obtenue dans le théorème 1.10.

L'algorithme comporte des fonctions qui peuvent être calculées explicitement selon les modèles ou être approximées avec la méthode Monte Carlo. Dans notre cas, la fonction génératrice des moments permet des formules explicites pour certaines de ces fonctions.

La méthode consiste à construire des grilles de valeurs possibles pour tous les intrants nécessaires à chaque pas dans le temps. Par exemple, la grille à $t = T$ nécessite :

- $\mathcal{L}_{T-1} \in [4, 100, 000 ; 8, 000, 000]$,
- $\omega_{T-1, T-1} \in [-0.85; 0.55]$,
- $\sigma_{0, T} \in [0; 0.6]$,
- $F_{T-1, T} \in [5; 110]$ et
- $F_{T-3, T} \in [5; 110]$.

Plus la grille contient de points, c'est-à-dire plus les intervalles sont courts, plus la méthode est précise. Chaque fonction de l'algorithme de programmation dynamique est calculée pour chaque point de la grille. La grille à $t = T - 1$ est calculée de la même manière avec 6 intrants. La valeur des fonctions explicites est facile à obtenir mais les valeurs des autres fonctions sont calculées avec la méthode Monte Carlo sur l'interpolation linéaire des valeurs de la grille $t = T$. Similairement, la compilation de la grille à $t = T - 2$ utilise des interpolations de la grille $t = T - 1$ précédemment calculée.

Ainsi, à partir des valeurs connues à $t = T - 3$, il est possible d'obtenir la position $(\theta_{T-2}^0, \theta_{T-2}^1)$ avec une interpolation des valeurs sur la grille $t = T - 2$. Le fournisseur d'électricité répètera cette méthode la semaine suivante et celle subséquente pour connaître $(\theta_{T-1}^0, \theta_{T-1}^1)$ et (θ_T^0, θ_T^1) , respectivement.

La complexité du modèle nécessitant jusqu'à 7 intrants dans la grille à $t = T - 3$, nous avons dû changer de méthode lorsqu'il est venu le temps d'augmenter la précision des grilles car ceci exigeait trop de mémoire. Par exemple, une grille à $t = T - 3$ de 7 intrants divisés en seulement 10 intervalles demande d'enregistrer la valeur de $10^7 = 10,000,000$ de données pour chaque fonction.

De plus, le modèle étant composé d'une copule variant dans le temps, ces grilles doivent être réestimées pour chaque stratégie de couverture, c'est-à-dire les grilles calculées pour la semaine du 7 janvier 2008 ne peuvent être utilisées pour la stratégie de couverture de la semaine du 14 janvier 2008.

ANNEXE B : MARCHÉ D'ÉLECTRICITÉ PJM

L'objectif initial de la recherche était de répliquer la stratégie de couverture globale présentée dans l'article de Dupuis et al. (2016) sur les données d'un autre marché d'électricité déréglementé. Il semblait intéressant de voir si les modèles présentés dans l'article pouvaient s'appliquer à un autre marché qui n'a pas nécessairement les mêmes sources de production d'électricité.

Le point Western Hub dans le marché Pennsylvania-New Jersey-Maryland (PJM) est un des points les plus actifs dans le monde du trading d'énergie. Il est référencé pour les transactions physiques d'électricité aussi bien que pour les transactions financières, dont les contrats à terme. Comme Dupuis et al. (2016) se basent sur les contrats les plus transigés, nous avons choisi des contrats liquides pour lesquels le règlement se fait en référence à ce point.

Intercontinental Exchange (ICE) est une place boursière spécialisée dans les produits dérivés, dont les produits financiers liés à l'électricité. Puisque ICE a modifié ses règlementations en lien avec les produits dérivés sur l'électricité en 2013, il semble cohérent d'utiliser les données qu'à partir de 2013.

Les contrats de NASDAQ OMX utilisés par Dupuis et al. (2016) ont un format 24h alors que les contrats de ICE pour PJM Western Hub se divisent en deux catégories : « *on-peak*¹ » et « *off-peak*² ». De plus, le règlement des contrats de NASDAQ OMX se fait sur le prix du système de Nord Pool qui est un prix « *day-ahead* » alors que dans le cas de PJM, nous avons le choix entre des contrats dont le règlement se fait sur le prix « *day-ahead* » ou sur le prix « *real-time* ».

Les statistiques des transactions de ICE depuis 2013 démontrent que les contrats journaliers ne sont pas liquides au-delà de 1 journée avant la maturité et que les contrats de fin de semaine sont sans intérêt au-delà d'une semaine. Nous pouvons considérer les contrats hebdomadaires couvrant du lundi au vendredi qui sont transigés entre 59% et 90% des jours selon la maturité du contrat variant d'une à trois semaines d'échéance.

1. Lundi au vendredi, excluant les fériés définis par la NERC, de 07 :00 à 23 :00 EPT.

2. Toutes les heures qui ne sont pas « *on-peak* ».

En catégorisant les contrats en termes de « *on-peak/off-peak* » et de prix « *day-ahead/real-time* », nous trouvons que les contrats « *off-peak* » sont extrêmement moins liquides que les contrats « *on-peak* » et que les contrats « *real-time* » sont plus liquides que les contrats « *day-ahead* ». Ces résultats sont probablement dus à un intérêt plus grand pour la gestion du risque de la volatilité présente en « *real-time* » lorsque la demande est plus élevée en période « *on-peak* ».

Le premier objectif de la recherche était d'établir si ces données de PJM de 2013 à la mi-2019 présentaient les mêmes caractéristiques que les données de Nord Pool.

Le test augmenté de Dickey-Fuller (ADF) est un test statistique permettant de décréter si une série temporelle est stationnaire ou pas. À titre de rappel, un processus x_t est dit fortement stationnaire si la suite $(x_{t_1+h}, \dots, x_{t_n+h})$ a la même loi de probabilité que la suite $(x_{t_1}, \dots, x_{t_n}) \forall n, h$. Dans la pratique, la définition de stationnarité au sens faible, aussi appelée stationnarité au second ordre, est généralement utilisée. Celle-ci dit qu'une série est stationnaire (au sens faible) si son espérance, sa variance ainsi que son autocorrélation ne varient pas dans le temps.

Le test le Ljung-Box est un test statistique qui vérifie la présence d'autocorrélation d'ordre supérieur à 1. Il détermine si le modèle présente un important manque d'ajustement en évaluant l'autocorrélation entre les résidus. Le test fournit par R mentionne que l'hypothèse nulle du test est l'indépendance de la série temporelle. Ainsi, une petite p – *value* signifie le rejet de l'hypothèse d'indépendance.

Suite aux tests de stationnarité et d'autocorrélation sur diverses définitions de rendements, nous observons dans la Table B1 que la série temporelle des rendements des contrats à terme d'échéance 3 semaines semble la plus prometteuse en terme de stationnarité et d'autocorrélation. Nous testerons la modélisation du log-rendement.

À l'aide de la fonction `auto.arima` de R qui identifie le modèle $ARIMA(p,d,q) \times (P,D,Q)_S$ le plus adéquat pour une série temporelle, nous trouvons pour les log-rendements un modèle $ARIMA(2,0,2)(2,0,0)[52]$ pourvue d'une moyenne non-nulle. Par contre, tel qu'illustré par la Figure B1, l'ajustement ne semble pas convenir, surtout en période de volatilité. De plus, les résidus observés dans la Figure B2 illustrent de

RENDEMENT	STATIONNARITÉ			AUTOCORRÉLATION		
	1 semaine	2 semaines	3 semaines	1 semaine	2 semaines	3 semaines
$\ln\left(\frac{F_T}{F_{T-1}}\right)$	OUI	OUI	OUI	NON	NON	OUI
$\frac{F_T}{F_{T-1}}$	NON	NON	NON	NON	OUI	OUI
$\left(\frac{F_T}{F_{T-1}}\right)^2$	NON	OUI	OUI	NON	OUI	OUI
$\frac{(F_T - F_{T-1})}{F_{T-1}}$	OUI	OUI	OUI	NON	OUI	OUI
$\left(\frac{(F_T - F_{T-1})}{F_{T-1}}\right)^2$	OUI	OUI	OUI	NON	OUI	OUI
$e^{\frac{F_T}{F_{T-1}}}$	NON	OUI	OUI	NON	OUI	OUI
$F_T - F_{T-1}$	OUI	OUI	OUI	OUI	NON	OUI

TABLE B1 – Tests de stationnarité et d'autocorrélation

l'hétéroscédasticité.

Nous poursuivons les tests en omettant la grande volatilité du début de 2014. Nous entreprenons l'analyse avec les données de la semaine commençant le 19 mai 2014. Le modèle proposé par auto.arima est ARIMA(0,0,0). En prenant l'ensemble des données commençant le 12 mai 2014, la fonction auto.arima nous donne un modèle ARIMA(2,0,1)(1,0,0)[52].

Les données semblent être instables. De plus, lorsque vient le temps de prédire une période plus tard, la prévision n'est pas satisfaisante et si on commence notre analyse après le début 2014, les données ne semblent plus autocorrélées.

Pour les contrats à terme d'une semaine d'échéance, le seul rendement qui était à la fois stationnaire et autocorrélé était la différence $F_T - F_{T-1}$. La fonction auto.arima de R nous donne un modèle ARIMA(3,0,3)(2,0,1)[52] pourvue d'une moyenne non-nulle dont l'ajustement laisse fortement à désirer tel qu'observé dans la Figure B3 et la Figure B4.

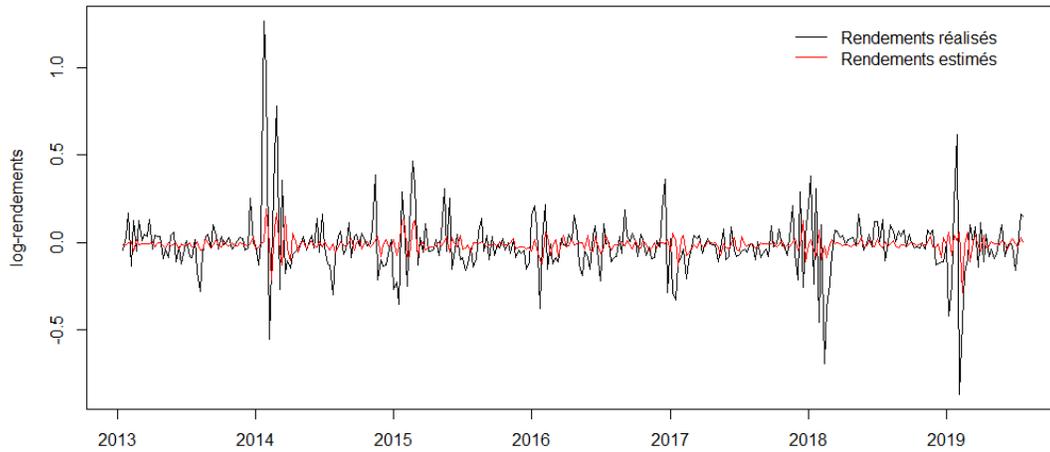


FIGURE B1 – Rendements réalisés vs estimés d'un modèle $ARIMA(2,0,2)(2,0,0)[52]$ pour les rendements des contrats à terme d'échéance 3 semaines

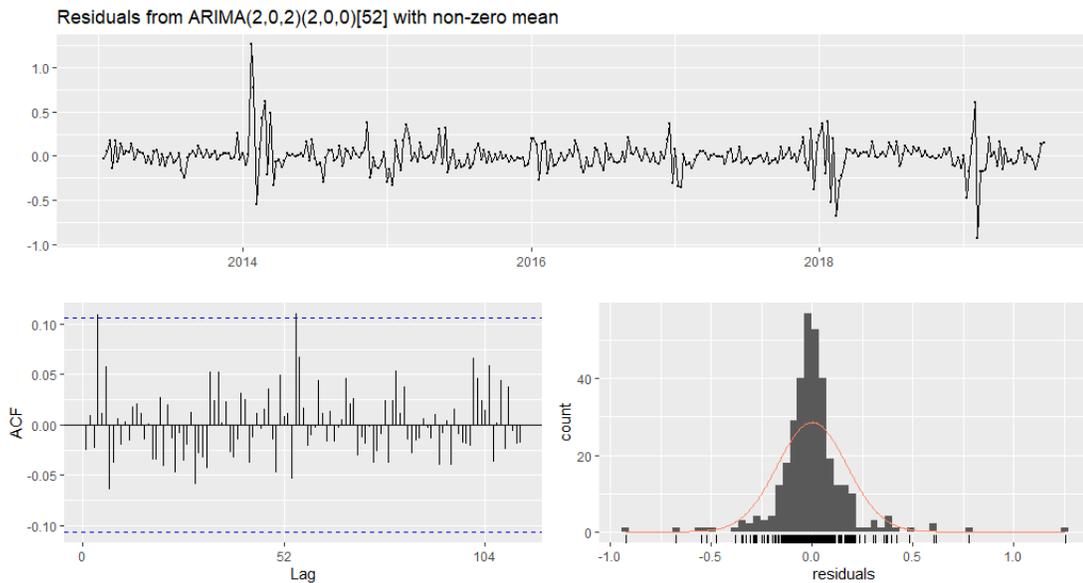


FIGURE B2 – Étude des résidus du modèle $ARIMA(2,0,2)(2,0,0)[52]$ pour les rendements des contrats à terme d'échéance 3 semaines

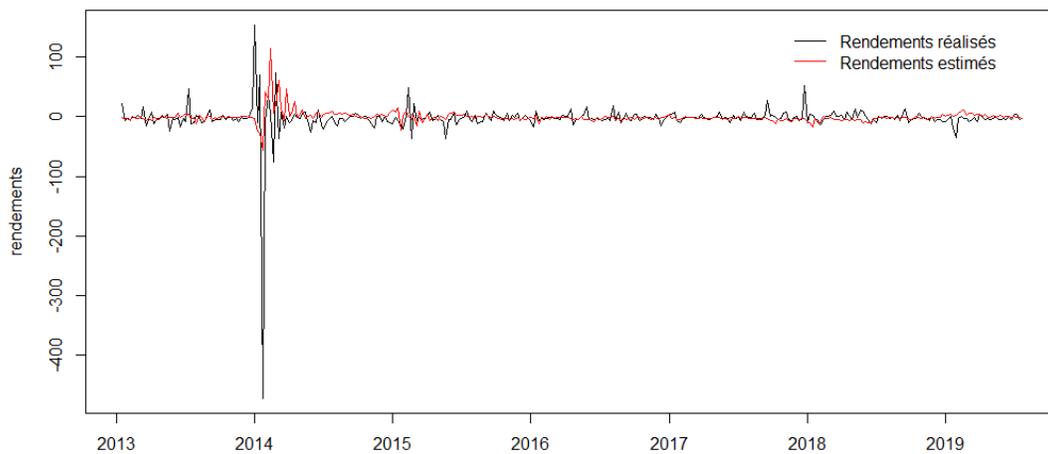


FIGURE B3 – Rendements réalisés vs estimés d'un modèle $ARIMA(3,0,3)(2,0,1)[52]$ pour les rendements des contrats à terme d'échéance 1 semaine

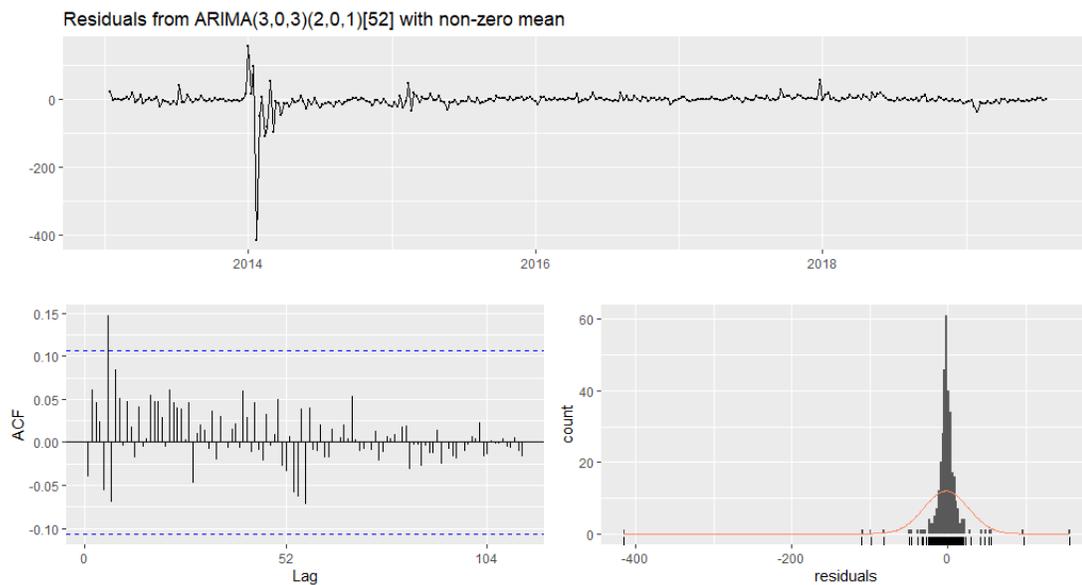


FIGURE B4 – Étude des résidus du modèle $ARIMA(3,0,3)(2,0,1)[52]$ pour les rendements des contrats à terme d'échéance 1 semaine

BIBLIOGRAPHIE

- M. Augustyniak, F. Godin, and C. Simard. Assessing the effectiveness of local and global quadratic hedging under garch models. *Quantitative Finance*, 17 (9) :1305–1318, 2017.
- F. Benth, J. Benth, and S. Koekebakker. *Stochastic Modelling of Electricity and Related Markets*. World Scientific, Singapore, 2008.
- N. Bouleau and D. Lamberton. Residual risks and hedging strategies in markovian markets. *Stochastic Processes and their Applications*, 33 :131–150, 1989.
- A. Černý and J. Kallsen. Hedging by sequential regressions revisited. *Math. Finance*, 19(4) :561–617, 2009.
- D. Duffie and H. R. Richardson. Mean-variance hedging in continuous time. *Annals of Applied Probability*, 1 :1–15, 1991.
- D. Dupuis, G. Gauthier, and F. Godin. Short-term hedging for an electricity retailer. *The Energy Journal*, 37(2) :31–59, 2016.
- N. El Karoui and M. Quenez. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*, 33(1) :29–66, 1995.
- H. Föllmer and A. Schied. *Stochastic Finance : An Introduction in Discrete Time*. Walter de Gruyter, Berlin, 2002.
- H. Föllmer and M. Schweizer. Hedging by sequential regression : An introduction to the mathematics of option trading. *ASTIN Bulletin*, 18 :147–160, 1988.
- H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. In *Contributions to mathematical economics*, pages 205–223. North-Holland, 1986.

- E. Gonzalez-Romera, M. A. Jaramillo-Moran, and D. Carmona-Fernandez. Monthly electric energy demand forecasting with neural networks and fourier series. *Energy Conversion and Management*, 49(11) :3135–3142, 2008.
- J. Harrison and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3) :381–408, 1979.
- M. Haugh. An introduction to copulas. Technical report, Columbia University, 2016.
- D. Heath, E. Platen, and M. Schweizer. A comparison of two quadratic approaches to hedging in incomplete markets. *Math. Finance*, 11(4) :385–413, 2001a.
- D. Heath, E. Platen, and M. Schweizer. Numerical comparison of local risk-minimisation and mean-variance hedging. In *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, pages 509–537. Cambridge University Press, Cambridge, 2001b.
- I. Karatzas and S. Shreve. *Methods of Mathematical Finance*. Number 39 in Applications of Mathematics. Springer, New York, 1998.
- N. E. Regulators. Nordic market report 2007. <http://www.nordicenergyregulators.org/wp-content/uploads/2013/02/market-report-2007.pdf>, 2007.
- M. Schweizer. *Hedging of Options in a General Semimartingale Model*. PhD thesis, ETH Zurich 8615, 1988.
- M. Schweizer. Mean-variance hedging for general claims. *Annals of Applied Probability*, 2 :171–179, 1992.
- M. Schweizer. Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, 20 :1–32, 1995.
- M. Schweizer. A guided tour through quadratic hedging approaches. In *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, pages 538–574. Cambridge University Press, Cambridge, 2001.

- I. Vehvilainen and T. Pyykkonen. Stochastic factor model for electricity spot price : The case of the nordic market. *Energy Economics*, 27(2) :351–367, 2005.
- T. Wiecki. An intuitive, visual guide to copulas. <https://twiecki.io/blog/2018/05/03/copulas/>, 2018.