Risk valuation for weather derivatives related to the energy market

École doctorale MSTIC

Discipline : Mathématiques

Thèse CIFRE, Ecole Doctorale MSTIC, n°532
CERMICS, Ecole des Ponts, Marne-la-Vallée, France.
AXA Climate, Paris, France.

Thèse soutenue le 11 Janvier 2024, par
Nerea VADILLO

Composition du jury:

Ying, JIAO	*Rapporteur*
Professeure des universités, Université Claude Bernard Lyon 1	
Carlo, SGARRA	*Rapporteur*
Professore, Università degli Studi di Bari	
Agnès, SULEM	*Examinatrice*
Directrice de recherche, Centre de Recherche Inria Paris	
Peter, TANKOV	*Examinateur*
Professeur, ENSAE, Institut Polytechnique de Paris	
René, AID	*Examinateur*
Professeur des universités, Université Paris-Dauphine, PSL Research University	
Aurélien, ALFONSI	*Directeur de thèse*
Directeur adjoint, CERMICS, Ecole des Ponts, MathRisk, Inria	
Amaury, DUFETEL	*Invité*
Directeur de l'assurance paramétrique, AXA Climate	

Préambule

Face à une demande croissante d'outils financiers permettant de transférer du risque climatique, les dérivés climatiques ont connu un essor important parmi les chercheurs et les professionnels. Cette thèse contribue à l'approfondissement des techniques d'évaluation de risques des produits dérivés liés à la température.

Dans le Chapitre 1, nous développons un nouveau modèle de volatilité stochastique pour la température journalière moyenne. Ce modèle constitue un élargissement du modèle classique d'Ornstein-Uhlenbeck proposé par Benth et Benth [16]. Il nous permet d'être plus conservateur en ce qui concerne les événements extrêmes tout en conservant des méthodes de calcul numérique efficaces. Nous estimons les paramètres du modèle à partir de la méthode des moindres carrés conditionnels sur une base de données incluant huit grandes villes européennes. Nous montrons ensuite comment obtenir efficacement la distribution des paiements des dérivés par des techniques de Monte-Carlo et de transformée de Fourier. Ce nouveau modèle permet de mieux capter le risque lié à la volatilité de la température.

Dans le Chapitre 2, nous nous concentrons sur les dérivés hybrides, appelés quantos, et liant prix de l'électricité et température moyenne journalière. Ces produits connaissent un grand succès puisqu'ils permettent de se couvrir à la fois contre les risques volumétriques et les risques de prix. Nous développons un modèle couplé. Les sous-jacents sont modélisés par des processus d'Ornstein-Uhlenbeck non homogènes entraînés par un mouvement Brownien et un processus de Lévy Normal Inverse Gaussien, permettant d'inclure la dépendance entre eux. Une méthode des moindres carrés conditionnels est développée pour estimer les différents paramètres du modèle et appliquée sur des données réelles. Ensuite, nous développons des formules explicites et semi-explicites des espérances de paiements des dérivés, y compris des options quantos. Ces résultats sont comparés à des simulations de Monte Carlo. Enfin, nous développons des formules explicites pour couvrir statiquement les options quantos simples et doubles par un portefeuille d'options d'électricité et d'options de température (CDD ou HDD).

Dans l'ensemble, cette étude contribue à l'établissement d'un cadre mathématique permettant de mieux comprendre le risque lié aux dérivés de température et d'énergie tout en répondant aux problématiques des professionnels et de la recherche.

Mots clés : Evaluation de risque, Dérivés climatiques, Température, Quantos, Décomposition du risque.

Preamble

With the increasing demand for climate risk transfer financial tools, climate derivatives have gained considerable popularity among both researchers and practitioners. This thesis contributes to deepening the understanding of the risk related to temperature and energy based derivatives.

In Chapter 1, we develop a new stochastic volatility model for the average daily temperature that is a natural extension of the Ornstein-Uhlenbeck model proposed by Benth and Benth [16]. This model allows to be more conservative regarding extreme events while keeping tractability. We give a method based on Conditional Least Squares to estimate the parameters on daily data and estimate our model on eight major European cities. We then show how to calculate efficiently the average payoff of weather derivatives both by Monte-Carlo and Fourier transform techniques. This new model allows to better assess the risk related to temperature volatility.

In Chapter 2, we focus on quanto derivatives. We develop a coupled model for day-ahead electricity prices and average daily temperature which allows to model quanto weather and energy derivatives. These products are quickly spreading as they enable to hedge against both volumetric and price risks. Electricity day-ahead prices and average daily temperatures are modelled through non homogeneous Ornstein-Uhlenbeck processes driven by a Brownian motion and a Normal Inverse Gaussian Lévy process, which allows to include dependency between them. A Conditional Least Square method is developed to estimate the different parameters of the model and used on real data. Then, explicit and semi-explicit formulas are obtained for derivatives including quanto options and compared with Monte Carlo simulations. Last, we develop explicit formulas to hedge statically single and double sided quanto options by a portfolio of electricity options and temperature options (CDD or HDD).

All in all, this study contributes to the establishment of a mathematical framework to better understand risk related to temperature and energy derivatives while answering to both business and research challenges.

Keywords: Risk valuation, Weather derivatives, Temperature model, Joint Temperature-Electricity model, Risk hedging.

List of papers

Here is a list of papers that were written during this thesis:

- [8] Alfonsi, A., & Vadillo, N. (2022). A stochastic volatility model for the valuation of temperature derivatives. *arXiv preprint arXiv:2209.05918*
(in revision).

- [9] Alfonsi, A., & Vadillo, N. (2023). Risk valuation of quanto derivatives on temperature and electricity. *arXiv preprint arXiv:2310.07692*
(submitted).

Remerciements

Ce manuscrit n'aurait pas pu voir le jour sans le soutien d'Aurélien Alfonsi, directeur de cette thèse et coauteur des deux papiers liés à ce travail. Un grand merci Aurélien pour ta persistence, accessibilité et grand sens pédagogique.

Je voudrais remercier ensuite tout l'écosystème qui a rendu ce travail possible : l'environnement de recherche du CERMICS (mes co-doctorants, l'équipe enseignante, Isabelle et Stéphanie) mais aussi l'écosystème bouillonnant d'AXA Climate (les souscripteurs, les capitaines et ma chère équipe PPP). Un merci spécial à Amaury Dufetel qui a soutenu ce projet depuis son origination et l'a également rendu possible.

Par ailleurs, je suis reconnaissante au jury de thèse de s'être rendu disponible pour assister à la soutenance (Agnès Sulem, Peter Tankov et Réné Aïd). Merci notamment à Ying Jiao et Carlo Sgarra d'avoir pris le temps de poser un regard critique sur ce manuscrit.

Pour finir, un grand merci à mon entourage personnel : à la familia qui est loin mais toujours présente, à mon relecteur de prédilection et à mes amis sur lesquels on peut toujours compter.

Table des matières

Résumé détaillé

Le besoin de renforcer la gestion des risques climatiques, et, en particulier, de développer des instruments efficaces de transfert de ces risques constitue un des piliers du Mécanisme international des pertes et les dommages (WIM) [106] établi lors de la COP19. Bien que ces forums se concentrent surtout sur l'impact des événements climatiques extrêmes sur les populations vulnérables, l'accélération du changement climatique a augmenté l'exposition de notre société et de nos économies aux événements climatiques. En effet, les professionnels estiment que le changement climatique pourrait faire perdre jusqu'à 10% de sa valeur à l'économie mondiale d'ici 2050 [134]. Dans ce contexte, les dérivés climatiques sont apparus comme des outils financiers efficaces de transfert du risque. Notre étude se concentre sur l'amélioration de la compréhension quantitative du risque associé à ces instruments.

Les dérivés climatiques sont des instruments financiers donnant lieu à un transfert de risque lié à un péril climatique. Ces dérivés sont indexés sur des indices météorologiques et génèrent un paiement à maturité selon des conditions précontractualisées. En particulier, notre recherche s'intéresse aux dérivés climatiques liés à la température pour le secteur de l'énergie. L'intérêt de notre analyse a été, tout d'abord, d'établir un cadre mathématique propre à ces produits. Nous avons pour cela puisé dans la littérature mais également fait converger des approches financières, actuarielles et des praticiens. L'introduction de la thèse constitue une mise en contexte de notre étude. Ensuite, nous nous sommes intéressés aux dérivés indexés à la température. Nous avons ainsi développé un modèle de volatilité stochastique pour des températures journalières moyennes. Le Chapitre 1 présente ce modèle et développe une tarification possible de ce type de produits. Nous nous sommes ensuite intéressés aux quantos, dérivés indexés sur un indice climatique et sur le prix de l'énergie. Le Chapitre 2 expose le développement d'un modèle couplé pour la température moyenne journalière et le prix de l'énergie. Nous utilisons ensuite ce modèle pour évaluer les risques liés aux quantos.

Cette recherche a été développée dans le cadre d'une *Convention Industrielle de Formation par la Recherche (CIFRE)*. L'étude a donc permis un transfert de connaissances entre un acteur industriel (AXA Climate) et la recherche scientifique. AXA Climate est l'entité d'AXA en charge des produits dérivés météorologiques et des produits d'assurance indéxés sur ces indices. L'unité conçoit, tarife et vend ces couvertures. Le sujet de la thèse est donc au coeur de son activité, en particulier, pour les dérivés liés à la température.

Introduction aux dérivés climatiques

Un dérivé climatique est un contrat qui décrit un service de transfert de risque climatique entre un acheteur qui souhaite se couvrir contre ce risque et un vendeur qui accepte de porter ce risque. Le contrat est défini par trois éléments principaux :

- Un paramètre météorologique sous-jacent qui correspond à une grandeur physique suivie pour déterminer s'il y a un paiement à maturité.

- Un indice climatique qui correspond à l'agrégation du paramètre météorologique sous-jacent sur une certaine période de temps contractuellement prédéfinie.

- Une fonction de paiement qui relie l'indice météorologique au paiement monétaire. Nous retrouvons, en général, les fonctions de paiement communes à d'autres marchés dérivés : des contrats à terme, des swaps et des options.

Le marché des dérivés climatiques est né en 1996 avec une première transaction entre Enron et Koch. En 1999, face à la demande croissante de solutions de transfert du risque météorologique, la Chicago Mercantile Exchange (CME), principal marché nord-américain des commodités, lance dix dérivés mensuels liés à la température sur des villes aux États-Unis. La même année, la Weather Risk Management Association (WRMA), organisation réunissant les acteurs du marché, est créée. Le marché connaît un grand succès dans les années 2000. En 2006, la CME rapporte des transactions dans 47 villes du monde entier et le volume des sommes en jeu s'élève à 45 milliards de dollars selon la WRMA [147]. Néanmoins, la crise de 2008 fait chuter la demande et il faut attendre les années 2020 pour voir le marché repartir. En 2020, le CME fait état d'une augmentation de 60% du volume sur les contrats à terme [53] et de nouveaux marchés nationaux s'y joignent comme le Zhengzhou Commodity Exchange [152].

Aujourd'hui, le marché est divisé en deux domaines : un marché ouvert avec des produits standardisés et un marché de gré à gré avec des contrats hautement ajustables. Le marché ouvert standardisé se réduit principalement aux produits dérivés basés sur la température du CME qui propose des contrats à terme et des options pour HDD, CDD et CAT pour 19 villes dans le monde. Malheureusement, les volumes d'échanges quotidiens restent considérablement limités [146] avec plusieurs jours sans transactions. La plupart des échanges se font de gré à gré. Plusieurs acteurs semblent être actifs sur ce marché, allant des courtiers spécialisés aux principaux réassureurs. Enfin, il est surtout difficile de recueillir des informations sur le volume des transactions de dérivés associés à cause de l'absence d'un marché ouvert.

Quant aux clients, la plupart des secteurs économiques sont exposés aux aléas climatiques. Les principaux acheteurs de ces produits sont certainement les énergéticiens. Dans ce secteur, les dérivés climatiques sont utilisés pour se couvrir contre les risques volumétriques : vagues de froid et de chaleur qui provoquent des pics de demande, manque d'intensité solaire ou de vent qui réduisent la production d'électricité... De même, dans d'autres secteurs, les dérivés climatiques sont utilisés pour se protéger contre les mauvaises récoltes, les dommages matériels ou les interruptions d'activité dus à des conditions météorologiques défavorables.

L'évaluation du risque lié aux dérivés climatiques

Le sujet de cette thèse est la compréhension des risques liés aux dérivés climatiques. L'objectif premier était de définir cette notion de risque afin d'aboutir à une tarification convaincante des dérivés climatiques.

Pour établir un cadre mathématique, nous avons confronté plusieurs cadres de compréhension de risque :

- Une approche assurantielle qui s'établit à partir des notions de mesures de risque et des principes de calcul de prime, le tout dans un cadre de probabilité historique.

- Une approche financière qui émane de la théorie d'évaluation financière et qui suppose un marché sans arbitrage et complet.

- L'approche des professionnels qui fait preuve de pragmatisme puisqu'elle consiste à travailler avec les indices annuels mais pêche par manque de robustesse.

Puisque les paramètres météorologiques ne peuvent pas être échangés, le marché des dérivés climatiques ne répond pas aux conditions de non-arbitrage et complétude du cadre classique des mathématiques financières. Nous avons donc emprunté des éléments à toutes les approches ci-dessus afin de proposer le cadre le plus complet possible de cette vision du risque. Nous avons, par la suite, surtout travaillé avec la distribution des paiements générés à partir de l'estimation des modèles de sous-jacents sous probabilité historique.

Nos deux chapitres sont organisés de manière similaire. Nous développons tout d'abord un modèle convaincant pour illustrer les dynamiques des sous-jacents ; température journalière pour le Chapitre 1 et température journalière et prix de l'énergie pour le Chapitre 2. Nous justifions et estimons les paramètres de ces modèles au vu des propositions de la littérature et de différentes techniques d'estimation. Nous travaillons ensuite sur les distributions des paiements pour plusieurs fonctions de paiement et sur notre capacité à obtenir des formules explicites et semi-explicites pour la tarification.

Chapitre 1 : Un modèle de volatilité stochastique pour les dérivés climatiques liés à la température

Le Chapitre 1 étudie la dynamique des températures journalières moyennes et son application dans le domaine des dérivés climatiques. En étendant un des modèles les plus populaires de la littérature [28], nous proposons un modèle autoregressif de volatilité stochastique pour le processus de température $(T_t)_{t\geq0}$. Notre modèle élargit les modèles existants tout en capturant les queues de distribution qui n'étaient pas bien prises en compte par les modèles précédents.

Soit $(T_t)_{t\geq0}$ la température journalière moyenne, $(\zeta_t)_{t\geq0}$ sa volatilité et $(\mathcal{F}_t)_{t\geq0}$ une filtration adaptée à ces deux processus. Nous supposons que $(T_t)_{t\geq0}$ et $(\zeta_t)_{t\geq0}$ suivent les dynamiques du

Modèle (M) :

$$
\begin{cases}
T_t & = s(t) + \tilde{T}_t, \\
d\tilde{T}_t & = -\kappa\tilde{T}_t dt + \sqrt{\zeta_t}(\rho dW_t + \sqrt{1-\rho^2}dZ_t), \\
d\zeta_t & = -K(\zeta_t - \sigma^2(t))dt + \eta\sqrt{\zeta_t}dW_t,
\end{cases}
\tag{M}
$$

où $(W_t)_{t\geq 0}$ et $(Z_t)_{t\geq 0}$ sont des mouvements Browniens indépendants, $\kappa, \eta, K > 0$, $\rho \in [-1, 1]$, σ^2 est une fonction non négative et les fonctions s et σ^2 sont définis en Equation (1.2). $(\mathcal{F}_t)_{t\geq 0}$ est alors la filtration générée par (W, Z). Ce modèle constitue une adaptation du célèbre modèle de Heston [84]. Il permet de dépasser le modèle classique d'Ornstein-Uhlenbeck [3] [28] en intégrant de la flexibilité au processus de volatilité. Ainsi, ce processus aborde la problématique des résidus non-Gaussiens tout en maintenant la continuité temporelle et les propriétés de la diffusion. En particulier, le Modèle (M) est plus conservateur que les modèles classiques vis-à-vis des queues de distribution.

En termes d'estimation, notre principale contribution consiste à étendre les travaux de Bolyog et Pap [34] sur l'estimation à partir des moindres carrés conditionnels (CLS). En particulier, nous étendons la démonstration d'Overbeck et Ryden [123] sur la convergence forte des estimateurs CLS aux processus CIR aux cas non-homogènes dans le temps. Par ailleurs, nous estimons le Modèle (M) sur des températures moyennes journalières dans huit grandes villes européennes du 1er janvier 1980 au 31 décembre 2020.

La structure affine du Modèle (M) permet d'effectuer une tarification efficace grâce aux techniques de la transformée de Fourier rapide et en combinaison avec la méthode de la variable de contrôle. L'utilisation de la tarification par transformée de Fourier rapide pour les dérivés climatiques est peu fréquente, même si elle a déjà été développée [27]. Nous explicitons une méthodologie complète incluant la résolution de l'équation de Riccati autonome non-homogène pour effectuer cette tarification. Nos résultats sont comparés à des simulations de Monte Carlo et aux méthodes de tarification des praticiens, ce qui permet de démontrer la précision et l'efficacité des techniques développées.

Ce premier chapitre a donné lieu à des applications industrielles concrètes. Tout d'abord, il a permis de donner un cadre conceptuel aux pratiques de tarification. De plus, il a approfondi la compréhension du risque lié aux dérivés de température à partir de la comparaison des méthodes ainsi que des différentes analyses de sensibilité. Finalement, une version simplifiée du Modèle (M) a été mobilisée pour l'adaptation de la tarification des couvertures contre le risque de gel.

Ce chapitre correspond à la prépublication [8] qui est actuellement en révision à *IMA Journal of Management Mathematics*.

Chapitre 2 : Evaluation du risque lié aux quantos indéxés sur la température et le prix de l'énergie

Dans le Chapitre 2, nous nous intéressons aux dérivés combinant un indice de température et le prix de l'énergie. Nous développons pour cela un modèle couplé pour la tarification de ces

dérivés.

Soit $(T_t)_{t \geq 0}$ la température journalière moyenne, $(X_t)_{t \geq 0}$ le logarithme du prix spot du jour précédent tel que le prix spot $S_t = e^{X_t}$ et $(\mathcal{F}_t)_{t \geq 0}$ une filtration adaptée à ces deux processus. Ces deux processus sont gouvernés par le Modèle (ETM) :

$$\begin{cases} d(X_t - \mu_X(t)) &= -\kappa_X(X_t - \mu_X(t)) + \lambda \sigma_T dW_t^T + dL_t^X \\ d(T_t - \mu_T(t)) &= -\kappa_T(T_t - \mu_T(t)) + \sigma_T dW_t^T \end{cases} \tag{ETM}$$

où L^X est un processus de Lévy Normal Inverse Gaussien (NIG) de paramètres $(\alpha^X, \beta^X, \delta^X, m^X)$ supposé centré. L^X est indépendant de W^T, un bruit Brownien. $\mu_X(\cdot)$ et $\mu_T(\cdot)$ représentent la tendance et la composante saisonnière déterministe de $(X_t)_{t \geq 0}$ et $(T_t)_{t \cdot 0}$ respectivement, κ_X et κ_T des paramètres autorégressifs et λ un coefficient illustrant la dépendance entre les deux bruits.

Ce modèle est confronté à des données réelles. Pour l'énergie, nous considérons les prix spot du marché en France et en Italie du Nord du 5 janvier 2015 au 31 décembre 2018 provenant de la plateforme de transparence ENTSO-E et du Gestore Mercati Energetici (GME). Les données des températures moyennes sont celles des stations météorologiques de l'aéroport de Paris-Charles de Gaulle et de l'aéroport de Milan Linate.

Les premières sections du Chapitre 2 se concentrent sur la sélection des modèles. Tout d'abord, nous discutons du choix des lois marginales. En particulier, nous confrontons les données sur le prix de l'énergie aux modèles de la littérature et nous sélectionnons un processus Normal Inverse Gaussien autorégressif. Ensuite, nous développons différentes méthodes d'estimation basées sur l'estimation des moindres carrés conditionnels appliquée à la fonction caractéristique. La discussion concernant la température est moins approfondie, le lecteur pouvant se référer au Chapitre 1. Nous estimons le paramètre de dépendance λ et montrons ensuite que la dépendance entre les deux dynamiques est correctement prise en compte.

Nous nous penchons ensuite sur la question de la tarification. Nous considérons différentes structures de quantos et définissons son espérance des paiements sous une probabilité historique :

$$\mathbb{E}\Big(\sum_{t=t_1}^{t_2} f_S(S_t) \times f_T(T_t) \mid \mathcal{F}_{t_0} \Big)$$

où f_S et f_T représentent les fonctions de paiement liées au prix de l'énergie et à la température respectivement. Nous considérons que le prix du contrat est fixé à $t_0 \leq t_1$ et qu'il engendre un paiement à l'échéance t_2 avec $t_1 \leq t_2$.

Nous considérons différentes fonctions f_S et f_T afin de représenter des contrats à terme, des swaps, des options unilatérales et bilatérales sur chaque sous-jacent. Nous obtenons des formules explicites de l'espérance des paiements pour les options à terme, les swaps et les options unilatérales ($\mathcal{E}$-HDD et $\mathcal{E}$-CDD). Pour les options bilatérales, nous proposons une décomposition de Taylor au premier ordre en λ faisant intervenir des espérances pouvant être calculées à partir des formules de Carr Madan [48]. Ces formules sont confrontées aux paiements simulés par

Monte Carlo. Nous vérifions que les deux méthodes fournissent des résultats similaires et que les calculs explicites permettent de gagner en temps de calcul. Nous montrons également le rôle de λ dans l'évaluation de risque des produits dérivés.

Finalement, nous travaillons sur la décomposition du risque des $\mathcal{E}$-HDD et des quantos bilatéraux dans le cadre d'un portefeuille autofinancé. Nous réussissons à optimiser ce portefeuille à partir de formules explicites ou semi-explicites. Nous montrons que pour 100 000 portefeuilles simulés, nous parvenons à couvrir le dérivé quanto en moyenne et à diminuer la variance du portefeuille. Nous analysons également l'impact du paramètre λ.

La contribution de ce chapitre est considérable puisqu'elle répond à un besoin des praticiens auquel la littérature scientifique apportait peu de réponses. En effet, très peu d'experts se sont penchés sur des modèles combinant température et prix de l'énergie [44] [17] et ont exploré des dérivés sous forme de double-options. De même, la décomposition du risque sur les quantos est clé pour les porteurs de risque puisqu'elle permet une meilleure gestion du portefeuille ainsi que de répondre aux exigences de solvabilité.

Ce chapitre correspond à la prépublication [9] qui a été soumise.

Introduction

The need for enhanced climate risk management, and, in particular, the need to develop effective climate risk transfer instruments, emerged as one of the conclusions of the Warsaw International Mechanism for Loss and Damage (WIM) [106] established during COP19. While these forums focus particularly on the impact of extreme weather on vulnerable populations, it is a reality that the acceleration of climate change has increased our society's and economy's exposure to unexpected weather events. In fact, it is estimated that climate change could decrease up to 10% of total economic value worldwide by 2050 [134]. In this context, weather derivatives have emerged as effective climate risk transfer tools. Our study focuses on improving the quantitative understanding of the risk associated to these instruments.

The below provides an introduction to weather derivatives. First, we define these instruments to understand their history and current success. Second, we present different approaches to assess risk, contrasting the actuarial and financial approaches and describing business practices. Third, we propose a deep dive into temperature-based derivatives with a description of Chapter 1's contribution to the related challenges. Finally, we focus on the relationship between weather derivatives and energy markets. This would be the occasion to introduce the contributions of Chapter 2, where we design and evaluate a hybrid temperature and energy quanto derivative.

This research has been developed in the framework of a business-applied PhD contract known as *Convention Industrielle de Formation par la Recherche (CIFRE)*. Therefore, the study has enabled a knowledge transfer between an industrial actor (AXA Climate) and mathematical research. AXA Climate is the AXA entity in charge of weather derivatives and index-based insurance products. The unit designs, prices and sells these covers. The subject of the thesis is therefore at the core of its business, in particular for temperature-related derivatives.

1 An overview of the weather risk transfer market

1.1 History of weather derivatives

The weather risk transfer market began in 1996 when a first contract between Aquila Energy and Consolidated Edison stated that Aquila Energy would sell electricity at a discounted price if August temperatures were milder [112]. That same year, the first real weather derivative transaction was reported between Enron and Koch. In 1999, faced with increasing demand for weather risk transfer solutions, the Chicago Mercantile Exchange (CME) introduced standard-

ized monthly temperature contracts for 10 locations in the United States. At the same time, the Weather Risk Management Association (WRMA) was formed. It brought together energy companies, insurers, reinsurers, brokers and weather data providers for biannual meetings. WRMA acts as a market monitor and communicator.

In the early 2000s, the market grew rapidly between power and gas traders as a way to hedge volumetric risk in the context of volatile energy trading markets. While the market expanded in the U.S., Europe and Japan joined the trend. In 2001, the London International Financial Futures and Options Exchange (LIFFE) launched six daily temperature contracts for London, Paris and Berlin. Between 2001 and 2002, the total notional value (maximum payoff) of contracts traded in Europe increased from $49 million to $600 million [39]. In Asia, the market grew also with many non-energy market participants entering the Japanese weather market and the first transactions taking place in Australia. In 2003, the WRMA reported more than $10-billion worth of transactions in weather markets [90].

By 2006, CME was reporting transactions in 47 cities around the world with increasing success [52] and the notional value of the market was $45 billion according to the WRMA survey. At the time, it was a market dominated by 95% temperature contracts, of which 50% were monthly degree-day futures [146] [147]. The following years showed a slowdown in the market with $19.2 billion notional traded in 2007, $32 billion in 2008 and only $11.8 billion in 2011 [148]. In 2014, snowfall derivatives were removed from the market. While the 2008 crisis is in part responsible for the market shrunk, Pérez-Gonzalez and Yun [125] argue that the birth of hybrid derivatives, combining weather and commodities, also explains part of this decline.

However, the beginning of 2020 and the increasing impact of extreme weather has led to a rebirth of the weather derivatives market. In 2020, the CME reported an increase of 60% in notional volume traded on futures and 143% on options compared to 2019, resulting in $750 million and $480 million respectively [53]. In parallel, China also joined the market in July 2021 by launching weather futures on the Zhengzhou Commodity Exchange [152]. In 2023, with the weather outlook, the CME reported increasing open interest (Figure 1) and expanded its weather derivative futures to new cities [51].

All in all, after a rapid growth and a slowdown at the beginning of the twenty-first century, the weather derivatives market is a market that responds to the need for a financial solution to the increasing exposure to weather risks in the global economy.

1.2 Market description

Nowadays the market is divided in two fields: an open market with standardized products and an over-the-counter (OTC) market with highly adjustable contracts.

The standardized open market which presents the advantage of decreased cost and increased transparency is mainly reduced to the CME temperature-based derivatives. The CME marketplace offers futures and options for HDD, CDD and CAT for 19 cities worldwide and two seasons which can vary from contract to contract. Unfortunately, daily trading volumes remain
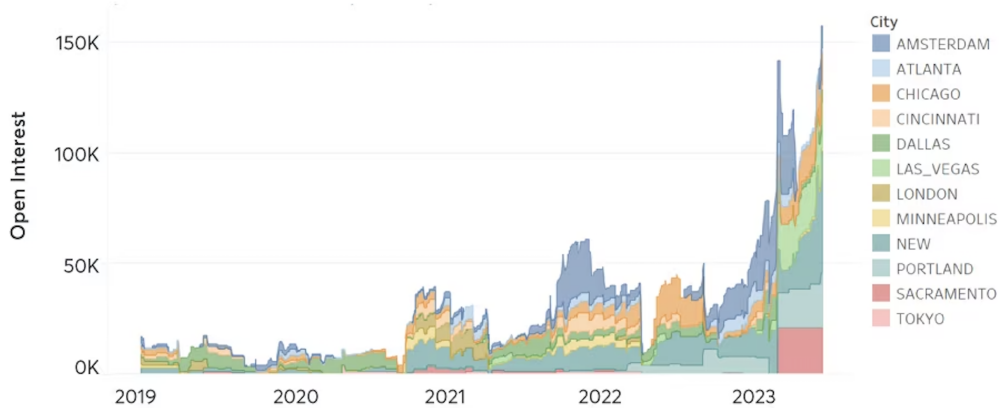
Figure 1: CME Group reported open interest in weather futures and options [51].

considerably limited [146] with several days without transactions. Other markets, like LIFFE, have practically removed weather derivatives from their products.

Most of the exchanges are made OTC. In 2004-2005, WRMA estimated that 69% of the weather exchanges were OTC [146]. Several actors seem to be active on this market from specialized brokers to top reinsurers as show the Annual Market Rankings [67]. The products sold in this market are however less standardized and structured by the broker or the reinsurer to meet the buyer's needs. Finally, it is difficult to gather information on the volume of weather derivative transactions.

1.3 Contract definition and characterisation

A weather derivative is a contract that describes a weather risk transfer service between parties that wish to hedge against this risk. The contract is defined by three main elements:

- An **underlying meteorological parameter** which corresponds to a physical quantity that is monitored to determine whether there will be a payout. In our study, although other parameters will be presented in the introduction, the main underlying meteorological parameter is the average daily temperature. The latter is defined as the average between the daily minimum and maximum temperatures.

- A **weather index** which corresponds to the aggregation of the underlying meteorological parameter over the risk or contract period. Different aggregation formulas are presented in the following sections. The reader should note that the aggregation method is contractual, as the calculation of the weather index from the measured meteorological parameter at maturity is not open to interpretation.

- A **payoff function** which links the weather index to the monetary payoff. This function

is also contractual, so the payoff at maturity is objective. In terms of payoff functions, we can cite the payoff functions common to other derivative markets; futures, swaps, and options. Some payoff functions are shown in Figure 5. First, it can be noted that for some payoff functions, such as futures and options, a buyer-seller relationship is established that results in a premium transaction at the time of contracting. For swaps, the premium payment depends on the symmetry of the payoff function. Second, the payoff functions shown in Figure 5 are all bounded by a positive value L. This corresponds to a common market practice where claims are capped to a payoff limit of L.
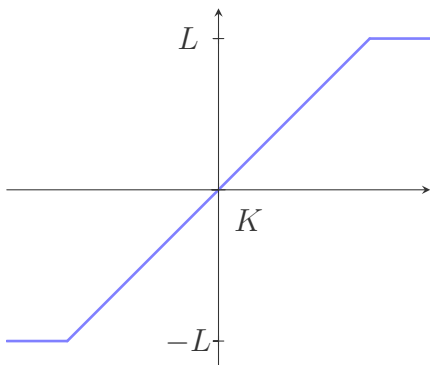
Figure 2: Swap $P(I) = \max(-L, \min(L, \alpha * (I - K)))$
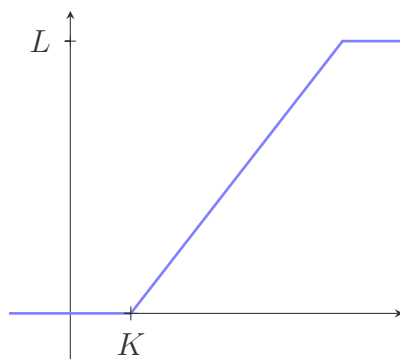
Figure 3: Call $P(I) = \min(L, \alpha * (I - K)^+)$
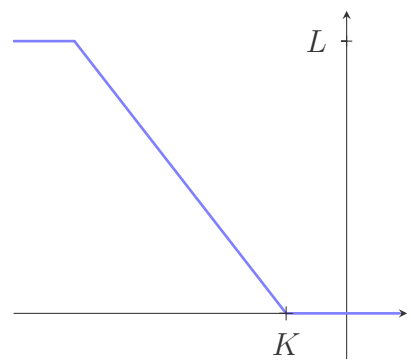
Figure 4: Put $P(I) = \min(L, \alpha * (K - I)^+)$

Figure 5: Payoff functions for three weather derivatives. Here I corresponds to the weather index, K to the strike, α to the notional or tick in (\$ per weather index unit) and L to the limit of payoff.

Apart from the above, other elements also characterize a weather derivative contract:

- The risk or contract period defines from which date to which date the weather parameter is aggregated to calculate the weather index. Similar to European derivatives, the weather derivatives are exercised at maturity. The index is accumulated through the whole contract period and the amount of the claim can only be computed, and hence paid, at maturity.

- The weather data source corresponds to the weather station or official website from which the weather measurements can be extracted. In general, for individual risks and in developed countries, it corresponds to the identification of the weather stations as reported by the World Meteorological Organization (WMO) or the country's public authority. For non-developed countries, where weather stations are not available, it can correspond to the satellite data provider, usually a spacial agency (NASA, ESA...) or a university (Columbia). For specific hazards such as tropical cyclones, the tropical cyclone track is monitored by national agencies (NOAA, JMA, HKO, BOM) which provide official tracks.

- The settlement data provider is a third party to the contract that delivers the data measurements at maturity. This figure certifies the quality and objectivity of the measured data. This data quality control can add some time to the claim settlement process. The data settlement period is also contractually specified so that the buyer is certified to receive the claim, if any, in a period after maturity. In addition, it should be noted that the agency providing the weather measurements and the settlement data provider can be two different entities. For example, the agency may be the national agency that manages the network of weather stations, while the settlement data provider may be a private company that certifies the readings of a particular weather station involved in the contract.

The above describes the main features of a weather derivative contract. Subsection 1.5 breaks these down later into different weather perils.

1.4 Advantages of weather derivatives

In the energy sector, weather derivatives are used to hedge against volumetric risks: cold and heat waves that cause peaks in demand, lack of solar or wind intensity that reduces electricity generation... Similarly, in other sectors, weather derivatives are used to hedge against crop failure, property damage or business interruption due to adverse weather conditions. In all of the above cases, the weather index is not the loss experienced by the buyer, but a proxy for the potential financial loss. There may be a difference between the financial losses predicted by the index and the actual losses incurred by the company.

These products present however several advantages:

- Transparency and objectivity: The underlying weather parameter is independent of human activity in the short term. Therefore, neither the buyer nor the seller can influence the course of this parameter. In addition, the weather data is often public and, in our case, certified by a third party data provider. This provides high transparency and prevents information asymmetries.

- Moral hazard: Similarly, the behavior of the buyer cannot, in principle, influence the outcome of the index at maturity. This reduces the risk of fraud and moral hazard for the seller of the contract. In addition, the role of the settlement data provider is to ensure that there has been no manipulation of the data or the instruments used to capture it. The buyer is not incentivized to increase its risk.

- Payment speed: Measured weather data is available shortly after maturity. Provided there is adequate time to monitor data quality, claims can be settled very quickly. While this is not different from other commodity derivatives, it can make an important difference for crop failure, property damage or business interruption. In these cases, recovery times are much shorter.

- Lower and shorter transaction costs: Index-based contracts rely on the index value at maturity to trigger a claim, unlike other risk transfer contracts that require a loss adjuster to confirm the loss level. This is particularly true for agricultural and property losses. In these cases, claims settlement can involve lengthy and costly negotiations, including field visits, additional intermediaries and litigation costs. Weather derivatives avoid these post-disaster transaction costs by providing a fully transparent claims process.

The above points are often used to explain the success of weather derivatives, particularly against alternative risk transfer products such as indemnity-based insurance. While they can partly explain the growing success of the weather derivatives market, other factors should not be forgotten. In particular, the increasing availability of weather data through the development of satellite data and the efforts of national meteorological agencies to share available data. The increasing frequency of extreme weather events and the greater volatility of seasonal weather have also raised the awareness of those exposed to these risks.

1.5 Who and what is covered by weather derivatives

Almost any weather risk can be covered by a weather derivative, provided there is a seller. But who are the main buyers of such products? We propose below a non-exhaustive list of actors that may be affected by weather perils:

- Energy companies, which manage the risk of reduced revenues in mild weather conditions, given that the amount of energy sold depends heavily on consumer demand, which is driven by temperature.

- Agricultural production and transportation companies that are particularly sensitive to meteorological conditions that result in reduced yields.

- Retailers whose sales are sensitive to weather conditions. For example, sales of beer in the summer may decline due to cooler weather [90].

- Leisure infrastructure such as hotels and amusement parks that depend on favorable weather conditions.

- Financial actors (investment banks, asset managers, insurers and reinsurers) that are exposed to the impact of weather conditions on their portfolios.

There are many weather adversities that can be covered by the above. Table 1 provides a non-exhaustive summary of the diversity of weather derivative underlyings.

However, some perils are more common, affect more weather derivatives market participants, or are more difficult to obtain coverage for. The most popular weather derivatives characterization are described below:

Peril	Parameter	Weather index	Source
Heat wave, cold wave, frost	Temperature	HDD, CDD, CAT, minimum temperature, maximum temperature	Weather station, satellite data
Drought, excess of rainfall	Rainfall	Cumulative rainfall, number of rainy days	Weather station, satellite data
Lack of snow	Snowfall	Cumulative snowfall	Weather station
Lack of wind energy production	Windspeed	Wind power production	Weather station, satellite data
Lack of solar energy production	Solar radiation	Solar power production	Satellite data
Cyclone	Cyclone intensity or track	Cyclone intensity or track	Public agency
Earthquake	Earthquake intensity	Seismic intensity, magnitude	Public agency
Drought, wildfire	Vegetation indices	NDVI, soil moisture, burned index	Satellite data

Table 1: Summary of the different weather parameters and indices defining a weather derivative contract

Temperature Temperature weather derivatives are, by far, the most popular type of weather derivatives. In fact, temperature fluctuations strongly affect all sectors of the economy. Garcia Léon [73] estimates that the heat waves of 2003, 2010, 2015 and 2018 affected 0.3 to 0.5% of European Gross Domestic Product (GDP) [73]. Similarly, the weather derivatives market is largely dominated by temperature-based products, which accounted for 95% of the market in 2006 [147]. The underlying parameter of these contracts is often the average daily temperature, defined as the average between the maximum and minimum daily temperatures, such that on day t,

$$T_t := \frac{T_t^{max} + T_t^{min}}{2}.$$

This average daily temperature is aggregated to calculate HDD (Heating Degree Days), CDD (Cooling Degree Days), and CAT (Cumulative Average Temperature) as defined below:

$$HDD := \sum_{t=t_1}^{t_2} \max(0, T_b - T_t), \ CDD := \sum_{t=t_1}^{t_2} \max(0, T_t - T_b), \ CAT := \sum_{t=t_1}^{t_2} T_t,$$

In the above, T_b is a base temperature, t_1 is the start of the contract period, and t_2 is the end of the contract period. T_b can vary, but is set by default to 65°F and 18°C for U.S. and non-U.S. based contracts, respectively [51]. Due to their volume, HDD, CDD, and CAT contracts have spawned numerous scientific studies and dominate the weather derivatives books, see for

example Jewson and Brix [90], Zapranis and Alexandridis [4], and Benth and Benth [24]. Some trades involve other indices such as minimum or maximum temperature, which can be used to identify frost or heat waves, particularly affecting the agricultural sector.

Rainfall One might expect rainfall covers to be as popular as temperature covers, as some sectors such as agriculture are very sensitive to rainfall patterns [144] [4]. However, these transactions are mainly OTC and often within the framework of insurance contracts. Consequently, less research has been devoted to modelling precipitation indices and parameters. In both cases, modeling challenges include a large number of zero observations, non-negative values, and the presence of long tails. For index modelling, the most widely used rainfall index is the cumulative rainfall over a risk period (CR):

$$CR := \sum_{t=t_1}^{t_2} DR_t$$

where DR is the daily rainfall. The gamma distribution has emerged as the most popular choice in practice to model rainfall data, as it addresses the first two issues cited above [90] [104]. Regarding daily modelling, some models similar to the temperature models presented below have also been developed for precipitation. We can cite the frequency severity model [120] [30] and autoregressive models integrating jumps [46].

Wind Derivatives based on wind intensity are often used to hedge against shortfalls in wind energy production. However, energy output depends not only on the input wind intensity, but also on the characteristics of the turbine. While wind direction is no longer an issue, the equivalent energy produced by a wind turbine is non-linearly proportional to the instantaneous wind [44] [90]. Therefore, daily models based on average wind speeds [5] [18] are often insufficient to proxy wind production. Various responses have been proposed to address this challenge by introducing hourly modeling [154], introducing a wind index [90], or working with polynomial wind based indices [109].

Solar Radiation Similar to wind related derivatives, solar irradiance based derivatives have emerged as powerful tools to hedge against cloud presence decreasing solar production [31]. Very little has been written on irradiance indices and daily modelling. We can cite Boyle and al. [36]'s model that integrates cloudy days.

Cyclone Intensity While there are some statistical approaches to modelling hurricane intensity [66], there is little literature between the above and the pricing of weather derivatives [50]. No clear scientific consensus has emerged for the pricing of such derivatives, and the business practice methods described in Subsection 2.3 prevail in practice.

1.6 Data Accessibility and Quality

Accessing weather data can be more complicated than expected. While satellite data is readily available at a reasonable level of granularity, access to weather station data varies by location. Some countries such as the USA and Australia provide free access to weather station data, while the same access to European or Japanese stations is expensive and difficult to obtain [90]. Given the above, it is not surprising that a number of private companies offering access to a large and cleaned network of weather stations have appeared in recent years. In our study, we used data from one of these private companies, Speedwell Climate. The company offers a network of public and private weather stations and provides data cleansing as described in the paragraph below.

The analysis of the quality of weather data is a field of research in itself [54] [1] [91]. A particular challenge is the detection of inhomogeneities in weather time series. Several approaches are possible [54]:

- Historical coherence check, which consists of testing the historical statistical and physical consistency of the observed time series.

- Spatial coherence check, which consists of comparing the observations with data from neighbouring stations.

- A geostatistical simulation approach that consists of comparing the observations with predicted data proxied by the neighbouring stations.

The third approach is particularly interesting because it not only allows outliers to be detected, but also provides replacement data that can be used to correct inhomogeneities or fill gaps. Another important key aspect of weather station data cleaning is the detection and correction of breakpoints. Indeed, a change in the weather station environment or the movement of the weather station can lead to permanent breaks that affect the consistency of the historical time series [1] [91]. Recalibration of the time series is then required.

Finally, the importance of data quality depends on the use of the data. Forecasting requires the most recent data available, regardless of whether that data has been robustly validated. Therefore, the data used for forecasting is usually uncleaned. It is also known as synoptic data. For pricing weather derivatives, data quality requirements are higher [3]. Homogeneity adjustments and recalibration are essential to avoid under- or overestimation of risk. Similarly, at maturity, it is important to rely on reliable data to settle the claim. The same former companies often provide weather data settlement services. They act as third-party data providers, providing both parties with cleaned and externally verified data.

1.7 Index-based insurance and its link to weather derivatives

In parallel with the weather derivatives market, an index-based insurance market has developed in recent years. This section provides a brief overview of the index-based insurance market and

compares the two products.

Index-based insurance products are insurance contracts that define a claim payment between an insured and an insurer when an index (usually a weather index) exceeds a specified threshold. Similar to weather derivatives, these products were born in the late 1990s when the Association of Small Island States (AOSIS) proposed the establishment of an international insurance scheme to compensate small islands and underdeveloped nations affected by sea level rise [111]. While the project grew in subsequent years, the index-based insurance scheme remained theoretical until the early 2000s [106]. In 2005, the World Bank launched the Weather Index-based Crop Insurance in Malawi, a rainfall-based insurance scheme for 50,000 smallholder farmers [76]. In the same year, the Mongolian government, with support from the World Bank, introduced the Index-Based Livestock Insurance Project [138], which is based on livestock mortality rates and has covered more than 14,000 herders in its first four years. Since early 2000, some index-based insurance programs have gained recognition:

- Uruguay Hydro Energy Insurance Program, launched in 2009 to protect Uruguay's public electricity company against the risks of drought and high oil prices. This program, financed by the World Bank, covered the company for losses up to $450 million [139].

- Since 2007, the Caribbean Catastrophe Risk Insurance Facility (CCRIF) has insured small governments in the Gulf of Mexico against tropical cyclones and excess rainfall. During Hurricane Irma in 2017, this facility paid out $31.2 million to affected countries [151].

- The African Risk Capacity (ARC), a sovereign risk pool treaty organization launched in 2012, provides rainfall-based insurance to 33 countries in Africa [86].

As we can see, index-based insurance programs have emerged as a natural response to the challenge of loss and damage. They have gained important recognition as successful public-private partnerships that can protect the most vulnerable from major climate disasters.

Returning to weather derivatives, one might wonder how they differ from the products described above. Contrary to the common belief that index-based insurance tends to cover more low frequency/high severity events such as natural catastrophes [137], nothing in the characterization of the risk transfer contract justifies this. In fact, both contracts are identically defined as described in Subsection 1.3. The main formal differences arise from the regulatory and legal framework which impose different tax and accounting obligations [90] [133]. Weather derivatives are also more prone to speculative behaviours [3].

2 Weather risk valuation practices

The previous section has introduced weather derivatives and their market. This section focuses on risk valuation of these products.

2.1 Actuarial vs financial approach

One of the challenges of our study is to define risk, and therefore price. In fact, there are two approaches for evaluating risk: an actuarial approach and a financial approach.

Actuarial approach Following Denuit [61] and Laeven, and Goovaerts [98], we define the framework of actuarial premium calculation. Let consider a probability measurable space $(\Omega, \mathcal{F})$ where Ω is the outcome space and $\mathcal{F}$ is a σ-algebra defined on it. Let S represent a risk corresponding to the final loss or claim. S is a random variable defined on $(\Omega, \mathcal{F})$.

Definition 2.1. *A premium calculation principle π is a functional assigning a real number to any random variable defined on $(\Omega, \mathcal{F})$.*

Under the actuarial approach, for an insurer exposed to a loss S, a premium calculation principle gives the premium $\pi(S)$ to be charged to the insured for a contract. Premium principles are the most common risk measures applied in the actuarial field. They can present some desirable properties: law invariance, no-ripoff, monotonicity, transitivity, positive homogeneity and subadditivity. In addition, some of premium calculation principles are related to utility theory as exposed by Goovaerts, De Vijlder and Hazendonck [79]. We list below most common principles:

- Expectation principle: $\pi(S) = (1 + \delta)\mathbb{E}_{\mathbb{P}}(S)$ where the loading $\delta > 0$ enables to integrate expenses and solvency margins.

- Variance principle: $\pi(S) = \mathbb{E}_{\mathbb{P}}(S) + \delta Var(S)$ where $\delta > 0$.

- Standard deviation principle: $\pi(S) = \mathbb{E}_{\mathbb{P}}(S) + \delta\sqrt{Var(S)}$ where $\delta > 0$.

- Quantile principle: $\pi(S) = F^{-1}(q)$ where q corresponds to the quantile such that $F^{-1}(q) = \inf\{s \in \mathbb{R}, F(s) \geq q\}$ and F the cumulative distribution function of S. This principle is also called Value at Risk when used as a risk measure.

We leave behind a bunch of other principles like the exponential principle and the Esscher principle. The choice of the actuarial principle depends on the desired properties as well as on market standards. The reader can also note that we introduced the probability $\mathbb{P}$ corresponding to the so-called historical or physical probability. One can also note that a discount rate $e^{-r(T-t)}$, where T corresponds to the time of maturity of the contract, r to the return of riskless asset and t to the buying or selling time, can also be added to the principles π.

While all precedent definitions hold for any probability law, the actuarial principles are often applied in the historical framework. Finally, we define the *Actuarial Fair Price* as the expectation of risk S under the historical probability $\mathbb{E}_{\mathbb{P}}(S)$.

Financial approach   Up to now we have worked in the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The financial principle corresponds to the shifting to a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t\geq 0}, \mathbb{P})$ where $(\mathcal{F})_{t\geq 0}$ corresponds to an increasing family of σ-subalgebras of $\mathcal{F}$ representing the information on the underlying process S (here corresponding to the claims process). The Fundamental Theorem of Asset Pricing states then that, in a market with no arbitrage opportunity, there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the discounted price of assets are $\mathbb{Q}$ martingales. Following Lamberton and Lapeyre [100], if the market is complete, we can define an admissible strategy replicating the contingent claim S that gives us the following price:

$$\pi_t = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}(S \mid \mathcal{F}_t)$$

Leaving aside the notion of discount we will address later, we can see that the price formula is based on the hypotheses that for any claim an admissible replicating strategy is attainable. This is only valid in complete markets. In addition, under this market completeness hypotheses, there exists a unique risk-neutral measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which S is a martingale.

The hypotheses of market completeness strongly depends on the dynamics of the underlying asset producing the claim S. Under desirable assumptions on the dynamics of the underlying, like the geometric Brownian motion in the Black-Scholes framework, converting the physical measure $\mathbb{P}$ into the risk-neutral measure $\mathbb{Q}$ can be done through the Girsanov theorem. In the geometric Brownian motion case this leads to the following measure transformation:

$$W^{\mathbb{Q}} = W^{\mathbb{P}} + \frac{\mu - r}{\sigma}t$$

where $W^{\mathbb{Q}}$ is the Brownian motion under the risk-neutral measure, $W^{\mathbb{P}}$ the Brownian motion under the historical measure, μ the drift and σ the standard deviation of the geometric Brownian motion and r to the return of riskless asset.

While there have been attempts to reconcile the actuarial and financial approaches [59], they derive from completely different initial postulates. Nevertheless, for certain actuarial principles and underlying asset dynamics, the pricing formulas can coincide. Taking again the Black-Scholes case with S following a geometric Brownian motion, we get the following price at maturity T:

$$\pi_t = e^{-r(T-t)}\mathbb{E}_{\mathbb{Q}}(S \mid \mathcal{F}_t) = e^{-\mu(T-t)}\mathbb{E}_{\mathbb{P}}(S \mid \mathcal{F}_t)$$

One can recognize the expectation insurance principle in the above formula in this very specific case.

2.2 Pricing challenges

While some experts use the previous approaches [142][28][3], we will see that they present significant limitations to the pricing of weather derivatives. The latter are explored below.

Existence of an arbitrage free market The difficulty to show completeness and no-arbitrage on the weather derivative market is the main reason why we cannot take a purely financial approach to price weather derivatives.

The weather derivatives market does not exhibit the characteristics of an arbitrage-free market. First, the risk-neutral framework assumes the existence of an underlying tradable market. However, since weather cannot be stored or traded, there is no such underlying weather market [5]. Muëller and Grandi [117] argue that we could not establish a unique monetary value for the underlying weather parameter because market participants may have different financial sensitivities to adverse weather conditions. Second, the weather derivatives market is a relatively illiquid market [5], with approximately 70 types of contracts traded and only 1,000 sold per month on the CME. Most of the contracts are traded over the counter and cannot be used to feed the market information. The main consequence is possible arbitrage opportunity which means enrichment without initial investment may exist and the results of the theory of asset pricing do not apply.

Let accept for one moment the hypotheses of no-arbitrage. For the same reasons as above, we are confronted to an incomplete market with potentially infinite risk-neutral measures. Again, we can not deploy the classic option theory for pricing weather derivatives. Finally, some authors defend a close-to-nill market price of risk. Analysing CME data, Weagley [146] finds that weather derivatives are usually priced close to their *Actuarial Fair Price*. Using an alternative approach, Cao and Wei [42] also find that the market price of risk for temperature derivatives is very small.

Discount rate Both the actuarial and financial approaches agree on the introduction of a discount rate r which takes into account the opportunity cost of not investing in a risk-free asset between the time the derivative is sold and maturity. While this assumption is theoretically justified, most authors bypass the challenge of combining risk rates of different markets. Some authors suggest using the interest rate on a treasury bill but do not implement it [28], others remove this factor by supposing premium transfer at maturity [16], others ignore this factor [90] or set it to 0 [38].

In our case, we also choose to ignore this factor. First, because contract selling time is only few months before maturity. The influence of the discount rate is therefore small. Second, because we consider we can hedge against rate this risk in liquid markets.

Recent works on weather derivative pricing The main proposal to address the challenge of lack of open market is the so-called indifference pricing [153]. This method does not give an exact price, but an interval of acceptable prices at which the seller and the buyer would like to make the transaction. In both cases, the minimum (or maximum) price of the buyer (or seller) is determined as the value at which the buyer (or seller) is indifferent between retaining or hedging the risk. Going one step further and under perfect information hypotheses, Barrieu and al. [13] argue that the seller would set the price that maximizes his expected utility. While

the above approach remains interesting, it has several drawbacks. The first is the somewhat arbitrary utility functions. The second is the ability to provide only intervals and not prices without constraining hypotheses. Other approaches have been considered. Davis [58] suggests using marginal rate substitution, which also integrates utility functions. Cao and Wei [42] propose an equilibrium framework based on the Lucas model. Benth and Sgarra advocate to use the Esscher transformation for pricing in such markets [29].

In summary, there is no clear agreement on the most appropriate approach to price weather derivatives. The actuarial and financial approaches offer avenues for convergence. However, there are still challenges that they do not fully address. Our position throughout this study has been to stick to our ability to estimate and simulate the distribution of losses. When "price" is mentioned, it always refers to the *Actuarial Fair Price* under historical probability. We leave this topic as a possible extension of the current study.

2.3 Risk valuation business practices

The following two subsections develop the risk valuation methods used in the business world. Both are based on the assumption that the weather index I follows a certain distribution and is the only random element of the contract. Using these methods, we compute the *Actuarial Fair Price* defined in subsection 2.1, which will be denoted π from now on.

Burned Analysis The burned analysis method takes a naive approach to calculate $\mathbb{E}_{\mathbb{P}}(S)$. It approximates the theoretical expectation of losses by the average historical payoffs. For further references see Schiller and al. [130], Jewson and al. [90] and Benth and Taib [135]. Suppose $(I_n)_{1 \leq n \leq N}$ corresponds to the historical seasonally aggregated past realisations of the index between season 1 and season N. Then the burned price is:

$$\pi_{BA} = \frac{1}{N} \sum_{n=1}^{N} P(I_n)$$

where P is the payoff function.

The main advantage of this method is that it is easy to implement. However, it presents an important drawback: missrepresentation. Indeed historical weather stations present, in the best cases, 40 years of good data. The length of the series $(I_n)_{1 \leq n \leq N}$ is therefore of $N = 40$. Such short samples may not be representative of the distribution of the index I, missing particularly events of high return periods. $\mathbb{E}_{\mathbb{P}}(S)$ computed through the burned analysis method is likely to be biased.

A second question arises: the time homogeneity of the index I. In fact, the burned analysis is based on the hypothesis that the index observed over 40 years ago was as lucky to happen today. However, in the context of climate change, global warming, seasonal shifts and increased climate volatility cannot be ignored.

We will not describe all the methods used in business practices to address these climate change issues, but only trend removal, which is by far the most widely used step. The principle is to consider that our index I can be decomposed into a deterministic part $s(n)$, the year-dependent trend, and a random part $\tilde{I}$. Since the risk comes from the random term, the risk is historically time-homogeneous such that:

$$I_n = s(n) + \epsilon_n \text{ and } \tilde{I}_n = I_n - s(n) + s(N) \tag{2.1}$$

$(\epsilon_n)_{0 \leq i \leq N}$ are considered independent and identically distributed, and $(\tilde{I}_n)_{0 \leq i \leq N}$ corresponds to the updated version at year N of the index I. The burnt analysis is now applied to the updated observations as follows:

$$\pi_{BA} = \frac{1}{N} \sum_{n=1}^{N} P(I_n - s(n) + s(N))$$

The reader may also note that the above detrending approach can be applied to both the index I and the underlying that defines the index. Sometimes both detrendings converge, and sensitivity to the choice of detrending is often tested.

Finally, the burned analysis approach may seem rather naive. It corresponds, especially for low-volatility covers, to a first and widely used pricing method.

Index modeling Index modeling is the second common method of pricing risk in business. For literature see Dorfleitner and al. [63], Schiller and al. [130] and Jewson and al. [90].

The principle is to model $(I_n)_{0 \leq i \leq N}$ or $(\tilde{I}_n)_{0 \leq i \leq N}$ by a classical continuous distribution (Normal, Gamma, Logistic...) and infer the value of the *Actuarial Fair Price* π_{IM}.

Below is a summary of the steps taken to implement this method.

1. From the underlying daily parameter, we compute the aggregate index $(I_n)_{0 \leq n \leq N}$.

2. If necessary, the data are updated by correcting the trend factor as in the equation (2.1).

3. We fit a probability distribution to the annual indices $(I_n)_{0 \leq n \leq N}$ using the method of moments or maximum likelihood. Given a certain visual or statistical criteria (Kolmogorov-Smirnoff, Anderson-Darling metric), we choose the best distribution.

4. We calculate the expected payoff under this fitted distribution by Monte Carlo simulations or numerical integration and obtain π_{IM}.

3 An introduction to temperature models and their application to weather derivatives risk valuation

As introduced in Subsection 1.5 most of our work has focused on temperature-based derivatives. This section offers an introduction to the main models and describes the contributions and potential developments of Chapter 1.

3.1 Literature review on temperature models

We consider $(T_t)_{t \geq 0}$ a stochastic process representing the daily average temperature defined as the average between the maximum and the minimum daily temperature.

The success of Ornstein-Uhlenbeck models In spite of some dissonances there seems to be an overall agreement on the modelling of temperature dynamics through Ornstein-Uhlenbeck equation. First introduced by Alaton et al. [3] and spread by different articles of Benth [28] [129], the dynamics are modelled as below:

$$\begin{cases} T_t & = s(t) + \tilde{T}_t, \\ \tilde{T}_t & = -\kappa \tilde{T}_t dt + \sigma(t) dW_t, \end{cases} \tag{3.1}$$

where $(W_t)_{t \geq 0}$ is an independent Brownian motion, s a seasonality and trend deterministic function, κ a nonnegative mean-reverting parameter and σ a deterministic nonnegative function corresponding to the volatility.

Deterministic function specifications The first source of divergence in the literature is the specification of the deterministic functions s and σ. The general specifications correspond to:

$$s(t) = \alpha_0 + \beta_0 t + \sum_{k=1}^{K_s} \alpha_k \sin(\xi_k t) + \sum_{k=1}^{K_s} \beta_k \cos(\xi_k t)$$

$$\sigma^2(t) = \gamma_0 + \sum_{k=1}^{K_{\sigma^2}} \gamma_k \sin(\xi_k t) + \sum_{k=1}^{K_{\sigma^2}} \delta_k \cos(\xi_k t). \tag{3.2}$$

where $\xi = \frac{2\pi}{365}$, $\xi_k = k\xi$ and $K_s \in \mathbb{N}^*$. Alaton [3] and Benth and al. [28] agree to keep $K_s = 1$ while they disagree on K_{σ^2} as Alaton [3] suggest to keep $K_{\sigma^2} = 1$ and Benth and al. [28] take $K_{\sigma^2} = 4$. In addition, Alaton [3] tests monthly volatilities and finally advocates to keep σ constant. Others suggest alternative modeling for s through splines [130] or through wavelet analysis [4].

Handling non-normal residuals While Model (1.1) is taken as base model, many authors contest the normality of the residuals and have developed alternative model specifications, see below a non-exhaustive enumeration of alternative noises:

- GARCH models: Campbell and Diebold suggest a AR(25) model followed by a GARCH(1,1) model for the residuals of average daily temperatures of four U.S. cities [41]. Even thought the complexity of this model is discussed by Benth and Benth [129], the article introduces an interesting approach to model discrete volatility.

- ARIMA models: Caballero and al. [40] also work on discrete time series analysis, propose more than one autoregressive term dynamics AR(3) and defend statistical significance of ARIMA(1,d,1) models for Central England Temperature (CET), Chicago and Los Angeles daily temperatures.

- Fractional Brownian motions: Brody and al.[38] suggest to use Fractional Brownian motions [74] to capture long-memory effects and apply the model to daily CET data.

- Lévy noises: An alternative model to Brownian residuals that keeps time continuity are Lévy noises. These have been explored by Benth and Benth [25]. The authors suggest to model average daily temperatures in 7 Norwegian cities with marginals following generalized hyperbolic distributions. Richard and al. [126] study temperature in Fresno (California, US) and defend, after a comparison of different statistical metrics, a mean-reverting Brownian motion with log-normal jumps and first-order autoregressive conditional heteroscedastic noises.

- Regime-switching models: Elias and al. [64] defend a two-state Markov regime-switching model with one regime governed by a mean-reverting process and the other one by a Brownian motion to describe daily temperature in Toronto (Canada).

3.2 An overview of Chapter 1

A stochastic volatility model for temperature dynamics   Our first Chapter contributes to the analysis of average temperature dynamics and their application in the field of weather derivatives. Inspired by some models explained in Subsection 3.1, we investigate the interest of applying a stochastic volatility model for the temperature process $(T_t)_{t \geq 0}$. Our model extends the existing models proposed in the literature while it enables to capture important fluctuations that were not illustrated by former models. Namely, we introduce Model (M):

$$\begin{cases} T_t & = s(t) + \tilde{T}_t, \\ d\tilde{T}_t & = -\kappa \tilde{T}_t dt + \sqrt{\zeta_t}(\rho dW_t + \sqrt{1-\rho^2}dZ_t), \\ d\zeta_t & = -K(\zeta_t - \sigma^2(t))dt + \eta\sqrt{\zeta_t}dW_t, \end{cases} \tag{M}$$

where $(W_t)_{t \geq 0}$ and $(Z_t)_{t \geq 0}$ are independent Brownian motions, $\kappa, \eta, K > 0$, $\rho \in [-1, 1]$, σ^2 is a nonnegative function and the functions s and σ^2 will be taken as in (1.2). We will denote $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by (W, Z), so that the processes T and ζ are adapted to it. This model is somehow an adaptation of the celebrated Heston model [84].

We apply Model (M) to the study of daily average temperatures in 8 major European cities from January 1st 1980 to December 31st 2020.

The following sections focus on key calibration challenges that are less explored in literature. As in Bolyog and Pap [34], we implement Conditional Least Square (CLS) estimation to calibrate Model (M). We study strong consistency of CLS estimators for the time-dependent

CIR volatility process. We complement the proof of Overbeck and Ryden [123] in a time inhomogeneous case.

With respect to pricing, the development of an extended Heston model allows for efficient pricing methods using Fourier transform techniques. We adapt the approach in Carr and Madan [48] to our model and implement Fast-Fourier Transform pricing. We combine simulation-based pricing and Fourier-transform-based pricing using control variables to reduce the computational time to 10^5.

3.3 Contributions of Chapter 1

Model (M) enables to go beyond the classical Ornstein-Uhlenbeck model [3] [28] by adding the last flexibility to the volatility process while keeping tractability. This process addresses the issue of non-Gaussian residuals while maintaining time continuity and diffusion properties. In particular, it is more conservative regarding extreme events than the corresponding Gaussian model. It also enables to better understand and model the risk related to the volatility.

In terms of estimation, our main contribution consists in leveraging Bolyog and Pap [34] work on Conditional Least Squares estimation. In particular, we extend Overbeck and Ryden [123] demonstration of strong consistency of CLS estimators for the time-dependent CIR processes to a time inhomogeneous case.

Keeping the affine structure of the dynamics enables to perform efficient pricing through Fast Fourier Transform techniques for some indices and combined with control variates for other indices. The use of Fast Fourier Transform pricing for weather derivatives is not as widespread even though it has already been developed [27]. We explicit a whole methodology including the resolution of time inhomogeneous autonomous Riccati equation to perform this pricing. All our results are compared with Monte Carlo simulation enabling to show accuracy and increased computational efficiency of the developed techniques.

Industrial applications The first major contribution of the study has been the mathematical formulation of pricing practices (Burned Analysis and Index Modelling), leading to a better understanding and adjustment of current practices. In particular, it has made it possible to extend simulation-based methods to calculate the *Actuarial Fair Price*. It has also helped to rationalize the view of risk according to actuarial principles and loadings.

The second contribution has been to deepen the understanding of daily temperature models. Given global warning and the low reliability of seasonal forecasts, understanding temperature trends is an important issue. This study and the related scientific literature have contributed to the discussion of some trend modelling choices and has led to a Research and Development project within AXA Climate.

Finally, this study has participated to the implementation of daily models dealing with cold waves and frost cover. Various adjustments have been made to the models studied in this report to facilitate computational feasibility and to adapt to extreme temperature modeling.

However, this study has brought the tools and capacity to stand back and improve the pricing of these products. This model has already been leveraged for pricing and is currently being put in production.

3.4 Further relevant issues

While we have chosen to move into energy and temperature derivative risk pricing in line with industry needs, the field of weather derivative pricing still presents challenges.

Compute a market price of risk As discussed above, pricing in incomplete markets has not yet been tackled. Various methods have emerged, such as indifference pricing [153], the marginal utility approach [58], quantile hedging and shortfall minimization [71] [70], and quadratic approaches [82]. Unfortunately, none of these approaches has emerged as the best method, and some require the introduction of additional hypotheses such as utility functions. Another alternative would have been to compute the market price of risk in complete markets where weather derivatives traders are present and generalize the market price of risk to the weather market. However, this would have required finding a complete market common to weather derivatives traders, which may not exist, knowing that most commodity markets, such as the energy market, are also considered incomplete. In addition, Weagly [146] finds that weather contracts are usually priced near their *Actuarial Fair Price*.

Expand to multiple locations Weather derivatives are often presented as a bucket of weather stations, where the index corresponds to the aggregation of a weighted average of daily temperatures by location. The corresponding derivative is therefore dependent on various marginals that follow the process studied in this thesis. In order to evaluate the risk with respect to this bucket, it is necessary to develop multidimensional daily temperature models that introduce dependence structures to account for risk diversification effects. Unfortunately, there is not much literature on this topic, while it corresponds to a real business need [121].

4 An introduction to derivatives for the energy market

Weather derivatives are closely related to the energy market. In fact, the first exchanges were between energy companies, and in 2004-2005 the WRMA estimated that 69% of OTC weather derivatives end users were energy companies [146]. This section explores these links and presents our research on coupled energy and weather models.

4.1 The role of weather derivatives in the energy market

An evident link between energy markets and weather derivatives The impact of climate drivers such as ENSO or the North Atlantic Oscillation (NAO) on the electricity market has

been demonstrated in the literature [145]. Curtis and al [57] show that the NAO phases have a statistically significant impact on both thermal generation costs and electricity prices in Ireland. While Morcillo and al [116] show that the El Nino phenomenon affects the water contributions of Colombian rivers used to power hydroelectric plants, affecting energy availability and prices. Ely and al. [65] defend that NAO has an impact on both electricity generation and demand. Rowinska an al. [127] show that the wind energy generation and the wind penetration index partly explains electricity spot prices. Overall, on the generation side, climate factors affect reservoir inflows and the availability of wind and solar for renewable generation. On the demand side, an unusually warm or cold premium can trigger peak demand for cooling or heating.

In the above cases, and because energy cannot be stored, we consider weather conditions to affect the volume of energy available. We call this volumetric risk because it results from a change in demand for goods due to a change in weather [117]. Weather derivatives are often used to hedge this volumetric risk. While weather derivatives can also be used to hedge against price risk, this is a less popular approach as other financial tools already exist for this hedging need [37].

The relationship between the energy market and temperature   Focusing now on temperature, the literature has also studied the impact of temperature variability on the energy market, where temperature is considered to mainly affect demand. Benth and Meyer-Brandis [23] suggest a negative correlation between temperature and energy prices in the Nordics, where cold conditions lead to increased heating demand and energy prices. Similarly, warm conditions can lead to increased demand for cooling. Another perspective is price expectations. Considering that energy cannot be stored and the variability of the energy market, forecasts are one of the pieces of information available to predict the day ahead and future electricity demand and prices. Temperature forecasts therefore play a key role in the construction of day-ahead and future electricity prices [83].

Hedging volumetric and price risk with weather derivatives   In recent years, hybrid products have been developed to hedge both price and volumetric risk. According to Pérez-González and Yun [125], the birth of these products is partly responsible for the shrinking popularity of classic weather derivative products since 2007. These hybrid products are known as quantos and have existed in the commodity derivative market for some time, allowing simultaneous hedging of commodity price and rate changes [14] [88].

On the field of climate derivatives different combinations have been studied. Benth and al.[22] use a Heath–Jarrow–Morton approach to price hybrid derivatives combining New York Mercantile Exchange-traded natural gas futures and Chicago Mercantile Exchange-traded heating degree days futures for New York. Matsumoto and Yamada study optimal design of mixed weather derivatives on wind indices and electricity prices [155]. Benth and Ibrahim [19] develop continuous-time models combining spot prices and logarithmic photovoltaic power production.

However, the literature becomes quite narrow if one wants to combine energy prices and

temperature [62]. We should mention Caporin and al [44] who develop a two-dimensional daily ARFIMA-FIGARCH model for energy price and temperature. They consider both an actuarial and a financial approach and perform simulation-based pricing that leads to important price differences [44]. Benth and al. [17] consider bivariate Markov-modulated additive processes with independent non-stationnary increments to model quantos combining temperature and energy prices and electricity and gas prices. Finally, Cucu and al. [56] develop a combined model for natural gas spot prices and temperature. They address calibration and pricing challenges for temperature-gas swaps.

4.2 Which underlying energy

It is easy to get lost when talking about energy commodities. Indeed, we are facing a market with multiple locations, multiple qualifications (spot, day-ahead, on-peak, off-peak), multiple commodities (gas, oil, electricity) and the fact that prices can be negative given the non-storability of some commodities [45]. In addition to the latter, there is an extensive derivatives market, including futures, options and other exotic derivatives. The futures market can be of particular interest, as we will see that futures with short maturities are used to proxy the spot price. Furthermore, the maturity of commodity derivatives is often a period of time rather than a specific date, which can be up to a month for some commodities such as crude oil or natural gas.

In the following study when we talk about the energy spot market, we mainly refer to the electricity day-ahead auction market clearing prices [149]. Since operators need advance notice to verify that the schedule is feasible and within the transmission constraints, continuous electricity price trading is not possible and prices are set in the day-ahead auction market. The price paid by each participant depends on the bidding rules, as some markets are *uniform price* auctions and others are *pay-as-bid* auctions. It should also be noted that although all price forecasts are made at the same time the day before, day-ahead auctions set hourly prices.

In terms of modeling, there are a number of choices to be made, and no clear consensus seems to have emerged for this complexity. First, the choice of an arithmetic or geometric model (spot versus log spot) has not been solved. Second, the granularity of the model can vary, as most models are time-continuous. However, they are usually calibrated in daily data, even if hourly data are available. Finally, there is a choice to make between modeling spot price [77] [107] versus futures price [22] as some experts argue that the derivatives market is established on the basis of futures markets [45].

4.3 Energy models

Before entering into coupled models, this section describes the main literature references for electricity spot price and other commodity models. For a more in-depth review, we encourage the reader to refer to Weron [149] and Dechatres and al. [62].

Convenience yield models In the field of commodity finance, the most extended model is the Gibson-Schwartz two-factor model [77]. This model includes two factors corresponding to the commodity spot price S and the convenience return δ. It assumes risk-neutral dynamics of the form:

$$\begin{cases} dS_t & = (r_t - \delta_t)S_t dt + \sigma S_t dW_t^1 \\ d\delta_t & = \kappa(\theta - \delta_t)dt + \sigma \delta_t dW_t^2 \end{cases}$$

where r is the risk-free rate, θ and σ are the long-run mean and volatility of the convenience yield process, κ is the mean-reverting parameter, and W^1 and W^2 are Brownian motions. Overall, the idea is that the commodity spot price is governed by a short-run noise W^1 and a process called convenience yield δ, which integrates notions of consumption time and storage costs [47].

Convenience yield models for electricity modelling Gibson-Schwartz models were extended to electricity modelling by Lucia and Schwartz [107], who propose a similar two-factor model to explain the spot electricity price S. These models are also known as Gaussian factor models. In Lucia and Schwartz [107], S can be decomposed into a seasonally deterministic function f and two random factors X and Y corresponding to the short-run and long-run random dynamics of the log price such that:

$$\begin{cases} S_t & = \exp(f(t) + X_t + Y_t) \\ dX_t & = -\kappa X_t dt + \sigma_X dW_t^1 \\ dY_t & = \mu dt + \sigma_Y dW_t^2 \end{cases}$$

where κ is the mean reverting parameter of X, σ_X and σ_Y are the volatilities of the processes X and Y, μ is the drift of the process Y, and W^1 and W^2 are two Brownian motions that can be correlated. This model was later extended to more than two factors [132] [89]. The Gaussian models are particularly convenient here and explain the success of this model by allowing the derivation of Black-Scholes formulas for the expected value of the derivatives.

Introduction of spikes Despite the convenience of Gaussian factor models, they are often challenged when confronted with real data. Market data exhibit non-Gaussian stylized facts such as spikes, heavy tails, and price clustering. Therefore, one of the proposals in the literature has been to replace Brownian motions with more general Lévy processes such that:

$$\begin{cases} S_t & = \exp(f(t) + \sum_n X_t^n) \\ dX_t^n & = -\kappa^n X_t^n dt + \sigma_n dL_t^n \end{cases}$$

where f is a seasonality function, X^n is the n-th factor with autoregressive parameter κ^n, volatility σ_n, and Levy noise L^n. Several candidates have been proposed as Levy processes: Compound Poisson [60], Normal Inverse Gaussian [26] and other jump processes [49] [115] [75].

Unfortunately, none of the models has clearly established itself as a reference. Here are the main criteria for selection:

- Ability to develop a robust estimation of parameters,

- Capacity to derive analytical or semi-analytical formulas to evaluate derivatives,

- Ability to compute sensitivity to parameters.

Stochastic volatility models Stochastic volatility models are also considered to address the challenge of non Gaussian residuals. Kellerhals [94] apply the below model to Californian power day-ahead market prices and calibrate it through Kalman filters:

$$
\begin{cases}
dS_t & = \mu_S S_t dt + S_t \sqrt{\zeta_t} dW_t^1 \\
d\zeta_t & = \mu_\zeta dt + \sigma \sqrt{\zeta_t} dW_t^2
\end{cases}
$$

where $(\zeta_t)_{t \geq 0}$ is the volatility process, μ_S and μ_ζ the drift parameters, σ the volatility of the volatility process and W^1 and W^2 two Brownian motions. Deng [60] and Benth [15] also explore similar models in a more complex framework as they include Levy processes. As shown in Chapter 1, stochastic volatility models present very interesting properties but raise important calibration and pricing challenges.

Structural Models In parallel to the above statistical models, alternative econometric models have also gained success: the so-called structural models. These models are based on the definition of supply and demand curves. The reader can refer to Carmona and Coulon [45] for an in-depth review. Following Barlow [11], these models postulate a supply and demand curve and establish a market price. They aim to better reproduce market behavior and allow the inclusion of other market fundamentals such as demand, marginal fuel, maximum capacity, gas price, or renewable production.

4.4 An overview of Chapter 2

Chapter 2 is dedicated to the development of a convincing coupled model for daily average temperature and log spot electricity price. We leverage this coupled model to price energy and temperature quantos and perform static portfolio hedging.

A coupled model for hybrid temperature and electricity derivatives pricing We consider two random processes: average daily temperature $(T_t)_{t \geq 0}$ and the day-ahead log spot price $(X_t)_{t \geq 0}$ such that the day-ahead spot price $S_t = e^{X_t}$. We suggest the following coupled model:

$$
\begin{cases}
d(X_t - \mu_X(t)) & = -\kappa_X(X_t - \mu_X(t)) + \lambda \sigma_T dW_t^T + dL_t^X \\
d(T_t - \mu_T(t)) & = -\kappa_T(T_t - \mu_T(t)) + \sigma_T dW_t^T
\end{cases}
$$

where $\mu_X(\cdot)$ and $\mu_T(\cdot)$ represent the trend and seasonality deterministic component of $(X_t)_{t \geq 0}$ and $(T_t)_{t \geq 0}$ correspondingly, κ_X and κ_T the autoregressive parameters, L^X and W^T two independent motions and λ is a coefficient linking the dependence between the two noises driving X and T. The process $(T_t)_{t \geq 0}$ follows an Orestein-Uhlebeck diffusion process and $(X_t)_{t \geq 0}$ follows a switching diffusion process with a Brownian motion W^T and Normal Inverse Gaussian (NIG) motion L_t. L_t is a centered Normal Inverse Gaussian process of parameters $(\alpha^X, \beta^X, \delta^X, m^X)$.

This model is confronted to real data. For energy, we consider day a-head market spot energy prices in France and North Italy from 5th January 2015 to 31st December 2018 coming from the ENTSO-E Transparency Platform and Gestore Mercati Energetici (GME). Average daily temperature time series are extracted from a weather data provider platform for Paris-Charles de Gaulle airport and Milano Linate airport weather stations.

Model estimation The first sections of the chapter concentrate on model selection. First, we argue the choice of marginals. In particular we confront the spot energy price to the literature models presented in Subsection 4.3 and select a autoregressive Normal Inverse Gaussian process based on our data. Second, we develop different estimation methods based on Conditional Least Square estimation applied to the characteristic function and Maximum Likelihood and EM-algorithm applied to an approximated distribution of the noise. Given the model coherence we keep the Conditional Least Square estimation applied to the characteristic function.

The discussion concerning the temperature marginal is less extensive as the reader can refer to Chapter 1 and Subsection 3.1. We perform different estimations and show that the Gaussian model as in Benth and Benth [16] is an acceptable model given the complexity already raised by the dependence structure. We estimate the dependence parameter λ and show the significativity of Model ETM by comparing real and simulated dynamics through χ^2 tests.

Pricing and static hedging We also address the challenge of pricing. We first consider different quanto derivative product structures and considers its average payoff under historical probability:

$$\mathbb{E}\Big(\sum_{t=t_1}^{t_2} f_S(S_t) \times f_T(T_t) \mid \mathcal{F}_{t_0} \Big),$$

where f_S and f_T represent the payoff function attached on the spot energy and temperature. The dates $t_1 < t_2$ indicates two days, and the summation is made on each day between t_1 and t_2 (including these days). We consider the contract is priced at $t_0 \leq t_1$ and claims at maturity t_2.

We consider different f_S and f_T to account for forward, swaps, single-sided and double sided options on each underlying. We explicit formulas of the expected payoff for forward, swaps and single-sided ($\mathcal{E}$-HDD and $\mathcal{E}$-CDD). For double sided options, we suggest first order Taylor decomposition integrating pricing through Carr Madan [48] formulas. These formulas are confronted with Monte Carlo simulated payoffs. We check that both methods provide

similar results and that the explicit computations enable to gain on computation time. We also reveal that the integration of λ enables to get within the significance intervals.

Finally we work on the risk decomposition of $\mathcal{E}$-HDD and double sided quantos in the framework of an autofinancing portfolio. We perform portfolio optimisation through explicit or semi-explicit formulas. We show that for 100,000 simulate portfolio and show that we succeed to hedge the quanto derivative on average and decrease the variance of the portfolio. We also analyse the impact of leveraging λ.

4.5 Contributions of Chapter 2

The first contribution of this paper is the development of an efficient model describing spot energy and average temperature dynamics. We justify this model in view of the previous literature contributions but also in confrontation with real data. We suggest two estimation approaches and confirm goodness of fit.

The tractability of the above combined model allows us to derive closed-form formulas for quanto derivatives. Only few papers have proposed such formulas [17][56] and often for linear payoffs. The inclusion of double sided options is a challenge we overpass thanks to the tractability of our model and confirm through Monte Carlo simulations.

The introduction of the dependency structure thanks to the parameter λ can also illustrate an important aspect of the energy prices under certain conditions. For instance, we could imagine that the weight of this factor would increase in tighter markets.

Finally, not much has been written about the risk hedging and portfolio management of electricity and temperature quantos. Matsumoto and Yamada [110] study the optimal payoff design and hedging effect of weather derivatives against the fluctuation of solar power production. Similarly, Oum and Oren [122] introduce an expected utility maximization problem that allows them to obtain the optimal payoff function at the time of contracting as well as the optimal hedging time. They first consider a static hedging problem where they minimize a mean-variance utility function over net profit, subject to a self-financing constraint, and acknowledge that they will develop dynamic hedging. Lee and Oren [103] take a portfolio equilibrium approach in a multi-commodity market and analyze the hedging effects brought by the inclusion of weather derivatives. Alternative econometric approaches demonstrating the effectiveness of including weather derivatives in the portfolio of electricity retailers can be found in Lai and al [99] and Masala and al. [108]. Our approach complements the above contributions by proving that an efficient risk management of the portfolio is possible through daily risk decomposition.

Industrial applications   Chapter 2 is more recent and has not yet led to concrete applications. However, it responds to a business need. First, it proposes a first combined model for quanto derivatives. This was definitely missing both in the academic literature and in business practice. Second, it proposes an application with open formulas for option pricing, while the literature

only addressed futures and swaps [56]. Finally, it provides a first portfolio approach that allows to hedge the quanto products. Understanding hedging capacity is also key to portfolio risk management and solvency requirements.

Chapter 1

A stochastic volatility model for temperature derivative pricing

Introduction

With the increased awareness on climate risk and its tremendous consequences on all economic sectors, the demand for financial tools that enable to hedge this weather-related perils has significantly increased in the last decades. The development of weather derivatives coincides with this trend. First launched in 1997 with over the counter (OTC) contracts, the Chicago Mercantile Exchange (CME) introduced standardised contracts for American cities in 1999. Until date, the open market remains relatively small and therefore lacks of liquidity. However, OTC market has developed between industrialists and large insurance and finance companies leading to the necessity to develop satisfactory risk valuation methodologies for those derivatives. These derivatives usually bring on a Heating Degree Day (HDD) index or a Cooling Degree Day (CDD) index or a Cumulative Average Temperature (CAT) index. Since most of the deals are OTC, the market is incomplete and thus no arbitrage free pricing. Weather derivative sellers are more interested in understanding the distribution of the payoff, and how this distribution may change under stressed conditions. These information allow them to determine their pricing depending on their risk appetite as well as to assess their maximum losses.

The literature on the valuation of weather derivatives mostly relies on the development of temperature models. On the one hand, there are continuous time models. Brody et al. [38] suggests to use an Ornstein-Uhlenbeck process driven by a fractional Brownian motion with periodic time-dependent parameters. Benth and Benth [16] consider the same dynamics with a classical Brownian motion with a further expansion in periodic functions and including a trend for the temperature. Benth et al. [129] proposes a continuous time version of autoregressive models. More recently, Groll et al. [80] have developed a continuous time model with factors for the temperature forecasting curve. All these papers use their model to calculate the average payoff of standard weather derivatives on HDD, CDD or CAT. On the other hand, Tol [140] and Franses et al. [72] have proposed discrete-time GARCH models for the temperature. Cao and Wei [43] and Campbell and Diebold [41] use this approach in view of pricing derivatives.

Benth and Benth [129] complete this latter study and show that a low order of autoregression is enough to fit well temperature data. Recently, Meng and Taylor [114] have proposed an extension of this family of models that gives a joint modelling of the daily minimum and maximum temperature.

In this study, we focus on the temperature related derivatives which have been, by far, the most studied tools in the literature. Following the studies of Benth and Benth [28] and many others on temperature dynamics, we develop a time inhomogeneous affine stochastic volatility model inspired from the celebrated Heston model [84] for equity. This model integrates additional flexibility to the temperature dynamics enabling a better time-continuous modelling of the temperature, while keeping tractability. In particular, it is more conservative regarding extreme events than the corresponding Gaussian model. Following the recent work of Bolyog and Pap [34], we develop a Conditional Least Squares estimation method for its parameters, which is easy to implement. Besides, we propose two different pricing algorithms. The first one based on simulation enables to sample the payoff distribution and then to compute empirically quantities such as the average payoff or quantiles of this distribution. The second one takes advantage of the affine structure and is an adaptation of the Fast Fourier Transform method introduced by Carr and Madan [48] to calculate the average payoff of some weather derivatives. Besides, this second method, combined with control variates, enables to reduce considerably the computational time (up to 10^5). We can then easily calculate the sensitivity of the average payoff with respect to the different parameters, including those on the temperature volatility. Thus, an important contribution of our model is to better assess the risk behind the volatility of the temperature. Finally, we compare the results of this approach to price weather derivatives with common business practices.

This study can also be leveraged by practitioners. First, our model contributes to a deeper understanding of daily temperature models. This topic is particularly sensitive in the context of global warming and increased extreme events. This study reviews different approaches to trend modeling and volatility capture. Second, the study describes the step-by-step pricing of temperature derivatives from calibration to computational implementation. This is particularly useful for reproductive purposes and has already been implemented, with some adjustments, in the industrial framework. Finally, the last section provides a comparison with standard business pricing practices to bridge the gap between research and application.

The study is organised as follows. The first section presents our model for the average daily temperature. In particular, we explain the rationale behind this model and how it goes beyond the dynamics studied in the literature up to date. The second section focuses on the estimation of the different parameters of our model. We develop a conditional least squares approach and check the robustness of the fitting on simulated data. The third section concentrates on the pricing of weather derivatives. It develops both Monte Carlo and Fast Fourier Transform algorithms and studies the sensitivity to model parameters of the average payoff. It also compares these pricing methods with current business approaches.

1 Temperature models

1.1 A stochastic volatility model for temperature dynamics

Temperature dynamics have largely been analysed in the statistics literature and they are of particular interest in the field of weather derivatives. These analysis apply to average daily temperature models that are defined as the average of maximum and minimum daily temperature, i.e. $T = \frac{T_{max} + T_{min}}{2}$. This choice comes from the insurance contracts that typically use this average for the daily temperature.

Different models have been suggested for the associated $(T_t)_{t \geq 0}$ process. In the present study, we investigate the interest of applying a stochastic volatility model for the temperature process $(T_t)_{t \geq 0}$. Our model extends the existing models proposed in the literature while it enables to capture important fluctuations that were not illustrated by former models. Namely, we introduce Model (M):

$$\begin{cases} T_t & = s(t) + \tilde{T}_t, \\ d\tilde{T}_t & = -\kappa \tilde{T}_t dt + \sqrt{\zeta_t}(\rho dW_t + \sqrt{1 - \rho^2} dZ_t), \\ d\zeta_t & = -K(\zeta_t - \sigma^2(t))dt + \eta \sqrt{\zeta_t} dW_t, \end{cases} \tag{M}$$

where $(W_t)_{t \geq 0}$ and $(Z_t)_{t \geq 0}$ are independent Brownian motions, $\kappa, \eta, K > 0$, $\rho \in [-1, 1]$, σ^2 is a nonnegative function and the functions s and σ^2 will be taken as in (1.2) and (1.5). We will denote $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by (W, Z), so that the processes T and ζ are adapted to it. This model integrates three components:

- First, the function s represents the trend and seasonality of the average temperature $(T_t)_{t \geq 0}$. This function is deterministic, bounded and continuously differentiable. In this study, we will consider the following parametric form

$$s(t) = \alpha_0 + \beta_0 t + \alpha_1 \sin\left(\frac{2\pi}{365}t\right) + \beta_1 \cos\left(\frac{2\pi}{365}t\right), \ t \geq 0,$$

  see Subsection 1.2.1 for a discussion on this.

- Second, the detrended and deseasonalised temperature process $(\tilde{T}_t)_{t \geq 0}$ follows a mean-reverting process which enables to include memory effects into the model. The parameter κ tunes the mean-reversion speed.

- Third, the volatility of the $(\tilde{T}_t)_{t \geq 0}$ process denoted $(\zeta_t)_{t \geq 0}$ follows a Cox-Ingersoll-Ross (CIR) process [55] with time-dependent parameters [87]. The process $(\zeta_t)_{t \geq 0}$ includes a seasonal deterministic component $\sigma^2(\cdot)$ which is supposed deterministic, bounded, continuously differentiable and nonnegative, so that the process ζ is well defined and remains

nonnegative if $\zeta_0 \geq 0$. In this study, we will the following parametric form for σ^2:

$$\sigma^2(t) = \gamma_0 + \sum_{k=1}^{2} \gamma_k \sin\left(k\frac{2\pi}{365}t\right) + \delta_k \cos\left(k\frac{2\pi}{365}t\right).$$

The parameter K tunes the mean-reversion of the volatility process and η corresponds to the volatility of the volatility.

Here and through the study, we note $f(t)$ (time $t \geq 0$ in parenthesis) a deterministic function and F_t (time in index) a stochastic process.

This model is somehow an adaptation of the celebrated Heston model [84]. It enables to go beyond Ornstein–Uhlenbeck models [16] [28] [136] [72] and is an alternative to GARCH volatility models [136] [72] [41] [129], while keeping some flexibility as we will show in the next sections.

In the following sections we study daily average temperature data series for 8 major European cities: Stockholm, Paris, Amsterdam, Berlin, Brussels, London, Rome and Madrid. The data series present daily data from January 1st 1980 to December 31st 2020. After removing 29 February of leap years, this gives a time series of 14,965 observations coming from weather stations. Table 1.12 in Appendix 4.1 summarizes the characteristics of the weather stations. These time series were extracted from Speedwell, the main historical and settlement data provider for weather derivatives. The company performs quality checks including physical consistency comparison, statistical consistency tests and comparison with neighboring sources and corrects the data if necessary. For instance, gap filling is mentioned for Rome Ciampino. The positive aspect of using these data is that it can be considered as cleaned and reliable and has been used to determine weather derivatives payoff in the past. The downside is that it is private with restricted access.

1.2 Some background on temperature models

1.2.1 Ornstein–Uhlenbeck models

Interest on temperature models for weather temperature derivative pricing have arisen in the last years. Even though there exists some alternative modeling [81] [64] [130], there seems to be an overall agreement on the capacity of Ornstein–Uhlenbeck processes to model daily average temperatures. These models have largely been studied by Benth and Benth [16] [28] and are written as follows:

$$\begin{cases} T_t & = s(t) + \tilde{T}_t, \\ \tilde{T}_t & = -\kappa\tilde{T}_t dt + \sigma(t)dW_t, \end{cases} \tag{1.1}$$

where $(W_t)_{t \geq 0}$ is an independent Brownian motion, s a seasonality deterministic function, κ a nonnegative mean-reverting parameter and σ^2 is a deterministic nonnegative function.

There is no clear agreement on the form of s. Different functions have been suggested from step functions [136] to polynomial functions [72]. However, there is a clear preference for Fourier decomposition ([28], [85], [129]) which gives the following expression

$$s(t) = \alpha_0 + \beta_0 t + \sum_{k=1}^{K_s} \alpha_k \sin(\xi_k t) + \sum_{k=1}^{K_s} \beta_k \cos(\xi_k t), \qquad (1.2)$$

where $\xi = \frac{2\pi}{365}$, $\xi_k = k\xi$ and $K_s \in \mathbb{N}^*$. Here and through the study, the time unit is the day.

In this model, s integrates a trend component which enables to introduce climate change phenomenon into our model and a periodic component through trigonometric functions. There is a large discussion on the convenient $K_s \in \mathbb{N}^*$ to choose as well as on the pertinence to introduce interaction terms [113]. For simplicity purposes, we take $K_s = 1$ which is the most commonly choice (see e.g. [28] [129]) and shows significance at 5% confidence level on our experiments.

The different parameters of the Ornstein-Uhlenbeck process can be determined by Conditional Least Squares Estimation (CLSE) developed by Klimko and Nelson [95], which boils down to minimise

$$\sum_{i=0}^{N-1} \left(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}] \right)^2. \qquad (1.3)$$

Since we have daily data, we mostly use $\Delta = 1$ through the study, unless specified. The quantity

$$Res_{i\Delta} = T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}] = T_{(i+1)\Delta} - s((i+1)\Delta) - e^{-\kappa\Delta}(T_{i\Delta} - s(i\Delta)) \qquad (1.4)$$

is called the residual of the regression at time $i\Delta$. Note that the same formula for the residual holds true for Model (M). We can equally note that the autoregressive factor $e^{-\kappa\Delta}$ derives from the integration of the Ornstein-Uhlenbeck dynamics which is explicit in Equation (4.5) in Appendix 6.2. The CLS estimators have been studied by Overbeck and Ryden [123] for the CIR process, Li and Ma [105] for the stable CIR process, and Bolyog and Pap [34] for Heston-like models. This approach has been used to estimate the parameters of Model (M), See Section 2.1.

Similarly, σ represents the deterministic volatility function. For flexibility and periodicity reasons, it can also be modeled thanks to a Fourier decomposition:

$$\sigma^2(t) = \gamma_0 + \sum_{k=1}^{K_{\sigma^2}} \gamma_k \sin(\xi_k t) + \sum_{k=1}^{K_{\sigma^2}} \delta_k \cos(\xi_k t). \qquad (1.5)$$

The coefficients γ's and δ's are assumed to be such that σ^2 is indeed a nonnegative function, and we also exclude the trivial case $\sigma^2(t) = 0$ for all $t \geq 0$. For $K_{\sigma^2} = 1$, a straightforward necessary and sufficient condition for having σ^2 nonnegative is $\gamma_0 \geq \sqrt{\gamma_1^2 + \delta_1^2}$. For $K_{\sigma^2} \geq 2$, a sufficient condition is to assume $\gamma_0 \geq \sum_{k=1}^{K_{\sigma^2}} \sqrt{\gamma_k^2 + \delta_k^2}$. This decomposition is used for example

by [16] or [114] in a more evolved model. We also consider it in Model (M) as the function to which the stochastic volatility mean reverts. The annual periodicity is a quite natural feature. Besides, this choice gives a bounded continuous function whose nonnegativity is easy to check, which is required for the definition of Model (M). In this study, K_{σ^2} will be taken equal to 2 but $K_{\sigma^2} = 1$ will also be considered.

1.2.2 Limits of Ornstein–Uhlenbeck models

Ornstein-Uhlenbeck processes, and their discrete form corresponding to Autoregressive Models, present some important limitations.

First, they remove any potential long memory effects as today's temperature will only depend on the previous day's. This weakness has been challenged through literature. In particular, Brody et al. [38] suggest the introduction of Fractional Brownian motions. They suppose an Ornstein-Uhlenbeck model driven by a Fractional Brownian motion of Hurst parameter H. Such models generate temperature paths with $(H - \varepsilon)$-Hölder regularity, $\varepsilon > 0$ being an arbitrary small real number. Following the work of Gatheral and al. [74], we perform an analysis on our data to check whether we could observe such regularity and consider the below metric.

$$m(q, \Delta) = \frac{1}{\lfloor N/\Delta \rfloor} \sum_{i=1}^{\lfloor N/\Delta \rfloor} \mid \tilde{T}_{(i+1)\Delta} - \tilde{T}_{i\Delta} \mid^q$$

For a H-Hölder dynamics, $m(q, \Delta)$ should behave as Δ^{qH}. Figure 1.1 shows $\log(m(q, \Delta))$ is well approximated by an affine function of $\log(\Delta)$ for different values of q. Therefore, we can compute the coefficient of the regression of $\log(m(q, \Delta))$ by $\log(\Delta)$ and estimate H, see Table 1.1. We observe that this parameter is very close to 0.5, which corresponds to the regularity of a diffusion driven by a standard Brownian Motion. This justifies why we still consider in Model (M) a diffusion model.

City	Stockholm	Paris	Amsterdam	Berlin	Brussels	London	Rome	Madrid
$\hat{H}$	0.498	0.476	0.497	0.477	0.456	0.510	0.525	0.539

Table 1.1: Parameter estimations for the Hurst coefficient H

Second, we also contemplated autoregressive models with higher than one autoregressive order. Figure 2.1 shows the partial autocorrelation plots of $(\tilde{T}_t)_{t \geq 0}$ and the residuals $(Res_t)_{t \geq 0}$ (see Eq. (1.4)) for the city of Paris in the first $1,000$ observed days. While we can consider a second significant correlation term, its coefficient should be pretty small. Higher than two autoregressive terms are clearly non significant. Therefore, we studied the possibility of having two time delay Ornstein-Uhlenbeck dynamics. However the second order autoregressive coefficients were small and unstable. We hence decided to keep a model with a unique autoregressive terms which is coherent with Franses ad al. [72], Diebold and Campbell [41] and Taylor and Buizza [136] findings.
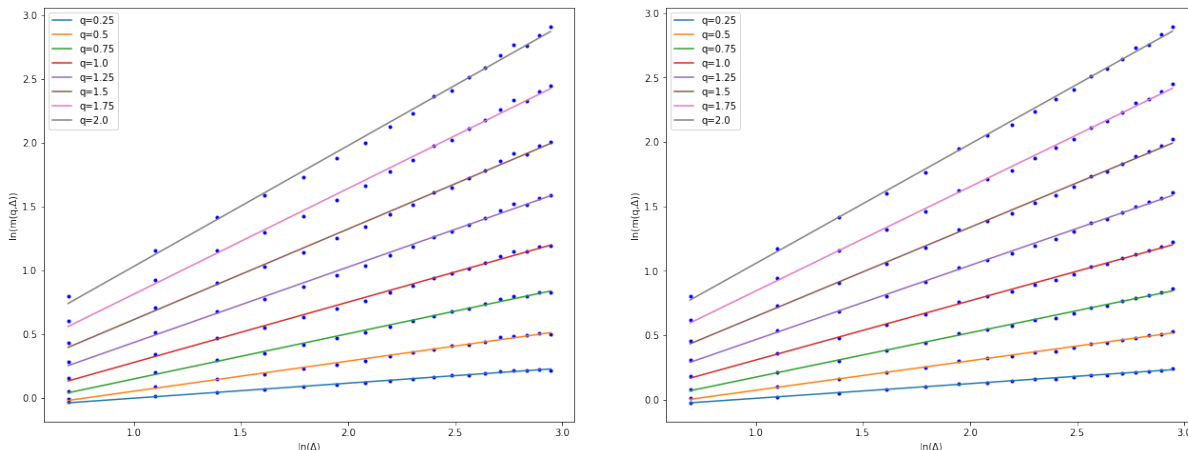
Figure 1.1: Smoothing plots of the temperature dynamics for Stockholm and Paris (left and right)

Third, as can be observed in Figure 1.3, daily temperature present erratic noises with possible volatility autoregression or clustering.

To analyse this, we focus on the residuals (1.4) and compare the historical ones to those obtained on 9 simulations of the same 40 years period. More precisely, for each simulation, we have plotted the curve $(Res_{(i)}^{sim}, Res_{(i)}^{obs})_{1 \leq i \leq 40 \times 365}$, where $Res_{(i)}^{sim}$ (resp. $Res_{(i)}^{obs}$) is the ordered statistic of the residuals obtained with the simulated (resp. observed) temperature. In the center of the distribution, the points are on the line $y = x$ in red, which indicates a very good fit by the model. Instead, we see that for small (resp. large) values, the curves are slightly below (resp. above) the red line which indicates that the extreme events produced by the Ornstein-Uhlenbeck model are smaller than the ones observed. The residuals given by Model (1.1) present noticeable deviations from the historical ones as can be seen in the qqplots: Figure 1.4 shows slight skewness but the main observation remains that the observed tails are heavier than the simulated ones, especially for the left tail. Different authors emphasise skewness deviation of residuals as well as volatility clustering [72] [113]. This skewness and heavy tail issues are coped with different methods either thanks to the fitting of a skew-t distribution on residuals [68] or through GARCH models that enable to capture dependencies on volatility and can lead to skewed residual qqplots. Here, anticipating on our estimation results, we have plotted in Figure 1.5 the qqplot on the residuals between those observed and 9 simulations of Model (M) with estimated parameters. We observe again a very good fit of the center distribution, but the curves are are slightly above (resp. below) the red line $y = x$ for small (resp. large) values, which means that Model (M) produces larger extreme events than the ones observed. Thus, compared with the Ornstein-Uhlenbeck model, we note that Model (M) produces heavier tails, that are slightly heavier to those observed on our 40 years data set. Thus, Model (M) is more conservative on extreme events, which is an interesting feature when dealing about risk quantification.
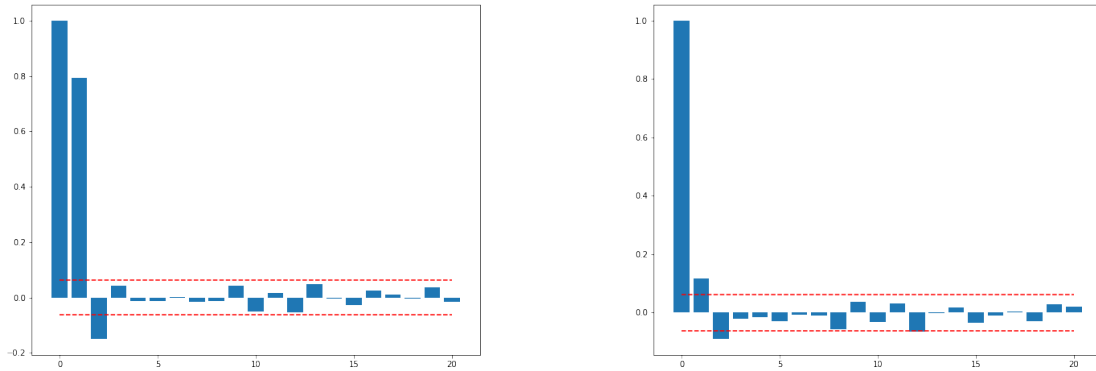
Figure 1.2: Partial autocorrelation plots of $(\tilde{T}_t)_{t\in\mathbb{N}}$ and of the residuals Res_t (see Eq. (1.4)) for the city of Paris in the first $1,000$ observations. The dashed red line corresponds to the 95% confidence interval from which we can consider the partial autocorrelation coefficient is significantly different from 0.
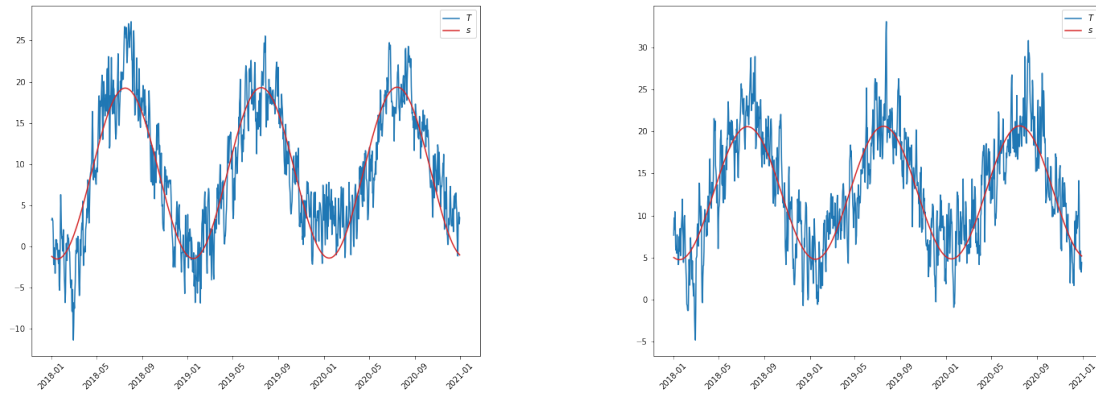
Figure 1.3: Average temperature and trend on temperature from 2018 to 2020 in Stockholm and Paris (from left to right respectively)

1.2.3 GARCH models

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models [33] represent a first response to these limits. They enable to integrate stylized features of temperature times series such as skewness, tail heaviness and volatility clustering. Let us define the following temperature dynamics for $t \in \mathbb{N}$:

Figure 1.4: Quantile quantile plots for observed and 9 simulated residuals of (1.3) for Stockholm, Paris and Rome for the Ornstein-Uhlenbeck model.

Figure 1.5: Quantile quantile plots for observed and 9 simulated residuals of (1.3) for Stockholm, Paris and Rome for Model (M).

$$
\begin{cases}
T_t & = s(t) - \mu(T_{t-1} - s(t-1)) + Z_t \\
Z_t & = \sqrt{\zeta_t}\epsilon_t, (\epsilon_t)_{t \geq 0} \sim \mathcal{N}(0,1) \text{ i.i.d.} \\
\zeta_t & = \sigma^2(t) + \sum_{k=1}^{p} \lambda_k (Z_{t-k} - \sigma^2(t-k))^2 + \sum_{l=1}^{q} \theta_l \zeta_{t-l},
\end{cases}
\tag{1.6}
$$

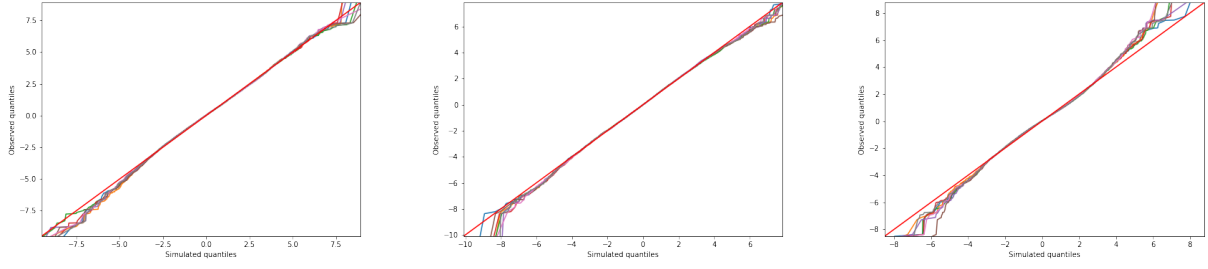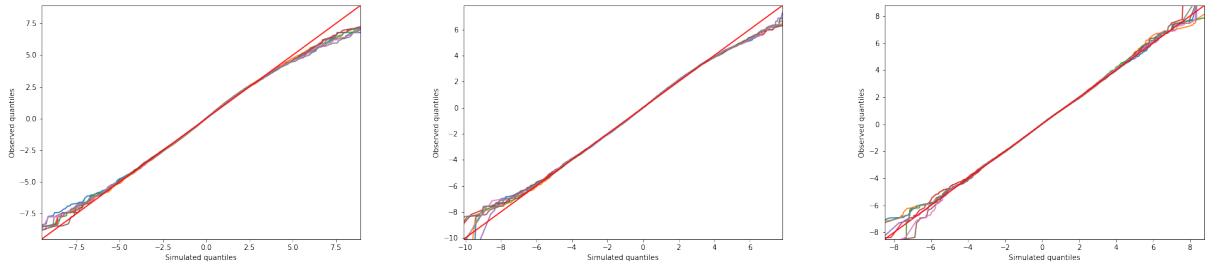with $\lambda_k \geq 0$ for $1 \leq k \leq p$, $\theta_l \geq 0$ for $1 \leq l \leq q$, and s and σ^2 are defined respectively by (1.2) and (1.5) with parameters γ's and δ's such that $\sigma^2(t) \geq 0$ for all $t \geq 0$. Thus, (ζ_t) is indeed a nonnegative process.

Different authors have worked on the application of these GARCH models particularly on GARCH(1,1) to temperature process fitting [113] [136], [72] [129] [41]. While these models enable to integrate the complexity of the series we will see that they also present considerable limits.

First, the analysis of autocorrelation and partial autocorrelation plots on squared corrected residuals show evidence of the necessity to integrate first order autocorrelation on the volatility model. This is coherent with the previously cited papers. However, it should be noted that the significance of the first order autocorrelation term remains small even for subsamples. We can therefore expect that it has a limited impact into the quality of multi-day forecasts and pricing.

Second, the GARCH volatility model and its continuous equivalents integrate an important hypotheses: they suppose that temperature and volatility are driven by the same noise (with a

time shift). In Model (M), this would formally correspond to have $\rho \in \{-1, 1\}$. Given the complexity of meteorological dynamics there is no reason explaining why temperature and volatility on temperature should be driven by the same noises. Model (M) represents a generalization of average temperature dynamics that enables to integrate more flexibility on the temperature's volatility.

Finally, Model (M) presents a clear advantage as it provides closed form formula for the pricing of temperature indices deriving from average temperature models. While these closed formulas can only be used for some derivatives depending on the payoff structure, they can be combined with Monte Carlo approaches to speed up derivative's pricing. In particular, this document will develop the use of control variates methods to reduce the variance of Monte Carlo approaches, see Subsection 3.3.4.

To sum up, Model (M) is aligned with the average temperature models that have been considered in the literature. First, from an Ornstein–Uhlenbeck process, it integrates a seasonal component corresponding to the natural climatology and climate change trend. The Ornstein–Uhlenbeck model also enables to include an autoregressive component upon which agree both literature and statistical tools. However, the residuals of this first model show deviation from normal hypotheses with skewness, tail heaviness and volatility clustering patterns. GARCH models partly answer to these limits and integrate an autoregressive component on the volatility process which is statistically observed. Our model is another natural extension to the Ornstein-Uhlenbeck dynamics that gives a larger flexibility on the volatility process. Model (M) presents two additional advantages. First, unlike ARMA and GARCH models it is a time-continuous model. While this might not be necessary when we only have daily data, it can be a significant advantage if this model is coupled with a model for energy or commodities that are traded continuously. We could foresee, for example, combining this model with other commodity models to identify hedging opportunities or price hybrid derivatives. Second, Model (M) is an affine model for which efficient pricing methods based on Fourier techniques can be implemented. In particular, we show in Section 3 how to use the Fast Fourier Transform for pricing, which translates into a significant competitive advantage.

2 Fitting Model (M) to historical data

The previous section has motivated the interest of the stochastic volatility model (M) for the temperature dynamics. The pertinence of this new model is however related to our capacity to well estimate its parameters. This section focuses on this challenge as well as on the robustness of the estimation. Contrary to financial derivatives, we do not consider the calibration of our model to market prices. First, as pointed by Weagley [146] the market volume of weather derivatives is quite low and the transactions are mostly Over The Counter (OTC). Market prices are thus arguable. Besides, the underlying of these contracts is the real temperature, which is not a traded asset. There is therefore no justification of a risk-neutral pricing of weather derivatives. For these reasons, we prefer to estimate our model to historical data and then use

it to determine the distributions of weather derivatives payoffs under the real probability.

2.1 Parameter estimation

Like Overbeck and Ryden [123] and Bolyog and Pap [34], we implement the conditional least squares estimation (CLSE) which consists in minimising the observed value against the predicted conditional expectation. Under general assumptions, Klimko and Nelson [95] have shown that CLS estimators are strongly consistent, with a speed of convergence close to $O(N^{-1/2})$, where N is the number of observations.

It can be noted that Bolyog and Pap suggest in [34], on a time-homogeneous model that is similar to (M), to simultaneously minimize the conditional temperature and volatility expectations, i.e. to minimize

$$\sum_{i=0}^{N-1} \left(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}, \zeta_{i\Delta}] \right)^2 + \left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}] \right)^2,$$

with respect to κ, K, and the (parameterised) functions $s(\cdot)$ and $\sigma^2(\cdot)$. As already remarked in [34], the estimators of the parameters $\kappa, K, s(\cdot), \sigma^2(\cdot)$ do not involve the values of η and ρ. We can thus do the estimation without knowing these values, and then estimate separately the parameters η and ρ. Besides, Model (M) corresponds to a special case of the dynamics considered by Bolyog and Pap where the volatility is absent from the mean-reverting term of the temperature process (this corresponds to $\beta = 0$ in [34]). In this case, $\mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}, \zeta_{i\Delta}] = \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}]$, and the minimisation problem is equivalent to minimise $\sum_{i=0}^{N-1} \left(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}] \right)^2$ and $\sum_{i=0}^{N-1} \left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}] \right)^2$ separately, which we do here.

Thus, Paragraph 2.1.1 deals with the estimation of κ and $s(\cdot)$, Paragraph 2.1.2 brings on the estimation of K and $\sigma^2(\cdot)$, while Paragraph 2.1.3 focuses on the estimation of η and ρ.

2.1.1 Parameter estimation for κ and $s(\cdot)$

To estimate κ and $s(\cdot)$, we are thus interested by the following minimisation problem:

$$\min_{(\kappa, \alpha_0, \beta_0, \alpha_1, \beta_1) \in \mathbb{R}^5} \sum_{i=0}^{N-1} \left(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}] \right)^2,$$

that brings on temperature process $(T_t)_{t \geq 0}$. Proposition 6.1 solves this problem and gives explicit formulas for $\hat{\kappa}$, $\hat{\alpha}_0$, $\hat{\beta}_0$, $\hat{\alpha}_1$ and $\hat{\beta}_1$. Table 1.2 shows the estimated parameters for the 8 European cities. All the parameters are significant at 5% confidence level (their 95% confidence interval do not cross zero) . We can also add that there is a certain coherence between the different European cities for the mean-reverting parameter. In addition, given the range of values close to 0.25 we can also derive that the temperature memory effect lasts for around 4 days. These results are aligned with the literature, see e.g. [28]. We also notice that the

estimated values of β_0 are around 0.00013, which corresponds to a warming of about 0.5°C every 10 years.

City	$\hat{\alpha}_0$	$\hat{\beta}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\kappa}$
Stockholm	6.678	0.00016	-4.564	-9.142	0.192
Paris	10.868	0.00013	-3.540	-6.993	0.230
Amsterdam	9.402	0.00013	-3.509	-6.426	0.228
Berlin	9.190	0.00013	-3.863	-8.834	0.203
Brussels	9.746	0.00012	-3.467	-6.761	0.195
London	10.670	0.00011	-3.345	-6.035	0.260
Rome	14.826	0.00013	-4.733	-7.522	0.228
Madrid	13.961	0.00010	-4.572	-8.608	0.221

Table 1.2: Parameter estimations for the seasonal function s and mean reverting κ.

2.1.2 Parameter estimation for K and $\sigma(\cdot)$

An important challenge we face when estimating the parameters of the process $(\zeta_t)_{t \geq 0}$ is that the instantaneous volatility process is, per se, unobservable. There is a large literature on this issue, we mention here [2] and [10] that deal with the Heston model. Following Azencott *et al.* approach [10], we approximate the unobservable volatility ζ by the series of realized volatilities $\hat{\zeta}$. These realized volatilities $\hat{\zeta}$ correspond to the observed volatility on a time window of Q-days such that we get:

$$\hat{\zeta}_{iQ\Delta} := \frac{1}{Q} \sum_{j=1}^{Q} \frac{2\hat{\kappa}}{1-e^{-2\hat{\kappa}\Delta}} \left(\tilde{T}_{(iQ+j)\Delta} - e^{-\hat{\kappa}\Delta} \tilde{T}_{(iQ+j-1)\Delta} \right)^2, \quad i \in \{0, \ldots, \lfloor N/Q \rfloor - 1\}. \quad (2.1)$$

Here, $\hat{\zeta}_{iQ\Delta}$ corresponds to the realized volatility on $[iQ\Delta, (i+1)Q\Delta]$ and we have thus $I = \lfloor N/Q \rfloor$ different values. The autoregressive factor comes from the integration of the temperature dynamics of Model (M) which is explicit in Equation (4.5) in Appendix 6.2. The correction factor $\frac{2\hat{\kappa}}{1-e^{-2\hat{\kappa}\Delta}}$, which is not present in [10], is related to the mean-reverting behaviour of the temperature and is justified by Remark 2.1. Since $\frac{2\hat{\kappa}}{1-e^{-2\hat{\kappa}\Delta}} \approx_{\Delta \to 0} \frac{1}{\Delta}$, $\hat{\zeta}_{iQ\Delta}$ is close to the usual quadratic variation, but the difference is not negligible as Δ cannot be smaller than one day on observed data.

Remark 2.1. *Let us suppose that $\zeta_t = \zeta_{iQ\Delta}$ for $t \in [iQ\Delta, (i+1)Q\Delta]$. Then, for $j \in \{1, \ldots, Q\}$, we have*

$$\tilde{T}_{(iQ+j)\Delta} - e^{-\hat{\kappa}\Delta} \tilde{T}_{(iQ+j-1)\Delta} = \sqrt{\zeta_{iQ\Delta}} \int_{(iQ+j-1)\Delta}^{(iQ+j)\Delta} e^{(iQ+j)\Delta-s} dW_s^\rho,$$

with $W^\rho = \rho W + \sqrt{1-\rho^2} Z$. Since $\int_{(iQ+j-1)\Delta}^{(iQ+j)\Delta} e^{(iQ+j)\Delta-s} dW_s^\rho \sim \mathcal{N}\left(0, \frac{1-e^{-2\hat{\kappa}\Delta}}{2\hat{\kappa}}\right)$ is independent of $\zeta_{iQ\Delta}$, we get that $\frac{1}{\zeta_{iQ\Delta}} \sum_{j=1}^{Q} \frac{2\hat{\kappa}}{1-e^{-2\hat{\kappa}\Delta}} \left(\tilde{T}_{(iQ+j)\Delta} - e^{-\hat{\kappa}\Delta} \tilde{T}_{(iQ+j-1)\Delta} \right)^2$ follows a chi-squared

distribution with Q degrees of freedom. Thus, if ζ_t were frozen for $t \in [iQ\Delta, (i+1)Q\Delta]$, $\hat{\zeta}_{iQ\Delta}$ would be an unbiased estimator, i.e. $\mathbb{E}[\hat{\zeta}_{iQ\Delta}|\mathcal{F}_{iQ\Delta}] = \zeta_{iQ\Delta}$.

Estimation of σ^2 and K. Once again, we use the conditional least squares method (CLSE) to simultaneously compute the parameters of the deterministic component of the volatility and the mean-reversion coefficient. For this, we would like to minimise

$$\sum_{i=0}^{N-1} \left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}]\right)^2$$

and use Proposition 4.2. The convergence of such kind of estimators for a time inhomogeneous CIR is given by Theorem 1. However, since the volatility is not directly observed, we minimise the difference of the realized volatility and its conditional expectation given by the previously realised volatilities. Namely, we apply Proposition 4.2 to minimise

$$\sum_{i=0}^{I-2} \left(\zeta_{(i+1)Q\Delta} - \mathbb{E}[\zeta_{(i+1)Q\Delta}|\zeta_{iQ\Delta}]\right)^2,$$

replacing the volatility $\zeta_{iQ\Delta}$ by the realized volatility $\hat{\zeta}_{iQ\Delta}$. This leads to the following estimators of the volatility dynamics of Model (M):

$$
\begin{cases}
\hat{\gamma}_0 = \dfrac{\hat{\theta}_0}{1 - \hat{\phi}_0} \\[2ex]
\hat{K} = -\dfrac{1}{Q\Delta} \ln(\hat{\phi}_0) \\[2ex]
\hat{\gamma}_k = \dfrac{\hat{\theta}_k D_k - \hat{\phi}_k B_k}{A_k D_k - C_k B_k} \\[2ex]
\hat{\delta}_k = \dfrac{\hat{\theta}_k C_k - \hat{\phi}_k A_k}{C_k B_k - A_k D_k}
\end{cases}
\tag{2.2}
$$

where $k \in \{1, 2\}$ and

$$
\begin{cases}
A_k = \hat{K}\dfrac{\hat{K}(\cos(\xi_k Q\Delta) - e^{-\hat{K}Q\Delta}) + \xi_k \sin(\xi_k Q\Delta)}{\hat{K}^2 + \xi_k^2} \\[2ex]
B_k = -\hat{K}\dfrac{\hat{K}\sin(\xi_k Q\Delta) - \xi_k(\cos(\xi_k Q\Delta) - e^{-\hat{K}Q\Delta})}{\hat{K}^2 + \xi_k^2} \\[2ex]
C_k = \hat{K}\dfrac{\hat{K}\sin(\xi_k Q\Delta) - \xi_k(\cos(\xi_k Q\Delta) - e^{-\hat{K}Q\Delta})}{\hat{K}^2 + \xi_k^2} \\[2ex]
D_k = \hat{K}\dfrac{\hat{K}(\cos(\xi_k Q\Delta) - e^{-\hat{K}Q\Delta}) + \xi_k \sin(\xi_k Q\Delta)}{\hat{K}^2 + \xi_k^2},
\end{cases}
$$

and

$$\hat{\vartheta} := (\hat{\theta}_0, \hat{\phi}_0, \hat{\theta}_1, \hat{\theta}_2, \hat{\phi}_1, \hat{\phi}_2)^T = \left(\sum_{i=0}^{I-2} \hat{X}'_{iQ\Delta} \hat{X}'^T_{iQ\Delta} \right)^{-1} \left(\sum_{i=0}^{I-2} \hat{X}'_{iQ\Delta} \hat{\zeta}_{(i+1)Q\Delta} \right), \quad (2.3)$$

with $\hat{X}'_{iQ\Delta} = (1, \hat{\zeta}_{iQ\Delta}, \sin(\xi_1 iQ\Delta), \sin(\xi_2 iQ\Delta), \cos(\xi_1 iQ\Delta), \cos(\xi_2 iQ\Delta))^T$.

Table 1.3 summarises the numerical implementation of the parameter estimation for our eight cities. We can again observe a coherence between the different cities. It can also be noted that $\hat{\gamma}_0$ has more importance in the σ^2 than the trigonometric components. Finally the mean reverting parameter K is more unstable than κ through the cities. Its influence can differ from one to 7 days depending on the city.

City	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{K}$
Stockholm	4.790	0.684	-0.450	1.401	0.704	0.147
Paris	5.603	0.201	-0.266	0.358	0.459	0.396
Amsterdam	4.690	0.503	-0.524	0.500	0.604	0.335
Berlin	5.857	0.646	-0.542	0.410	0.578	0.255
Brussels	4.490	0.266	-0.337	0.298	0.431	0.255
London	4.387	-0.014	-0.314	0.479	0.174	0.774
Rome	3.086	0.212	-0.391	1.335	0.373	0.332
Madrid	4.164	0.418	-0.267	0.746	0.443	0.269

Table 1.3: Parameter estimations for the seasonal function σ^2 and volatility mean reverting parameter K.

We also tested the choice of setting $K_{\sigma^2} = 1$. Table 1.4 summarizes the estimation of the different parameters for $K_{\sigma^2} = 1$. We can see that the impact on the estimations of γ_0, γ_1, δ_1 and K is small or nill. Hence, while all the parameters γ_2 and δ_2 are significant, we find that setting $K_{\sigma^2} = 1$ or 2 has a rather small impact in the reasoning that follows.

City	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\delta}_1$	$\hat{K}$
Stockholm	4.790	0.674	1.408	0.137
Paris	5.603	0.198	0.359	0.332
Amsterdam	4.691	0.496	0.509	0.256
Berlin	5.857	0.644	0.416	0.226
Brussels	4.491	0.265	0.300	0.233
London	4.387	-0.022	0.479	0.430
Rome	3.086	0.198	1.338	0.273
Madrid	4.164	0.414	0.749	0.243

Table 1.4: Parameter estimations for the seasonal function σ^2 and mean reverting K.

In addition, Figure 1.6 shows the plots corresponding to the observed volatility and estimated seasonality on volatility. We can see that while the seasonality does not seem negligible σ^2 is far from completely explaining the observed volatility. Indeed, we observe important fluctuations around $\sigma^2(t)$ on the dynamics of ζ.
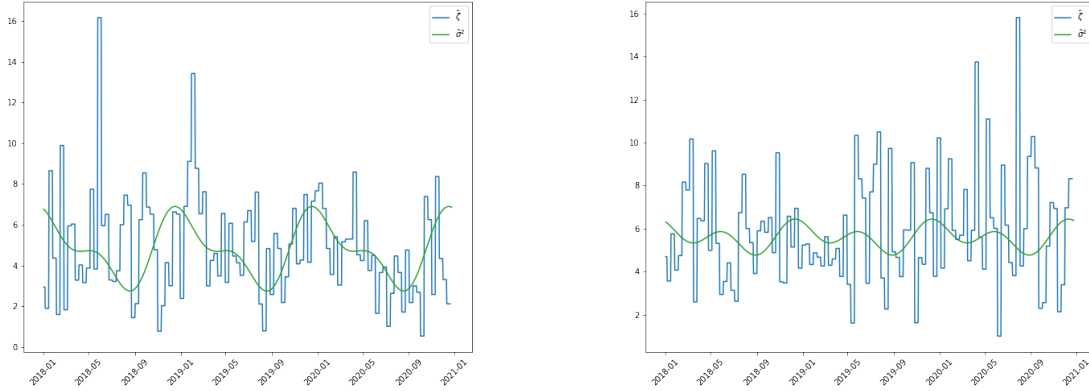

Figure 1.6: Realized volatility and trend on observed volatility from 2018 to 2020 in Stockholm and Paris (from left to right respectively).

2.1.3 Estimation of the parameters η^2 and ρ

Estimation of the volatility of volatility η^2 While in the previous section we use the conditional expectation to estimate σ^2 and K, here the idea is to implement a similar approach based on conditional variance. Namely, we want to solve:

$$\min_{\eta^2} \quad \sum_{i=0}^{N-1} \left((\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}])^2 - \mathbb{E}[(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}])^2|\zeta_{i\Delta}] \right)^2$$

Note that Li and Ma [105] or Bolyog and Pap [34] do not study the properties of such conditional second moment estimators. Here, we show in Theorem 2 the convergence of such estimator for a time-dependent CIR process, when the values $\zeta_{i\Delta}$ are directly observed.

Again, we have to work with the estimated volatility $\hat{\zeta}_{iQ\Delta}$ since we do not directly observe the process ζ. We then apply Proposition 4.3 with the previously estimated parameters $\hat{K}, \hat{\gamma}, \hat{\delta}$. This leads to the following estimator

$$\widehat{\eta^2} = \frac{\sum_{i=0}^{I-2} \hat{Y}_{iQ\Delta}(\hat{\zeta}_{(i+1)Q\Delta} - \hat{\vartheta}^T \hat{X}'_{iQ\Delta})^2}{\sum_{i=0}^{I-2} \hat{Y}_{iQ\Delta}^2}, \tag{2.4}$$

where $\hat{\vartheta}^T$ and $\hat{X}'_{iQ\Delta}$ are defined by (2.3) and

$$\hat{Y}_{iQ\Delta} = \theta'_0 + \phi'_0 \hat{\zeta}_{iQ\Delta} + \sum_{k=1}^{2} \theta'_k \sin(\xi_k iQ\Delta) + \sum_{k=1}^{2} \phi'_k \cos(\xi_k iQ\Delta)$$

with

$$\begin{cases} \theta'_0 = \hat{\gamma}_0 \dfrac{(1 - e^{-\hat{K}Q\Delta})^2}{2\hat{K}} \\[2mm] \phi'_0 = \dfrac{e^{-\hat{K}Q\Delta}}{\hat{K}}(1 - e^{-\hat{K}Q\Delta}) \\[2mm] \theta'_k = \hat{\gamma}_k \hat{K} \dfrac{\hat{K}[A'_k - \phi'_0] + \xi_k B'_k}{\hat{K}^2 + \xi_k^2} - \hat{\delta}_k \hat{K} \dfrac{\hat{K}B'_k - \xi_k[A'_k - \phi'_0]}{\hat{K}^2 + \xi_k^2} \\[2mm] \phi'_k = \hat{\gamma}_k \hat{K} \dfrac{\hat{K}B'_k - \xi_k[A'_k - \phi'_0]}{\hat{K}^2 + \xi_k^2} + \hat{\delta}_k \hat{K} \dfrac{\hat{K}[A'_k - \phi'_0] + \xi_k B'_k}{\hat{K}^2 + \xi_k^2}, \end{cases}$$

and

$$\begin{cases} A'_k = \dfrac{2\hat{K}[\cos(\xi_k Q\Delta) - e^{-2\hat{K}Q\Delta}] + \xi_k \sin(\xi_k Q\Delta)}{4\hat{K}^2 + \xi_k^2} \\[3mm] B'_k = \dfrac{2\hat{K}\sin(\xi_k Q\Delta) - \xi_k(\cos(\xi_k Q\Delta) - e^{-2\hat{K}Q\Delta})}{4\hat{K}^2 + \xi_k^2} \end{cases}$$

and $(\hat{\gamma}_0, \hat{K}, \hat{\gamma}_k, \hat{\delta}_k)$ defined by Equation (2.2).

city	Stockholm	Paris	Amsterdam	Berlin	Brussels	London	Rome	Madrid
$\widehat{\eta^2}$	0.629	1.043	0.929	0.884	0.713	1.605	0.988	0.737

Table 1.5: Parameter estimations for the volatility of volatility η^2.

The numerical application of the above formula applied to our dataset gives the values collected in Table 1.5. First, comparing Tables 1.3 and 1.5 enables to observe the two components of the volatility dynamics, the mean reversion component and the pure noise. They both have comparable magnitude and therefore both contribute significantly to the volatility dynamics. The mean reverting coefficients are rather small, which is consistent with the fluctuations observed in Figure 1.6. Second, Table 1.5 shows a certain coherence in terms of magnitude for all the European cities.

Estimation of the correlation ρ The last step consists in estimating the correlation ρ. The idea this time is to use conditional covariance to estimate this parameter, and Proposition 4.4

gives the minimiser of the following problem:

$$\min_{\rho} \sum_{i=0}^{N-1} \Big((T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|\mathcal{F}_{i\Delta}])(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\mathcal{F}_{i\Delta}])$$

$$- \mathbb{E}\Big[(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|\mathcal{F}_{i\Delta}])(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\mathcal{F}_{i\Delta}]\Big|\mathcal{F}_{i\Delta}\Big] \Big)^2.$$

Again, since we do not observe the volatility, we use Proposition 4.4 with the estimated volatility $\hat{\zeta}_{iQ\Delta}$ and the previously estimated parameters $\hat{\kappa}, \hat{\alpha}, \hat{\beta}, \hat{K}, \hat{\gamma}, \hat{\delta}, \widehat{\eta^2}$ given by Proposition 6.1, (2.2) and (2.4). This leads to

$$\hat{\rho} = \frac{\sum_{i=0}^{I-2} \hat{Y}'_{iQ\Delta}(T_{(i+1)Q\Delta} - \hat{\lambda}^T X_{iQ\Delta})(\hat{\zeta}_{(i+1)Q\Delta} - \hat{\vartheta}^T \hat{X}'_{iQ\Delta})}{\sum_{i=0}^{I-2}(\hat{Y}'_{iQ\Delta})^2},$$

where $\hat{\lambda}$ and $X_{i\Delta}$ are defined by (4.3) with $N := I - 1$, $\hat{\vartheta}$ and $\hat{X}'_{iQ\Delta}$ are defined by (2.3), and $\hat{Y}'_{iQ\Delta} = \theta''_0 + \phi''_0 \hat{\zeta}_{iQ\Delta} + \sum_k \theta''_k \sin(\xi_k iQ\Delta) + \sum_k \phi''_k \cos(\xi_k iQ\Delta)$, with

$$\begin{cases} \theta''_0 = \hat{\eta}\hat{\gamma}_0 \left(\dfrac{1 - e^{-(\hat{\kappa}+\hat{K})Q\Delta}}{\hat{\kappa} + \hat{K}} + \dfrac{e^{-(\hat{\kappa}+\hat{K})Q\Delta} - e^{-\hat{K}Q\Delta}}{\hat{\kappa}} \right) \\[3mm] \phi''_0 = \hat{\eta} e^{-\hat{K}Q\Delta} \dfrac{1 - e^{-\hat{\kappa}Q\Delta}}{\hat{\kappa}} \\[3mm] \theta''_k = \hat{\eta}\hat{\gamma}_k \hat{K} \dfrac{\hat{K}(A''_k - \phi''_0) + \xi_k B''_k}{\hat{K}^2 + \xi_k^2} - \hat{\eta}\hat{\delta}_k \hat{K} \dfrac{\hat{K}B''_k - \xi_k(A''_k - \phi''_0)}{\hat{K}^2 + \xi_k^2} \\[3mm] \phi''_k = \hat{\eta}\hat{\gamma}_k \hat{K} \dfrac{\hat{K}B''_k - \xi_k(A''_k - \phi''_0)}{\hat{K}^2 + \xi_k^2} + \hat{\eta}\hat{\delta}_k \hat{K} \dfrac{\hat{K}(A''_k - \phi''_0) + \xi_k B''_k}{\hat{K}^2 + \xi_k^2}, \end{cases}$$

and

$$\begin{cases} A''_k = \dfrac{(\hat{K} + \hat{\kappa})(\cos(\xi_k Q\Delta) - e^{-(\hat{K}+\hat{\kappa})Q\Delta}) + \xi_k \sin(\xi_k Q\Delta)}{(\hat{K} + \hat{\kappa})^2 + \xi_k^2} \\[3mm] B''_k = \dfrac{(\hat{K} + \hat{\kappa})\sin(\xi_k Q\Delta) - \xi_k(\cos(\xi_k Q\Delta) - e^{-(\hat{K}+\hat{\kappa})Q\Delta})}{(\hat{K} + \hat{\kappa})^2 + \xi_k^2}. \end{cases}$$

City	Stockholm	Paris	Amsterdam	Berlin	Brussels	London	Rome	Madrid
$\hat{\rho}$	-0.000	-0.006	-0.010	-0.013	-0.014	-0.011	0.023	-0.005

Table 1.6: Parameter estimations for the correlation ρ.

The numerical application of the above formula applied to our data sets gives the values collected in Table 1.6. We observe that the correlation is close to zero for all the cities. Therefore,

for simplification purposes, we will consider on the following of this document that ρ equals zero. This finding also questions the pertinence of GARCH model (1.6) that corresponds to $\rho \in \{-1, 1\}$ since the temperature and its volatility are driven by the same noise.

2.2 Robustness of estimators

In the previous subsection, we have obtained conditional least squares estimators for the different parameters of Model (M). All the above expressions have been computed by discretizing the processes $(\tilde{T}_t)_{t \geq 0}$ and $(\zeta_t)_{t \geq 0}$. Theoretically speaking, Overbeck and Ryden [123] and Bolyog and Pap [34] have proven the convergence of the CLS estimators for the Cox-Ingersoll-Ross and a generalized Heston model. Their proof is mainly based in ergodic arguments. In Appendix 4.6, we prove the consistency of CLSE for a time inhomogeneous Cox-Ingersoll-Ross process.

However, to estimate the parameters of the volatility dynamics, we have approximated the unobservable volatility ζ by the realized volatility $\hat{\zeta}$. Azencott, Ren and Timofeyev [10] have deeply studied the convergence modes of the volatility process $\hat{\zeta}$ to the instantaneous process ζ as well as the estimated realized estimators (K, σ, η) under a classic Heston framework. Under boundary and continuity hypothesis on T and ζ, uniform convergence of $\hat{\zeta}$ to ζ over $[0, T]$ in L^2 is proven. Probability convergence of estimators is also proven for moments based estimators. The extension of these convergence properties to our particular model is left as a further work. In this section, we test numerically the robustness of our estimators and check their accuracy on simulated data.

2.2.1 Methodology

Robustness of the estimators is checked through simulated data. Essentially, we simulate data series with the model that have the same length as our data set (40 years) and we check that we find back the parameters by using the CLS estimators. The detailed methodology is presented in Appendix 4.2.

As a first qualitative check, we have represented in Figure 1.7 an example of simulated temperature for Paris. The simulated paths looks similar to the observed ones. In the following paragraphs, we will test our capacity to estimate the values of the parameters and discuss the choice of averaging-time windows Q.

2.2.2 Estimation of parameters $\kappa, s(\cdot)$ related to the simulated temperature T

The estimation of κ and $s(\cdot)$ is a priori easy since it relies on the temperature that is directly observable. Table 1.8 summarizes the estimators of the parameters related to the temperature. Figure 1.8 corresponds to the related temperature plot for Stockholm and Paris. We can see that all the estimated parameters remain very close to the original values. We can thus conclude that the estimation of the parameters of the temperature is robust enough to be reliable.
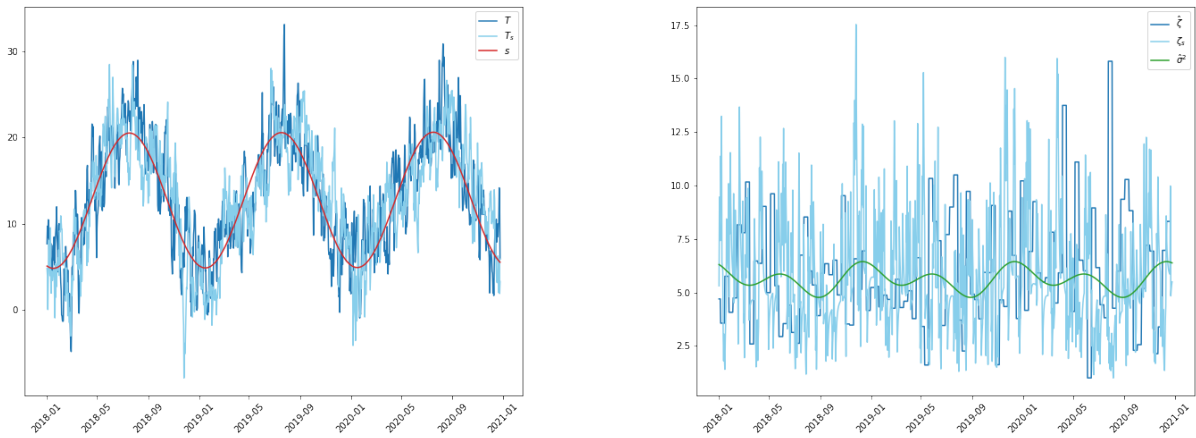
Figure 1.7: Plots of simulated temperature and volatility processes for Paris and $(K, \eta^2) = (0.396, 1.043)$. On the left, we plot the observed temperature T (blue), the simulated temperature T_s (light blue) and the trend and seasonal function s. On the right, we plot the observed volatility $\hat{\zeta}$ (blue), defined here as the 10-lag moving average of ζ, the simulated volatility ζ_s (light blue) and the seasonal volatility function σ^2.

City	α_0	$\hat{\alpha}_0$	β_0	$\hat{\beta}_0$	α_1	$\hat{\alpha}_1$	β_1	$\hat{\beta}_1$	κ	$\hat{\kappa}$
Stockholm	6.678	6.966	0.00016	0.00016	-4.564	-4.564	-9.142	-9.142	0.192	0.192
Paris	10.868	10.733	0.00013	0.00013	-3.540	-3.540	-6.993	-6.993	0.230	0.235
Amsterdam	9.402	9.353	0.00013	0.00013	-3.509	-3.509	-6.426	-6.426	0.228	0.220
Berlin	9.190	9.472	0.00013	0.00013	-3.863	-3.863	-8.834	-8.834	0.203	0.200
Brussels	9.746	9.350	0.00012	0.00012	-3.467	-3.467	-6.761	-6.761	0.195	0.192
London	10.670	10.659	0.00011	0.00011	-3.345	-3.345	-6.035	-6.035	0.260	0.270
Rome	14.826	14.751	0.00013	0.00013	-4.733	-4.733	-7.522	-7.522	0.228	0.224
Madrid	13.961	13.763	0.00010	0.00010	-4.572	-4.572	-8.608	-8.608	0.221	0.232

Table 1.7: Estimation of temperature parameters from the simulated temperature path.

2.2.3 Estimation of K, $\sigma^2(\cdot)$ and η^2 on the simulated realized volatility $\hat{\zeta}$

Let us recall that instantaneous volatility ζ is not observable and we estimate it by $\hat{\zeta}$ defined in Equation (2.1). In this paragraph, we apply the same process as in Section 2.1 to simulated data.

We first focus on the effect of Q, i.e. the size of the averaging window (2.1), on the estimation of parameters. For simplicity purposes, we start by setting the trigonometric coefficients of σ^2 to 0 and study the impact on K and η^2. For each time window Q and city, we perform $50,000$ simulations of daily volatility ζ, by using (4.1). We average this series trough the time window Q and then estimate the corresponding (K, η^2). This exercise enables to analyse the influence

City	α_0	$\hat{\alpha}_0$	β_0	$\hat{\beta}_0$	α_1	$\hat{\alpha}_1$	β_1	$\hat{\beta}_1$	κ	$\hat{\kappa}$
Stockholm	6.678	6.939	0.00016	0.00015	-4.564	-5.492	-9.142	-8.355	0.192	0.185
Paris	10.868	10.712	0.00013	0.00015	-3.54	-4.276	-6.993	-6.577	0.230	0.229
Amsterdam	9.402	9.647	0.00013	0.00009	-3.509	-4.110	-6.426	-6.107	0.228	0.225
Berlin	9.190	9.429	0.00013	0.00011	-3.863	-4.804	-8.834	-8.325	0.203	0.197
Brussels	9.746	10.011	0.00012	0.0001	-3.467	-3.954	-6.761	-6.232	0.195	0.195
London	10.670	10.830	0.00011	0.0001	-3.345	-4.059	-6.035	-5.439	0.260	0.259
Rome	14.826	14.781	0.00013	0.00013	-4.733	-5.522	-7.522	-6.933	0.228	0.235
Madrid	13.961	14.081	0.00010	0.00008	-4.572	-5.377	-8.608	-7.95	0.221	0.224

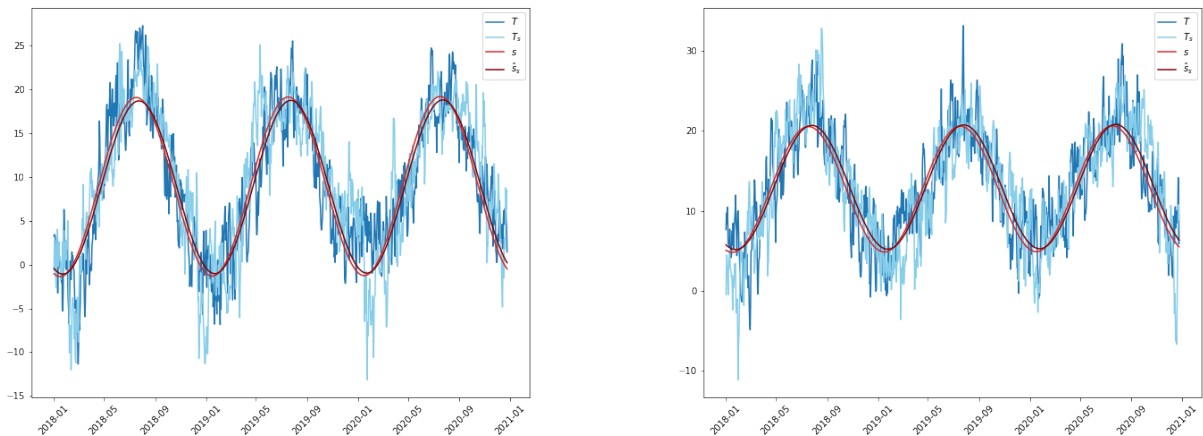Table 1.8: Estimation of temperature parameters from the simulated temperature path.

Figure 1.8: Estimation of the trend and seasonal function s on both real temperature T (blue) and simulated temperature T_s (light blue) for Stockholm (left) and Paris (right) and $Q = 10$. This function is plotted in red (resp. brown) when estimated on observed (resp. simulated) data.

of the time window Q on the capacity to well estimate (K, η^2).

Table 1.9 and 1.10 represent different estimates of (K, η^2) depending on window width Q. First, we can observe that for small Q, (K, η^2) is overestimated due to the preponderance of the noise related to volatility estimation. On the contrary, for large values of Q, (K, η^2) is underestimated due to the averaging effect. Therefore there is a need to find a trade-off between the effect of the noise and the effect of the window averaging. Depending on the city the most efficient Q can range from 5 to 10.

We now focus on the impact of Q on the estimation of the function σ^2. We have plotted in Figure 1.9 the estimated trend with the original one for $Q = 5$ and $Q = 12$. We see that σ^2 is correctly estimated, independently of Q. From this numerical study, the choice of $Q = 10$ appears to be quite reliable. Of course, as one may expect, the parameters are not as well

City	$\hat{K}$	$\hat{K}_{Q=1}$	$\hat{K}_{Q=2}$	$\hat{K}_{Q=5}$	$\hat{K}_{Q=8}$	$\hat{K}_{Q=10}$	$\hat{K}_{Q=12}$
Stockholm	0.147	2.261	0.886	0.301	0.190	0.157	0.140
Paris	0.396	2.853	1.336	0.552	0.403	0.286	0.265
Amsterdam	0.335	2.578	1.159	0.463	0.345	0.266	0.260
Berlin	0.255	2.590	1.083	0.408	0.278	0.243	0.208
Brussels	0.255	2.540	1.042	0.401	0.262	0.228	0.208
London	0.774	3.363	1.637	0.880	0.651	0.459	0.464
Rome	0.332	2.381	1.059	0.433	0.303	0.264	0.216
Madrid	0.269	2.495	1.067	0.407	0.278	0.260	0.219

Table 1.9: Estimation of K for different averaging time windows Q.

City	$\widehat{\eta^2}$	$\widehat{\eta^2}_{Q=1}$	$\widehat{\eta^2}_{Q=2}$	$\widehat{\eta^2}_{Q=5}$	$\widehat{\eta^2}_{Q=8}$	$\widehat{\eta^2}_{Q=10}$	$\widehat{\eta^2}_{Q=12}$
Stockholm	0.629	56.229	12.288	2.123	0.896	0.644	0.499
Paris	1.043	56.429	13.385	2.506	1.156	0.690	0.531
Amsterdam	0.929	46.609	11.364	2.013	1.010	0.625	0.528
Berlin	0.884	61.162	13.930	2.377	1.070	0.795	0.580
Brussels	0.713	47.390	10.652	1.921	0.834	0.599	0.474
London	1.605	45.348	11.535	2.608	1.236	0.723	0.588
Rome	0.988	35.879	8.962	1.686	0.787	0.561	0.407
Madrid	0.737	43.787	10.316	1.862	0.809	0.657	0.430

Table 1.10: Estimation of η^2 for different averaging time windows Q

estimated as for κ and $s(\cdot)$. They still however give the correct magnitude of the parameters, which is acceptable for risk management. Since we are then interested in evaluating derivative products, we will then analyse the sensitivity to these parameters, see Section 3.4, which can be done efficiently in Model (M).

3 Application to pricing weather derivatives

The previous sections concentrate on modeling the daily average temperature and on the estimation of the parameters of the model. However, the final objective of our model is to better assess the risk related to weather temperature derivatives. This section will focus on how we evaluate the average payoff of these derivatives and how Model (M) improves our capacity to understand their risk.
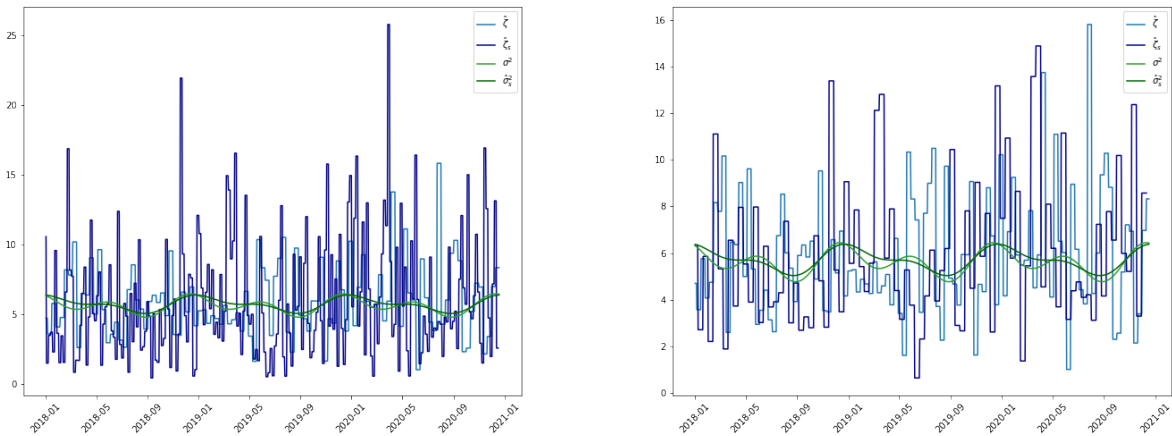
Figure 1.9: Plots of observed volatility process $\hat{\zeta}$ (blue) and simulated volatility processes $\hat{\zeta}_s$ (dark blue) for Paris for averaging windows Q equals 5 (left) and 12 (right). The function σ^2 is in green while the estimated functions for $Q = 5$ and $Q = 12$ are in dark green.

3.1 Temperature derivatives

3.1.1 Average temperature indices

Temperature derivatives are financial products used to hedge weather risk. The covers are often based on an index corresponding to a proxy of the buyer's financial risk. This index corresponds to an aggregate of a more granular meteorological parameter which in this document is the average daily temperature. There exist different possible indices, the main ones being HDD (Heating Degree Days), CDD (Cooling Degree Days) and CAT (Cumulative Average Temperature):

$$HDD := \sum_{t=t_1}^{t_2} \max(0, T_b - T_t), \ CDD := \sum_{t=t_1}^{t_2} \max(0, T_t - T_b), \ CAT := \sum_{t=t_1}^{t_2} T_t,$$

where T_t corresponds to the average daily temperature on day t, T_b to a base temperature, t_1 to the inception date and t_2 the exit date of the contract. We will call risk period the time period between t_1 and t_2.

Physically speaking, the HDD corresponds to the cumulative degrees needed to heat a given building. Hence the base temperature T_b corresponds to the temperature at which heating is probably switched on and the cumulative HDD measures the demand on energy of the building. The base temperature T_b varies depending on the country. In the EU, T_b is often taken as equal to $15.5°C$ while in the US it corresponds to $65°F$. Symmetrically, CDD corresponds to the energy demand for air conditioning. Finally, CAT corresponds to cumulative temperature

degrees which is related to the energy demand between t_1 and t_2.

In the following of this document, we will focus on the HDD index however the methodology presented can be applied to all average temperature related indices. Particularly for options on the CAT, the Fast Fourier Transform approach would enable to get a very efficient pricing method.

3.1.2 Payoff function

Weather derivatives are used to hedge weather risk. They trigger a payment depending on an aggregate temperature index. The payment is defined given a payoff structure. Standard payoff structures correspond to capped put or call options applied to the aggregate index. For simplification purposes we will suggest the payoff structure:

$$\min((HDD - HDD_{strike})^+, L). \tag{3.1}$$

As we want to price this kind of instruments our objective is to understand the characteristics of the payment distribution (expectation, VaR and CVaR) under the real world probability. In particular, we consider the average payoff of the derivative

$$\mathbb{E}[D(t_0, t_2) \min((HDD - HDD_{strike})^+, L)], \tag{3.2}$$

where $D(t_0, t_2)$ is a discount rate that will be taken equal to 1 in this study. Note that this is not a fair price: there is no market dealing HDD continuously and therefore the classical pricing theory of Black and Scholes does not apply. The calculation of the average payoff (3.2), as well as other indicators on the distribution of $\min((HDD - HDD_{strike})^+, L)$ such as the variance and quantiles, is used in practice to propose a price over the counter. Thus, a very accurate evaluation of (3.2) for a given model is not really at stake: one is more interested in evaluating risk and how the average payoff may change under stressed parameters.

Finally, there is no consensus on how to choose HDD_{strike}. However, it is a market practice to use quantiles, and particularly historical quantiles of the index, to define this strike. In the present study, we consider the 90% quantile which is within market practices.

3.2 Monte-Carlo Approach

A first pricing approach to identify the distribution of payments is to simulate temperature paths based on the discretization schemes in (4.1) for $\Delta = 1$. We proceed as follows:

1. Simulate temperature paths starting from the pricing date t_0, the day until which we can observe temperature data, to the expiration date t_2.

2. Compute simulated HDD between t_1 and t_2 for each of the paths and obtain an HDD distribution.

3. Either fix an arbitrary HDD_{strike} or choose a quantile to select the moneyness of the structure.

4. Deduce the payment distribution.

5. Compute payment distribution characteristics: mean, VaR and CVaR.

Figure 1.10 shows the results of this method for Paris temperature in 2019. Contracts last one month and are computed 30 days in ahead i.e. $t_1 - t_0 = 30$. We consider the payoff function (3.1) with $L = +\infty$ and HDD_{strike} set to the 90% empirical quantile of the HDD distribution obtained with Model (M). We perform $50,000$ Monte Carlo simulations for the Ornstein–Uhlenbeck model and for Model (M). All these choices are challenged in the following sections.
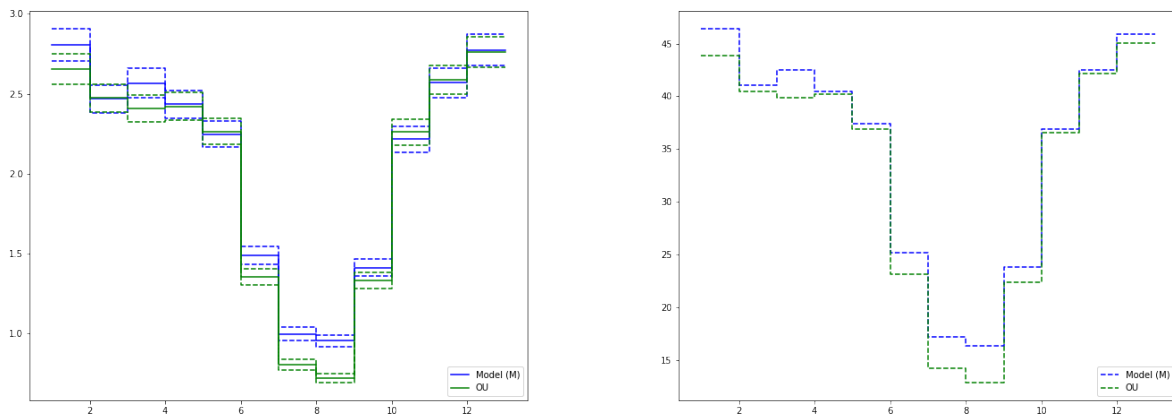
Mean with 95% confidence interval   Conditional Value at Risk at 95%

Figure 1.10: Different metrics of the payment distribution for $50,000$ Monte Carlo simulations, Paris, a cumulation period of a month, a forecast 30 days ahead and HDD_{strike} corresponding to a 90% quantile of the monthly HDD. Monte Carlo simulations are performed for both Model (M) and the Ornstein–Uhlenbeck model (1.1).

From Figure 1.10 we can see that both Model (M) and the Ornstein–Uhlenbeck model lead to similar expected payoffs for winter months while Model (M) states higher expected payoffs for summer months. Model (M) tends to have heavier left tails which are particularly visible in summer months. This explains slightly higher mean payoffs for Model (M) during these months. Nevertheless, it should be noted that these derivatives are mainly sold for winter months to cover against cold waves. During these months both models give similar mean payoffs.

In terms of risk metrics, we compute Conditional Value at Risks at 95% for both models. We can see that Model (M) presents again heavier tails. Thus, as already noticed from Figures 1.4 and 1.5, Model (M) is more conservative and reduces the problem of underestimating rare events related to the Ornstein–Uhlenbeck Gaussian framework.

3.3 Fast Fourier Transform Approach

This section explores an alternative methodology for HDD pricing thanks to the Fast Fourier Transform (FFT) approach developed by Carr and Madan [48]. This method has been widely used in the literature for pricing, we just mention here the recent work of Benth et al. [17] for an application close to ours.

3.3.1 The characteristic function

In order to apply FFT pricing, we first calculate the characteristic function of $(\tilde{T}_t, \zeta_t, \int_0^t \tilde{T}_s ds)$. A semi explicit formula is available because of the affine structure of Model (M).

Proposition 3.1. *Let $0 \leq t \leq t'$. Let $(\tilde{T}, \zeta)$ be the solution of (M) with $\rho = 0$. The characteristic function of $(\tilde{T}_{t'}, \zeta_{t'})$ given $\mathcal{F}_t$ is, for $u_1, u_2, u_3 \in \mathbb{R}$,*

$$\mathbb{E}\left[\exp\left(i[u_1\tilde{T}_{t'} + u_2\zeta_{t'} + u_3 \int_t^{t'} \tilde{T}_s ds]\right)|\mathcal{F}_t\right] = \exp(a_0(t,t') + a_1(t'-t)\tilde{T}_t + a_2(t'-t)\zeta_t), \quad (3.3)$$

where a_2 is the unique solution on $\mathbb{R}_+$ of the time inhomogeneous autonomous Riccati equation

$$a_2' = -Ka_2 - \frac{1}{2}\left[u_1\exp(-\kappa t) + u_3\frac{1-\exp(-\kappa t)}{\kappa}\right]^2 + \frac{1}{2}\eta^2 a_2^2, \quad a_2(0) = iu_2, \quad (3.4)$$

$a_1(t) = iu_1\exp(-\kappa t) + iu_3\frac{1-\exp(-\kappa t)}{\kappa}$ and $a_0(t,t') = K\int_t^{t'}\sigma^2(s)a_2(t'-s)ds$. Besides, the real part of $a_2(t)$ remains nonpositive for all $t \geq 0$.

Proof. Let us first check that Equation (3.4) admits a unique solution, which is well defined for all $t \geq 0$. When $u_1 = u_2 = u_3 = 0$, $a_2(t) = 0$ is the unique solution and we get then $a_1(t) = 0$, $a_0(t,t')$ so that (3.3) holds. We now exclude this case, and observe that

$$\begin{cases} \mathfrak{R}(a_2') = -K\mathfrak{R}(a_2) - \frac{1}{2}\left[u_1\exp(-\kappa t) + u_3\frac{1-\exp(-\kappa t)}{\kappa}\right]^2 + \frac{1}{2}\eta^2(\mathfrak{R}(a_2)^2 - \mathfrak{I}(a_2)^2), \quad \mathfrak{R}(a_2(0)) = 0, \\ \mathfrak{I}(a_2') = -K\mathfrak{I}(a_2) + \eta^2\mathfrak{R}(a_2)\mathfrak{I}(a_2), \quad \mathfrak{I}(a_2(0)) = u_2, \end{cases}$$

with $\mathfrak{R}(z)$ and $\mathfrak{I}(z)$ denoting the real and imaginary parts of a complex number z. Let $\bar{t} = \inf\{t \geq 0 : \mathfrak{R}(a_2(t)) > 0\}$. Since $\mathfrak{R}(a_2(0)) = 0$, we have $\mathfrak{R}(a_2'(0)) = -\frac{1}{2}(u_1^2 + \eta^2 u_2^2) < 0$ when $u_1 \neq 0$ or $u_2 \neq 0$ and thus we have $\bar{t} > 0$. If $u_1 = u_2 = 0$ and $u_3 \neq 0$, we have $\mathfrak{I}(a_2(t)) = 0$, $\mathfrak{R}(a_2'(0)) = 0$, $\mathfrak{R}(a_2''(0)) = 0$ and $\mathfrak{R}(a_2'''(0)) = -u_3^2 < 0$ and thus again $\bar{t} > 0$. Then, we have $\mathfrak{I}(a_2(t)) = u_2\exp\left(-Kt + \eta^2\int_0^t \mathfrak{R}(a_2(s))ds\right)$ and thus $|\mathfrak{I}(a_2(t))| \leq |u_2|$ for $t \in [0,\bar{t})$. We now observe that $\bar{t}$ cannot be finite. If it were finite, we would have $\mathfrak{R}(a_2(\bar{t})) = 0$ by continuity and then

$$\mathfrak{R}(a_2'(\bar{t})) = -\frac{1}{2}\left[u_1\exp(-\kappa\bar{t}) + u_3\frac{1-\exp(-\kappa\bar{t})}{\kappa}\right]^2 - \frac{1}{2}\eta^2\mathfrak{I}(a_2(\bar{t}))^2 \leq 0.$$

If $\Re(a_2'(\bar{t})) < 0$, we get $\Re(a_2(t)) > 0$ in a left neighbourhood of $\bar{t}$ which is impossible. If $\Re(a_2'(\bar{t})) = 0$, we then have $u_1 \exp(-\kappa \bar{t}) + u_3 \frac{1-\exp(-\kappa \bar{t})}{\kappa} = 0$ and $\Im(a_2(\bar{t})) = 0$. The latter gives $u_2 = 0$. We check then that $\Re(a_2''(\bar{t})) = 0$ and $\Re(a_2'''(\bar{t})) = -(u_3 - \kappa u_1)^2 e^{-2\kappa \bar{t}} < 0$ since $u_1 \exp(-\kappa \bar{t}) + u_3 \frac{1-\exp(-\kappa \bar{t})}{\kappa} = 0$ and $(u_1, u_3) \neq (0,0)$. Again, this gives that $\Re(a_2(t)) > 0$ in a left neighbourhood of $\bar{t}$ which is impossible. Thus, $\bar{t} = +\infty$ and the ODE is then clearly well defined for all $t \geq 0$.

We now check that we indeed have (3.3). Let $\mathcal{E}_t = \exp(a_0(t,t') + a_1(t'-t)\tilde{T}(t) + a_2(t'-t)\zeta_t + iu_3 \int_0^t \tilde{T}_s ds)$. By Itô's formula, we get for $t \in [0, t']$,

$$
d\mathcal{E}_t = \mathcal{E}_t \left[\partial_t a_0(t,t') - a_1'(t'-t)\tilde{T}_t - a_2'(t'-t)\zeta_t - \kappa a_1(t'-t)\tilde{T}_t + a_2(t'-t)K(\sigma^2(t) - \zeta_t) \right.
$$
$$
\left. + \frac{1}{2}a_1^2(t'-t)\zeta_t + \frac{1}{2}\eta^2 a_2^2(t'-t)\zeta_t + iu_3\tilde{T}_t \right] dt + \mathcal{E}_t \sqrt{\zeta_t}[a_1(t'-t)dZ_t + \eta a_2(t'-t)dW_t].
$$

The first term vanishes, and we get

$$
\mathcal{E}_{t'} = \mathcal{E}_t + \int_t^{t'} \mathcal{E}_s \sqrt{\zeta_s}[a_1(t'-s)dZ_s + \eta a_2(t'-s)dW_s]
$$

We note that $0 \leq |\mathcal{E}_t| \leq \exp(a_0(t,t'))$ for $t \in [0,t']$ since $a_1 \in i\mathbb{R}$ and $\Re(a_2) \leq 0$ and that $\mathbb{E}[\zeta_t] = \zeta_0 e^{-Kt} + \int_0^t \sigma^2(s)e^{-K(t-s)}ds$ is integrable with respect to t. Thus, the integrand of the stochastic integral is square integrable, and we get

$$
\mathcal{E}_t = \mathbb{E}[\mathcal{E}_{t'}|\mathcal{F}_t] = \mathbb{E}\left[\exp\left(i[u_1\tilde{T}_{t'} + u_2\zeta_{t'} + u_3 \int_0^{t'} \tilde{T}_s ds] \right) \bigg| \mathcal{F}_t \right],
$$

which gives the claim. □

Remark 3.1. *Formula* (3.3) *can be extended easily to* $u_2 \in \mathbb{R} + i\mathbb{R}_+$. *We then have* $\Re(a_2(0)) = -\Im(u_2) \leq 0$, *and the proof of Proposition* 3.1 *can be repeated step by step.*

3.3.2 Approximation of the characteristic function

We now discuss the approximation of the characteristic function (3.3). To do so, we consider a time step $\delta > 0$, and we will assume that $t = t_k = k\delta$ and $t' = t_l = l\delta$. Note that the function a_1 is fully explicit and does not need to be approximated. We use the trapezoidal rule to integrate the function a_0:

$$
a_0(t_k, t_l) \approx K \sum_{j=k}^{l-1} \frac{1}{2}[\sigma^2(t_j)a_2(t_l-t_j) + \sigma^2(t_{j+1})a_2(t_l-t_{j+1})]\delta, \ k < l.
$$

The main issue may come from the discretization of the Riccati equation which may lead to instabilities if it is not well handled. Here, we take advantage of the fact that an explicit solution of (3.4) is known for $\kappa = 0$ and $u_3 = 0$, see e.g. [6, p. 101],

$$a_2(t) = \Psi + \frac{2\sqrt{D}(\Psi - iu_2)}{\left(\eta^2(\Psi - iu_2) - 2\sqrt{D}\right)\exp(-\sqrt{D}t) - \eta^2(\Psi - iu_2)},$$

with

$$D = K^2 + \eta^2 u_1^2, \quad \Psi = \frac{K + \sqrt{D}}{\eta^2}.$$

Thus, to solve (3.4), we freeze on each interval $[t_k, t_{k+1}]$ the value of the time inhomogenous term to its value at $t = \frac{t_k + t_{k+1}}{2}$, and use the explicit formula. This is the midpoint method that leads formally to a convergence of order $O(\delta^2)$. This leads to:

$$a_2(t_{k+1}) = \Psi_k + \frac{2\sqrt{D_k}(\Psi_k - a_2(t_k))}{\left(\eta^2(\Psi_k - a_2(t_k)) - 2\sqrt{D_k}\right)\exp(-\sqrt{D_k}\delta) - \eta^2(\Psi_k - a_2(t_k))} \tag{3.5}$$

where

$$D_k = K^2 + \eta^2\left(u_1\exp\left(-\kappa\frac{t_k + t_{k+1}}{2}\right) + u_3\frac{1 - \exp\left(-\kappa\frac{t_k + t_{k+1}}{2}\right)}{\kappa}\right)^2, \quad \Psi_k = \frac{K + \sqrt{D_k}}{\eta^2}.$$

We implement the three functions in (3.5) which enable us to deduce the characteristic function of $(u_1\tilde{T}(t) + u_2\zeta_t)_{t \geq 0}$ where $(u_1, u_2) \in \mathbb{C}^2$. Figure 1.11 shows the characteristic function of $(T_t)_{t \geq 0}$ calculated with the approximation (3.5). It is compared with the Monte-Carlo estimator obtained with simulated path using (4.1) (we have used here independent simulations for each values of t'). We can see that both methods give close results, which validates the relevance of the approximation.

3.3.3 Fast Fourier Transform for pricing HDD and related options

Once we have the characteristic function of $(T_t)_{t \geq 0}$, we can price HDD by using Fourier inverse transform techniques. Here, we adapt the approach of Carr and Madan [48] that uses the Fast Fourier Transform in order to calculate the cumulative distribution function of T_t. This allows us to calculate then easily the average value different types of bespoke options.

We first focus on the calculation of $\mathbb{E}[(T_b - T_t)^+]$ (resp. $\mathbb{E}[\min((T_b - T_t)^+, L)]$), that can be seen as the average price of a "daily HDD" (resp. capped daily HDD). We will use this naming
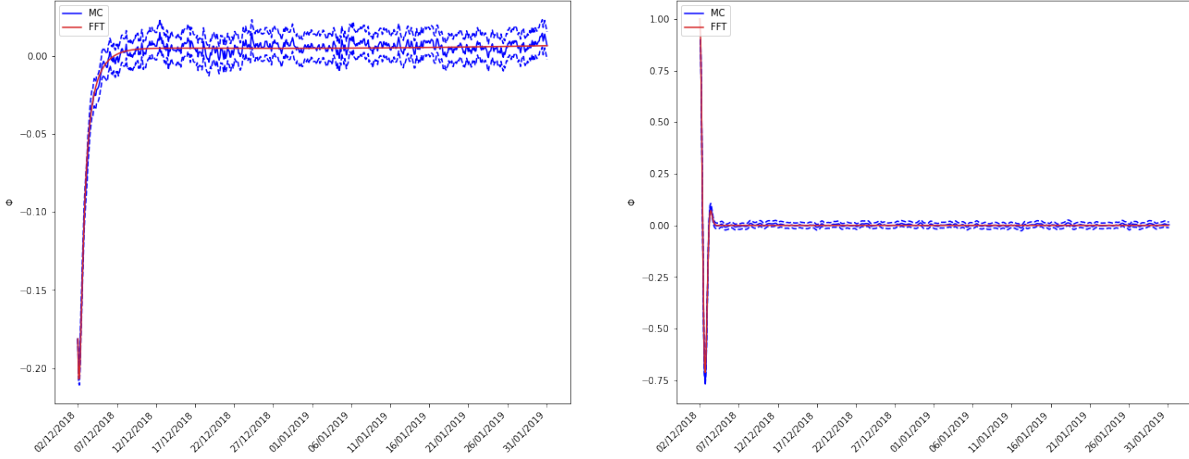
Figure 1.11: Characteristic function $\mathbb{E}\left[\exp\left(iu_1\tilde{T}_{t'}\right)\right]$ (left) and $\mathbb{E}\left[\exp\left(iu_3\int_t^{t'}\tilde{T}_s ds\right)|\mathcal{F}_t\right]$ (right) for Paris temperature during January 2019 for an observation time 30 days ahead and $\delta = 0.1$ day.

later on. The characteristic function $\Phi(u) = \mathbb{E}[e^{iu\tilde{T}_t}]$ is given by Proposition 3.1. We have

$$\mathbb{E}[(T_b - T_t)^+] = \int_0^\infty \mathbb{P}(T_b - T_t \geq x)dx = \int_{-\infty}^{T_b - s(t)} \mathbb{P}(\tilde{T}_t \leq x)dx \qquad (3.6)$$

$$\mathbb{E}[\min((T_b - T_t)^+, L)] = \int_0^L \mathbb{P}(T_b - T_t \geq x)dx = \int_{T_b - s(t) - L}^{T_b - s(t)} \mathbb{P}(\tilde{T}_t \leq x)dx$$

We approximate this cumulative density function by using the Gil-Pelaez inversion formula [6, Theorem 4.2.3 p. 104]:

$$\mathbb{P}(\tilde{T}_t \leq x) = \frac{1}{2} - \frac{1}{\pi} \lim_{m \to 0^+, M \to +\infty} \int_m^M \mathfrak{R}\left(\frac{e^{-ivx}\Phi(v)}{iv}\right) dv, \ \Phi(v) = \mathbb{E}[e^{iv\tilde{T}_t}].$$

Let $\delta_x, \delta_v > 0$ be such that $\delta_x \delta_v = \frac{2\pi}{N}$ (one can take for example $\delta_x = \delta_v = \sqrt{\frac{2\pi}{N}}$ or $\delta_x = \frac{L}{N-1}$ for the capped option so that x_0 defined below is equal to $T_b - s(t) - L$). We define:

$$v_{j+1/2} = (j + 1/2)\delta_v, \ x_k = T_b - s(t) + (k - N + 1)\delta_x, j \in \{0, ..., N-1\}, k \in \{0, \ldots, N-1\},$$

so that $x_{N-1} = T_b - s(t)$, and use the following approximation:

$$\mathbb{P}(\tilde{T} \leq x_k) \approx \frac{1}{2} - \frac{\delta_v}{\pi}\mathfrak{R}\left(\sum_{j=0}^{N-1} \frac{e^{-iv_{j+1/2}x_k}\Phi(v_{j+1/2})}{iv_{j+1/2}}\right)$$

$$= \frac{1}{2} - \frac{\delta_v}{\pi}\mathfrak{R}\left(e^{-\frac{1}{2}i\delta_v k\delta_x}\sum_{j=0}^{N-1} e^{-2i\pi\frac{jk}{N}}\frac{e^{-i(j+1/2)\delta_v x_0}\Phi(v_{j+1/2})}{iv_{j+1/2}}\right), \qquad (3.7)$$

since $v_{j+1/2}x_k = 2\pi\frac{jk}{N} + (j+1/2)\delta_v x_0 + \frac{1}{2}\delta_v k\delta_x$. This amounts to use the midpoint rule and to truncate the integral at $M = N\delta_v$. Other choices of quadrature are possible but have to be taken in compliance with the FFT. Using (3.7), we can obtain $(\mathbb{P}(\tilde{T} \leq x_k), 0 \leq k \leq N-1)$ by applying the FFT to $\left(\frac{e^{-i(j+1/2)\delta_v x_0}\Phi(v_{j+1/2})}{iv_{j+1/2}}, 0 \leq j \leq N-1\right)$: these N values are obtain with a time complexity of $O(N\log(N))$ (instead of $O(N^2)$ with the naive calculation of the sums). We finally approximate the expectation of the daily HDD by:

$$\mathbb{E}[(T_b - T_t)^+] \approx \delta_x \left(\sum_{k=0}^{N-2} \mathbb{P}(\tilde{T}_t \leq x_k) + \frac{1}{2}\mathbb{P}(\tilde{T}_t \leq x_{N-1})\right).$$

Figure 1.12 shows the characteristic function and the expected HDD in Paris during January 2019 comparing Monte Carlo simulation and FFT approach. Both graphs display a clear coherence between Monte Carlo simulations and the FFT approach. In this case Monte Carlo simulations show precision given that we simulate $50,000$ scenarios. However, FFT pricing is more precise, smooth and faster. We perform FFT with $N = 2^{17}$.
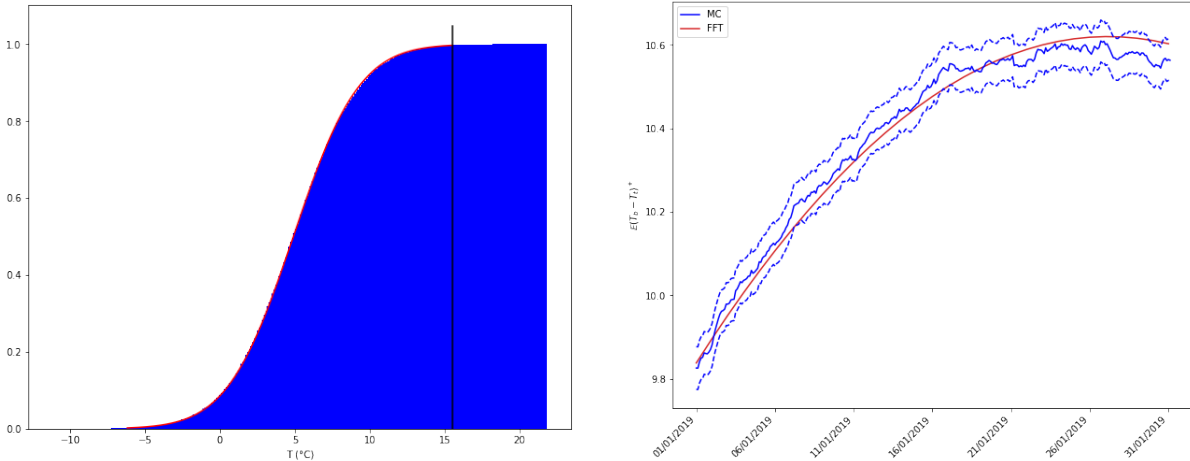

Figure 1.12: Cumulative distribution function of 31st January 2019's daily temperature (left) and expected daily HDD during January 2019 days computed 30 days ahead of the month (right) by the FFT method (red) and Monte-Carlo (blue). The black vertical line corresponds to T_b.

We now focus on the pricing on options on HDD. We observe from the distribution in Figure 1.12 that we almost always have $T_t \leq T_b$ in January, otherwise we would notice a Dirac mass at 0. In fact, with the standard strike $T_b = 15.5°C$, we mostly have $T_t \leq T_b$ during winter, and therefore $HDD \approx (t_2 - t_1 + 1)T_b - CAT$, so that the average value of the option (3.1) can be approximated by

$$\mathbb{E}[\min((HDD - HDD_{strike})^+, L)] \approx \mathbb{E}[\min(((t_2 - t_1 + 1)T_b - HDD_{strike} - CAT)^+, L)]. \quad (3.8)$$

The problem of computing the right hand side is then similar to the pricing of daily HDD (3.6): the underlying is now CAT instead of T_t. We can thus calculate the average payoff, provided that we know the characteristic function of the CAT $\Phi(u) = e^{iu \sum_{t=t_1}^{t_2} s(t)} \mathbb{E}[e^{iu \sum_{t=t_1}^{t_2} \tilde{T}_t}]$. To do so, it is possible to use formula (3.3) inductively with $u_3 = 0$ (using Remark 3.1) in order to calculate $\mathbb{E}[e^{iu \sum_{t=t_1}^{t_2} \tilde{T}_t} | \mathcal{F}_{t_2 - \ell}]$ for $\ell = 1, \dots, t_2 - t_1$ and then Φ. This is however cumbersome, and we prefer to make the following approximation

$$\Phi(u) \approx e^{iu \sum_{t=t_1}^{t_2} s(t)} \mathbb{E}[e^{iu \int_{t_1}^{t_2+1} \tilde{T}_t dt}].$$

We apply Proposition 3.1 with $u_1 = u_2 = 0$ and $u_3 = u$, and get $\mathbb{E}[e^{iu \int_{t_1}^{t_2+1} \tilde{T}_t dt} | \mathcal{F}_{t_1}] = \exp(a_0(t_1, t_2 + 1) + iu \frac{1 - e^{-\kappa(t_2+1-t_1)}}{\kappa} \tilde{T}_{t_1} + a_2(t_2 + 1 - t_1)\zeta_{t_1})$. Hence, for $t_0 \le t_1 \le t_2$,

$$\begin{aligned}
\mathbb{E}[e^{iu \int_{t_1}^{t_2+1} \tilde{T}_t dt} | \mathcal{F}_{t_0}] &= \mathbb{E}[\mathbb{E}[e^{iu \int_{t_1}^{t_2+1} \tilde{T}_t dt} | \mathcal{F}_{t_1}] | \mathcal{F}_{t_0}] \\
&= \mathbb{E}[\exp(a_0(t_1, t_2 + 1) + iu \frac{1 - e^{-\kappa(t_2+1-t_1)}}{\kappa} \tilde{T}_{t_1} + a_2(t_2 + 1 - t_1)\zeta_{t_1}) | \mathcal{F}_{t_0}] \\
&= \exp(a_0(t_1, t_2 + 1)) \exp(\check{a}_0(t_0, t_1) + \check{a}_1(t_1 - t_0)\tilde{T}_{t_0} + \check{a}_2(t_1 - t_0)\zeta_{t_0})
\end{aligned}$$

To obtain $\check{a}_0, \check{a}_1, \check{a}_2$ and the above characteristic function, we apply a second time Proposition 3.1 with $u_1 = u \frac{1 - e^{-\kappa(t_2+1-t_1)}}{\kappa}$, $u_2 = -ia_2(t_2 + 1 - t_1)$ and $u_3 = 0$. Figure 1.13 compares CAT distribution obtained with Monte Carlo and FFT inverse methods for the month of January 2019. We can observe a good fit between both methods.

This section has focused on the capacity of Fast Fourier Transform method to compute explicit formulas for options on T and CAT. This methodology enables to get rid of the computation burden of Monte Carlo simulations. However, not all the indices or payoff functions can be explicited with FFT. In particular, derivatives that integrate double non-linearities, like put or call payoff functions applied to HDD, cannot be explicitly computed with the FFT method. The next section will focus on how to use FFT results to increase the performance of Monte Carlo simulation for such cases.

3.3.4 Control variates method for Monte-Carlo

In practice the approximation (3.8) is precise when $\mathbb{P}(T_t > T_b)$ is close to zero. However, when this probability is small but not negligible, the approximation may not be enough precise. However, we can use the calculation above to run a Monte-Carlo method with the control variable $\min((HDD - HDD_{strike})^+, L) - \lambda \min(((t_2 - t_1 + 1)T_b - HDD_{strike} - CAT)^+, L)$ in order to calculate the average payoff $\mathbb{E}[(HDD - HDD_{strike})^+]$. Namely, we write (we take here $L = +\infty$ for simpler notation)

$$\begin{aligned}
\mathbb{E}[(HDD - HDD_{strike})^+] = &\lambda \mathbb{E}[((t_2 - t_1 + 1)T_b - HDD_{strike} - CAT)^+] \qquad (3.9) \\
&+ \mathbb{E}\left[(HDD - HDD_{strike})^+ - \lambda((t_2 - t_1 + 1)T_b - HDD_{strike} - CAT)^+\right],
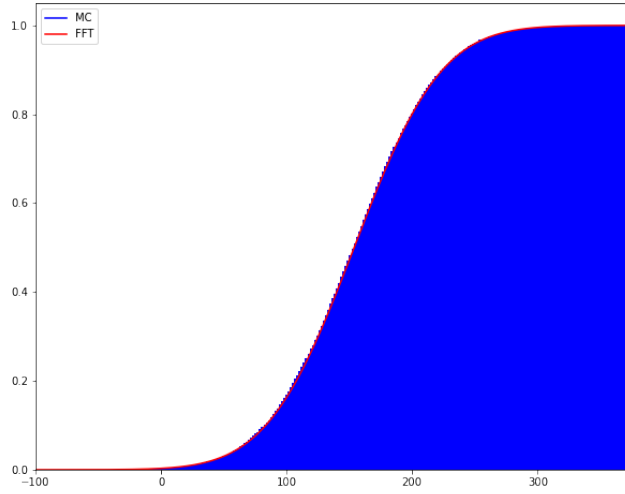\end{aligned}$$

Figure 1.13: Cumulative distribution function of CAT for January 2019 and 30 days observation in advance computed by the FFT method (red) and Monte-Carlo with $50,000$ simulations (blue).

and we chose λ that minimizes $Var\left[(HDD - HDD_{strike})^+ - \lambda((t_2 - t_1 + 1)T_b - HDD_{strike} - CAT)^+\right]$, i.e.

$$\lambda^* = \frac{Cov((HDD - HDD_{strike})^+, ((t_2 - t_1 + 1)T_b - HDD_{strike} - CAT)^+)}{Var(((t_2 - t_1 + 1)T_b - HDD_{strike} - CAT)^+)}.$$

The first term of the right hand side of (3.9) is calculated by using the FFT while the second one is calculated by Monte-Carlo.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Corr	1.00	1.00	1.00	1.00	0.94	0.66	0.38	0.33	0.66	0.97	1.00	1.00
VR	2.41e5	5.24e4	4.73e3	2.22e2	5.08	1.19	1.01	1.01	1.20	9.84	3.92e2	1.40e4

Table 1.11: Correlation and variance reduction (VR) brought by the control variates method for options computed during each month of 2019. Variance ratio corresponds to the variance of $(\sum_{t=t_1}^{t_2}(T_b - T_t)^+ - HDD_{strike})^+$ divided by the variance of the control variable.

Figure 1.14 and Table 1.11 show the results of the implementation of the control variates method. First, from Table 1.11 we can see that the control variates method enables to decrease the variance up to 2.41×10^5 times, leading to a price computation 10^5 time faster. Second, we can observe that the performance of the method depends on the correlation between $(\sum_{t=t_1}^{t_2}(T_b - T_t)^+ - HDD_{strike})^+$ and the control variable: the more correlated they are, the more variance
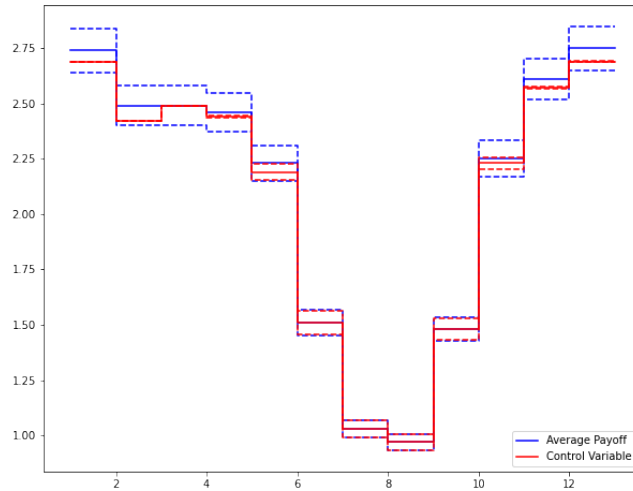
Figure 1.14: Expected payoffs forecasted 30 days ahead for the HDD derivative (3.1) (with $L = +\infty$) on each month of 2019 (blue) and for the control variable (red). We performed $50,000$ Monte Carlo simulations, dotted lines indicate the 95% confidence interval.

reduction we obtain (and the more approximation (3.8) is valid). Hence for the winter months the computation is more effective, which coincides with the months for which such options are sold. In Figure 1.14, we see that the confidence interval for winter months is considerably narrower.

To sum up, this section has focused on exploring alternative pricing methodologies. The Fast Fourier Transform pricing method enables to bypass Monte Carlo simulations and to get analytical expressions for the expected payoffs of some derivatives. This enables direct expectation computations for some derivatives like CAT. For derivatives integrating non linear indices and payoffs, this method can be combined with the control variates method to decrease the computational cost of the Monte Carlo simulations. In our case, this method enables to considerably decrease the number of required simulations.

3.4 Sensitivity study

This section studies the sensitivity of the pricing to the different parameters that were either imposed or estimated in the previous sections.

Sensitivity to κ We first analyse the sensitivity to the mean reverting parameter of the temperature dynamics. Figure 1.15 shows that increasing κ has an important effect on the average payoff and the simulated HDD distribution and therefore on the pricing. Indeed when

κ increases the volatility loses its importance, the HDD distribution becomes more certain and therefore peaks around an expected average value. Similarly, the quantiles are less spread and therefore the strikes based on initially estimated κ are less frequent and mean payoffs shrink.
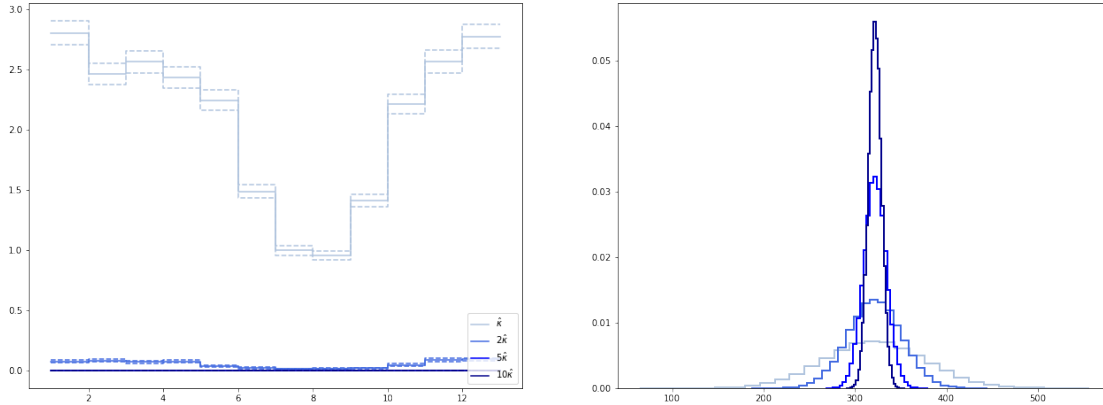

Figure 1.15: Average payoffs for values of $\kappa \in \{\hat{\kappa}, 2\hat{\kappa}, 5\hat{\kappa}, 10\hat{\kappa}\}$ (left) and HDD distribution for a derivative on the month of January 2019, forecasted 30 days ahead and based on $50,000$ Monte Carlo simulations (right). HDD_{strike} is kept at 90% empirical quantile of $\kappa = \hat{\kappa}$ for all simulations.

Sensitivity to η^2   Figure 1.16 shows different average payoffs and confidence intervals for $50,000$ simulations and different values of the volatility of the volatility η^2. We can see that increasing η^2 creates more peaked distributions for HDD and heavier tails. This also leads to usually higher prices. Nevertheless, it should be noted that the impact of η^2 is marginal in winter months when this product is meant to be sold. In summer months, HDDs only capture extreme temperature left tails and this is when we can see a real impact of the volatility of Model (M). Besides, let recall that the estimation of η^2 is sensitive to the choice of Q. Wrongly estimating this parameter would therefore mainly impact the pricing of derivatives on summer months where the demand of such derivatives is much lower.

Sensitivity to K   Figure 1.17 shows different average payoffs and confidence intervals for $50,000$ simulations. We can see that increasing K creates less peaked distributions for HDD and lighter tails. This leads to usually lower mean payoffs when K increases. Likewise, the impact is marginal in winter months when this product is meant to be sold. This phenomenon is intuitive as we increase the mean reverting term of the volatility, the volatility of the volatility losses weight in the dynamics and the extreme HDDs decrease.
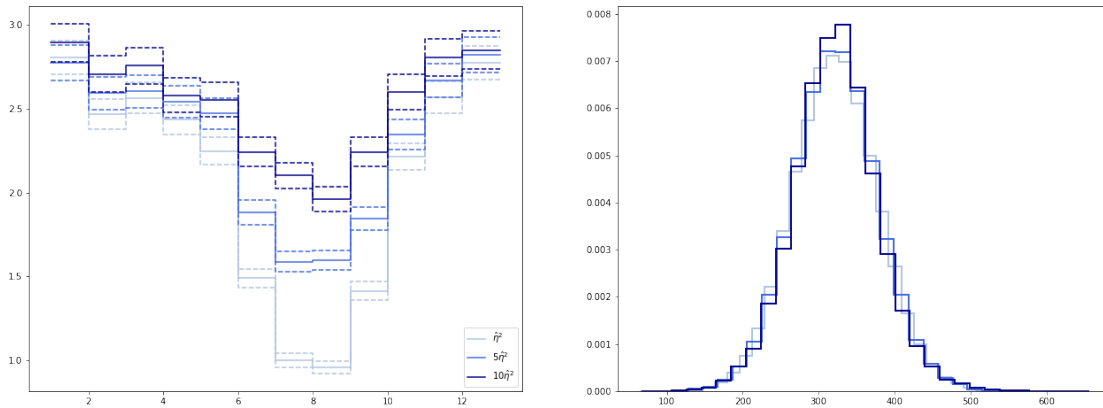
Figure 1.16: Average payoffs for different values of $\eta^2 \in \{\widehat{\eta^2}, 5\widehat{\eta^2}, 10\widehat{\eta^2}\}$ (left) and HDD distribution for a derivative on the month of January 2019, forecasted 30 days ahead and based on $50,000$ Monte Carlo simulations (right). HDD_{strike} is kept at 90% empirical quantile of $\eta^2 = \widehat{\eta^2}$ for all simulations.
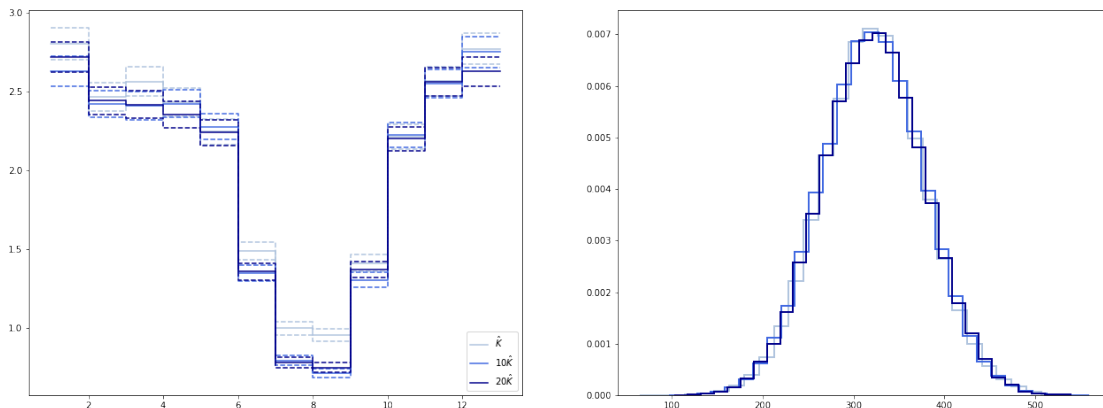

Figure 1.17: Average payoffs for different values of $K \in \{\hat{K}, 10\hat{K}, 20\hat{K}\}$ (left), HDD distribution distribution starting from HDD_{strike} (right) for a derivative on the month of January 2019, forecasted 30 days ahead and based on $50,000$ Monte Carlo simulations (right). HDD_{strike} is kept at 90% empirical quantile of $K = \hat{K}$ for all simulations.

Sensitivity to $t_1 - t_0$ Now we suppose we compute the price of the derivative different possible times ahead. Figure 1.18 shows temperature paths for different observation times t_0 but the same observed temperature and volatility at t_0. The derivative applies between the black

vertical lines. We can observe that all the paths end up following the seasonality s.
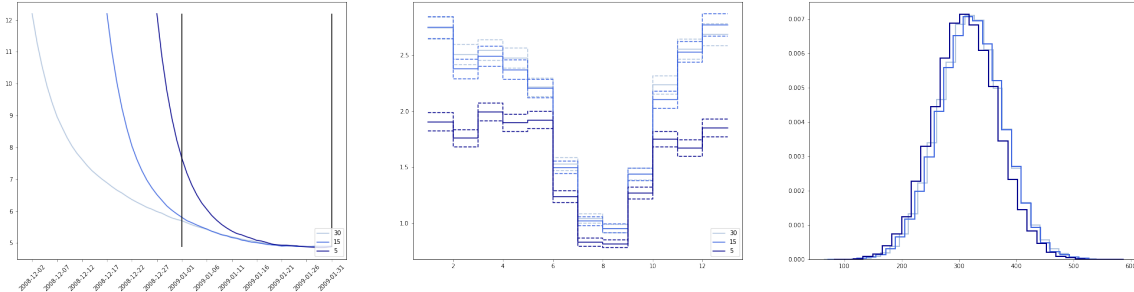

Figure 1.18: Average temperature paths for different values of $t_1 - t_0$ (left), average payoffs for different values of $t_1 - t_0$ for a derivative on each month of 2019 (center) and HDD distribution for a derivative on the month of January 2019, forecasted 5, 15 and 30 days ahead and based on $50,000$ Monte Carlo simulations (right). Here, $T_{t_0} = s(t_0) + 2\sigma(t_0)$ and $\zeta_{t_0} = \sigma(t_0)$ for $t_0 = 30$. HDD_{strike} is fixed at the 90% quantile of the 30 days ahead simulation. In the left plot, the two black vertical lines represent times t_1 and t_2.

From simulated densities in Figure 1.18 we can first observe shifts depending on $t_1 - t_0$. This significantly impacts the quantiles of these densites and therefore the HDD_{strike}. Second, we can note that the more ahead we forecast the less information we have. In this case, we can observe that while pricing 20 and 30 days ahead lead to similar average payouts during the risk period, forecasting 5 days ahead significantly impacts the average payoffs and therefore pricing. This element is key to answer the risk of antiselection.

Sensitivity to the moneyness of the product   The moneyness of the product has a direct impact on the payoffs distribution as can be observed in Figure 1.19. The lowest the HDD_{strike}, the more HDD we capture in the payoff and the higher the mean payoffs becomes.

To sum up, this section has focused on the sensitivity of the pricing to the different parameters. We particularly show that the parameters related to the temperature dynamics like κ as well as the moneyness of the payoff function are the ones affecting the most the mean payoffs. Parallely, the distribution of the payoffs and the strike based on quantiles show important sensitivity to the time interval $t_1 - t_0$. Finally, the parameters related to the volatily have a relatively lower impact on the payoffs distribution and hence on the pricing.

3.5 Comparison of our pricing methodology with business practices

This section aims to make the bridge between the pricing methodology exposed in this document and the current market practices. In particular, we will compare a pricing based on modeling of the underlying meteorological parameter with a pricing based on historical index modeling.
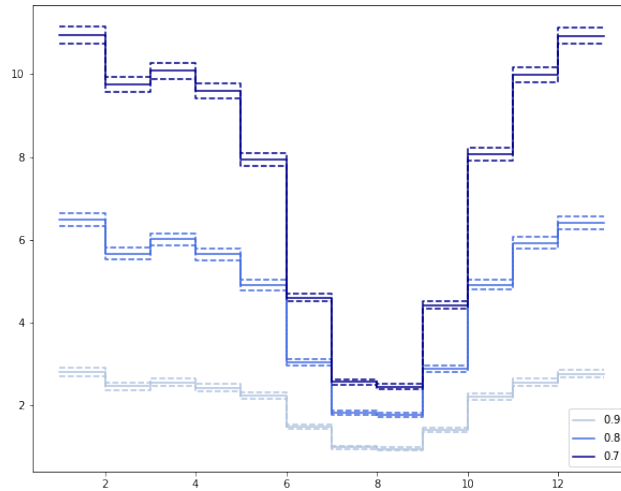
Figure 1.19: Average payoffs for different quantiles defining HDD_{strike} in $\{0.7, 0.8, 0.9\}$ for each month of 2019 and based on the $50,000$ Monte Carlo simulation 30 days ahead.

As described in Schiller et al. [130] and Jewson and Brix [90], the index based pricing methodology consists in modeling the independent yearly indices, in this case cumulative HDD. For computation ease, the index pricing is automatised following the below algorithm:

1. Compute historical index, cumulative HDD, from 1980 to 2018.

2. Remove the linear trend in this time series.

3. Fit a gamma distribution to these observations through a maximum likelihood method.

4. Compute the expected payoff of the fitted index distribution.

While this approach can be simplistic, as we can consider other probability distributions for the index, it syntheses the common market practices.

Figure 1.20 represents the expected payoffs computed with the two methodologies. HDD_{strike} corresponds to the 90% simulated quantile with the Monte Carlo method and to the 90% historical quantile for the index modeling method. First, we can observe there is a coherence between the approaches that give expected payoffs in the same ranges. However, the index modeling approach introduces important instabilities. These instabilities affect both the average payoffs as well as the strike HDD_{strike}, which is estimated from less than 40 observations.

We also studied the possibility of using the same strikes with both methods. First, using the 90% quantile of $50,000$ Monte Carlo simulations leads to slightly more volatile average payoffs for the index model method. However, we feel it is counter-intuitive to use a yearly index modeling for pricing and a more cumbersome daily index modeling just to get the strikes.
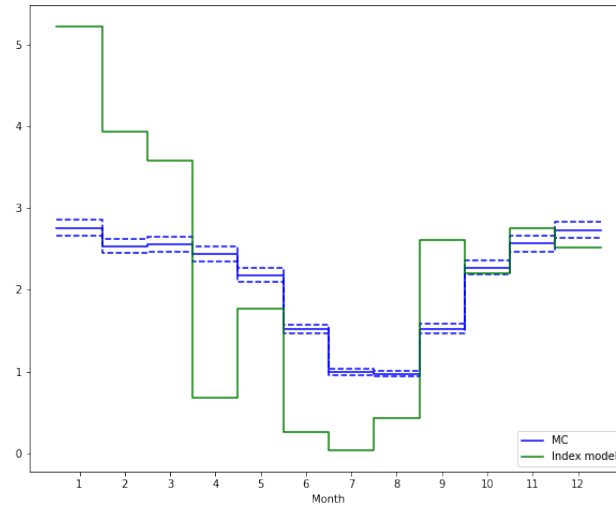
Figure 1.20: Expected payoffs forecasted 30 days ahead for a derivative on each month of 2019. HDD_{strike} is defined differently with two methodologies. For the Monte Carlo approach it corresponds to the 90% quantile of 50,000 simulations while for the Index model approach the HDD_{strike} corresponds to the historical quantile.

Second, we can use historical quantiles for both methodologies as on the left of Figure 1.21. In this case we can see that the winter months seem to be completely overpriced by the business practice while the summer months are underpriced. On the right of Figure 1.21, we compare simulated and historical quantiles. We can see that the more difference we have between these quantiles the higher the risk of over or underpricing.

To sum up, there exist a clear coherence between the pricing methodology followed in this document and the current market practices as both give prices in the same ranges. However, our approach enables to better quantify the sources of risk, which is crucial in this growing market.
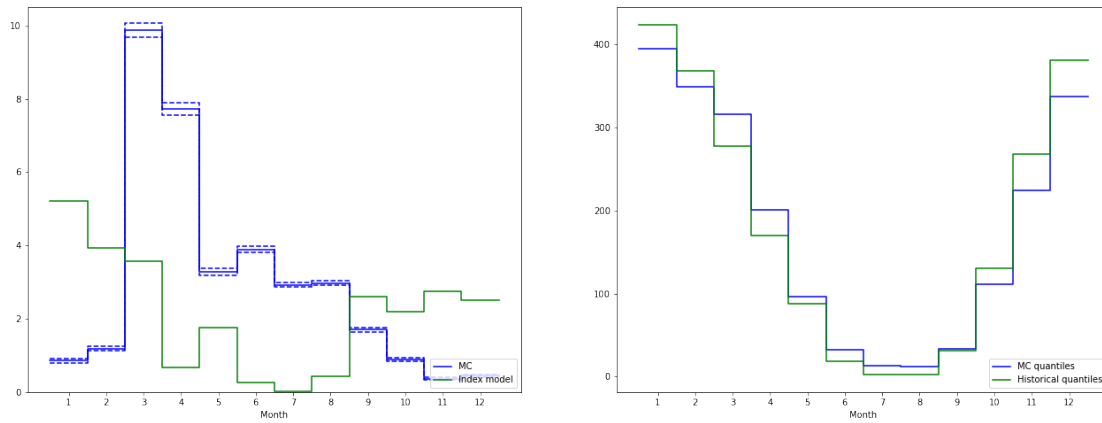
Figure 1.21: On the right, expected payoffs forecasted 30 days ahead for a derivative on each month of 2019 and HDD_{strike} defined as the 90% historical quantile for both approaches. On the left, a comparison of simulated and historical 90% quantile.

4 Appendices

4.1 Weather station data description

Table 1.12 summarizes the weather station data used in the above study.

City	WMO	Latitude	Longitude	Elevation	Original Source
Stockholm	2485	59.34	18.05	43 m	Swedish Meteorological & Hydrological Institute, SMHI
Paris Charles de Gaulle	7157	49.02	2.53	109 m	ASOS-METAR
Amsterdam AP Schiphol	6240	52.3	4.78	-4 m	Royal Netherlands Meteorological Institute, KNMI
Berlin Tempelhof	10384	52.47	13.4	50 m	Deutsche Wetterdienst, DWD
Brussels National	6451	50.9	4.53	58 m	Royal Meteorological Institute of Belgium
London Heathrow	3772	51.48	-0.45	25 m	ASOS-METAR
Rome Ciampino	16239	41.78	12.58	105 m	ASOS-METAR
Madrid Barajas	8221	40.5	-3.58	633 m	ASOS-METAR

Table 1.12: Characteristics of the weather stations providing temperature data. Speedwell explicitly states the meteorological agency the data is extracted from except for ASOS-METAR data. ASOS-METAR is in charge of the monitoring of airport weather stations following the standards of the International Civil Aviation Organization (ICAO) and the World Meteorological Organization (WMO).

4.2 Model (M) simulation and estimator testing

This section explains the algorithm to simulate Model (M). Simulation is first performed to test robustness of the estimation through the following steps:

1. Estimate κ and the seasonality s from temperature data.

2. Estimate the parameters K and σ^2 and then η^2.

3. Fix these parameters at the estimated values.

4. Generate a simulated instantaneous volatility series ζ based on a generalized Ninomiya-Victoir scheme for Cox-Ingersoll-Ross (CIR) processes, and the corresponding temperature series T.

5. Estimate $(\alpha_0, \beta_0, \alpha_1, \beta_1, \kappa)$ on the simulated data and compare with the fixed values.

6. Compute realized volatility $\hat{\zeta}$ for different time lags Q.

7. Estimate $(\gamma_0, \gamma_1, \delta_1, \gamma_2, \delta_2, K)$ then $\widehat{\eta^2}$, and compare with the fixed values.

The generation of the combined discrete series $(T_{i\Delta}, \zeta_{i\Delta})_{i\in\mathbb{N}}$ is performed thanks to the recurrence formula:

$$\begin{cases} T_{(i+1)\Delta} = s((i+1)\Delta) + e^{-\kappa\Delta}(T_{i\Delta} - s(i\Delta)) + \sqrt{\dfrac{1 - e^{-2\kappa\Delta}}{2\kappa}} \dfrac{\zeta_{i\Delta} + \zeta_{(i+1)\Delta}}{2} Z_i \\ \zeta_{(i+1)\Delta} = \phi(\zeta_{i\Delta}, \Delta, \sqrt{\Delta}Y_i), \end{cases} \tag{4.1}$$

where $(Y_i, Z_i)_{i\geq 0}$ is an i.i.d. sequence of two independent standard normal variables. Remember that we assume here and in the sequel of the study that $\rho = 0$. The first row of (4.1) corresponds to a discretization of the integral following the temperature dynamics in Model (M) for a step Δ. The use of the trapezoidal rule comes from the operator splitting method and allow to get a second order scheme as in [6, Eq. (4.31)]. The second row of (4.1) is the Ninomiya-Victoir scheme for CIR processes [119] when freezing the time-dependent coefficients at time $(i+1/2)\Delta$, which preserves the convergence of order 2 of this scheme, see [6, Paragraph 3.3.4]. In general (i.e. when $K\sigma^2((i+1/2)\Delta) \geq \frac{\eta^2}{4}$), ϕ corresponds to:

$$\phi(\zeta_{i\Delta}, \Delta, \sqrt{\Delta}Y_i) = e^{-\frac{K\Delta}{2}}\left(\sqrt{\left(K\sigma^2((i+1/2)\Delta) - \frac{\eta^2}{4}\right)\psi_K\left(\frac{\Delta}{2}\right) + \zeta_{i\Delta}e^{-\frac{K\Delta}{2}}} + \frac{\eta}{2}\sqrt{\Delta}Y_i\right)^2$$
$$+ \left(K\sigma^2((i+1/2)\Delta) - \frac{\eta^2}{4}\right)\psi_K\left(\frac{\Delta}{2}\right)$$

where $\psi_K(t) = \frac{1 - e^{-Kt}}{K}$. The case where $K\sigma^2((i+1/2)\Delta) < \frac{\eta^2}{4}$ is handled as in Alfonsi [7], so that (4.1) is a second order scheme for the weak error, and thus an accurate approximation of the exact law.

Remark 4.1. *With this simulation method, it is possible to introduce additional granularity into our simulated processed having more than one point per day, that is $\Delta < 1$. We have analysed numerically if this extra-granularity has an incidence on the parameter estimation. Namely, we have calculated the estimators on simulated paths with different values of Δ but with the same number of points. We have noticed that Δ has a small influence on the estimators, and we do not reproduce these experiments in the study.*

4.3 CLS estimators of the temperature process

Let us consider the following dynamics for the temperature $(T_t)_{t\geq 0}$:

$$\begin{cases} T_t & = s(t) + \tilde{T}_t, \\ d\tilde{T}_t & = -\kappa\tilde{T}_t dt + \sqrt{\zeta_t}(\rho dW_t + \sqrt{1-\rho^2}dZ_t) \end{cases} \qquad (4.2)$$

where $\kappa > 0$, $s(t) = \alpha_0 + \beta_0 t + \alpha_1 \sin(\xi t) + \beta_1 \cos(\xi t)$, W and Z are two independent Brownian motions and ζ_t is a nonnegative adapted process such that $\mathbb{E}[\int_0^t \zeta_s de] < \infty$ for all $t > 0$. The goal of this appendix is to compute the conditional least squares estimators of $(\kappa, \alpha_0, \beta_0, \alpha_1, \beta_1)$ and to prove the next proposition.

Proposition 4.1. *Let $X_{i\Delta} = (1, i\Delta, T_{i\Delta}, \sin(\xi i\Delta), \cos(\xi i\Delta))^T \in \mathbb{R}^5$ for $i \in \mathbb{N}$ with $(T_t)_{t\geq 0}$ following the dynamics of* (4.2) *and $\Delta > 0$. We assume that $\sum_{i=0}^{N-1} X_{i\Delta}X_{i\Delta}^T$ is invertible and define*

$$\hat{\lambda} = (\hat{\lambda}_0, \ldots \hat{\lambda}_4)^T = \left(\sum_{i=0}^{N-1} X_{i\Delta}X_{i\Delta}^T\right)^{-1}\left(\sum_{i=0}^{N-1} X_{i\Delta}T_{(i+1)\Delta}\right). \qquad (4.3)$$

If $\hat{\lambda}_2 \in (0,1) \cup (1, +\infty)$, the solution of the minimisation problem

$$\min_{(\kappa, \alpha_0, \alpha_1, \beta_0, \beta_1) \in \mathbb{R}^5} \sum_{i=0}^{N-1}\left(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}]\right)^2 \qquad (4.4)$$

is given by

$$\begin{cases} \hat{\kappa} & = -\frac{1}{\Delta}\ln\hat{\lambda}_2 \\ \hat{\alpha}_0 & = \frac{\hat{\lambda}_0}{1-\hat{\lambda}_2} - \frac{\hat{\lambda}_1\Delta}{(1-\hat{\lambda}_2)^2} \\ \hat{\beta}_0 & = \frac{\hat{\lambda}_1}{1-\hat{\lambda}_2} \\ \hat{\alpha}_1 & = \frac{\hat{\lambda}_3(\cos(\xi\Delta)-e^{-\hat{\kappa}\Delta})+\hat{\lambda}_4\sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}\Delta})^2+\sin^2(\xi\Delta)} \\ \hat{\beta}_1 & = \frac{\hat{\lambda}_4(\cos(\xi\Delta)-e^{-\hat{\kappa}\Delta})-\hat{\lambda}_3\sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}\Delta})^2+\sin^2(\xi\Delta)}. \end{cases}$$

Proof. Recall that $T_t = s(t) + \tilde{T}_t$. Applying Ito's formula for $(e^{\kappa t}\tilde{T}_t)_{t\geq 0}, we$ have

$$\tilde{T}_{t+\Delta} = \tilde{T}_t e^{-\kappa\Delta} + \rho\int_t^{t+\Delta} e^{-\kappa(t+\Delta-s)}\sqrt{\zeta_s}dW_s + \sqrt{1-\rho^2}\int_t^{t+\Delta} e^{-\kappa(t+\Delta-s)}\sqrt{\zeta_s}dZ_s \qquad (4.5)$$

From the martingale property of the stochastic integral (we have $\int_t^{t+\Delta}\mathbb{E}[\zeta_s]ds < \infty$ since $\mathbb{E}[\zeta_s] = \zeta_0 e^{-Ks} + \int_0^s e^{-K(s_u)}\sigma^2(u)du$), we have $\mathbb{E}[\tilde{T}_{t+\Delta}|\mathcal{F}_t] = \tilde{T}_t e^{-\kappa\Delta}$ and thus

$$\mathbb{E}[T_{t+\Delta}|\mathcal{F}_t] = T_t e^{-\kappa\Delta} + s(t+\Delta) - s(t)e^{-\kappa\Delta}.$$

We now use trigonometric identities to get

$$
\begin{aligned}
s(t+\Delta) - e^{-\kappa\Delta}s(t) &= \alpha_0 + \beta_0(t+\Delta) - \alpha_0 e^{-\kappa\Delta} - \beta_0 e^{-\kappa\Delta}t + \alpha_1\sin(\xi(t+\Delta)) - \alpha_1 e^{-\kappa\Delta}\sin(\xi t) \\
&\quad + \beta_1\cos(\xi(t+\Delta)) - \beta_1 e^{-\kappa\Delta}\cos(\xi t) \\
&= \lambda_0 + \lambda_1 t + \lambda_3\sin(\xi t) + \lambda_4\cos(\xi t),
\end{aligned}
$$

with

$$
\begin{cases}
\lambda_0 &= \alpha_0(1 - e^{-\kappa\Delta}) + \beta_0\Delta \\
\lambda_1 &= \beta_0(1 - e^{-\kappa\Delta}) \\
\lambda_2 &= e^{-\kappa\Delta} \\
\lambda_3 &= \alpha_1(\cos(\xi\Delta) - e^{-\kappa\Delta}) - \beta_1\sin(\xi\Delta) \\
\lambda_4 &= \alpha_1\sin(\xi\Delta) + \beta_1(\cos(\xi\Delta) - e^{-\kappa\Delta}),
\end{cases}
\tag{4.6}
$$

where λ_2 is set to have $\mathbb{E}[T_{(i+1)\Delta}|\mathcal{F}_t] = \lambda^T X_{i\Delta}$. The minimization problem (4.4) is then equivalent to

$$
\min_{\lambda\,\in\,\mathbb{R}^5}\quad \sum_{i=0}^{N-1}\left(T_{(i+1)\Delta} - \lambda^T X_{i\Delta}\right)^2.
$$

This corresponds to a linear regression, whose solution is given by (4.3). When $\lambda_2 \in (0,1)$, the system (4.6) can be inverted, and the claim follows easily.

$\square$

Let us note here that $\hat{\lambda}^T X_{i\Delta}$ can then be seen as the estimation of $\mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}]$.

4.4 CLS estimators of the volatility process

Let consider the volatility of the temperature $(\zeta_t)_{t\geq 0}$ follows the below dynamics:

$$
d\zeta_t = -K(\zeta_t - \sigma^2(t))dt + \eta\sqrt{\zeta_t}dW_t
\tag{4.7}
$$

where $K > 0$, σ^2 is a nonnegative function with the parametric form given by (1.5) and W is a Brownian motion. The goal of this appendix is to compute the conditional least squares estimators of the parameters $(\gamma_0, K, \gamma_1, \ldots, \gamma_{K_{\sigma^2}}, \delta_1, \ldots, \delta_{K_{\sigma^2}})$.

Proposition 4.2. *Let* $X'_{i\Delta} = (1, \zeta_{i\Delta}, \sin(\xi_1 i\Delta), \ldots, \sin(\xi_{K_{\sigma^2}} i\Delta), \cos(\xi_1 i\Delta), \ldots, \cos(\xi_{K_{\sigma^2}} i\Delta))^T$ *with* $(\zeta_t)_{t\geq 0}$ *following the dynamics* (4.7) *and* $\Delta > 0$. *We assume that* $\sum_{i=0}^{N-1} X'_{i\Delta} X'^T_{i\Delta}$ *is invertible and define*

$$
(\hat{\theta}_0, \hat{\phi}_0, \hat{\theta}_1, \ldots, \hat{\theta}_{K_{\sigma^2}}, \hat{\phi}_1, \ldots, \hat{\phi}_{K_{\sigma^2}})^T = \left(\sum_{i=0}^{N-1} X'_{i\Delta} X'^T_{i\Delta}\right)^{-1}\left(\sum_{i=0}^{N-1} X'_{i\Delta}\zeta_{(i+1)\Delta}\right).
$$

If $\hat{\phi}_0 \in (0,1) \cup (1,\infty)$, the solution of the minimisation problem

$$\min_{K \in \mathbb{R}, \gamma \in \mathbb{R}^{K_{\sigma^2}+1}, \delta \in \mathbb{R}^{K_{\sigma^2}}} \sum_{i=0}^{N-1} \left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta} | \zeta_{i\Delta}] \right)^2 \tag{4.8}$$

is given by

$$\begin{cases} \hat{\gamma}_0 = \dfrac{\hat{\theta}_0}{1 - \hat{\phi}_0} \\[2mm] \hat{K} = -\dfrac{1}{\Delta} \ln(\hat{\phi}_0) \\[2mm] \hat{\gamma}_k = \dfrac{\hat{\theta}_k D_k - \hat{\phi}_k B_k}{A_k D_k - C_k B_k} \\[2mm] \hat{\delta}_k = \dfrac{\hat{\theta}_k C_k - \hat{\phi}_k A_k}{C_k B_k - A_k D_k} \end{cases} \tag{4.9}$$

where, for $k \in \{1, \ldots, K_{\sigma^2}\}$,

$$\begin{cases} A_k = \hat{K} \dfrac{\hat{K}(\cos(\xi_k \Delta) - e^{-\hat{K}\Delta}) + \xi_k \sin(\xi_k \Delta)}{\hat{K}^2 + \xi_k^2} \\[3mm] B_k = -\hat{K} \dfrac{\hat{K}\sin(\xi_k \Delta) - \xi_k(\cos(\xi_k \Delta) - e^{-\hat{K}\Delta})}{\hat{K}^2 + \xi_k^2} \\[3mm] C_k = \hat{K} \dfrac{\hat{K}\sin(\xi_k \Delta) - \xi_k(\cos(\xi_k \Delta) - e^{-\hat{K}\Delta})}{\hat{K}^2 + \xi_k^2} \\[3mm] D_k = \hat{K} \dfrac{\hat{K}(\cos(\xi_k \Delta) - e^{-\hat{K}\Delta}) + \xi_k \sin(\xi_k \Delta)}{\hat{K}^2 + \xi_k^2}. \end{cases}$$

For the proof of Proposition 4.2 we will need Lemma 4.1, whose proof is straightforward.

Lemma 4.1. *For $K^2 + \xi^2 > 0$ and $k \in \mathbb{N}^*$, we have*

$$\int_t^{t+\Delta} e^{-K(t+\Delta-s)} \cos(\xi_k s) ds = \cos(\xi_k t) \frac{K[\cos(\xi_k \Delta) - e^{-K\Delta}] + \xi_k \sin(\xi_k \Delta)}{K^2 + \xi_k^2}$$

$$- \sin(\xi_k t) \frac{K\sin(\xi_k \Delta) - \xi_k(\cos(\xi_k \Delta) - e^{-K\Delta})}{K^2 + \xi_k^2},$$

$$\int_t^{t+\Delta} e^{-K(t+\Delta-s)} \sin(\xi_k s) ds = \sin(\xi_k t) \frac{K[\cos(\xi_k \Delta) - e^{-K\Delta}] + \xi_k \sin(\xi_k \Delta)}{K^2 + \xi_k^2}$$

$$+ \cos(\xi_k t) \frac{K\sin(\xi_k \Delta) - \xi_k(\cos(\xi_k \Delta) - e^{-K\Delta})}{K^2 + \xi_k^2}.$$

Proof of Proposition 4.2. Applying Ito's formula for $(e^{Kt}\zeta_t)_{t\geq 0}$, we have:

$$\zeta_{t+\Delta} = \zeta_t e^{-K\Delta} + K\int_t^{t+\Delta} e^{-K(t+\Delta-s)}\sigma^2(s)ds + \eta\int_t^{t+\Delta} e^{-K(t+\Delta-s)}\sqrt{\zeta_s}dW_s \qquad (4.10)$$

Hence, we get

$$\mathbb{E}[\zeta_{t+\Delta}|\mathcal{F}_t] = \zeta_t e^{-K\Delta} + K\int_t^{t+\Delta} e^{-K(t+\Delta-s)}\sigma^2(s)ds.$$

From (1.5) and Lemma 4.1, we then obtain for $t \geq 0$ and $\zeta_0 \geq 0$,

$$\mathbb{E}[\zeta_{t+\Delta}|\mathcal{F}_t] = \theta_0 + \phi_0\zeta_t + \sum_{k=1}^{K_{\sigma^2}} \theta_k \sin(\xi_k t) + \sum_{k=1}^{K_{\sigma^2}} \phi_k \cos(\xi_k t), \qquad (4.11)$$

with

$$\begin{cases} \theta_0 = \gamma_0(1 - e^{-K\Delta}) \\ \phi_0 = e^{-K\Delta} \\ \theta_k = \gamma_k K\dfrac{K[\cos(\xi_k\Delta) - e^{-K\Delta}] + \xi_k\sin(\xi_k\Delta)}{K^2 + \xi_k^2} - \delta_k K\dfrac{K\sin(\xi_k\Delta) - \xi_k(\cos(\xi_k\Delta) - e^{-K\Delta})}{K^2 + \xi_k^2} \\ \phi_k = \gamma_k K\dfrac{K\sin(\xi_k\Delta) - \xi_k(\cos(\xi_k\Delta) - e^{-K\Delta})}{K^2 + \xi_k^2} + \delta_k K\dfrac{K[\cos(\xi_k\Delta) - e^{-K\Delta}] + \xi_k\sin(\xi_k\Delta)}{K^2 + \xi_k^2} \end{cases}$$

We can invert the above system when $\phi_0 \in (0,1) \cup (1, +\infty)$ to get the formulas of $K, \gamma_0, \gamma_k, \delta_k$.

$$\begin{cases} \gamma_0 = \dfrac{\theta_0}{1 - \phi_0} \\ K = -\dfrac{1}{\Delta}\ln(\phi_0) \\ \gamma_k = \dfrac{\theta_k D_k - \phi_k B_k}{A_k D_k - C_k B_k} \\ \delta_k = \dfrac{\theta_k C_k - \phi_k A_k}{C_k B_k - A_k D_k} \end{cases} \qquad (4.12)$$

where

$$
\begin{cases}
A_k = K \dfrac{K(\cos(\xi_k \Delta) - e^{-K\Delta}) + \xi_k \sin(\xi_k \Delta)}{K^2 + (\xi_k)^2} \\[4mm]
B_k = -K \dfrac{K \sin(\xi_k \Delta) - \xi_k(\cos(\xi_k \Delta) - e^{-K\Delta})}{K^2 + (\xi_k)^2} \\[4mm]
C_k = K \dfrac{K \sin(\xi_k \Delta) - \xi_k(\cos(\xi_k \Delta) - e^{-K\Delta})}{K^2 + (\xi_k)^2} \\[4mm]
D_k = K \dfrac{K(\cos(\xi_k \Delta) - e^{-K\Delta}) + \xi_k \sin(\xi_k \Delta)}{K^2 + (\xi_k)^2}
\end{cases}
$$

Let us observe that $A_k D_k - C_k B_k \geq 0$ as a sum of squares. We even have $A_k D_k - C_k B_k > 0$ since $(K(\cos(\xi_k \Delta) - e^{-K\Delta}) + \xi_k \sin(\xi_k \Delta), K \sin(\xi_k \Delta) - \xi_k(\cos(\xi_k \Delta) - e^{-K\Delta})) \neq (0,0)$ as

$$
\det \begin{bmatrix} (\cos(\xi_k \Delta) - e^{-K\Delta}) & \sin(\xi_k \Delta) \\ \sin(\xi_k \Delta) & (\cos(\xi_k \Delta) - e^{-K\Delta}) \end{bmatrix} = 1 - 2\cos(\xi_k \Delta)e^{-K\Delta} + e^{-2K\Delta} \geq (1 - e^{-K\Delta})^2 > 0,
$$

since $K \neq 0$ by the assumption $\phi_0 \neq 1$. The minimization problem (4.8) is then equivalent to

$$
\min_{\vartheta} \quad \sum_{i=0}^{N-1} \left(\zeta_{(i+1)\Delta} - \vartheta^T X'_{i\Delta} \right)^2
$$

where $\vartheta = (\theta_0, \phi_0, \theta_1, \ldots, \theta_{K_{\sigma^2}}, \phi_1, \ldots, \phi_{K_{\sigma^2}})^T$ and

$$
X'_{i\Delta} = (1, \zeta_{i\Delta}, \sin(\xi_1 i\Delta), \ldots, \sin(\xi_{K_{\sigma^2}} i\Delta), \cos(\xi_1 i\Delta), \ldots, \cos(\xi_{K_{\sigma^2}} i\Delta))^T, \ 0 \leq i \leq N-1. \quad (4.13)
$$

This problem corresponds to a simple multilinear regression problem. Its solution is given by:

$$
\hat{\vartheta} = \left(\sum_{i=0}^{N-1} X'^T_{i\Delta} X'_{i\Delta} \right)^{-1} \left(\sum_{i=0}^{N-1} X'^T_{i\Delta} \zeta_{(i+1)\Delta} \right)
$$

Combined with Equation (4.12), we get the estimators (4.9) of the volatility parameters of of Model (M). $\qquad \square$

Let us note here that $\hat{\vartheta}^T X'_{i\Delta}$ can be seen as the estimation of $\mathbb{E}[\zeta_{(i+1)\Delta} | \zeta_{i\Delta}]$.

4.5 Computation of the CLS estimators of η^2 and ρ

Let consider the volatility of the temperature $(\zeta_t)_{t \geq 0}$ follows the below dynamics for $t \geq 0$ and $\zeta_0 \geq 0$:

$$
d\zeta_t = -K(\zeta_t - \sigma^2(t))dt + \eta \sqrt{\zeta_t} dW_t, \qquad (4.14)
$$

as in Model (M) with $\sigma^2(t) = \gamma_0 + \sum_{k=1}^{K_{\sigma^2}} \gamma_k \sin(\xi_k t) + \sum_{k=1}^{K_{\sigma^2}} \delta_k \cos(\xi_k t)$. We first focus on the conditional least squares estimator of the volatility of the volatility η^2, and assume that the coefficients K and $\sigma^2(\cdot)$ are known.

Proposition 4.3. *Let $(\zeta_t)_{t\geq 0}$ follow the dynamics* (4.14) *with $\sigma^2(t)$ being a nonnegative function of the form* (1.5) *and $\Delta > 0$. Then, we have for $i \in \mathbb{N}$*

$$Y_{i\Delta} = \mathbb{E}[(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}])^2 | \zeta_{i\Delta}] = \theta_0' + \phi_0'\zeta_{i\Delta} + \sum_{k=1}^{K_{\sigma^2}} \theta_k' \sin(\xi_k i\Delta) + \sum_{k=1}^{K_{\sigma^2}} \phi_k' \cos(\xi_k i\Delta) > 0,$$

with θ' and ϕ' defined by (4.17). *The solution of the minimisation problem*

$$\min_{\eta^2 \geq 0} \sum_{i=0}^{N-1} \left((\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}])^2 - \mathbb{E}[(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}])^2 | \zeta_{i\Delta}] \right)^2 \qquad (4.15)$$

is given by

$$\widehat{\eta^2} = \frac{\sum_{i=0}^{N-1} Y_{i\Delta}(\zeta_{(i+1)\Delta} - \vartheta^T X'_{i\Delta})^2}{\sum_{i=0}^{N-1} Y_{i\Delta}^2}, \qquad (4.16)$$

where ϑ and $X'_{i\Delta}$ are defined by (4.13).

For the proof of Proposition 4.3, we first state Lemma 4.2, which is a straightforward generalisation of (4.11).

Lemma 4.2. *For all $s \geq t$,*

$$\mathbb{E}[\zeta_s|\mathcal{F}_t] = \zeta_t e^{-K(s-t)} + \gamma_0(1 - e^{-K(s-t)}) + \sum_{k=1}^{K_{\sigma^2}} \Theta_k(s-t)\sin(\xi_k t) + \sum_{k=1}^{K_{\sigma^2}} \Phi_k(s-t)\cos(\xi_k t),$$

with

$$\Theta_k(v) = \gamma_k K \frac{K[\cos(\xi_k v) - e^{-Kv}] + \xi_k \sin(\xi_k v)}{K^2 + \xi_k^2} - \delta_k K \frac{K\sin(\xi_k v) - \xi_k(\cos(\xi_k v) - e^{-Kv})}{K^2 + \xi_k^2},$$

$$\Phi_k(v) = \gamma_k K \frac{K\sin(\xi_k v) - \xi_k(\cos(\xi_k v) - e^{-Kv})}{K^2 + \xi_k^2} + \delta_k K \frac{K[\cos(\xi_k v) - e^{-Kv}] + \xi_k \sin(\xi_k v)}{K^2 + \xi_k^2}.$$

Proof of Proposition 4.3. Let recall Equation (4.10) and compute the conditional variance of $(\zeta_t)_{t\geq 0}$:

$$\zeta_{t+\Delta} = \zeta_t e^{-K\Delta} + K \int_t^{t+\Delta} e^{-K(t+\Delta-s)}\sigma^2(s)ds + \eta \int_t^{t+\Delta} e^{-K(t+\Delta-s)}\sqrt{\zeta_s}dW_s$$

$$\mathbb{E}\left[(\zeta_{t+\Delta} - \mathbb{E}[\zeta_{t+\Delta}|\zeta_t])^2 | \mathcal{F}_t\right] = \mathbb{E}\left[\left(\eta \int_t^{t+\Delta} e^{-K(t+\Delta-s)}\sqrt{\zeta_s}dW_s\right)^2 | \mathcal{F}_t\right]$$

$$= \eta^2 \int_t^{t+\Delta} e^{-2K(t+\Delta-s)}\mathbb{E}[\zeta_s|\zeta_t]ds$$

by Ito's isometry, and using that $\mathbb{E}[\zeta_s|\mathcal{F}_t] = \mathbb{E}[\zeta_s|\zeta_t]$. From Lemmas 4.2 and 4.1, we deduce that

$$\int_t^{t+\Delta} e^{-2K(t+\Delta-u)}\mathbb{E}[\zeta_u|\zeta_t]du = \theta_0' + \phi_0'\zeta_t + \sum_{k=1}^{K_{\sigma^2}}\theta_k'\sin(\xi_k t) + \sum_{k=1}^{K_{\sigma^2}}\phi_k'\cos(\xi_k t),$$

where

$$\begin{cases} \theta_0' = \gamma_0 \dfrac{(1-e^{-K\Delta})^2}{2K} \\ \phi_0' = \dfrac{e^{-K\Delta}}{K}(1-e^{-K\Delta}) \\ \theta_k' = \gamma_k K \dfrac{K[A_k'-\phi_0']+\xi_k B_k'}{K^2+\xi_k^2} - \delta_k K \dfrac{KB_k'-\xi_k[A_k'-\phi_0']}{K^2+\xi_k^2} \\ \phi_k' = \gamma_k K \dfrac{KB_k'-\xi_k[A_k'-\phi_0']}{K^2+\xi_k^2} + \delta_k K \dfrac{K[A_k'-\phi_0']+\xi_k B_k'}{K^2+\xi_k^2}, \end{cases} \tag{4.17}$$

and

$$\begin{cases} A_k' = \dfrac{2K[\cos(\xi_k\Delta)-e^{-2K\Delta}]+\xi_k\sin(\xi_k\Delta)}{4K^2+\xi_k^2} \\ B_k' = \dfrac{2K\sin(\xi_k\Delta)-\xi_k(\cos(\xi_k\Delta)-e^{-2K\Delta})}{4K^2+\xi_k^2}. \end{cases}$$

For example, $\theta_k' = \int_t^{t+\Delta} e^{-2K(t+\Delta-u)}\Theta_k(u-t)du = \int_0^\Delta e^{-2K(\Delta-u)}\Theta_k(u)du$, and we use then Lemma 4.1 to get the formulas for $A_k' = \int_0^\Delta e^{-2K(\Delta-u)}\cos(\xi_k u)du$ and $B_k' = \int_0^\Delta e^{-2K(\Delta-u)}\sin(\xi_k u)du$. The calculation of ψ_k' works in the same way.
Hence, we get

$$\mathbb{E}[(\zeta_{t+\Delta} - \mathbb{E}[\zeta_{t+\Delta}|\zeta_t])^2\,|\zeta_t] = \eta^2\left(\theta_0' + \phi_0'\zeta_t + \sum_{k=1}^{K_{\sigma^2}}\theta_k'\sin(\xi_k t) + \sum_{k=1}^{K_{\sigma^2}}\phi_k'\cos(\xi_k t)\right),$$

and hence the value of $Y_{i\Delta}$. We also observe that $Y_{i\Delta} > 0$ since $e^{-2K((i+1)\Delta-u)} > 0$ and $\mathbb{E}[\zeta_u|\zeta_{i\Delta}] = \zeta_{i\Delta}e^{-K(u-i\Delta)} + K\int_{i\Delta}^u e^{-K(u-s)}\sigma^2(s)ds \geq K\int_{i\Delta}^u e^{-K(u-s)}\sigma^2(s)ds > 0$ for $u \in (i\Delta, (i+1)\Delta)$.
Now, let come back to the considered minimisation problem (4.15). We consider ϑ and $X_{i\Delta}'$ defined by (4.13) and define $Y_{i\Delta}$ as in Proposition 4.3 $= \theta_0' + \phi_0'\zeta_{i\Delta} + \sum_k \theta_k'\sin(\xi_k i\Delta) + \sum_k \phi_k'\cos(\xi_k i\Delta)$. Thus, from Proposition 4.2, we have $\mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}] = \vartheta^T X_{i\Delta}'$ and $\mathbb{E}[(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}])^2|\zeta_{i\Delta}] = \eta^2 Y_{i\Delta}$. Problem (4.15) is then equivalent to

$$\min_{\eta^2} \sum_{i=0}^{N-1} \left((\zeta_{(i+1)\Delta} - \vartheta^T X'_{i\Delta})^2 - \eta^2 Y_{i\Delta} \right)^2,$$

whose solution is given by (4.16). $\qquad\square$

We now focus on the conditional least squares estimator of the correlation ρ for Model (M) and assume that the coefficients κ, $s(\cdot)$, K, $\sigma^2(\cdot)$ and η^2 are known.

Proposition 4.4. *Let $(T_t, \zeta_t)_{t\geq 0}$ follow* (M) *with $\sigma^2(t)$ being a nonnegative function of the form* (1.5) *and $\Delta > 0$. Then, we have for $i \in \mathbb{N}$*

$$Y'_{i\Delta} = \theta''_0 + \phi''_0 \zeta_{i\Delta} + \sum_k \theta''_k \sin(\xi_k i\Delta) + \sum_k \phi''_k \cos(\xi_k i\Delta) > 0,$$

with θ'' and ϕ'' given by (4.19). *The solution of the minimisation problem*

$$\min_{\rho \in \mathbb{R}} \sum_{i=0}^{N-1} \Big((T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|\mathcal{F}_{i\Delta}])(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\mathcal{F}_{i\Delta}])$$

$$- \mathbb{E}[(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|\mathcal{F}_{i\Delta}])(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}]|\mathcal{F}_{i\Delta}]) \Big)^2$$

is given by

$$\hat{\rho} = \frac{\sum_{i=0}^{N-1} Y'_{i\Delta} (T_{(i+1)\Delta} - \lambda^T X_{i\Delta})(\zeta_{(i+1)\Delta} - \vartheta X'_{i\Delta})}{\sum_{i=0}^{N-1} (Y'_{i\Delta})^2}, \tag{4.18}$$

where $X_{i\Delta}$ and λ are defined in Proposition 4.1 (resp. $X'_{i\Delta}$ and ϑ in Proposition 4.2).

Let us note that we do not know a priori that $\hat{\rho} \in [-1, 1]$.

Proof. We first calculate the covariance between the temperature $(T_t)_{t\geq 0}$ and the volatility $(\zeta_t)_{t\geq 0}$:

$$\mathbb{E}[(T_{t+\Delta} - \mathbb{E}[T_{t+\Delta}|T_t])(\zeta_{t+\Delta} - \mathbb{E}[\zeta_{t+\Delta}|\zeta_t])|\mathcal{F}_t]$$

$$= \mathbb{E}\left[\left(\int_t^{t+\Delta} e^{-\kappa(t+\Delta-s)}\sqrt{\zeta_s}(\rho dW_s + \sqrt{1-\rho^2}dZ_s)\right)\left(\eta \int_t^{t+\Delta} e^{-K(t+\Delta-s)}\sqrt{\zeta_s}dW_s\right)|\mathcal{F}_t\right]$$

$$= \rho\eta\mathbb{E}\left[\int_t^{t+\Delta} e^{-(\kappa+K)(t+\Delta-s)}\zeta_s ds|\zeta_t\right] = \rho\eta \int_t^{t+\Delta} e^{-(\kappa+K)(t+\Delta-s)}\mathbb{E}[\zeta_s|\zeta_t]ds,$$

by using the Ito isometry, the independence between W and Z and $\mathbb{E}[\zeta_s|\mathcal{F}_t] = \mathbb{E}[\zeta_s|\zeta_t]$.

From Lemma 4.2, we get by standard calculations

$$\eta \int_t^{t+\Delta} e^{-(K+\kappa)(t+\Delta-s)} \mathbb{E}[\zeta_s|\mathcal{F}_t]ds = \theta_0'' + \phi_0''\zeta_t + \sum_{k=1}^{K_{\sigma^2}} \theta_k'' \sin(\xi_k s) + \sum_{k=1}^{K_{\sigma^2}} \phi_k'' \cos(\xi_k s),$$

with

$$\begin{cases} \theta_0'' = \eta\gamma_0 \left(\dfrac{1 - e^{-(\kappa+K)\Delta}}{\kappa + K} + \dfrac{e^{-(\kappa+K)\Delta} - e^{-K\Delta}}{\kappa} \right) \\[2ex] \phi_0'' = \eta e^{-K\Delta} \dfrac{1 - e^{-\kappa\Delta}}{\kappa} \\[2ex] \theta_k'' = \eta\gamma_k K \dfrac{K(A_k'' - \phi_0'') + \xi_k B_k''}{K^2 + \xi_k^2} - \eta\delta_k K \dfrac{KB_k'' - \xi_k(A_k'' - \phi_0'')}{K^2 + \xi_k^2} \\[2ex] \phi_k'' = \eta\gamma_k K \dfrac{KB_k'' - \xi_k(A_k'' - \phi_0'')}{K^2 + \xi_k^2} + \eta\delta_k K \dfrac{K(A_k'' - \phi_0'') + \xi_k B_k''}{K^2 + \xi_k^2}, \end{cases} \tag{4.19}$$

and

$$\begin{cases} A_k'' = \dfrac{(K + \kappa)(\cos(\xi_k\Delta) - e^{-(K+\kappa)\Delta}) + \xi_k \sin(\xi_k\Delta)}{(K + \kappa)^2 + \xi_k^2} \\[2ex] B_k'' = \dfrac{(K + \kappa)\sin(\xi_k\Delta) - \xi_k(\cos(\xi_k\Delta) - e^{-(K+\kappa)\Delta})}{(K + \kappa)^2 + \xi_k^2}. \end{cases}$$

These calculations are similar to the ones of Proposition 4.3, and we get that $Y_{i\Delta}' > 0$ exactly as we have obtained $Y_{i\Delta} > 0$ in this proposition. Now, let come back to the considered minimisation problem and define $Y_{i\Delta}$ as in Proposition 4.3. We also consider $X_{i\Delta} = (1, i\Delta, T_{i\Delta}, \sin(\xi i\Delta), \cos(\xi i\Delta))^T$ and λ defined by (4.6) (resp. $X_{i\Delta}'$ and ϑ defined by (4.13)), so that $\mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}] = \lambda^T X_{i\Delta}$ (resp. $\mathbb{E}[\zeta_{(i+1)\Delta}|\zeta_{i\Delta}] = \vartheta^T \zeta_{i\Delta}$). The minimisation problem can be rewritten as follows,

$$\min_\rho \sum_{i=0}^{N-1} \left((T_{(i+1)\Delta} - \lambda^T X_{i\Delta})(\zeta_{(i+1)\Delta} - \vartheta^T \zeta_{i\Delta}|\mathcal{F}_{i\Delta}]) - \rho Y_{i\Delta}' \right)^2,$$

and the minimum is clearly given by (4.18). $\qquad\qquad\square$

4.6 Strong consistency of CLS estimators for the time-dependent CIR processes

We study in this appendix the strong consistency of CLS estimators of a time-dependent CIR process. This process is implemented in this study to represent the temperature volatility dynamics. Let us consider the following process

$$d\zeta_t = K(\gamma\theta(t) - \zeta_t)dt + \eta\sqrt{\zeta_t}dW_t, \ \zeta_0 \geq 0, \tag{4.20}$$

with $K, \gamma, \eta > 0$ and $\theta : \mathbb{R}_+ \to \mathbb{R}_+$. We assume that process is observed at discrete times $(\zeta_{k\Delta})_{k \in \mathbb{N}}$.

The goal of this appendix is twofold. First, we prove in Theorem 1 the consistency of the CLS estimator of γ when other parameters are known and give the rate of convergence. This result complements the one of Overbeck and Ryden [123] in a time inhomogeneous case. This is a simplification with respect to the estimation of K, γ's and δ's in model (M) given by Proposition 4.2: we only estimate one drift parameter instead of $2(K_{\sigma^2} + 1)$ drift parameters. This avoids cumbersome calculations, but the same behaviour is expected for the CLS estimators of these $2(K_{\sigma^2} + 1)$ parameters. Second, we prove in Theorem 2 the consistency of the CLS estimator of η^2 when other parameters are known. This result complements the results of Bolyog and Pap [34] that only focus on the CLS estimation of the drift part.

By straightforward calculations, we have for $0 \le s \le t$,

$$\zeta_t = \zeta_s e^{-K(t-s)} + \int_s^t K\gamma\theta(u)e^{-K(t-u)}du + \eta \int_s^t e^{-K(t-u)}\sqrt{\zeta_u}dW_u, \qquad (4.21)$$

$$\mathbb{E}[\zeta_t | \zeta_s] = \zeta_s e^{-K(t-s)} + \int_s^t K\gamma\theta(u)e^{-K(t-u)}du.$$

The CLS estimator of γ consists in minimizing $\sum_{i=0}^{N-1} \left(\zeta_{i\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta} | \zeta_{i\Delta}] \right)^2$, i.e.

$$\sum_{i=0}^{N-1} \left(\zeta_{(i+1)\Delta} - \zeta_{i\Delta} e^{-K\Delta} - K\gamma \int_{i\Delta}^{(i+1)\Delta} \theta(u)e^{-K((i+1)\Delta-u)}du \right)^2,$$

which leads to

$$\hat{\gamma}_{N,\Delta} = \frac{\sum_{i=0}^{N-1} (\zeta_{(i+1)\Delta} - \zeta_{i\Delta} e^{-K\Delta}) \int_{i\Delta}^{(i+1)\Delta} \theta(u)e^{-K((i+1)\Delta-u)}du}{K \sum_{i=0}^{N-1} \left(\int_{i\Delta}^{(i+1)\Delta} \theta(u)e^{-K((i+1)\Delta-u)}du \right)^2}. \qquad (4.22)$$

In this appendix, we note $\hat{\gamma}_{N,\Delta}$ instead of $\hat{\gamma}$ to remind the dependence on N and Δ. This makes clearer the statements of Theorems 1 and 2 that involve these two quantities.

In the particular case $\theta \equiv 1$, we have

$$\hat{\gamma}_{N,\Delta} = \frac{1}{N(1 - e^{-K\Delta})} \sum_{i=0}^{N-1} (\zeta_{(i+1)\Delta} - \zeta_{i\Delta} e^{-K\Delta}) = \frac{1}{N} \sum_{i=1}^{N-1} \zeta_{i\Delta} + \frac{\zeta_{N\Delta} - \zeta_0 e^{-K\Delta}}{N(1 - e^{-K\Delta})}.$$

The second term is negligible and, following Overbeck and Ryden [123], we get that the estimator $\hat{\gamma}_{N,\Delta}$ is strongly consistent (i.e. $\hat{\gamma}_{N,\Delta} \to \gamma$ a.s.) and asymptotically normal (i.e. $\sqrt{N}(\hat{\gamma}_{N,\Delta} - \gamma)$ converges in law to a normal random variable) by using the ergodic theorem.

When θ is not constant, we can no longer use the ergodic theorem. We will make the proof of consistency under the assumption that θ is a bounded function. We lose the asymptotic normality but still have a convergence rate of $\sqrt{N}$. We will use the following lemma.

Lemma 4.3. *Let $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ be a bounded measurable function and $K > 0$. Then, the process* (4.20) *is well defined, nonnegative, and we have*

$$\forall p > 0, \ \sup_{t \geq 0} \mathbb{E}[\zeta_t^p] < \infty.$$

Proof. By using the well-known result of Yamada and Watanabe (see e.g. Karatzas and Shreve [92, Proposition 2.13 p. 291]), there exists a pathwise unique strong solution to (4.20). From the comparison result [92, Proposition 2.18 p. 293], ζ_t is greater than $\tilde{\zeta}_t = \zeta_0 - \int_0^t K\tilde{\zeta}_s ds + \eta \int_0^t \sqrt{\tilde{\zeta}_s} dW_s$, since the initial values are the same and the drift of $\tilde{\zeta}$ is below the one of ζ. Since $\tilde{\zeta}$ is a Cox-Ingersoll-Ross process, it is nonnegative. We thus have $\zeta_t \geq \tilde{\zeta}_t \geq 0$.

Now let us turn to the moments. It is sufficient to check the result for $p \in \mathbb{N}^*$. For $p = 1$, we have

$$\mathbb{E}[\zeta_t] = \zeta_0 e^{-Kt} + \int_0^t K\gamma\theta(u)e^{-K(t-u)}du \leq \zeta_0 + \gamma\bar{\theta},$$

with $\bar{\theta} = \sup_{u \geq 0} \theta(u) < \infty$. We then prove $\sup_{t \geq 0} \mathbb{E}[\zeta_t^p] < \infty$ by induction on p.

By Itô's formula, we have $d\zeta_t^p = p\zeta_t^{p-1}K(\theta(t) - \zeta_t)dt + p\eta\zeta_t^{p-1/2}dW_t + p(p-1)\frac{\eta^2}{2}\zeta_t^{p-1}dt$ and thus

$$\mathbb{E}[\zeta_t^p] = \zeta_0^p e^{-Kpt} + \int_0^t e^{-Kp(t-u)}pK\left(\theta(u) + \frac{p-1}{2K}\eta^2\right)\mathbb{E}[\zeta_t^{p-1}]du,$$

since the stochastic integral has a zero expectation (note that the SDE (4.20) has finite moments of any order by [92, Problem 3.15 p. 306]). This leads to $\mathbb{E}[\zeta_t^p] \leq \zeta_0^p + \left(\bar{\theta} + \frac{p-1}{2K}\eta^2\right)\sup_{t \geq 0}\mathbb{E}[\zeta_t^{p-1}]$, and to the claim by induction on p. $\square$

Theorem 1. *Let us assume that $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ is a bounded measurable function such that $0 < \underline{\theta} \leq \theta(u) < \bar{\theta}$ for some $\underline{\theta}, \bar{\theta} \in \mathbb{R}_+^*$. Then, for all $\Delta > 0$, the estimator $\hat{\gamma}_{N,\Delta}$ is strongly consistent (i.e. converges to γ a.s. as $N \to \infty$) and such that $N^\alpha(\hat{\gamma}_{N,\Delta} - \gamma) \to 0$ as $N \to \infty$ a.s. for any $\alpha \in (0, 1/2)$.*

Proof. From (4.21), we get

$$\zeta_{(i+1)\Delta} - \zeta_{i\Delta}e^{-K\Delta} = \gamma \int_{i\Delta}^{(i+1)\Delta} K\theta(u)e^{-K((i+1)\Delta-u)}du + \eta \int_{i\Delta}^{(i+1)\Delta} e^{-K((i+1)\Delta-u)}\sqrt{\zeta_u}dW_u.$$

Using this in (4.22), we obtain

$$\hat{\gamma}_{N,\Delta} = \gamma + \eta\frac{\sum_{i=0}^{N-1}\int_{i\Delta}^{(i+1)\Delta}e^{-K((i+1)\Delta-u)}\sqrt{\zeta_u}dW_u \int_{i\Delta}^{(i+1)\Delta}\theta(u)e^{-K((i+1)\Delta-u)}du}{K\sum_{i=1}^{N-1}\left(\int_{i\Delta}^{(i+1)\Delta}\theta(u)e^{-K((i+1)\Delta-u)}du\right)^2}$$

$$= \gamma + \eta\frac{\sum_{i=0}^{N-1}\Theta_i(M_{i+1} - M_i)}{\sum_{i=0}^{N-1}\Theta_i^2},$$

with $\Theta_i = K \int_{i\Delta}^{(i+1)\Delta} \theta(u) e^{-K((i+1)\Delta - u)} du$, $M_{i+1} - M_i = \int_{i\Delta}^{(i+1)\Delta} e^{-K((i+1)\Delta - u)} \sqrt{\zeta_u} dW_u$.

We have $\sum_{i=0}^{N-1} \Theta_i (M_{i+1} - M_i) = \int_0^{N\Delta} \Theta_{i(u)} e^{-K(i(u)\Delta - u)} \sqrt{\zeta_u} dW_u$. By Burkholder-Davis-Gundy inequality and then Jensen inequality, we get for $p \geq 2$,

$$\mathbb{E}\left[\left|\sum_{i=0}^{N-1} \Theta_i (M_{i+1} - M_i)\right|^p\right] \leq C_p \mathbb{E}\left[\left|\int_0^{N\Delta} \Theta_{i(u)}^2 e^{-2K(i(u)\Delta - u)} \zeta_u du\right|^{p/2}\right]$$

$$\leq C_p (N\Delta)^{p/2-1} \mathbb{E}\left[\int_0^{N\Delta} \Theta_{i(u)}^p e^{-pK(i(u)\Delta - u)} \zeta_u^{p/2} du\right]$$

$$\leq C_p (N\Delta)^{p/2} \bar{\theta}^p \sup_{t \geq 0} \mathbb{E}[\zeta_t^{p/2}],$$

where $i(u) = i$ for $u \in [i\Delta, (i+1)\Delta]$. Here, we have used $\Theta_i \leq (1 - e^{-K\Delta})\bar{\theta} \leq \bar{\theta}$.

On the other hand, we have $\Theta_i \geq (1 - e^{-K\Delta})\underline{\theta}$, and therefore

$$\mathbb{E}[|\epsilon_N|^p] \leq \left(\frac{1}{N(1 - e^{-K\Delta})^2 \underline{\theta}^2}\right)^p \times C_p (N\Delta)^{p/2} \bar{\theta}^p \sup_{t \geq 0} \mathbb{E}[\zeta_t^{p/2}],$$

with $\epsilon_N = \frac{\sum_{i=0}^{N-1} \Theta_i (M_{i+1} - M_i)}{\sum_{i=0}^{N-1} \Theta_i^2}$. This gives $\mathbb{E}[|\epsilon_N|^p] = O(N^{-p/2})$ by Lemma 4.3.

Therefore, for any $\alpha \in (0, 1/2)$, we can take $p > 2$ such that $p(1/2 - \alpha) > 1$ and thus $\mathbb{E}[\sum_{N=1}^{\infty} |N^\alpha \epsilon_N|^p] < \infty$, which gives that $N^\alpha \epsilon_N \to 0$, a.s. $\qquad \square$

We now turn to the Conditional Least Squares estimation of η^2 for the process (4.20). We now assume that $K, \gamma > 0$ and $\theta(\cdot)$ are known. Without loss of generality, we assume that $\gamma = 1$ and consider the minimization problem of

$$\sum_{i=0}^{N-1} \left[\left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\mathcal{F}_{i\Delta}]\right)^2 - \mathbb{E}\left[\left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\mathcal{F}_{i\Delta}]\right)^2 |\mathcal{F}_{i\Delta}\right]\right]^2,$$

with respect to η^2. By using Equation (4.21) and Fubini theorem, we get

$$\mathbb{E}\left[\left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\mathcal{F}_{i\Delta}]\right)^2 |\mathcal{F}_{i\Delta}\right] = \eta^2 \int_{i\Delta}^{(i+1)\Delta} e^{-2K((i+1)\Delta - u)} \mathbb{E}[\zeta_u|\mathcal{F}_{i\Delta}] du$$

$$= \eta^2 \left(\int_{i\Delta}^{(i+1)\Delta} e^{-2K((i+1)\Delta - u)} \zeta_{i\Delta} e^{-K(u - i\Delta)} du + \int_{i\Delta < v < u < (i+1)\Delta} e^{-2K((i+1)\Delta - u)} \theta(v) e^{-K(u-v)} du dv\right)$$

$$= \eta^2 \left(e^{-K\Delta} \frac{1 - e^{-K\Delta}}{K} \zeta_{i\Delta} + \int_{i\Delta}^{(i+1)\Delta} \theta(v) e^{-K((i+1)\Delta - v)}(1 - e^{-K((i+1)\Delta - v)}) dv\right). \qquad (4.23)$$

The minimization of

$$\sum_{i=0}^{N-1} \left[\left(\zeta_{(i+1)\Delta} - \mathbb{E}[\zeta_{(i+1)\Delta}|\mathcal{F}_{i\Delta}]\right)^2 - \eta^2 \left(e^{-K\Delta} \frac{1 - e^{-K\Delta}}{K} \zeta_{i\Delta} + \int_{i\Delta}^{(i+1)\Delta} \theta(v) e^{-K((i+1)\Delta - v)}(1 - e^{-K((i+1)\Delta - v)}) dv\right)\right]^2$$

then leads to the following estimator

$$\widehat{\eta^2}_{\Delta,N} = \frac{\sum_{i=0}^{N-1} \left(\zeta_{(i+1)\Delta} - (\zeta_{i\Delta} e^{-K\Delta} + \Theta_i^1) \right)^2 \left(e^{-K\Delta} \frac{1-e^{-K\Delta}}{K} \zeta_{i\Delta} + \Theta_i^2 \right)}{\sum_{i=0}^{N-1} \left(e^{-K\Delta} \frac{1-e^{-K\Delta}}{K} \zeta_{i\Delta} + \Theta_i^2 \right)^2},$$

with $\Theta_i^1 = K\gamma \int_{i\Delta}^{(i+1)\Delta} \theta(v) e^{-K((i+1)\Delta-v)} dv$ and $\Theta_i^2 = \int_{i\Delta}^{(i+1)\Delta} \theta(v) e^{-K((i+1)\Delta-v)}(1-e^{-K((i+1)\Delta-v)}) dv$.

Theorem 2. *Let us assume $\gamma = 1$ and that $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ is a bounded measurable function such that $0 < \underline{\theta} \leq \theta(u) < \bar{\theta}$ for some $\underline{\theta}, \bar{\theta} \in \mathbb{R}_+^*$. Then, for all $\Delta > 0$, the estimator $\widehat{\eta^2}_{N,\Delta}$ is strongly consistent (i.e. converges to η^2 a.s. as $N \to +\infty$) and such that $N^\alpha(\widehat{\eta^2}_{N,\Delta} - \eta^2) \to 0$ a.s. for any $\alpha \in (0, 1/2)$.*

Proof. The proof follows the same arguments as the one of Theorem 1, and we give the main lines. We have $\zeta_{(i+1)\Delta} - (\zeta_{i\Delta} e^{-K\Delta} + \Theta_i^1) = \eta \int_{i\Delta}^{(i+1)\Delta} e^{-K((i+1)\Delta-u)} \sqrt{\zeta_u} dW_u$ by (4.21) and define $a_i = e^{-K\Delta} \frac{1-e^{-K\Delta}}{K} \zeta_{i\Delta} + \Theta_i^2$ which is nonnegative. We can rewrite

$$\widehat{\eta^2}_{\Delta,N} = \eta^2 + \eta^2 \frac{\sum_{i=0}^{N-1} a_i \left[\left(\int_{i\Delta}^{(i+1)\Delta} e^{-K((i+1)\Delta-u)} \sqrt{\zeta_u} dW_u \right)^2 - a_i \right]}{\sum_{i=0}^{N-1} a_i^2}$$

We set $M_0 = 0$ and $M_{i+1} - M_i = \left(\int_{i\Delta}^{(i+1)\Delta} e^{-K((i+1)\Delta-u)} \sqrt{\zeta_u} dW_u \right)^2 - a_i$ for $i \in \mathbb{N}$. The process M is a $\mathcal{F}_{i\Delta}$-martingale by (4.23), and since a_i is $\mathcal{F}_{i\Delta}$-adapted, $\sum_{i=0}^{N-1} a_i(M_{i+1} - M_i)$ is also a martingale. Applying Burkholder-Davis-Gundy inequality and then Jensen inequality, we get for $p \geq 2$

$$\mathbb{E}\left[\left| \sum_{i=0}^{N-1} a_i(M_{i+1} - M_i) \right|^p \right] \leq C_p \mathbb{E}\left[\left(\sum_{i=0}^{N-1} a_i^2(M_{i+1} - M_i)^2 \right)^{p/2} \right] \leq C_p N^{p/2-1} \mathbb{E}\left[\sum_{i=0}^{N-1} a_i^p |M_{i+1} - M_i|^p \right].$$

Now, we check easily from Lemma 4.3 that $\mathbb{E}[a_i^p |M_{i+1} - M_i|^p] \leq C_p' < \infty$ for all i, and thus $\mathbb{E}\left[\left| \sum_{i=0}^{N-1} a_i(M_{i+1} - M_i) \right|^p \right] = O(N^{p/2})$.

On the other hand, we have $a_i \geq \Theta_i^2 \geq \underline{\theta} \frac{(1-e^{-K\Delta})^2}{2K}$ and thus $\sum_{i=0}^{N-1} a_i^2 \geq N\underline{\theta}^2 \left(\frac{(1-e^{-K\Delta})^2}{2K} \right)^2$. Setting $\epsilon_N = \frac{\sum_{i=0}^{N-1} a_i(M_{i+1}-M_i)}{\sum_{i=0}^{N-1} a_i^2}$, we get $\mathbb{E}[|\epsilon_N|^p] = O(N^{-p/2})$, and we conclude as in the proof of Theorem 1. $\square$

Chapter 2

Risk valuation of quanto derivatives for temperature and electricity

Introduction

The increasing impact of climate change on businesses has led to a growing demand for risk transfer instruments to hedge against its consequences. The energy sector is particularly affected by such weather variability. On the one hand, weather variability affects energy production. The availability of wind and solar radiation impacts the production of renewable electricity [19]. Similarly, experienced and predicted temperatures influence demand, as cold snaps increase heating demand in winter and heat waves increase cooling demand in summer [23]. This exposure to weather variability is often referred to as volumetric risk [117]. On the other hand, weather forecasts can have a direct impact on energy prices as actors anticipate demand increases and act in advance. This less frequently discussed risk is referred to as price risk [37].

Weather derivatives emerged in the 1990s as a response to this need for risk transfer. These financial instruments are based on an underlying weather index and trigger a claim depending on the value of the index at maturity, similar to other financial market derivatives. These instruments experienced significant success in the early 2000s, reaching $45 billion in notional volume traded in the market in 2006 according to the World Risk Management Association [147]. Mainly dominated by temperature-based derivatives, up to 95% of the market, the weather market remained illiquid with small volumes traded in the standardized open market and most of the volume traded OTC [146]. It also led to extensive research into the modeling of weather derivatives and best pricing methodologies [90] [24] [28] [3] [42] [41] [38].

By 2008, the weather market experienced a significant slowdown, with trading volumes declining to $11.8 billion in 2011 [148]. This corresponded to a general slowdown of financial markets, but also, according to Pérez-González and Yun [125], to the birth of new hybrid derivatives that could combine both volumetric and price risk. These new products, also called quantos, were indexed to two underlying parameters, one proxying the volumetric risk, typically a weather parameter, and one proxying the price risk, typically the spot price of electricity, gas

or oil. These double-indexed products already existed in the market for other financial assets (foreign exchange, bonds, commodities) [14] [88]. They are technically challenging because they require a convincing model of the joint distribution of the underlyings. Our analysis will focus on finding a model to price temperature and spot electricity price quantos.

Unfortunately, the literature exploring weather quantos is thin. Benth and al. [22] use a Heath-Jarrow-Morton approach to price hybrid derivatives combining New York Mercantile Exchange-traded natural gas futures and Chicago Mercantile Exchange-traded heating degree day futures for New York. Matsumoto and Yamada study the optimal design of mixed weather derivatives on wind indices and electricity prices [155]. Benth and Ibrahim [19] develop continuous-time models combining spot prices and logarithmic photovoltaic power production. For quantos combining energy prices and temperature, we should mention Caporin and al [44], who develop a two-dimensional daily ARFIMA-FIGARCH model for energy prices and temperature. They consider both an actuarial and a financial approach and perform simulation-based pricing that leads to important price differences [44]. Cucu and al. [56] develop a combined natural gas spot price and temperature model. They address calibration and pricing challenges for temperature-gas swaps. Finally, Benth and al. [17] consider bivariate Markov-modulated additive processes with independent non-stationary increments to model quantos combining temperature and energy and electricity and gas prices. Given a known analytical joint characteristic function for the logarithmic futures prices, they derive quanto pricing formulas for the Fast Fourier Transform (FFT) technique.

We begin our analysis by exploring various marginal models for spot energy price and daily temperature. In particular, we dive deep into a large literature on energy and commodity modeling [62] [150]. First, we examine mean-reverting diffusion models. Pioneering models by Gibson and Schwartz [77], Schwartz [132], and Lucia and Schwartz [107] propose two- or three-factors Gaussian diffusion dynamics to model commodity assets. However, the presence of non-Gaussian behaviors, including spikes, jumps, and heavy tails, has led to a refinement of these initial models. One proposal is to extend mean-reverting diffusion processes to Levy noises. Thus, compound Poisson processes have been studied by Geman and Roncoroni [75], Cartea and Figueroa [49], and Meyer-Brandis and Tankov [115]. A second widespread proposal is to move to multi-factor models with Brownian [107] [17] or Levy increments [20] [32]. Finally, Benth and Benth explore the relevance of mean-reverting diffusion processes with Normal Inverse Gaussian (NIG) increments [26]. We compare these models and consider different estimation and process characterization challenges for day-ahead auction market clearing prices [149] for the French and Northern Italian electricity markets. Finally, we propose to model the daily day-ahead log spot prices with mean-reverting processes and NIG increments.

For daily temperature models, our analysis is less extensive as the reader can refer to Alfonsi and Vadillo [8] for a more detailed presentation of daily temperature modeling applied to temperature derivatives pricing. We mainly suggest using a simple mean-reverting Gaussian model as in Benth and Benth [28] to model the daily average temperature for Charles de Gaulle and Milano Linate weather station data.

Third, we address the challenge of the joint temperature and log spot energy price distribution by proposing a coupled model on the dynamics. In particular, we introduce the Brownian noise of the temperature dynamics into the energy process. This allows the integration of weather information available at the time of price formation, as suggested by Benth and Meyer-Brandis [23], while maintaining flexibility and tractability in both processes. We estimate the marginals and dependence parameters of the joint model using Condition Least Square estimation applied to the characteristic function. χ^2 tests comparing the simulated and observed joint distributions confirm the goodness of fit of the combined model for both French and Northern Italian datasets.

Next, we introduce the pricing of quanto derivatives. Contrary to Benth and al [17], we do not consider quanto and temperature derivatives market as arbitrage-free complete markets. As noted above, most exchanges are OTC and CME standardized weather derivatives lack daily trading volume [146]. Temperature and energy quantos do not exist in any open market. Therefore, risk-neutral pricing is arguable and we stick to analyse the risk under the historical probability. Given our combined model, we derive explicit and semi-explicit formulas for the average payoff of futures, swaps and single sided options, here called $\mathcal{E}$-options, and double-sided options on temperature indices (HDD and CDD) and spot electricity price. These formulas are compared with payoff distributions derived from Monte Carlo simulations. Finally, we discuss the static hedging of $\mathcal{E}$-HDD and double sided quanto options in an self-financing portfolio framework, where the option is hedged by HDD and energy spot derivatives. We show that by using our model we can hedge most of the risk of quanto options and reduce their variances.

Hence the contributions of this study are multiple. First, it develops a convincing joint model for spot energy prices and daily average temperatures. Second, it proposes a method to estimate all the parameters of the model and assess the goodness of fit. Third, it develops pricing formulas under historical probability for futures, swaps, single and double-sided quanto options. Finally, it shows the hedging capability of single and double-sided quanto options.

The study is organized as follows. Section 1 presents the models for the univariate and combined dynamics of the logarithmic day-ahead spot price and the average daily temperature. Section 2 explores different dynamics for the log day-ahead energy spot price and justifies the modeling choice. Section 3 discusses the estimation challenges. Section 4 introduces the combined model and confirms its goodness of fit. Section 5 addresses the risk valuation of quantos that depend on both energy and temperature and develops a framework for static hedging of $\mathcal{E}$-HDDs and quanto options.

1 Model and data description

In this section we introduce our combined models to describe the dynamics of daily day-ahead energy log spot price $(X_t)_{t\geq0}$ and the temperature $(T_t)_{t\geq0}$. In the following, we will also note $S_t = e^{X_t}$, $t \geq 0$, the daily day-ahead energy spot price. We will consider time-continuous models with the time unit of one day ($\Delta = 1$), which follows literature practices as noted

by Deschatre [62]. Thus, $T_{i\Delta}$ will model the average daily temperature of the i-th day, usually defined in financial contracts as the average between the hourly minimum and maximum temperature.

We will consider the below Model (ETM) as the combined model for daily day-ahead energy log spot price and average daily temperature.

$$\begin{cases} d(X_t - \mu_X(t)) & = -\kappa_X(X_t - \mu_X(t)) + \lambda\sigma_T dW_t^T + dL_t^X \\ d(T_t - \mu_T(t)) & = -\kappa_T(T_t - \mu_T(t)) + \sigma_T dW_t^T \end{cases} \quad \text{(ETM)}$$

When $\lambda = 0$, the dynamics of X and T are independent. The elements characterizing the dynamics of the log-price X are:

- The deterministic function $\mu_X : \mathbb{R}_+ \to \mathbb{R}$ represents the trend and seasonality component. We assume that

$$\mu_X(t) = \beta_0^X t + \alpha_1^X \sin(\xi t) + \beta_1^X \cos(\xi t) + \alpha_{DoW(t)}^{X,DoW} \quad (1.1)$$

  where $\xi = \frac{2\pi}{365}$ and $DoW(t) = \lfloor \frac{t}{\Delta} \rfloor \mod p$ where $p \in \mathbb{N}^*$. In practice, $p = 7$ and $\alpha_{DoW(\cdot)}^{X,DoW}$ corresponds to the constant depending of the day in the week.

- The parameter $\kappa_X > 0$ corresponds to the mean-reverting (or autoregressive) behaviour.

- L^X is a Normal Inverse Gaussian distribution of parameters $(\alpha^X, \beta^X, \delta^X, m^X)$ which properties are described in Appendix 6.1. We will assume that this process is centered ($\mathbb{E}[L_t^X] = 0$), which means

$$m^X + \delta^X \frac{\beta^X}{\gamma^X} = 0.$$

Similarly, the elements characterizing the dynamics of the temperature T are:

- The function μ_T represents the trend and seasonality component. We assume that

$$\mu_T(t) = \alpha_0^T + \beta_0^T t + \alpha_1^T \sin(\xi t) + \beta_1^T \cos(\xi t), \text{ where } \xi = \frac{2\pi}{365}.$$

- The parameter κ_T corresponds to the mean-reverting behaviour.

- W^T is a Brownian motion independent of L^X, and $\sigma_T > 0$ to the standard deviation of the noise.

Last, the parameter $\lambda \in \mathbb{R}$ allows for some dependence between both processes. For the temperature, the Ornstein-Uhlenbeck (OU) model with Brownian noise corresponds to a well established model developed by Benth et al. [28] and largely spread on literature. We refer to Alfonsi and Vadillo [8] for a recent discussion on temperature models. Section 2 presents

different models for the electricity spot price and justifies the choice of dynamics of X in (ETM) when $\lambda = 0$. Then, Section 4 explores the pertinence of Model (ETM) and shows that it reproduces well the features of our data.

As it will be often useful in the calculations, we write here the integrated version of Model (ETM)

$$
\begin{cases}
X_t - \mu_X(t) & = e^{-\kappa_X(t-s)}(X_s - \mu_X(s)) + \lambda\sigma_T \int_s^t e^{-\kappa_X(t-u)}dW_u^T + \int_s^t e^{-\kappa_X(t-u)}dL_u^X \\
T_t - \mu_T(t) & = e^{-\kappa_T(t-s)}(T_s - \mu_T(s)) + \sigma_T \int_s^t e^{-\kappa_T(t-u)}dW_u^T,
\end{cases}
\tag{1.2}
$$

and introduce the notation $\tilde{X}_t = X_t - \mu_X(t)$ and $\tilde{T}_t = T_t - \mu_T(t)$ that will be used through the study.

Data description

The above model is tested in real world data. In particular, we study day-ahead log spot energy prices in France and North Italy from 5th January 2015 to 31st December 2018. This data is extracted from the ENTSO-E Transparency Platform and Gestore Mercati Energetici (GME) and are available hourly until 31st December 2022. We decided to average hourly data into daily data to avoid additional intra-day noise and follow literature practices [62]. This granularity choice will equally enable us to match the granularity of other dynamics like the temperature data's. Additionally, we exclude 2019 to 2022 years as energy price time series show considerably erratic paths due two major macroeconomic shocks: the COVID-19 pandemic and the Ukrainian war.

For temperature data, we choose to extract average daily temperature time series for Paris-Charles de Gaulle airport and Milano-Linate airport weather stations. These weather stations are referenced in WMO with the following identification numbers 7157 and 16080. Daily average temperature is defined as the average between the maximum and minimum hourly temperatures. Data is extracted from a private data provider platform. This latter is in charge of the removal of outliers. The data is therefore considered as cleaned in the following of this study.

2 Overview of different energy models

The literature on energy modeling is large. We would particularly recommend the surveys of Weron [150] and Deschatre et al.[62]. Although there can be exceptions, experts usually focus on either day ahead daily spot or forwards prices. The granularity kept is hence the day and is seen as the average of hourly spot prices. While forward price market modeling has been explored successfully through HJM-modeling paradigm [21], we will focus on spot or log spot price modeling. In the study of this section, we do not consider structural models nor neural networks models, but we focus rather on stochastic models. Indeed, our objective is to combine

energy dynamic modeling with temperature modeling to handle the risk of hybrid options and have a clear understanding of the model parameters. In the following we will consider log spot price to ensure positivity of the energy dynamics.

2.1 Mean-reverting diffusion models

The first models describing electricity dynamics are mean-reverting diffusion models. They were first developed by Gibson and Schwartz [77], Schwartz [132] and Lucia and Schwartz [107]. They do not focus only on electricity but apply these models to wider range of energy commodities (crude oil, on-peak electricity spot prices). They are built around the concept of *convenience yield* and model daily commodity through a Ornstein-Uhlenbeck (OU) process with Brownian noise [143] as follows:

$$\begin{cases} X_t &= \mu(t) + \tilde{X}_t \\ \tilde{X}_t &= -\kappa \tilde{X}_t dt + \sigma dW_t \end{cases} \tag{2.1}$$

where $\mu(\cdot)$ corresponds to a deterministic function including trend and seasonality and W to a Brownian noise.

There is not much discussion on the form of the deterministic function $\mu(\cdot)$. While several papers reduce this function to a simple constant [132] [77] [60] [96], other suggest different order Fourier expansions [26] [49] or piece-wise stepped functions [107]. These latter enable to include annual seasonality and capture differences in winter and summer prices, a phenomenon agreed upon literature [96] [69]. Pawlowsky and Nowak [124] justifies the presence of a trend component on the deterministic component. We suggest to keep this constant, trend and seasonality deterministic components and test their significance such that we define:

$$\mu(t) = \beta_0 t + \alpha_1 \sin(\xi t) + \beta_1 \cos(\xi t), \text{ where } \xi = \frac{2\pi}{365} \tag{2.2}$$

This first deterministic equation is implemented and tested on our data. For this we consider the results of the following regression function.

$$\sum_{i=0}^{N-1} \left(X_{i+1} - \mu(i+1) - e^{-\kappa}(X_i - \mu(i)) \right)^2, \tag{2.3}$$

where $\mu(\cdot)$ is defined as in Equation (2.2).

Following minimisation of Equation (2.3), we check significance of the coefficients and residual plots. While all coefficients show 5% significance, residual plots are less satisfactory. Indeed, as shown in Figure 2.1, residuals show important weekly dependencies. This phenomenon has already been observed by several papers. Following this observation, some suggest to distinguish week and week-end effects [107] [115] while others show statistical significance of daily dummy integration [35]. We also considered alternatives such as additional weekly seasonality terms.
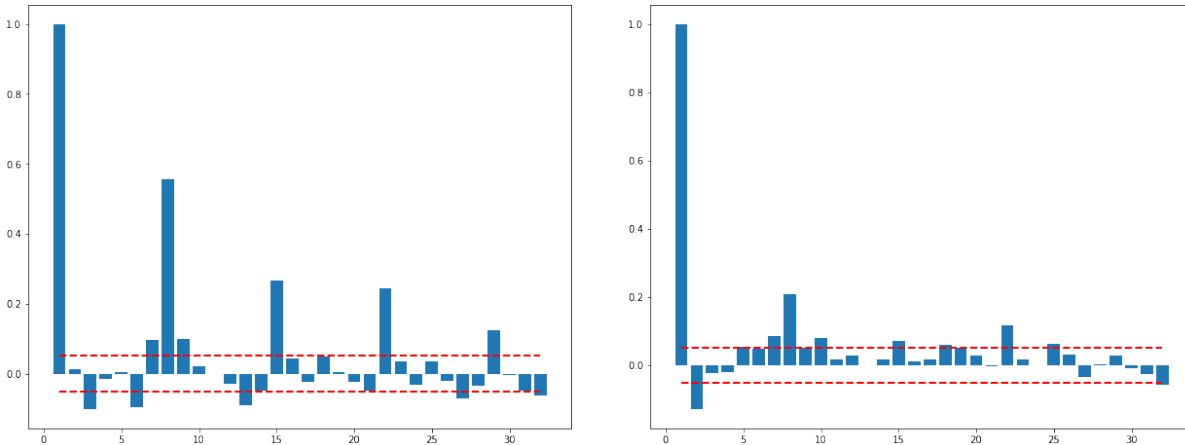
Figure 2.1: Partial autocorrelation plots of residuals of the Regression (2.3) where $\mu(\cdot)$ defined as in Equation (2.2) (left) and as in Equation (1.1) (right). The dashed red line corresponds to the 95% confidence interval from which we can consider the partial autocorrelation coefficient is significantly different from 0.

The form of $\mu(\cdot)$ minimising the AIC criteria turned to be the defined in Equation (1.1). Now let turn to the noise W, in our case we consider the residuals of Regression (2.3) where $\mu(\cdot)$ as defined in (1.1) to assess the characteristics of the distribution of W. Initially, the first models suggested Brownian dynamics for such noise following the model of Vasicek [143]. However, this proposal has been considerably challenged. Indeed, as can be seen in Figure 2.2, the qqplot of the residuals show significant deviation from normal theoretical quantiles. The residuals seem indeed to present heavier tails than normal residuals.

Different answers have been giving to this limitations. Some authors suggest to introduce price spikes thanks to jump-diffusion processes [60] [49] [75] while others explore multi-factor jump-diffusion models [115] or alternative distributions for the residuals [26]. Next sections will concentrate on these proposals and on their estimation power.

2.2 Mean-reverting jump-diffusion models (MRJD)

Several papers have studied the possibility to consider non-Gaussian increments. A particularly popular one is to combine Brownian motion with a compound Poisson process [60] that would capture the price spikes usually observed in energy prices, extending (2.1) as follows

$$X_t = \mu(t) + \tilde{X}_t$$
$$\tilde{X}_t = -\kappa \tilde{X}_t dt + \sigma dW_t + dJ_t,$$

where J_t is a Poisson process of intensity λ such that $J_t = \sum_{i=1}^{N_t} \xi_i$ where ξ_i are i.i.d. jump magnitudes that can follow distributions such as log-Normal [49], exponential [60] or mixture of exponential distributions [124]. One of the main drivers for distribution selection is the ability
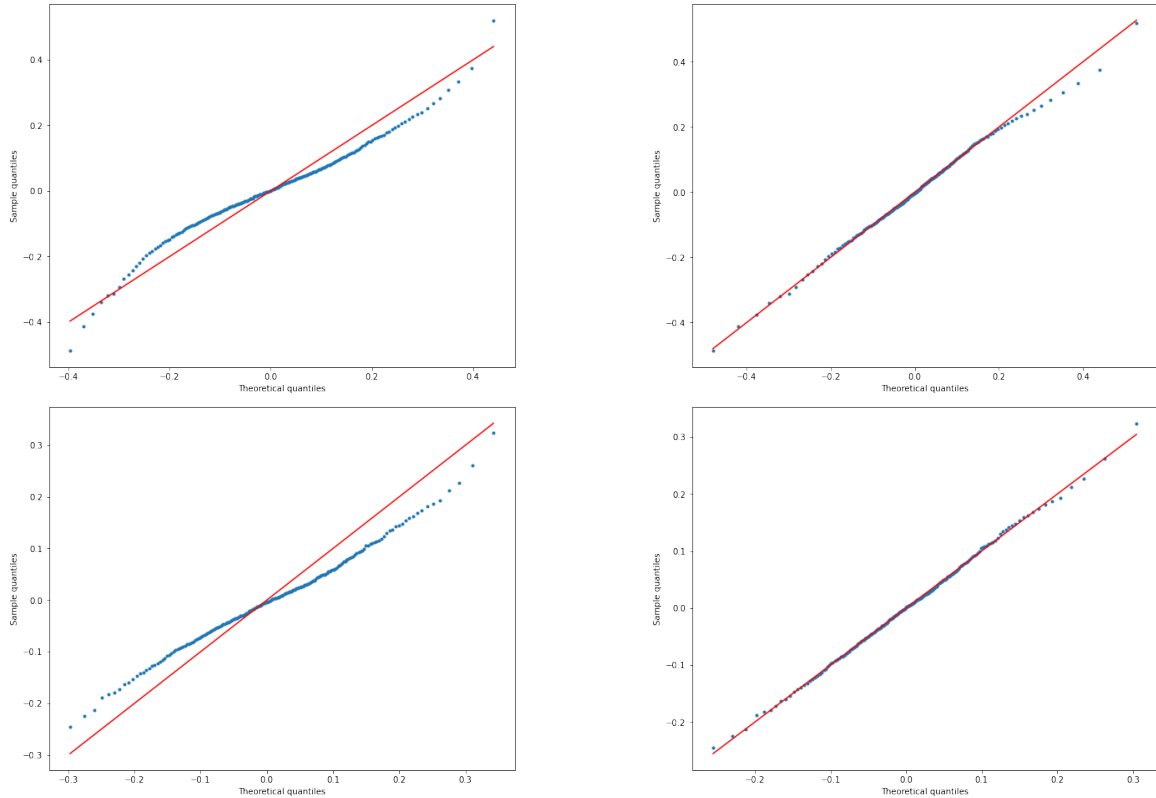
Figure 2.2: Quantile quantile plots for residuals of Regression (2.3) compared with a theoretical quantiles of a normal distribution (left) and of a normal inverse gaussian distribution (right) for French energy (first row) and North Italian Energy (second row).

to obtain explicit formulas for forward prices which mainly depends on the jump magnitude assumptions and the time-dependence of the other parameters. Indeed more flexible models, such us Geman and Roncoroni [75]'s, offer more flexibility on the properties of the Poisson process but do not enable explicit formulas of the forward prices.

However, only few papers analyse the challenge of estimating the parameters of such dynamics. Indeed, in order to estimate the parameters of both the continuous and the spiked noise, we need first to be able to distinguish them. There exist different methods of jump filtering. A first intuitive one is to settle a threshold, for example 3 standard deviations, such that data points within this threshold are considered to belong to the continuous part while the data points above correspond to the jumps. Cartea and Figueroa [49] and Pawlowski and Nowak [124] use an iterative method of filtering based on such threshold. However, the choice of the threshold seems arbitrary and integrates a standard deviation that itself combines continuous and spiked noises. Figure 2.3 shows the residuals of Regression (2.3) filtered through the above method. The residuals categorized as continuous suit well the normal quantiles however the jumps are pretty sparse and, hence, difficult to fit. We explored this method but we hardly could estimate

the jump parameters convincingly and found the filtering criteria rather arbitrary. We also con-
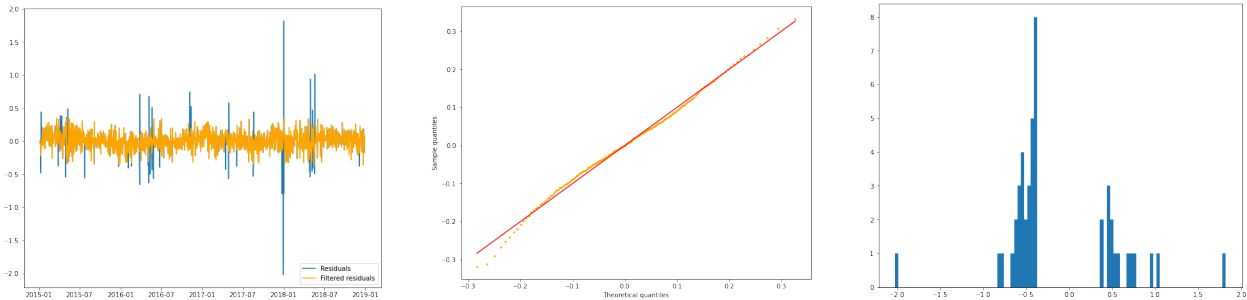

Figure 2.3: Jump filtering method as described by Cartea and Figueroa [49] (left). Filtered continuous residuals qqplot against a normal distribution (center). Histogram of jumps (right).

sidered alternative jump estimation methods. Meyer-Brandis and Tankov [115] challenge the threshold method in their two factor model and suggest two alternative filtering algorithms. Their implementation involves a hypothesis on the stochastic nature of the processes and the number of spikes. Alternatively, Deng [60] suggests to implement the method of moments which can introduce bias. Finally, Geman and Roncoroni [75] apply a maximum likelihood estimation applied to an unknown process using a prior reference process. This method is highly dependent on the underlying distribution and choice of priors.

2.3 Multi-factor mean-reverting models

Multi-factor models with non-Gaussian increments represent another popular alternative to model erratic dynamics. Two factors and three factors models with Gaussian increments were developed by Schwartz through different collaborations [132] [131] [77] [107]. The idea behind is that the spot prices could be driven by a long-term and a short-term dynamics, so that the spot price would integrate long-run terms and potential circumstantial tensions on the energy market. The estimation of such models can be performed through classic Kalman filtering. However, they do not answer the issue of bad fitting of the residuals with normal distribution as can be observed on the right figure of Figure 2.4.

Taking inspiration of Schwartz's models, several papers have explored the possibility to combine multi-factor models with Levy processes [32] [115]. The adaptation of (2.1) to a multi-factor model of n factors takes the following form:

$$\begin{cases} X_t & = \mu(t) + \sum \tilde{X}_t^n \\ \tilde{X}_t^n & = -\kappa_n \tilde{X}_t^n dt + dL_t^n \end{cases}$$

where L_t^n corresponds to a Levy process.
While Björk and Landén [32] deduce analytical expression for forward contracts with compound

Poisson processes, they do not address the estimation challenges. On their side, Meyer-Brandis and Tankov [115] study two factor models with non-defined Levy processes and suggest a calibration based on the autocorrelation function. Figure 2.4 shows the autocorrelation function and the exponential fitting enabling to compute the autoregressive parameters. It can be observed that the autocorrelation function does not present a clear exponential shape. Furthermore the fitting is extremely sensitive to the time horizon considered for the autocorrelation function.
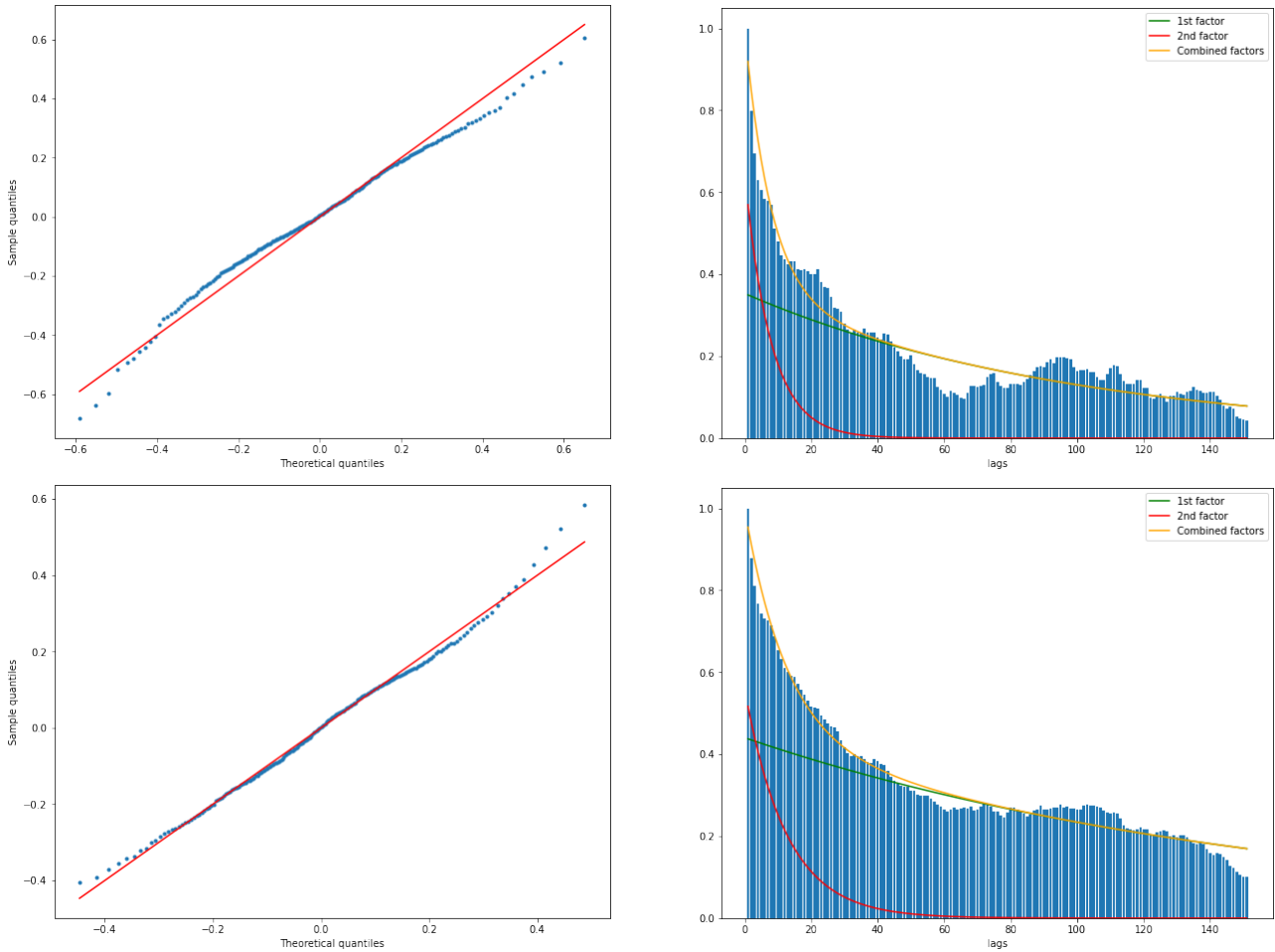

Figure 2.4: On the left, quantile quantile plots for residuals after removal of the two autocorrelation dynamics compared with a theoretical quantiles of a normal distribution. On the right, two factor model fitted through the autocorrelation function as described in [115]. First row corresponds to French energy data while second row to North Italy.

Finally, the inclusion of more than one autoregressive factors blocks from obtaining an integrant form of the dynamics. Given the latter and the lack of robust fitting method for

multi-factor models with Levy processes, we prefer to focus on one factor models.

2.4 Mean-reverting diffusion models with NIG noise

As suggested by Benth and Benth [26], we explore non-Gaussian Ornstein-Uhlenbeck process. The Normal Inverse Gaussian distribution was first introduced by Barndorff-Nielsen [12] and presents the distribution and properties in Appendix 6.1. The first motivation to keep this model is a rather good fitting of the residuals as can been seen in Figure 2.2. While residuals corresponding to $\int_s^t e^{-\kappa(t-u)}dL_u$ are not exactly distributed in NIG, the estimation of tinniness κ suggested a very close distribution that we find convincing. The second motivation to keep this model consisted in the relative easy estimation of the parameters. Section 3 will focus on these challenges.

Finally, this model can be easily generalized to multivariate distributions which will enable us to foresee combinations of energy and weather parameters dynamics in order to compute hybrid options.

As rightly noted by Benth and Benth [26], OU-NIG dynamics correspond to a first step towards stochastic volatility models that are common models for commodity derivative modelling [60] [141] [15]. While these models are usually developed for commodity forwards under the Heath–Jarrow–Morton (HJM) framework, the methodology can be replicated to spot price modeling [101]. Nevertheless, stochastic volatility models imply observing daily volatility on the modeled variable. While this is possible for energy spot prices, our objective is to combine this dynamic with daily average temperature dynamics for which the volatility is unobserved [8]. The following section will therefore focus on a two dimensional Ornstein-Uhlenbeck dynamic with NIG noise.

3 Separate parameter estimation of the two marginal processes

This section focuses on the estimation challenges of our two marginals: the daily day-ahead energy log spot price $(X_t)_{t\geq0}$ and the average daily temperature $(T_t)_{t\geq0}$. We assume that $\lambda = 0$ in Section 3, which gives the independence of these processes and allow to estimate their parameters separately. The joint estimation when $\lambda \neq 0$ will be discussed in the next section.

3.1 Estimation of κ_X and $\mu_X(\cdot)$

Following Klimko and Nelson [95], the objective of this section is to estimate κ_X and $\mu_X(\cdot)$ using Conditional Least Squares Estimation (CLSE). For this, we first write the conditional

expectation of $(X_t)_{t\geq 0}$. From

$$X_{t+\Delta} - \mu_X(t+\Delta) = e^{-\kappa_X \Delta}(X_t - \mu_X(t)) + \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dL_u^X,$$

we get

$$\mathbb{E}[X_{t+\Delta} - \mu_X(t+\Delta)|\mathcal{F}_t] = e^{-\kappa_X \Delta}(X_t - \mu_X(t)),$$

since we consider that L^X is centered, that is $m^X + \delta^X \beta^X / \gamma^X = 0$. We then get the following expression for the conditional expectation:

$$\mathbb{E}[X_{t+\Delta}|\mathcal{F}_t] = \mu_X(t+\Delta) + e^{-\kappa_X \Delta}(X_t - \mu_X(t)) \tag{3.1}$$

where $\mu_X(t) = \beta_0^X t + \alpha_1^X \sin(\xi t) + \beta_1^X \cos(\xi t) + \alpha_{DoW(t)}^{X,DoW}$ where $\xi = \frac{2\pi}{365}$ and $DoW(t) = \lfloor \frac{t}{\Delta} \rfloor$ mod 7.

We can now apply CLSE to the discrete for form of Equation (3.1) which boils down to minimise

$$\sum_{i=0}^{N-1} \left(X_{(i+1)\Delta} - \mathbb{E}[X_{(i+1)\Delta}|X_{i\Delta}] \right)^2. \tag{3.2}$$

This can be solved through linear regression, and Proposition 6.1 gives:

$$\begin{cases} \hat{\kappa}_X &= -\ln \hat{\eta}_2 \\ \hat{\beta}_0^X &= \frac{\hat{\eta}_1}{1-\hat{\eta}_2} \\ \hat{\alpha}_1^X &= \frac{\hat{\eta}_3(\cos(\xi\Delta)-e^{-\hat{\kappa}_X\Delta})+\hat{\eta}_4 \sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}_X\Delta})^2+\sin^2(\xi\Delta)} \\ \hat{\beta}_1^X &= \frac{\hat{\eta}_4(\cos(\xi\Delta)-e^{-\hat{\kappa}_X\Delta})-\hat{\eta}_3 \sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}_X\Delta})^2+\sin^2(\xi\Delta)} \\ \hat{\alpha}_j^{X,DoW} &= \frac{1}{1-e^{-7\hat{\kappa}_X\Delta}} \sum_{k=0}^{6}(\hat{\eta}_{j+k}^{DoW} - \hat{\beta}_0)e^{-(6-k)\hat{\kappa}_X\Delta}, \end{cases}$$

where

$$\hat{\eta} = \left(\sum_{i=0}^{N-1} \Xi_{i\Delta}\Xi_{i\Delta}^\top \right)^{-1} \left(\sum_{i=0}^{N-1} \Xi_{i\Delta}X_{(i+1)\Delta} \right), \tag{3.3}$$

with $\Xi_{i\Delta} = (i\Delta, X_{i\Delta}, \sin(\xi i\Delta), \cos(\xi i\Delta), (\mathbb{1}_{\{DoW(i\Delta)=j\}})_{0\leq j\leq 6}) \in \mathbb{R}^4 \times \{0,1\}^7$ for $i \in \mathbb{N}$ and $(\hat{\eta}_0^{DoW}, \ldots, \hat{\eta}_6^{DoW}) = (\hat{\eta}_5, \ldots, \hat{\eta}_{11})$ and $\hat{\eta}_j^{DoW} = \hat{\eta}_{\tilde{j}}^{DoW}$, with $\tilde{j} \in \{0, \ldots 6\}$ such that $j = \tilde{j} \mod 7$.

Figure 2.5 represents the fitted trend and seasonality component $\mu_X(\cdot)$ in the original logarithm of spot prices. We can observe an important difference of $\mu_X(\cdot)$ between week-ends and week days. In particular, Saturdays and Sundays are particularly cheaper days. Although energy still is traded on weekends, the volume is smaller which explains the different behaviours also noted by Meyer-Brandis and Tankov [115]. Finally, Table 2.1 shows estimated parameters for French and North Italian data.
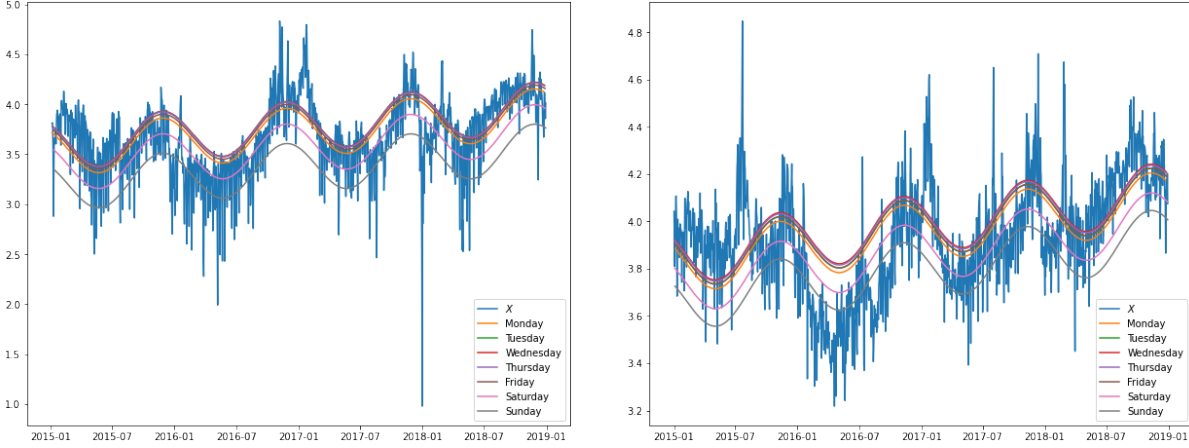
Figure 2.5: Fitted deterministic curve $\mu_X(\cdot)$ for France (left) and Italy (right).

	$\hat{\kappa}$	Residuals mean	Residuals SD
France	0.226	9.614×10^{-16}	0.171
North Italy	0.129	5.056×10^{-16}	0.099

Table 2.1: Fitted $\hat{\kappa}$ and residuals of Regression (3.2) for France (left) and Italy (right)

3.2 Parameter estimation for the NIG noise

We now move to the estimation of the parameters of the NIG process. Let first recall that:

$$\tilde{X}_{t+\Delta} = e^{-\kappa_X \Delta} \tilde{X}_t + \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dL_v^X \text{ where } \tilde{X}_t = X_t - \mu_X(t)$$

We can first observe that the residuals we study correspond to $\int_s^t e^{-\kappa_X(t-v)} dL_v^X$ and not the Levy noise per se. There is some study on the distribution of this integral particularly as Tempered Stable Processes, see Sabino [128]. However, given the inability to compute an explicit formula, we decided to work with approximations of this integrant. The following section will concentrate in three estimation methodologies: CLS, maximum likelihood and EM algorithm applied to a second order approximation of the characteristic function.

Let first study the form of the characteristic function of our process $(\tilde{X}_t)_{t\geq 0}$. We have

$$\mathbb{E}(e^{iu\tilde{X}_{t+\Delta}}) = \mathbb{E}\left(\exp(iu[e^{-\kappa_X \Delta} \tilde{X}_t + \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dL_v^X])\right)$$

If we focus on the second term and use Lemma 4.1. on Benth and Benth [26]. We call φ the characteristic function and we can write as follows:

$$\varphi(u; \Delta) = \mathbb{E}(\exp(iu \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dL_v^X)) = \exp\left(\int_t^{t+\Delta} \psi(iue^{-\kappa_X(t+\Delta-v)}) dv\right) \quad (3.4)$$

where ψ corresponds to the cumulant function of NIG distributions and is given by

$$\psi(x) = xm^X + \delta^X(\gamma^X - \sqrt{(\alpha^X)^2 - (\beta^X + x)^2})$$

Finally, we have:

$$\varphi(u; \Delta) = \exp\left(ium^X\frac{1 - e^{-\kappa_X\Delta}}{\kappa_X} + \delta\gamma^X\Delta - \delta^X\int_t^{t+\Delta}\sqrt{(\alpha^X)^2 - (\beta + iue^{-\kappa_X(t+\Delta-v)})^2}dv\right) \tag{3.5}$$

Estimation through CLS In this paragraph, we apply the CLS method developed by Klimko and Nelson [95] to the characteristic function φ. Our objective is to minimise the below function for $u \in \mathbb{R}$.

$$\left|e^{iu(\tilde{X}_{t+\Delta} - e^{-\kappa_X\Delta}\tilde{X}_t)} - \varphi(u; \Delta)\right|^2 \tag{3.6}$$

Now let consider a discrete time interval, the objective function in Equation (3.6) becomes:

$$\sum_{t=0}^{N-1}\left|e^{iu(\tilde{X}_{t+\Delta} - e^{-\kappa_X\Delta}\tilde{X}_t)} - \varphi(u; \Delta)\right|^2$$

In our case, we minimise for different values of u the below objective function:

$$\sum_u\sum_{t=0}^{N-1}\left|e^{iu(\tilde{X}_{t+\Delta} - e^{-\kappa_X\Delta}\tilde{X}_t)} - \varphi(u; \Delta)\right|^2 \tag{3.7}$$

We compute the characteristic function through numerical integration using the function `quad` in `Python` and u is taken in $\{-5, -4, ..., 5\}$.

Remark 3.1. *We contemplated to approximate the characteristic function through a Simpson's integration method. We tested how the choice of this approximation impacts the estimation by comparing with with the exact numerical integration. This method lead to similar results (see Table 2.2) and not particularly quicker (it takes 0.53s (resp. 0.63s) to minimize (3.7) with Simpson's on France (resp. North Italy) data) instead of 1.14s (resp. 1.25s) with exact integration). Therefore, on the following we keep using the exact numerical integration.*

The minimisation algorithm we apply is the Nelder–Mead algorithm [118]. We choose this method because it enables to integrate boundary constraints such as $\alpha^X > 0$ and $\alpha^X \geq |\beta^X|$ and shows good convergence. Table 2.2 summarizes the results of the minimisation.

The reader can also note that the above methodology was validated with simulated data. We simulated $100,000$ Normal Inverse Gaussian simulations with predefined parameters, computed the corresponding integral and verify that the minimisation of metric (3.7) lead to the correct parameter estimates. The sensitivity of this method was also tested with convincing results. We leave the mathematical proof of the characteristics of these estimators as a possible further research.

MLE and EM-algorithm estimation through second order approximation   An alternative approach for Normal Inverse Gaussian parameter estimation is to implement the Expectation-Maximization algorithm (EM) or maximum likelihood estimation (MLE). Unfortunately we do not know explicitly the density of $\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dL_v^X$. We hence proceed to an approximation of this density.

For this, let consider the characteristic function $\varphi(\cdot, \Delta)$ in Equation (3.5). We can see that:

$$\varphi(u; \Delta) \approx \exp\left(\Delta\left(ium^X e^{-\kappa_X\Delta/2} + \delta^X\gamma^X - \delta^X\sqrt{(\alpha^X)^2 - (\beta^X + iue^{-\kappa_X\Delta/2})^2}\right)\right)$$

Let apply the transformation,

$$\begin{cases} \tilde{m}^X & = \Delta m^X \\ \tilde{\delta}^X & = \Delta\delta^X \\ \tilde{u} & = ue^{-\kappa_X\Delta/2} \end{cases}$$

to obtain the below approximation for the characteristic function:

$$\mathbb{E}(\exp(i\tilde{u}e^{\kappa_X\Delta/2}\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dL_v^X)) \approx \exp\left(i\tilde{u}\tilde{m}^X + \tilde{\delta}^X\gamma^X - \tilde{\delta}^X\sqrt{(\alpha^X)^2 - (\beta^X + i\tilde{u})^2}\right)$$

Therefore, $\left(e^{\kappa_X\Delta/2}\int_{\ell\Delta}^{(\ell+1)\Delta} e^{-\kappa_X((\ell+1)\Delta-v)}dL_v^X\right)_{\ell\geq 1}$ behaves, at second order approximation, as a NIG independent distribution, the EM algorithm or maximum likelihood estimation is therefore applied to this time series. The Maximum Likelihood Estimation (MLE) is applied through the function pre-implemented in the Python library `scipy.stats` while the EM-algorithm is implemented following Karlis [93].

Method	$\hat{\alpha}^X$	$\hat{\beta}^X$	$\hat{m}^X$	$\hat{\delta}^X$	Method	$\hat{\alpha}^X$	$\hat{\beta}^X$	$\hat{m}^X$	$\hat{\delta}^X$
CLS Simp	4.222	-0.361	0.011	0.128	CLS Simp	13.736	0.319	0.004	0.152
CLS Num	4.222	-0.361	0.011	0.128	CLS Num	13.736	0.319	0.004	0.152
EM	4.066	-0.340	0.011	0.131	EM	12.576	0.533	-0.006	0.140
MLE	4.013	-0.203	0.007	0.130	MLE	12.575	0.533	-0.006	0.140

Table 2.2: Parameter estimation through the CLS, EM and maximum likelihood estimation for French (left) and Italian (right) energy log spot price.

Table 2.2 shows the parameter estimations through the CLS method, EM algorithm and maximum likelihood minimisation. We choose to move forward with the CLS Numerical method as all method lead to similar results and EM and MLE method cannot be directly applied to the combined model. From here onwards, MLE estimates are mainly used to initialize the CLS minimisation problem (3.7).

3.3 Estimation of κ_T and $\mu_T(\cdot)$

As introduced in Section 1, we consider that $(T_t)_{t \geq 0}$ follows an Ornstein Uhlenbeck with Gaussian residuals such that:

$$d(T_t - \mu_T(t)) = -\kappa_T(T_t - \mu_T(t))dt + dW_t^T$$

where $\mu_T(\cdot)$ represents the trend and seasonality component such that $\mu_T(t) = \alpha_0^T + \beta_0^T t + \alpha_1^T \sin(\xi t) + \beta_1^T \cos(\xi t)$, κ_T the autoregressive parameter of the OU and W_t^T is a Brownian motion.

First, similarly to Subsection 3.1, we implement CLS estimation to $(T_t)_{t \geq 0}$ by minimising:

$$\sum_{i=0}^{N-1} \left(T_{(i+1)\Delta} - \mathbb{E}[T_{(i+1)\Delta}|T_{i\Delta}] \right)^2. \tag{3.8}$$

This can be solved through linear regression to obtain (see [8, Proposition C.1])

$$\begin{cases} \hat{\kappa}_T &= -\ln \hat{\zeta}_2 \\ \hat{\alpha}_0^T &= \frac{\hat{\zeta}_0}{1-\hat{\zeta}_2} - \frac{\hat{\zeta}_1}{(1-\hat{\zeta}_2)^2} \\ \hat{\beta}_0^T &= \frac{\hat{\zeta}_1}{1-\hat{\zeta}_2} \\ \hat{\alpha}_1^T &= \frac{\hat{\zeta}_3(\cos(\xi\Delta)-e^{-\hat{\kappa}_T\Delta})+\hat{\zeta}_4\sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}_T\Delta})^2+\sin^2(\xi\Delta)} \\ \hat{\beta}_1^T &= \frac{\hat{\zeta}_4(\cos(\xi\Delta)-e^{-\hat{\kappa}_T\Delta})-\hat{\zeta}_3\sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}_T\Delta})^2+\sin^2(\xi\Delta)}, \end{cases}$$

where

$$\hat{\zeta} = \left(\sum_{i=0}^{N-1} \Pi_{i\Delta} \Pi_{i\Delta}^\top \right)^{-1} \left(\sum_{i=0}^{N-1} \Pi_{i\Delta} T_{(i+1)\Delta} \right) \tag{3.9}$$

with $\Pi_{i\Delta} = (1, i\Delta, T_{i\Delta}, \sin(\xi i\Delta), \cos(\xi i\Delta)) \in \mathbb{R}^5$ for $i \in \mathbb{N}$.

Figure 2.6 represents the daily average temperature and fitted function $\mu_T(\cdot)$. Table 2.3 presents the estimates of κ_T and residuals of Regression (3.8) for Paris and Milan. We can observe that the residuals are well centered.

	$\hat{\kappa}_T$	Residuals mean	Residuals SD
France	0.254	-1.741×10^{-15}	2.138
North Italy	0.250	-1.138×10^{-16}	1.638

Table 2.3: Fitted $\hat{\kappa}$ and residuals of Regression (3.8) for France (left) and Italy (right).

From here onwards, we will note $\tilde{T}$ the deseasonalised temperature such that $\tilde{T}_t = T_t - \mu_T(t)$.
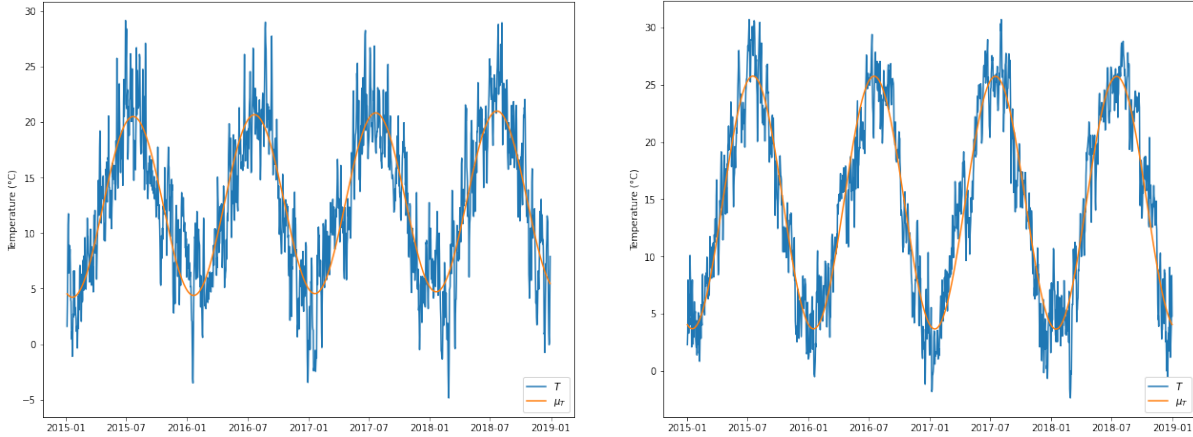
Figure 2.6: Fitted deterministic curve $\mu_T(\cdot)$ for France (left) and Italy (right).

3.4 Parameter estimation for the temperature residuals

Our first idea was to implement a multivariate NIG for $(\tilde{X}_t, \tilde{T}_t)$. We first then test if $(\tilde{T}_t)_{t \geq 0}$ can also follow a Ornstein Uhlenbeck process with NIG noise and finally show that Gaussian noises are more stable and reliable.

To estimate NIG parameters on $(\tilde{T}_t)_{t \geq 0}$, we implement the CLS approach as in Subsection 3.2.

Method	$\hat{\alpha}^T$	$\hat{\beta}^T$	$\hat{m}^T$	$\hat{\delta}^T$		Method	$\hat{\alpha}^T$	$\hat{\beta}^T$	$\hat{m}^T$	$\hat{\delta}^T$
CLS	14.095	10^{-3}	-0.002	98.610		CLS	38.152	10^{-9}	10^{-9}	111.119

Table 2.4: Parameter estimation through the CLS for French (left) and North Italian (right) temperature.

Table 2.4 shows the parameter estimations through CLS initiated with the MLE estimates. The results are quite unstable and particularly because β shrinks towards 0. Indeed, we are here confronted to the special case where $(\tilde{T}_t)_{t \geq 0}$ nearly follows a Normal distribution. The analysis of quantile quantile plots in Figure 2.7 show indeed that a Normal regression fits considerably well to the residuals regression 3.8. This result is aligned with Larsson's conclusions on German mean temperatures [102].

Hence, we choose to stick to a Normal distribution for $(T_t)_{t \geq 0}$. To estimate its parameters, we have that W^T is a Brownian noise $\sim \mathcal{N}(m^T, \sigma_T^2)$. Hence, we have $\int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)} dW_v^T \sim \mathcal{N}(m^T \sqrt{\frac{1-e^{-2\kappa_T \Delta}}{2\kappa_T}}, \sigma_T^2 \frac{1-e^{-2\kappa_T \Delta}}{2\kappa_T})$. We can obtain the parameters of the Normal distribution by correcting the residuals of Regression (3.8) by the factor $\sqrt{\frac{2\kappa_T}{1-e^{-2\kappa_T \Delta}}}$. Table 2.5 summarized the results.
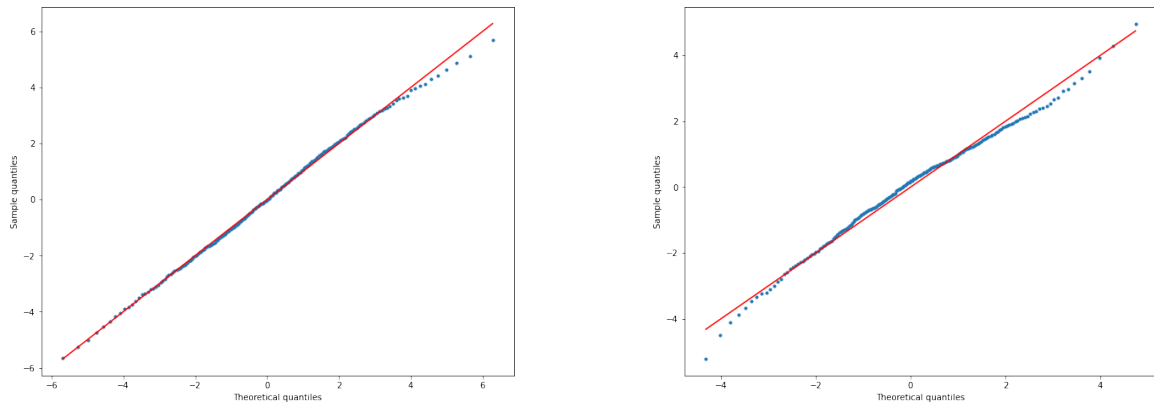
Figure 2.7: Quantile quantile plots for residuals of regression 3.8 compared with a theoretical quantiles of a normal distribution for Paris temperatures (left) and Milan temperatures (right).

Method	$\hat{m}^T$	$\hat{\sigma}_T^2$
MLE	10^{-15}	2.413

Method	$\hat{m}^T$	$\hat{\sigma}_T^2$
MLE	10^{-16}	1.846

Table 2.5: Parameter estimation through the maximum likelihood estimation for dynamic of temperature normally distributed for Paris (left) and Milan (right) temperature.

4 Towards a combined model for $(\tilde{X}_t, \tilde{T}_t)$

In the previous section, we have considered separate models for the electricity spot price and the temperature. This has enabled us to estimate the trend functions $\mu_X(\cdot)$ and $\mu_T(\cdot)$ and the speeds of mean-reversion κ_X and κ_T. We now consider the joint model (ETM) with $\lambda \neq 0$. First, we show empirical results on the dependence between electricity prices and the temperature. Then, we propose an estimation procedure of the different parameters. Since the temperature follows an autonomous dynamics in Model (ETM), the estimation of κ_T, $\mu_T(\cdot)$ and σ_T is unchanged from the previous section. By Proposition 6.1, the least square estimation of κ_X and $\mu_X(\cdot)$ is also unchanged when $\lambda \neq 0$. Therefore, this section focuses on the estimation of λ and then on the estimation of the NIG parameters when $\lambda \neq 0$.

4.1 Test of dependence

This section analyses the significance of the dependence structure between $(\tilde{X}_t)_{t \geq 0}$ and $(\tilde{T}_t)_{t \geq 0}$. For this we first estimate the Pearson correlation between the residuals $(\tilde{X}_{(i+1)\Delta} - e^{-\kappa_X \Delta} \tilde{X}_{i\Delta})$ and $(\tilde{T}_{(i+1)\Delta} - e^{-\kappa_T \Delta} \tilde{T}_{i\Delta})$. We obtain a correlation equal to -0.087 for France and -0.043 for north Italy, which suggests a small dependence. Figure 2.8 shows standardized residuals and ranked residuals plots for France and Italy. We cannot observe a clear dependence structure through these plots which supports the low correlations that we have obtained.

To better analyse the dependence, Table 2.6 shows the frequencies of ranked residuals given

a tercile classification. This time we can observe a slight anti-correlation as left top and bottom low corners are more populated than right top and bottom left corners in both cases.
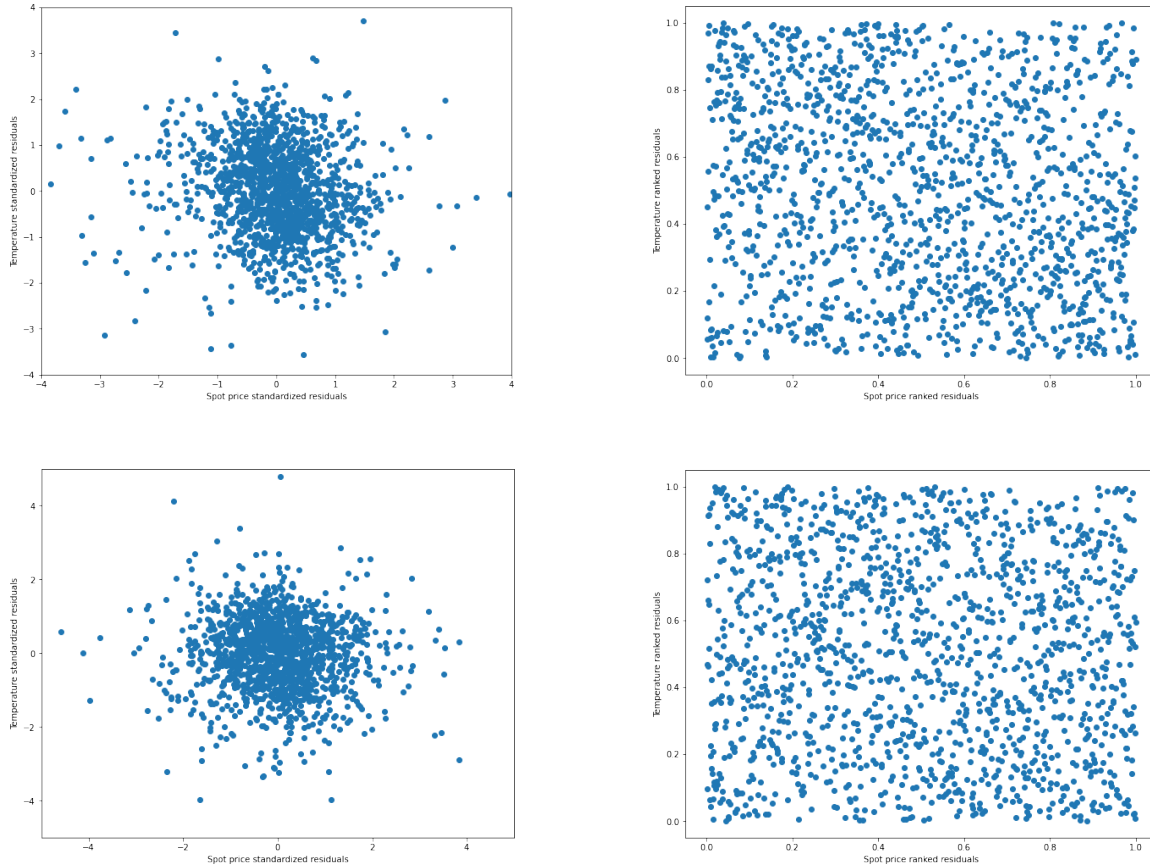

Figure 2.8: Standardized residuals (left) and ranked residuals (right) plots for France (top) and North Italy (bottom).

In order to move forward, we perform chi-square independence tests on the ranked residuals. For this we classify the ranked residuals based on quantiles, compute contingency tables with frequencies per coupled quantile classification and perform a chi-square independence test on these frequencies compared to a binomial distribution.

Figure 2.9 represents the results of the χ^2 independence test performed on residuals. We can see that the two datasets do not show same results: the French dataset clearly rejects the independence hypotheses while the North Italian dataset only rejects the independence hypothesis for 16 categories. This motivates us to propose a combined model for $(\tilde{X}_t, \tilde{T}_t)$ that allows dependence on the residuals.

197	162	126
163	160	162
124	163	197

165	173	149
176	147	163
145	166	175

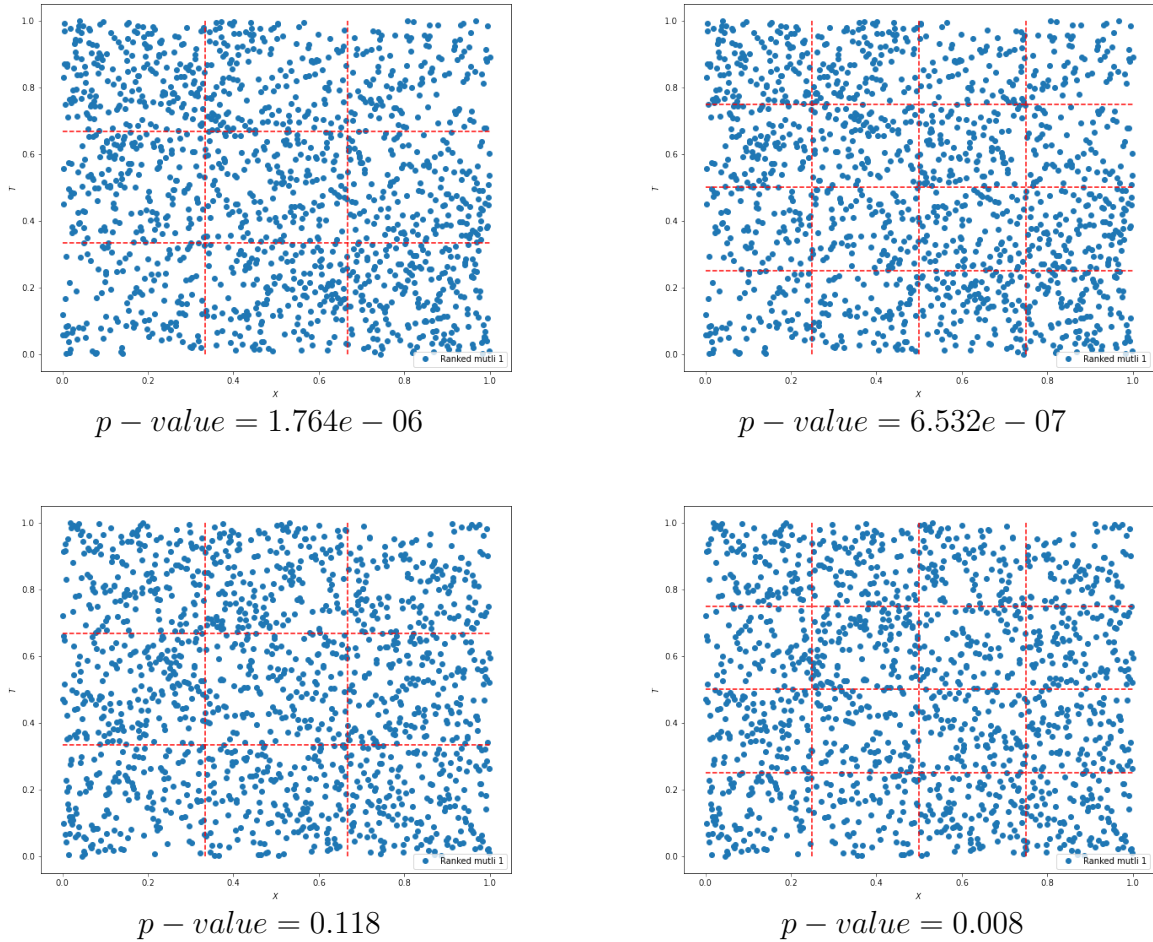Table 2.6: Observed frequencies by couple tercile for French (left) and Italian (right) coupled data.

$p - value = 1.764e - 06$

$p - value = 6.532e - 07$

$p - value = 0.118$

$p - value = 0.008$

Figure 2.9: χ^2-test performed on ranked residuals for 9 (left) and 16 (right) category classification for France (top) and Italy (bottom).

4.2 Estimation of λ and NIG parameters

Let us recall that the parameters of the Temperature diffusion κ_T, $\mu_T(\cdot)$ and σ_T can be estimated as in Section 3, as well κ_X and $\mu_X(\cdot)$ by Proposition 6.1. We assume these parameters estimated, and focus first on the estimation of λ.

Estimation of λ From (1.2), we compute the covariance of the residuals:

$$Cov\left(X_{t+\Delta} - \mu_X(t+\Delta) - e^{-\kappa_X\Delta}(X_t - \mu_X(t)), T_{t+\Delta} - \mu_T(t+\Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t))\right)$$

$$= Cov\left(\lambda\sigma_T\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dW_v^T + \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dL_v^X, \sigma_T\int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-s)}dW_v^T\right)$$

$$= \int_t^{t+\Delta} \sigma_T^2\lambda e^{-(\kappa_X+\kappa_T)(t+\Delta-v)}dv = \sigma_T^2\lambda\frac{1 - e^{-(\kappa_X+\kappa_T)\Delta}}{\kappa_X + \kappa_T}.$$

Hence, discretizing for a time period Δ, we get the following estimator

$$\hat{\lambda} = \frac{\hat{\kappa}_X + \hat{\kappa}_T}{\hat{\sigma}_T^2(1 - e^{-(\hat{\kappa}_X+\hat{\kappa}_T)\Delta})}\widehat{Cov},$$

where $\widehat{Cov}$ is the usual covariance estimator between residuals. Using values in Table 2.5, we get the estimated value of λ in Table 2.7.

Market	$\hat{\lambda}$
France	-0.007
North Italy	-0.002

Table 2.7: Estimated λ of Model (ETM) for France and North Italy.

Estimation of NIG parameters We will use, as in Section 3 the CLS estimation method. For this, we write the (conditional) characteristic function of the log-prices:

$$\psi_{X_t}(u;\Delta) := \mathbb{E}(e^{iuX_{t+\Delta}}|X_t) = e^{iu(\mu_X(t+\Delta)+e^{-\kappa_X\Delta}(X_t-\mu_X(t))-\frac{1}{2}\lambda^2\sigma_T^2\frac{1-e^{-2\kappa_X\Delta}}{2\kappa_X}u^2}\varphi(u;\Delta). \qquad (4.1)$$

This gives immediately the characteristic function of the residuals

$$\mathbb{E}[e^{iu(\tilde{X}_{t+\Delta}-e^{-\kappa_X\Delta}\tilde{X}_t)}|\mathcal{F}_t] = e^{-\frac{1}{2}\lambda^2\sigma_T^2\frac{1-e^{-2\kappa_X\Delta}}{2\kappa_X}u^2}\varphi(u;\Delta).$$

Then, to apply Conditional Least Square Estimation, we update the objective function in Equation (3.7) accordingly and seek to minimise the following quantity:

$$\sum_u\sum_{t=0}^{N-1}\left|e^{iu(\tilde{X}_{t+\Delta}-e^{-\kappa_X\Delta}\tilde{X}_t)} - e^{-\frac{1}{2}\lambda^2\sigma_T^2\frac{1-e^{-2\kappa_X\Delta}}{2\kappa_X}u^2}\varphi(u;\Delta)\right|^2.$$

Table 2.8 summarizes the NIG parameter estimations through CLS. We can observe that Table 2.8 is very close to Table 2.2 which is expected as the parameter λ is quite small.

We then turn to the goodness of fit of the estimated combined Model (ETM). Figure 2.10 and 2.11 represent χ^2 test performed between the historical and simulated distributions on the

Method	$\hat{\alpha}^X$	$\hat{\beta}^X$	$\hat{m}^X$	$\hat{\delta}^X$		Method	$\hat{\alpha}^X$	$\hat{\beta}^X$	$\hat{m}^X$	$\hat{\delta}^X$
CLS	4.189	-0.379	0.011	0.125		CLS	13.621	0.003	10^{-3}	0.151

Table 2.8: Parameter estimation through the CLS for French (left) and Italian (right) energy log spot price.

(2-dimensional) empirical copula between temperature and electricity spot price residuals. To ensure reliability of the results, we perform $1,000,000$ simulations and rescale the frequencies to compare with observed frequencies. We can see that the test does not globally reject the null hypotheses which means that the dependence is correctly reproduced by Model (ETM) for both French and North Italian data.

Finally we analyse the standard deviation explained by temperature component in Model (ETM). We use the ratio below:

$$\frac{|\sigma_T \lambda|}{\sqrt{\sigma_T^2 \lambda^2 + \frac{\delta \alpha^2}{\gamma^3}}}$$

we found that 9.43% and 4.34% of the standard deviation of the random term of the log energy spot price is explained by the temperature component for French and North Italian data correspondingly. This is small but not negligible, especially for handling the risk of derivatives as shown in the next section.

5 Handling the risk of quanto derivatives

On the previous sections we have developed a combined model for energy spot price and temperature in order to price financial derivatives combining both parameters. The objective of this section is to apply this model for quanto valuation and hedging.

5.1 An overview on quanto design and risk valuation

Quantos are derivative contracts of which payoff depends a double trigger, meaning the claim depends on the value at maturity of two indices. Their main interest relies on their capacity to hedge simultaneously volumetric risk, linked to weather conditions, and price risk, represented by the energy price. Quanto derivatives are defined, like weather and energy derivatives, over a time period $[t_1, t_2]$ such that the payoff of the contract will depend on the value of the underlying during all this risk period. While there exist large studies on quanto risk valuation involving different commodities [156], there exist little literature on the structuration and pricing of quanto products mixing commodity and weather inputs [44] [22] [97].

On the one hand, there is no clear consensus on the structuring of the products. For Benth and al. [22] and Kafakunusu [97], the payoff structure is applied to an aggregate of the two
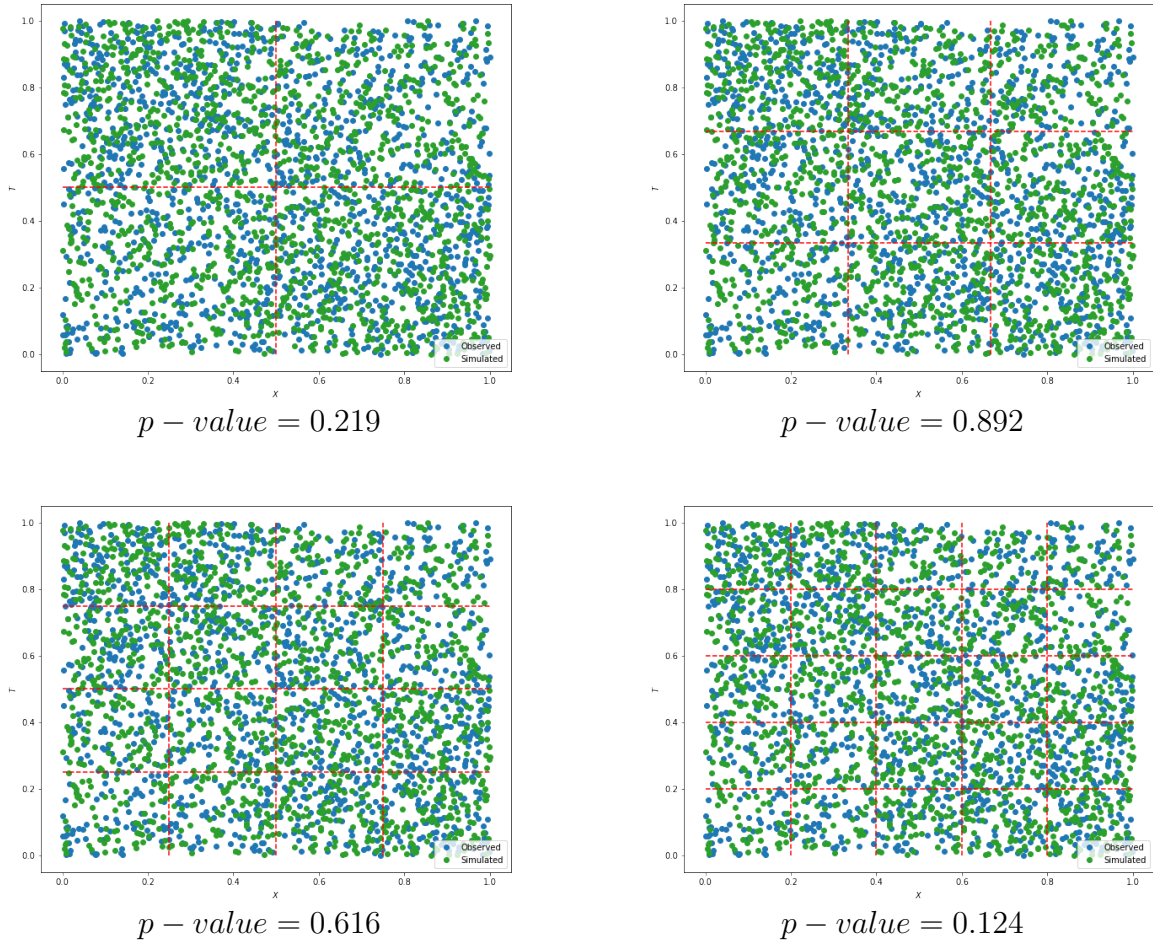
Figure 2.10: From top left to bottom right, χ^2 test performed on the distributions of real (blue) and simulated (green - based on $1,000,000$ simulations) ranked residuals for 4 and 25 categories for French data.

underlyings:

$$Payoff := f\left(\sum_{t=t_1}^{t_2} g_S(S_t), \sum_{t=t_1}^{t_2} g_T(T_t)\right)$$

where f represents the payoff function and g_S and g_T represent transformations of the initial inputs. The dates $t_1 < t_2$ indicates two days, and the summation is made on each day between t_1 and t_2 (including these days). In particular, for Caporin [44], f takes the form of a product and integrates common derivative payoff functions such that:

$$Payoff := f_S\left(\sum_{t=t_1}^{t_2} g_S(S_t)\right) \times f_T\left(\sum_{t=t_1}^{t_2} g_T(T_t)\right)$$
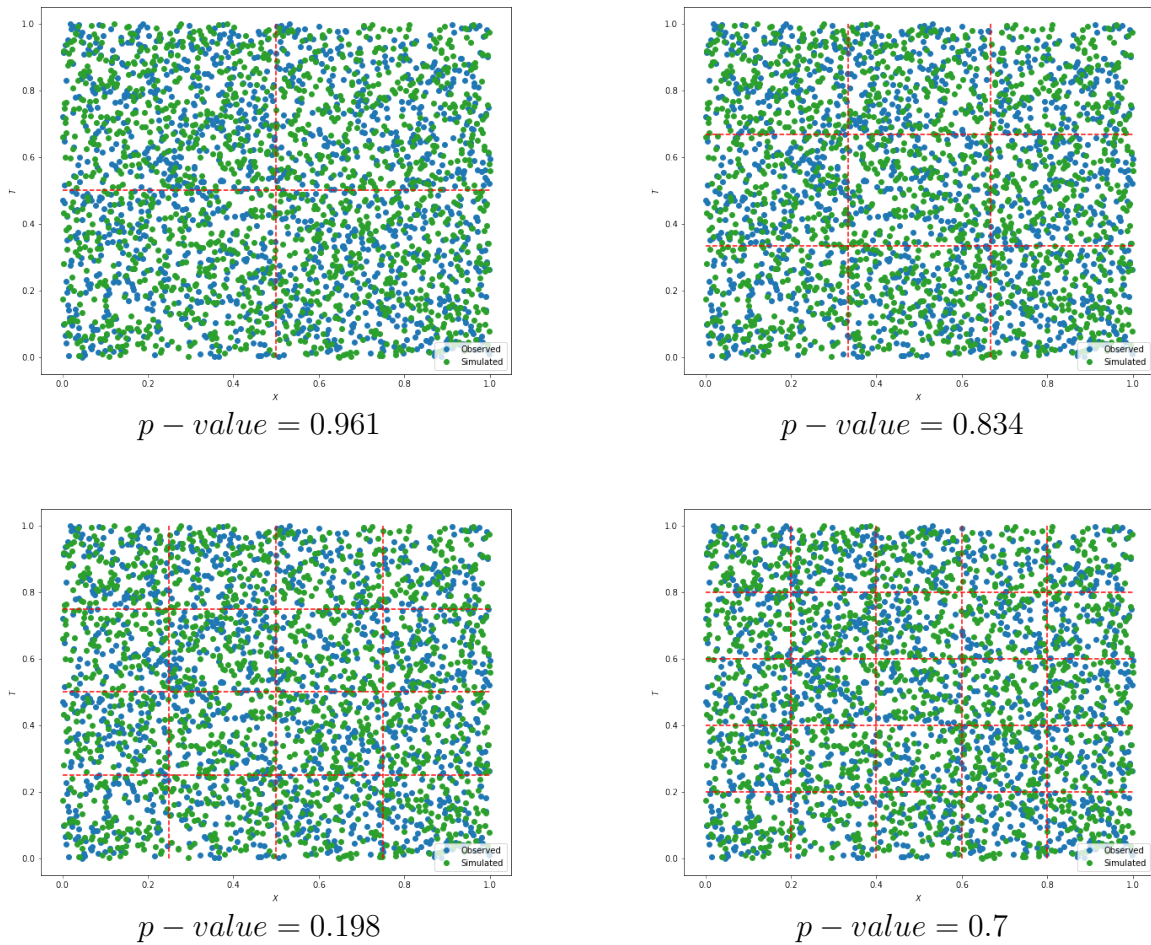
Figure 2.11: From top left to bottom right, chi-square test performed on the distributions of real (blue) and simulated (green - based on $1,000,000$ simulations) ranked residuals for 4 and 25 categories for North Italian data.

Here, f_S and f_T can correspond to the payoff function of put and call options, capped linear for swaps and identity for futures. Commonly f_T can correspond to the formula enabling to compute the Heating Degree Days (HDD) such that $f_T(\cdot) = (\bar{T} - T_t)^+$, with $\bar{T} = 18°C$. Finally, some practitioners seem to favour another definition [56]. In these cases the payoff function is directly applied to the daily values such that:

$$Payoff := \sum_{t=t_1}^{t_2} f_S(S_t) \times f_T(T_t) \tag{5.1}$$

We will focus on this latter definition as it answers to practitioner's needs. In addition, we consider price settlement takes place at t_0 and payoff payment at t_2.

We now discuss briefly the valuation of these products. Under classic risk-neutral pricing theory for financial derivatives we would like to write the price as follows:

$$\mathbb{E}_{\mathbb{Q}}\left(D(t_0, t_2) \sum_{t=t_1}^{t_2} f_S(S_t) \times f_T(T_t)\right)$$

where $\mathbb{Q}$ corresponds to the risk-neutral probability and $D(t_0, t_2)$ is a discount factor between t_0 and t_2.

However, temperature is not an asset traded on markets, and hence risk neutral theory cannot be applied. A possible way to get around this is to work with payoffs written on futures contracts, as in the works of Benth and al. [22] and Kafakunusu [97]. However, futures on temperature are still not liquid and usual quanto payoffs are written on the temperature itself, not on the future contracts. Here, we will rather work on the real-world probability world and analyse the payoff distribution in this framework. We will be particularly focus on the average payoff, which we discuss in the next subsection. In addition, we consider a unit discount rate (i.e. $D(t_0, t_2) = 1$) as the time span $t_2 - t_0$ is rather short and there is no clear alignment on which rate should be used, see e.g. [28] [38]. Note also that if the discount rate is deterministic or independent of (S, T), the formulas below are still valid up to a constant factor.

In the following section we develop explicit formulas for different payoff and compare then with Monte Carlo simulations. We also consider the possibility to hedge statically (5.1) with electricity and temperature derivatives, and analyse numerically the hedging error distribution.

5.2 Expected values of some standard payoffs

This section concentrates on our ability to get explicit values for the average payoff of different financial instruments. Namely, we will consider the payoff (5.1) for different choices of functions f_S and f_T. Let t_0 denote the present date, we consider two dates t_1 and t_2 such that $t_0 < t_1 < t_2$ and want to determine:

$$\mathbb{E}\left(\sum_{t=t_1}^{t_2} f_S(S_t) \times f_T(T_t) \mid \mathcal{F}_{t_0}\right).$$

The dates t_1 and t_2 indicate days, and the summation is made on all days between t_1 and t_2 including them. We also apply these formulas and compare the results to prices computed through Monte Carlo simulation. The Monte Carlo discretization schemes can be found in Appendix 6.3 and are applied to parameters in Table 2.9 and μ_X and μ_T as estimated for France with Formulas (3.3) and (3.9).

Forwards/Futures First, let consider $t_0 < t_1 < t_2$ and a future derivative between times t_1 and t_2. The payoff is given by (5.1) with $f_T(\cdot) = f_S(\cdot) = \cdot$. We consider the average payoff at

κ_X	β_0^X	α_1^X	β_1^X	α_{DoW0}^X	$\alpha_1^{X,DoW}$	$\alpha_2^{X,DoW}$	$\alpha_3^{X,DoW}$	$\alpha_4^{X,DoW}$	$\alpha_5^{X,DoW}$	$\alpha_6^{X,DoW}$
0.226	0.0003	-0.165	0.187	3.523	3.594	3.594	3.589	3.566	3.370	3.175

κ_T	α_0^T	β_0^T	α_1^T	β_1^T	α^X	β^X	m^X	δ^X	σ_T	λ
0.254	6.578	0.00004	-4.139	-6.959	4.189	-0.379	0.011	0.125	2.413	-0.007

Table 2.9: Parameters used for numerical experiments. These parameters are the ones obtained with the estimation on French electricity and temperature data.

time t_0 under the historical probability:

$$\mathcal{F}(t_1, t_2) = \mathbb{E}\Big(\sum_{t=t_1}^{t_2} S_t \times T_t \mid \mathcal{F}_{t_0} \Big)$$

Proposition 5.1. *Under Model* (ETM) *and for $t \in [t_1, t_2]$, we have:*

$$\mathcal{F}(t_1, t_2) = \sum_{t=t_1}^{t_2} \Bigg[\exp\Big(\mu_X(t) + e^{-\kappa_X(t-t_0)}(X_{t_0} - \mu_X(t_0)) \Big) \varphi(-i; t - t_0)$$

$$\Big((\mu_T(t) + e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0)))e^{\frac{1}{2}k_X(t-t_0)^2\lambda^2\sigma_T^2} \qquad (5.2)$$

$$+\lambda\sigma_T^2 k_{XT}^2(t-t_0)e^{\frac{1}{2}\lambda^2\sigma_T^2 k_X(t-t_0)^2} \Big) \Bigg]$$

where φ is the characteristic function defined in Equation (3.5) *and*

$$k_T(\Delta) = \sqrt{\frac{1 - e^{-2\kappa_T\Delta}}{2\kappa_T}}, \ k_X(\Delta) = \sqrt{\frac{1 - e^{-2\kappa_X\Delta}}{2\kappa_X}} \ and \ k_{XT}(\Delta) = \sqrt{\frac{1 - e^{-(\kappa_X+\kappa_T)\Delta}}{\kappa_X + \kappa_T}}. \qquad (5.3)$$

This result is a direct consequence of Proposition 6.4 that calculates explicitly $\mathbb{E}(S_t \times T_t \mid \mathcal{F}_{t_0})$. Figure 2.12 shows the price of monthly futures computed 30 days in advance for France data. The prices have been computed with 100,000 Monte Carlo simulations and with explicit formula in Equation (5.2). We can see that both methodologies provide same results. However, computation with Equation (5.2) is around 650 times faster. Additionally we compute the price with $\lambda = 0$. Although it is close to the precedent case, the price computed with $\lambda = 0$ is usually not within the confidence intervals of the Monte Carlo simulations.

Swap Swaps are derivatives where the payoff is given by (5.1), that is, $f_T(\cdot) = (\bar{T} - \cdot)$ and $f_S(\cdot) = (\cdot - \bar{S})$. Here we suppose strikes $\bar{S}$ and $\bar{T}$ are given. For applications, and on the following of the study $\bar{T} = 18°C$ and $\bar{S} = 50$ EUR/MWh. While taking $\bar{T} = 18°C$ is a market

standard, taking $\bar{S} = 50$ is our choice. We decide to take a strike which is relatively into the money and repeat the exercise for other strikes ($\bar{S} = 40, 60$), the below conclusions remain unchanged.

We define the swap's average payoff under historical probability $\mathcal{S}(t_1, t_2)$ as follows:

$$\mathcal{S}(t_1, t_2) = \mathbb{E}\Big(\sum_{t=t_1}^{t_2} (S_t - \bar{S})(\bar{T} - T_t) \mid \mathcal{F}_{t_0} \Big)$$

Proposition 5.2. *Under Model* (ETM) *and for* $t \in [t_1, t_2]$, *we have*

$$\mathcal{S}(t_1, t_2) = \sum_{t=t_1}^{t_2} \bar{T} \psi_{X_{t_0}}(-i; t - t_0) - \mathcal{F}(t_1, t_2) - \bar{S}\bar{T}(t_2 - t_1 + 1)$$

$$+ \bar{S} \sum_{t=t_1}^{t_2} (\mu_T(t) + e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0)))$$

(5.4)

where ψ *is the characteristic function defined in Equation* (4.1), $\mathcal{F}(t_1, t_2)$ *as defined in Equation* (5.2) *and* $k_T(\cdot)$, $k_X(\cdot)$ *and* $k_{XT}(\cdot)$ *are defined in Equation* (5.3).

This result is a direct consequence of Proposition 6.5 that calculates $\mathbb{E}((S_t - \bar{S})(\bar{T} - T_t) \mid \mathcal{F}_{t_0})$. Figure 2.12 shows the price of monthly swaps computed 30 days in advance for France data. The prices have been computed with 100,000 Monte Carlo simulations and with explicit formula in Equation (5.4). As for forwards, we can see that both methods provide similar results while using Formula (5.4) is around 300 times faster. Again, prices computed for $\lambda = 0$ are close but out of the confidence interval of the Monte Carlo simulations.

Single sided options $\mathcal{E}$-HDD and $\mathcal{E}$-CDD We now focus on put and call options on temperature. These are defined by the payoff (5.1) with $f_S(x) = x$ and either $f_T(\cdot) := (\bar{T} - \cdot)^+$ for a put option or $f(\cdot) := (\cdot - \bar{T})^+$ for a call option. We call these products single sided options, since the option brings only on the temperature. We define the average payoff under historical probability of a single sided option $\mathcal{E}-HDD(t_1, t_2)$ as follows:

$$\mathcal{E}-HDD(t_1, t_2) = \mathbb{E}\Big(\sum_{t=t_1}^{t_2} S_t (\bar{T} - T_t)^+ \mid \mathcal{F}_{t_0} \Big) = \sum_{t=t_1}^{t_2} \mathbb{E}(S_t (\bar{T} - T_t)^+ \mid \mathcal{F}_{t_0})$$

Let first consider each term separately. We have:

$$\mathbb{E}(S_t (\bar{T} - T_t)^+ \mid \mathcal{F}_{t_1}) = \int_{T^0}^{\bar{T}} \mathbb{E}(e^{X_t} \mathbb{1}_{T_t \leq u} \mid \mathcal{F}_{t_1}) du,$$
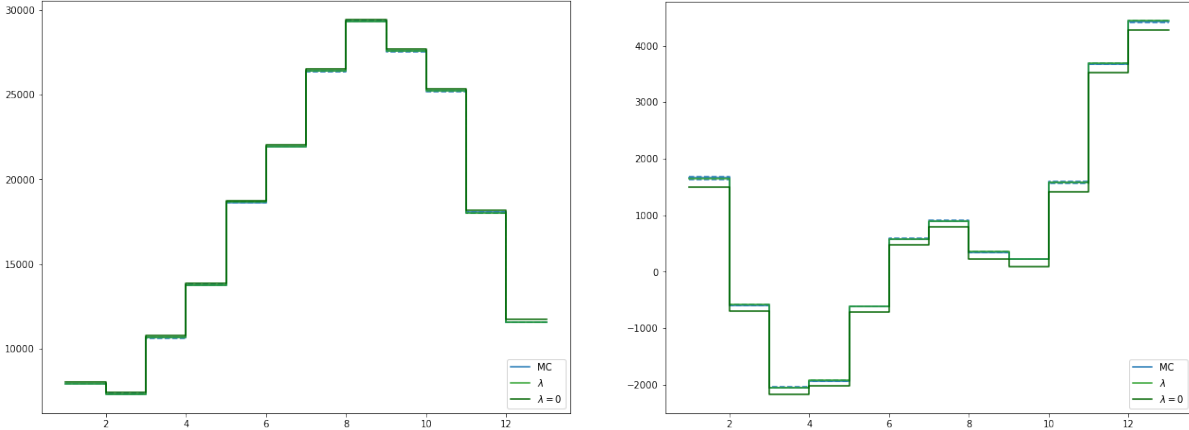
(5.5)

Figure 2.12: Forward (left) and Swap (right) prices computed with 100,000 simulation-Monte Carlo (blue) and Equations (5.2) and (5.4) (green) methods. Each contract lasts one month of 2018. Time t_0 corresponds to 30 days ahead of the first day of the month, t_1 to the first day of the month and t_2 to the last day of the month.

where $T^0 = -\infty$. Note that the Gaussian model for the temperature allows in principle any real temperature. In practice, with the estimated parameters, the probability of having extreme temperatures is very infinitesimal. We can thus use (5.5) with $T^0 = -273.15$ (the absolute zero temperature) or $T^0 = -100$ with a negligible error.

Proposition 5.3. *Under Model* (ETM) *and for* $t \in [t_1, t_2]$, *we have*

$$
\mathcal{E}-HDD(t_1, t_2) = \sum_{t=t_1}^{t_2} \Bigg[\psi_{X_{t_0}}(-i; t - t_0) \times
$$
$$
\int_{T^0}^{\bar{T}} \Phi\Big(\frac{u - (\mu_T(t) + e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0)))}{\sigma_T k(t - t_0)} - \lambda \sigma_T \frac{k_{XT}^2(t - t_0)}{k_T(t - t_0)} \Big) du \Bigg]
$$

(5.6)

where ψ *is the characteristic function defined in Equation* (4.1), Φ *is the cumulative function of the standard Gaussian distribution and* $k_T(\cdot)$, $k_X(\cdot)$ *and* $k_{XT}(\cdot)$ *are as in Equation* (5.3).

Proof. For $t \in [t_1, t_2]$, we use (5.5) and apply then Proposition 6.9. □

Similarly, we can calculate explicitly the average payoff of a $\mathcal{E}$-CDD defined as:

$$
\mathcal{E}-CDD(t_1, t_2) = \mathbb{E}\Big(\sum_{t=t_1}^{t_2} S_t(T_t - \bar{T})^+ \mid \mathcal{F}_{t_0} \Big).
$$

Proposition 5.4. *Under Model* (ETM) *and for* $t \in [t_1, t_2]$, *we have:*

$$
\mathcal{E}-CDD(t_1, t_2) = \sum_{t=t_1}^{t_2} \left[\psi_{X_{t_0}}(-i; t-t_0) \times \right.
$$
$$
\left. \int_{\bar{T}}^{T^m} \Phi\left(\lambda \sigma_T \frac{k_{XT}^2(t-t_0)}{k_T(t-t_0)} - \frac{u - (\mu_T(t) + e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0)))}{\sigma_T k_T(t-t_0)}\right) du \right]
$$

(5.7)

where $T^m = +\infty$, ψ *is the characteristic function defined in Equation* (4.1), Φ *is the cumulative function of the standard Gaussian distribution,* $k_T(\cdot)$, $k_X(\cdot)$ *and* $k_{XT}(\cdot)$ *are as in Equation* (5.3).

In practice, (5.7) can be used with $T_m = 100$ with a negligible error, the probability of having temperature above 100°C being infinitesimal.

Figures in 2.13 show the price of monthly $\mathcal{E}$-HDD and $\mathcal{E}$-CDD computed 30 days in advance for French data. The prices have been computed with 100,000 Monte Carlo simulations and with explicit formula in Equations (5.6) and (5.7). We can see that both methodologies provide same results. However, computation with formulas is, again, around 40 times faster for $\mathcal{E}$-HDD and 3 times faster for $\mathcal{E}$-CDD. In addition, we compute the price with $\lambda = 0$, the impact on prices is visual for $\mathcal{E}$-CDD as the prices are out of the confidence interval of the Monte Carlo simulations. This shows the significance of λ on the valuation of derivatives.
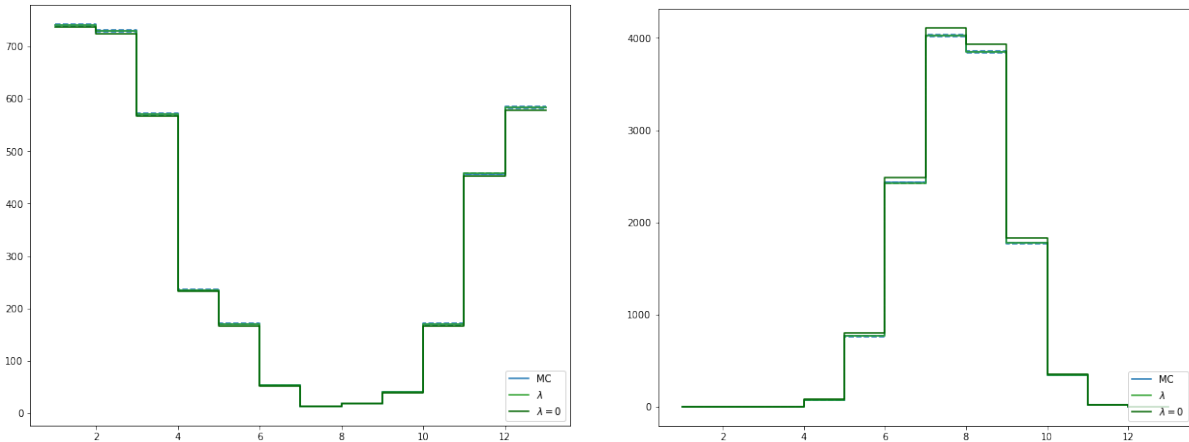

Figure 2.13: $\mathcal{E}$-HDD (left) and $\mathcal{E}$-CDD (right) prices computed with 100,000 simulation-Monte Carlo (blue) and Equations (5.6) and (5.7) (green) methods. Each contract lasts a month of 2018. t_0 corresponds to 30 days ahead of the first day of the month, t_1 to the first day of the month and t_2 to the last day of the month. The computation of the derivatives through the formulas around 40 times faster for $\mathcal{E}$-HDD and 3 times faster for $\mathcal{E}$-CDD.

Quanto options Let now consider double sided options given by Equation (5.1) with the following payoff function specifications $f_S(\cdot) = (\cdot - \bar{S})^+$ and $f_T(\cdot) = (\bar{T} - \cdot)^+$. We consider then the average payoff function:

$$\mathcal{Q}(t_1, t_2) = \mathbb{E}\Big(\sum_{t=t_1}^{t_2} (S_t - \bar{S})^+ (\bar{T} - T_t)^+ \mid \mathcal{F}_{t_0} \Big)$$

Given our Model (ETM), there is no explicit formula for $\mathcal{Q}$ up to our knowledge. However, we suggest to perform a Taylor's series expansion to the first order on λ given that λ is quite small. The next proposition gives semi-explicit formulas for this expansion.

Proposition 5.5. *Under Model (ETM) and for $t \in [t_1, t_2]$, we have the following Taylor's series expansion:*

$$
\begin{aligned}
\mathcal{Q}(t_1, t_2) = \sum_{t=t_1}^{t_2} \bigg(& \mathbb{E}_{\lambda=0}((S_t - \bar{S})^+ \mid \mathcal{F}_{t_0}) \times \bigg(\Big(\bar{T} - \mu_T(t) - e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0))\Big) \times \\
& \Phi\Big(\frac{\bar{T} - \mu_T(t) - e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0))}{\sigma_T k_T(t - t_0)}\Big) \\
& + \frac{\sigma_T k_T(t - t_0)}{\sqrt{2\pi}} \exp\Big(-\frac{1}{2}\Big(\frac{\bar{T} - \mu_T(t) - e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0))}{\sigma_T k_T(t - t_0)}\Big)^2\Big)\bigg) \\
& - \Big(\mathbb{E}_{\lambda=0}((S_t - \bar{S})^+ \mid \mathcal{F}_{t_0}) + \bar{S}\mathbb{P}_{\lambda=0}\big(S_t \geq \bar{S} \mid \mathcal{F}_{t_0}\big)\Big) \times \\
& \sigma_T^2 k_{XT}(t - t_0)^2 \Phi\Big(\frac{\bar{T} - \mu_T(t) - e^{-\kappa_T(t-t_0)}(T_{t_0} - \mu_T(t_0))}{-\sigma_T k_T(t - t_0)}\Big)\Big) \lambda \bigg) + o(\lambda)
\end{aligned}
$$
(5.8)

where Φ is the cumulative distribution function of the standard Gaussian distribution, $k_T(\cdot)$, $k_X(\cdot)$ and $k_{XT}(\cdot)$ are as in Equation (5.3).

This result is a direct application of Proposition 6.7. Note that $\mathbb{E}_{\lambda=0}((S_t - \bar{S})^+ \mid \mathcal{F}_{t_0})$ (resp. $\mathbb{P}_{\lambda=0}\big(S_t \geq \bar{S} \mid \mathcal{F}_{t_0}\big)$) can be computed efficiently as in Equation (6.3) (resp. Equation (6.4)).

Figure 2.14 shows quanto prices computed with Equation (5.8) and 100,000 simulation-Monte Carlo simulations. We can see that the first order Taylor development is sufficient as prices are always within Monte Carlo confidence intervals. We compare this with a quanto approached keeping only the first term of Taylor development which is equivalent to $\lambda = 0$. This time, average payoff values are close but out of the Monte Carlo confidence intervals. Explicit formula computation remains faster than Monte Carlo simulations however it gets less attractive than for previous derivatives as it includes several numerical integrations.

To sum up, we have developed in this subsection explicit or semi-explicit formulas for futures, swaps, single-sided and double sided options given Model (ETM). These formulas are verified
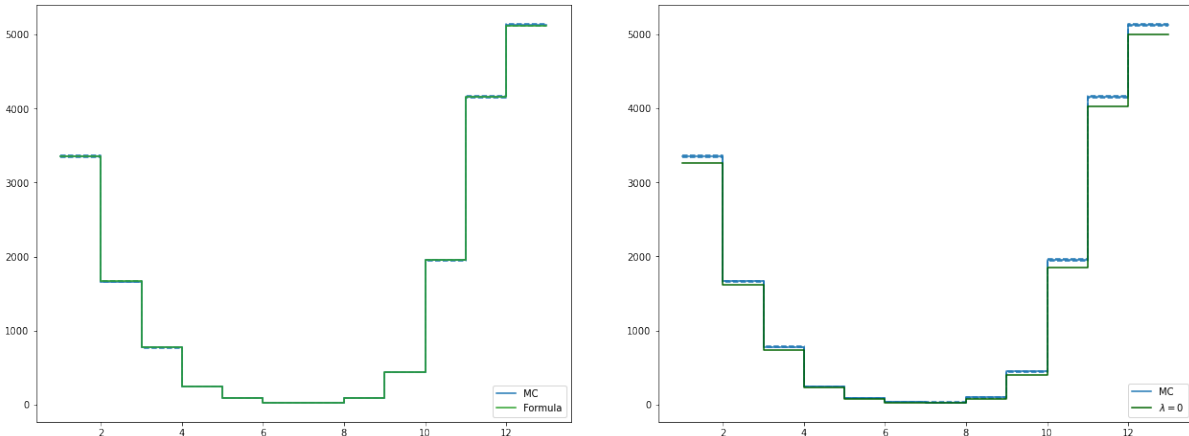
Figure 2.14: Quanto prices computed with 100,000 simulation-Monte Carlo (blue) and Equation (5.8) (green) methods. On the left the price corresponds to Formula (5.8). On the right only the first term of the Taylor development in Equation (5.8) is considered. This is equivalent to consider $\lambda = 0$. Each contract lasts a month of 2018. Time t_0 corresponds to 30 days ahead of the first day of the month, t_1 to the first day of the month and t_2 to the last day of the month. The computation of the derivatives through the formulas is around 6 times faster than using Monte Carlo simulations.

through Monte Carlo simulations and remain faster to use than Monte Carlo techniques. Finally, approaching the formulas with $\lambda = 0$ provides results close to the simulated ones but out of their confidence intervals showing the significance of the dependence between electricity and temperature on the risk associated to these derivatives.

5.3 Static hedging of hybrid derivatives

In the following section, we leverage explicit formulas developed in the above subsection and discuss potential statistic hedging strategies for $\mathcal{E}$-HDD and double-sided quantos.

We remind that hedging challenges are key for portfolio and risk managers to handle risk aggregation questions and meet solvency constraints. The aim of this subsection is to show that while quantos are somehow exotic derivatives, they can be hedged through more common derivatives increasing their attractiveness and risk understanding.

5.3.1 Risk decomposition of $\mathcal{E}$-HDD

Let first focus on $\mathcal{E}$-HDD and take a self-financing portfolio approach. We suppose that we are at time t and that we want to find the static portfolio made with single HDD, future on electricity and cash that minimizes the square hedging error at $t + \Delta$ where $\Delta > 0$:

$$\mathbb{E}[(S_{t+\Delta}(\bar{T} - T_{t+\Delta})^+ - c^0_{t,t+\Delta} - c^1_{t,t+\Delta}(\bar{T} - T_{t+\Delta})^+ - c^2_{t,t+\Delta}S_{t+\Delta})^2|\mathcal{F}_t]. \tag{5.9}$$

Remark 5.1. *The risk related to the future on the electricity spot can in principle be hedged dynamically on the electricity market. Instead, the risk related to the elementary HDD $(\bar{T} - T_t)^+$ cannot be hedged.*

This is a quadratic function with respect to $(c_{t,t+\Delta}^0, c_{t,t+\Delta}^1, c_{t,t+\Delta}^2)$ and the first order condition leads to:

$$
\begin{bmatrix}
1 & \mathbb{E}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] & \mathbb{E}[S_{t+\Delta}|\mathcal{F}_t] \\
\mathbb{E}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] & \mathbb{E}[((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t] & \mathbb{E}[S_{t+\Delta}(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] \\
\mathbb{E}[S_{t+\Delta}|\mathcal{F}_t] & \mathbb{E}[S_{t+\Delta}(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] & \mathbb{E}[S_{t+\Delta}^2|\mathcal{F}_t]
\end{bmatrix}
\begin{bmatrix}
c_{t,t+\Delta}^0 \\
c_{t,t+\Delta}^1 \\
c_{t,t+\Delta}^2
\end{bmatrix}
$$

$$(5.10)$$

$$
= \begin{bmatrix}
\mathbb{E}[S_{t+\Delta}(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] \\
\mathbb{E}[S_{t+\Delta}((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t] \\
\mathbb{E}[S_{t+\Delta}^2(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]
\end{bmatrix}
$$

Proposition 5.6. *Under Model* (ETM) *and for $\Delta > 0$, the vector $(c_{t,t+\Delta}^0, c_{t,t+\Delta}^1, c_{t,t+\Delta}^2)$ minimising the function* (5.9) *is the unique solution of the linear equation* (5.10), *whose components can be explicitly or semi-explicitly calculated.*

Proof. The fact that the linear quadratic problem (5.9) boils down to (5.10) is standard, one only has to check that the matrix on the left-hand side is invertible so that there is a unique solution. This matrix is a (conditional) covariance matrix: it is invertible, otherwise we could find $\mathcal{F}_t$-measurable coefficients $(c_{t,t+\Delta}^0, c_{t,t+\Delta}^1, c_{t,t+\Delta}^2)$ such that $c_{t,t+\Delta}^0 + c_{t,t+\Delta}^1(\bar{T} - T_{t+\Delta})^+ + c_{t,t+\Delta}^2 S_{t+\Delta} = 0$, which is clearly impossible from (1.2).

We now recall how to calculate explicitly or semi-explicitly: $\mathbb{E}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]$ is given by Proposition 6.6, $\mathbb{E}[((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$ by Proposition 6.8, $\mathbb{E}[S_{t+\Delta}|\mathcal{F}_t]$ and $\mathbb{E}[S_{t+\Delta}^2|\mathcal{F}_t]$ can be calculated with the characteristic function in Equation (3.5), $\mathbb{E}[S_{t+\Delta}(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]$ by Proposition 6.9 and $\mathbb{E}[S_{t+\Delta}^2(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]$ by Proposition 6.10. Finally we compute $\mathbb{E}[S_{t+\Delta}((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$ by using Proposition 6.11. □

Figures 2.15 show the results of the daily portfolio optimisation during 31 days starting the 1st January 2018. First, we can observe that c^0 and c^1 present a seasonality. This is explained by the integration of weekend days where there is small energy trade while the constant c^0 and c^1 are related to non-seasonal tools. This seasonality is not present in c^2 as the instrument hedged by the c^2 has the same seasonality as the output. Second, we can also comment on the signs of c^1 and c^2 which are both positive. This shows that both instruments are used to hedge the output $S_{t+\Delta}(\bar{T} - T_{t+\Delta})^+ - c_{t,t+\Delta}^0 - c_{t,t+\Delta}^1(\bar{T} - T_{t+\Delta})^+ - c_{t,t+\Delta}^2 S_{t+\Delta}$. The cash quantity c^0 accommodates to respond to the minimisation. Third, we can observe that for λ non-zero a small share of c^0 is reported to c^2 as we are leveraging the dependence structure between the energy and the temperature.

Figure 2.16 shows the empirical density of the portfolio only composed by $S_{t+\Delta}(\bar{T} - T_{t+\Delta})^+$ and including hedging for the month of January and May 2018. We compare the $100,000$ Monte
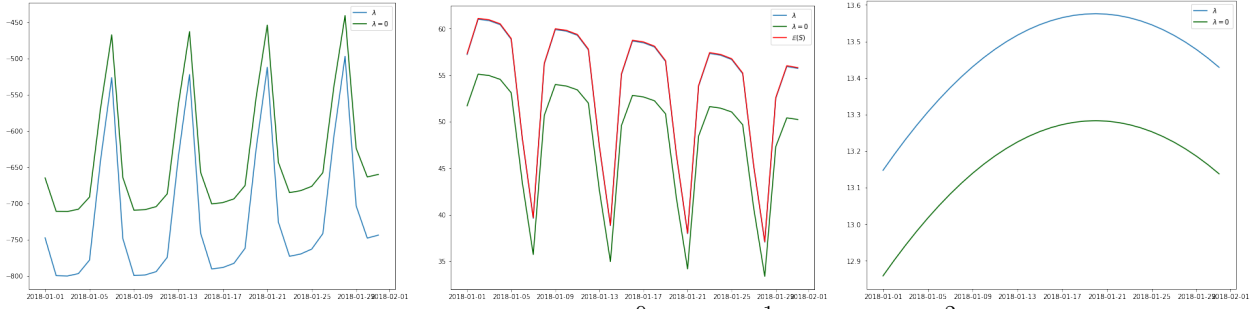
Figure 2.15: From top left to bottom right, $c^0_{t_0,t_1+i\Delta}$, $c^1_{t_0,t_1+i\Delta}$ and $c^2_{t_0,t_1+i\Delta}$ starting from 1st January 2018 (t_1) and with $t_0 = t_1 - 30$, $\Delta = 1$ and $i = 0, \ldots, 30$.

Carlo simulations of the portfolio with and without hedging. We can see that the hedging is efficient as the average of the hedge portfolio is 0 while the average of the portfolio without hedging is negative. In addition, the hedging strategy also decreases the variances of a portfolio payoff as the portfolio without hedging is clearly more spread than the one hedged. In the example of January (resp. May), the average payoff of the $\mathcal{E}$-HDD is $22,055$ (resp. $3,014$) and its standard deviation is $4,127$ (resp. $1,487$). In contrast, the average PnL of the static hedging portfolio is -0.754 (resp. -0.389 in May) and its standard deviation is 534 (resp. 276).

Furthermore, we analyse the impact of $\lambda < 0$ supposing a portfolio hedged by using the model with $\lambda = 0$. While the hedging of the portfolio still is effective the average PnL of the portfolio in January (resp. May) is -137.638 (resp. -67.734) with $\lambda = 0$ instead of -0.754 (resp. -0.389) with the correct value of λ, and the corresponding standard deviation is 597 (resp. 282) with $\lambda = 0$ instead of 534 (resp. 276) with the correct value of λ. This shows that λ has some influence on the quality of the hedge, and the portfolio hedging effectiveness when using options on temperature and energy as hedging instruments.

5.3.2 Risk decomposition of quantos

We replicate the above exercise for portfolios including quantos. We suppose that we are at time t and that we want to find the static portfolio made with single HDD, puts on electricity and cash that minimizes the square hedging error:

$$\mathbb{E}[((S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+ - d^0_{t,t+\Delta} - d^1_{t,t+\Delta}(\bar{T} - T_{t+\Delta})^+ - d^2_{t,t+\Delta}(S_{t+\Delta} - \bar{S})^+)^2|\mathcal{F}_t]. \quad (5.11)$$

This is again a quadratic function with respect to $(d^0_{t,t+\Delta}, d^1_{t,t+\Delta}, d^2_{t,t+\Delta})$ and the first order
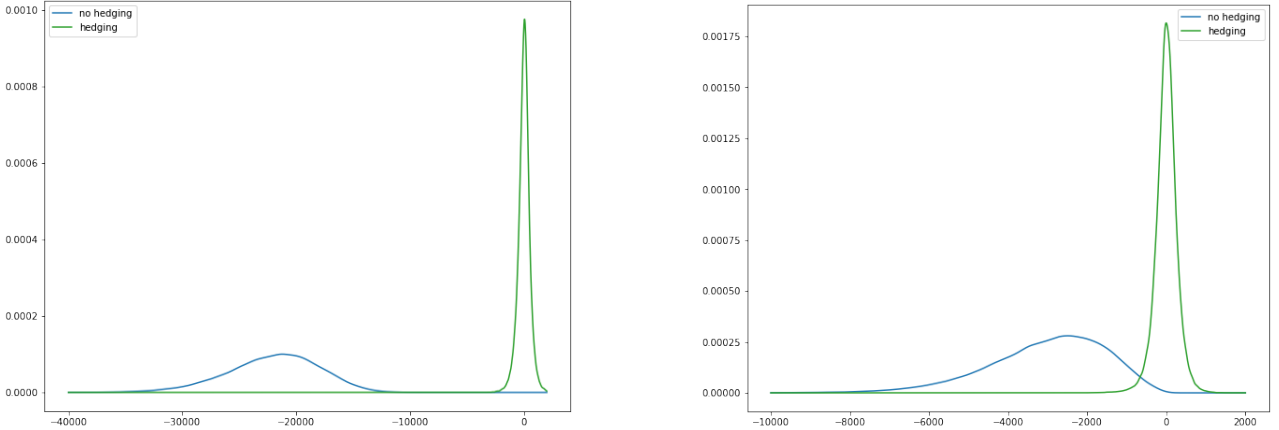
Figure 2.16: Empirical density of $\sum_{i=1}^{31} -S_{t+i\Delta}(\bar{T} - T_{t_1+i\Delta})^+$ (blue) and $\sum_{i=1}^{31} c^0_{t_0,t_1+(i-1)\Delta} + c^1_{t_0,t_1+(i-1)\Delta}(\bar{T} - T_{t_1+i\Delta})^+ + c^2_{t_0,t_1+(i-1)\Delta}S_t - S_{t_1+i\Delta}(\bar{T} - T_{t_1+i\Delta})^+$ (green) for portfolio optimisation starting on 1st January 2018 (for t_1 on the left) and 1st May 2018 (for t_1 on the right), with $t_0 = t_1 - 30$ and lasting the whole month.

condition leads to:

$$
\begin{bmatrix}
1 & \mathbb{E}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] & \mathbb{E}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t] \\
\mathbb{E}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] & \mathbb{E}[((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t] & \mathbb{E}[(S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] \\
\mathbb{E}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t] & \mathbb{E}[(S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] & \mathbb{E}[((S_{t+\Delta} - \bar{S})^+)^2|\mathcal{F}_t]
\end{bmatrix}
\begin{bmatrix}
d^0_{t,t+\Delta} \\
d^1_{t,t+\Delta} \\
d^2_{t,t+\Delta}
\end{bmatrix}
$$
$$
=
\begin{bmatrix}
\mathbb{E}[(S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] \\
\mathbb{E}[(S_{t+\Delta} - \bar{S})^+((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t] \\
\mathbb{E}[((S_{t+\Delta} - \bar{S})^+)^2(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]
\end{bmatrix}
\tag{5.12}
$$

Proposition 5.7. *Under Model* (ETM) *and for* $\Delta > 0$, *the vector* $(d^0_{t,t+\Delta}, d^1_{t,t+\Delta}, d^2_{t,t+\Delta})$ *minimising the quadratic criterion* (5.11) *is the unique solution of the linear system* (5.12). *The first order Taylor development when* $\lambda \to 0$ *of all components of this linear system can be explicitly or semi-explicitly calculated.*

Proof. The arguments assuring the existence of a unique minimizer are the same as in Proposition 5.6. All the terms above have already been implemented in the Subsection 5.2 except for $\mathbb{E}[((S_{t+\Delta} - \bar{S})^+)^2|\mathcal{F}_t]$, $\mathbb{E}[(S_{t+\Delta} - \bar{S})^+((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$ and $\mathbb{E}[((S_{t+\Delta} - \bar{S})^+)^2(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]$. The calculation of the first one can be made by using the Carr-Madan approach as presented in Proposition 6.3, while the Taylor developments of the two other terms are given respectively by Proposition 6.12 and Proposition 6.13. □

Figure 2.17 shows the coefficients $d^0_{t,t+\Delta}$, $d^1_{t,t+\Delta}$ and $d^2_{t,t+\Delta}$ evolution on the month of January 2018. We can observe that $d^0_{t,t+\Delta}$ and $d^1_{t,t+\Delta}$ are, as before, weekly seasonal. In addition, in this case $d^0_{t,t+\Delta}$ and $d^1_{t,t+\Delta}$ are very close for both λ nill and negative. For $d^2_{t,t+\Delta}$, we only get the weekly phenomenon when λ is negative.

Figure 2.18 shows the empirical density of the PnL of the portfolio for January and May 2018. In both cases we observe a significant hedging effect. The hedging effect is more important in January as the double condition on the options are hit more frequently. In May, the quanto does not claim so often and the PnL Monte Carlo simulations are closer to 0. In both cases, the PnL of the portfolio is reduced on average from $-3,358$ to 0.208 in January (from -89.174 to -0.113 in May) and on standard deviation from $2,197$ to 391 in January (from 177 to 98 in May).

We again compare the hedging obtained with the exact value of $\lambda < 0$ (coupled model) and the one obtained with $\lambda = 0$ (independent dynamics). In January (resp. May), the average PnL of the portfolio with λ is 0.208 (resp. -0.113) instead of -93.072 (resp. -12.283) for $\lambda = 0$ for January. The standard deviation with λ is 391 (resp. 99) instead of 394 (resp. 100) for $\lambda = 0$ respectively. This again shows that the parameter λ has some notable influence on the portfolio hedging: while λ is small, the hedging constructed with this value performs better than the one using $\lambda = 0$.
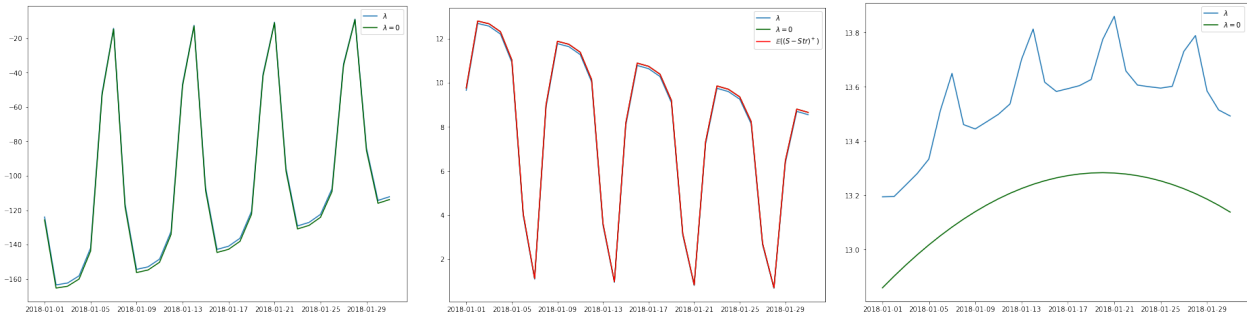

Figure 2.17: From top left to bottom right, $d^0_{t_0,t_1+i\Delta}$, $d^1_{t_0,t_1+i\Delta}$ and $d^2_{t_0,t_1+i\Delta}$ starting from 1st January 2018 (t_1), with $t_0 = t_1 - 30$ and with $t_0 = t_1 - 30$, $\Delta = 1$ and $i = 0, \ldots, 30$.

To sum up, the portfolio study above shows that one can effectively hedge $\mathcal{E}$-HDD and double-sided option quantos through single index-based derivatives. This single-index-based derivatives are traded in open market which ease accessibility and decrease operational costs. For energy derivatives, we can even consider these markets as liquid and suppose a perfect hedging of this risk component. Understanding this risk decomposition is key for risk managers and risk transfer businesses as it enables to gain comfort on the product and ensure meeting their own solvency constraints. Given this is a highly regulated economic sector, regulators are also concerned about this hedging capacity.

Finally, in this section we address the pricing of quanto derivatives on temperature and electricity. We explore different payoff functions and develop explicit pricing formulas for swaps, futures, single sided quanto options and double sided quanto options. These formulas are verified through Monte Carlo simulations. Finally we explore the possibility to statistically hedge single and double sided quanto options. We obtain an efficient daily risk decomposition of these derivatives leading to a averaged-simulated complete hedging of these derivatives. This
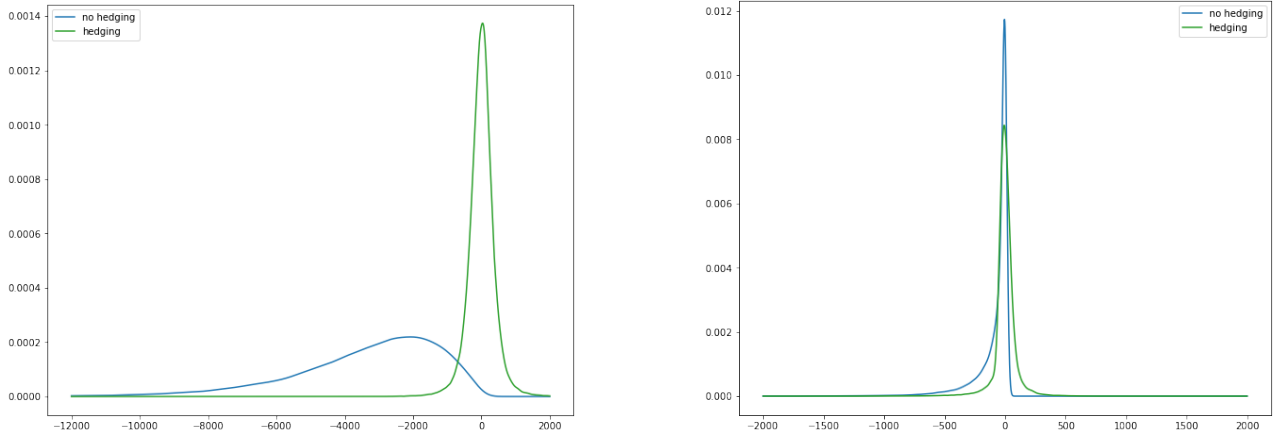
Figure 2.18: Empirical density of $\sum_{i=1}^{31} -(S_{t_1+i\Delta} - \bar{S})^+(\bar{T} - T_{t_1+i\Delta})^+$ (blue) and $\sum_{i=1}^{31} d^0_{t_0,t_1+(i-1)\Delta} + d^1_{t_0,t_1+(i-1)\Delta}(\bar{T} - T_{t_1+i\Delta})^+ + d^2_{t_0,t_1+(i-1)\Delta}(S_{t_1+i\Delta} - \bar{S})^+ - (S_{t_1+i\Delta} - \bar{S})^+(\bar{T} - T_{t_1+i\Delta})^+$ (green) for portfolio optimisation starting on 1st January 2018 (for t_1 on the left) and 1st May 2018 (for t_1 on the right), with $t_0 = t_1 - 30$ and lasting the whole month.

capacity to hedge these derivatives is key to confirm the efficiency and market viability of these products.

6 Appendices

6.1 The Normal Inverse Gaussian (NIG) distribution

This paragraph recalls the parametrization of NIG distributions and some elementary properties. The NIG distribution is a generalised hyperbolic distribution introduced by Barndorff-Nielsen [12]. Its density function is defined as follow:

$$f(x; \alpha, \beta, \delta, m) = \frac{\alpha \delta K_1 \left(\alpha \sqrt{\delta^2 + (x - m)^2} \right)}{\pi \sqrt{\delta^2 + (x - m)^2}} \, e^{\delta\gamma + \beta(x-m)}, x \in \mathbb{R}$$

where $m \in \mathbb{R}$ is the location of the density, $\beta \in \mathbb{R}$, $\alpha > |\beta|$, and $\delta \in \mathbb{R}$ the scale and K_1 denotes a modified Bessel function of the second kind. We denote by $NIG(\alpha, \beta, \delta, m)$ this law and set $\gamma = \sqrt{\alpha^2 - \beta^2} > 0$. The moments and the characteristic function are known explicitly: for $X \sim NIG(\alpha, \beta, \delta, m)$, we have

$$\mathbb{E}(X) = m + \frac{\delta\beta}{\gamma}, \qquad\qquad Var(X) = \frac{\delta\alpha^2}{\gamma^3}$$

$$Skewness(X) = \frac{3\beta}{\alpha\sqrt{\gamma\delta}}, \qquad\qquad Ex.Kurtosis(X) = 3\frac{1 + 4\beta^2/\alpha^2}{\delta\gamma}.$$

The characteristic function is given by

$$\mathbb{E}[e^{iuX}] = e^{ium + \delta\left(\gamma - \sqrt{\alpha^2 - (\beta + iu)^2}\right)}, \ u \in \mathbb{R}.$$

6.2 CLS estimator of the drift parameters of the log spot price process

We consider Model (ETM), and we want to estimate the mean-reversion parameter κ_X as well as the parameters defining the function $\mu_X(\cdot)$ given by (1.1).

The goal of this appendix is to compute the conditional least squares estimator of $(\kappa_X, \beta_0^X, \alpha_1^X, \beta_1^X, \alpha_0^{X,DoW}, \dots, \alpha_6^{X,DoW})$ and to prove the next proposition. We note $\alpha^{X,DoW} = (\alpha_0^{X,DoW}, \dots, \alpha_6^{X,DoW})$ and define, for $j \in \mathbb{N}$, $\alpha_j^{X,DoW} = \alpha_{\tilde{j}}^{X,DoW}$, with $\tilde{j} \in \{0, \dots 6\}$ such that $j = \tilde{j} \mod 7$.

Proposition 6.1. *Let $\Xi_{i\Delta} = (i\Delta, X_{i\Delta}, \sin(\xi i\Delta), \cos(\xi i\Delta), \left(\mathbb{1}_{\{DoW(i\Delta)=j\}}\right)_{j=0,\dots,6}) \in \mathbb{R}^4 \times \{0, 1\}^7$ for $i \in \mathbb{N}$ with $(X_t)_{t\geq 0}$ following the dynamics of (ETM) and $\Delta > 0$. We assume that $\sum_{i=0}^{N-1} \Xi_{i\Delta}\Xi_{i\Delta}^\top$ is invertible and define*

$$\hat{\eta} = (\hat{\eta}_1, \dots \hat{\eta}_{11})^\top = \left(\sum_{i=0}^{N-1} \Xi_{i\Delta}\Xi_{i\Delta}^\top \right)^{-1} \left(\sum_{i=0}^{N-1} \Xi_{i\Delta} X_{(i+1)\Delta} \right). \tag{6.1}$$

If $\hat{\eta}_2 \in (0,1) \cup (1,+\infty)$, the solution of the minimisation problem

$$\min_{\kappa_X, \beta_0^X, \alpha_1^X, \beta_1^X, \alpha^{X,DoW}} \sum_{i=0}^{N-1} \left(X_{(i+1)\Delta} - \mathbb{E}[X_{(i+1)\Delta}|X_{i\Delta}] \right)^2 \tag{6.2}$$

is given by

$$\begin{cases} \hat{\kappa}_X & = -\ln \hat{\eta}_2 \\ \hat{\beta}_0^X & = \frac{\hat{\eta}_1}{1-\hat{\eta}_2} \\ \hat{\alpha}_1^X & = \frac{\hat{\eta}_3(\cos(\xi)-e^{-\hat{\kappa}_X\Delta})+\hat{\eta}_4\sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}_X\Delta})^2+\sin^2(\xi\Delta)} \\ \hat{\beta}_1^X & = \frac{\hat{\eta}_4(\cos(\xi)-e^{-\hat{\kappa}\Delta})-\hat{\eta}_3\sin(\xi\Delta)}{(\cos(\xi\Delta)-e^{-\hat{\kappa}\Delta})^2+\sin^2(\xi\Delta)} \\ \hat{\alpha}_j^{X,DoW} & = \frac{1}{1-e^{-7\hat{\kappa}_X\Delta}} \sum_{k=0}^{6}(\hat{\eta}_{j+k}^{DoW}-\hat{\beta}_0)e^{-(6-k)\hat{\kappa}_X\Delta}, \end{cases}$$

with $(\hat{\eta}_0^{DoW},\ldots,\hat{\eta}_6^{DoW}) = (\hat{\eta}_5,\ldots,\hat{\eta}_{11})$ and $\hat{\eta}_j^{DoW} = \hat{\eta}_{\tilde{j}}^{DoW}$, with $\tilde{j} \in \{0,\ldots 6\}$ such that $j = \tilde{j}$ mod 7.

Proof. From (1.2), we get

$$\mathbb{E}[X_{t+\Delta}|\mathcal{F}_t] = X_t e^{-\kappa_X\Delta} + \mu_X(t+\Delta) - \mu_X(t)e^{-\kappa_X\Delta}$$

by using the martingale property of the stochastic integral and the fact that L^X is centered. We now use trigonometric identities to get

$$\begin{aligned} \mu_X(t+\Delta) - e^{-\kappa_X\Delta}\mu_X(t) =& \beta_0(t+\Delta) - \beta_0 e^{-\kappa_X\Delta}t + \alpha_1 \sin(\xi(t+\Delta)) - \alpha_1 e^{-\kappa_X\Delta}\sin(\xi t) \\ &+ \beta_1 \cos(\xi(t+\Delta)) - \beta_1 e^{-\kappa_X\Delta}\cos(\xi t) + \sum_{j=0}^{6}\alpha_j^{DoW}\mathbb{1}_{\{DoW(t+\Delta)=j\}} \\ &- \alpha_j^{DoW}e^{-\kappa_X\Delta}\mathbb{1}_{\{\lfloor DoW(t)=j\}} \\ =& \eta_1 t + \eta_3\sin(\xi t) + \eta_4\cos(\xi t) + \sum_{j=0}^{6}\eta_j^{DoW}\mathbb{1}_{\{\lfloor DoW(t)=j\}}, \end{aligned}$$

with

$$\begin{cases} \eta_1 & = \beta_0(1-e^{-\kappa_X\Delta}) \\ \eta_2 & = e^{-\kappa_X\Delta} \\ \eta_3 & = \alpha_1(\cos(\xi\Delta)-e^{-\kappa_X\Delta}) - \beta_1\sin(\xi\Delta) \\ \eta_4 & = \alpha_1\sin(\xi\Delta) + \beta_1(\cos(\xi\Delta)-e^{-\kappa_X\Delta}) \\ \eta_j^{DoW} & = \alpha_{j+1}^{DoW} - \alpha_j^{DoW}e^{-\kappa_X\Delta} + \beta_0\Delta \text{ with convention } \alpha_7^{DoW} = \alpha_0^{DoW}, \end{cases}$$

where η_2 is set to have $\mathbb{E}[X_{(i+1)\Delta}|\mathcal{F}_{i\Delta}] = \eta^\top \Xi_{i\Delta}$, i.e. is the regression coefficient with respect to $X_{i\Delta}$. The minimization problem (6.2) is then equivalent to

$$\min_{\eta \in \mathbb{R}^{11}} \sum_{i=0}^{N-1} \left(X_{(i+1)\Delta} - \eta^\top \Xi_{i\Delta}\right)^2.$$

This corresponds to a linear regression, whose solution is given by (6.1). When $\eta_2 \in (0, 1)$, the system can be inverted, and the claim follows easily. □

Let us note here that $\hat{\eta}^\top \Xi_{i\Delta}$ can then be seen as the estimation of $\mathbb{E}[X_{(i+1)\Delta}|X_{i\Delta}]$.

6.3 Simulation of Model (ETM) and associated characteristic function

Let recall (1.2):

$$\begin{cases} X_{t+\Delta} - \mu_X(t+\Delta) &= e^{-\kappa_X \Delta}(X_t - \mu_X(t)) + \lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dW_v^T + \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dL_v^X \\ T_{t+\Delta} - \mu_T(t+\Delta) &= e^{-\kappa_T \Delta}(T_t - \mu_T(t)) + \sigma_T \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)}dW_v^T. \end{cases}$$

The simulation algorithm is the following:

1. Simulate $N_1 \sim \mathcal{N}(0, 1)$ and $N_2 \sim \mathcal{N}(0, 1)$.

2. Simulate $Z^X \sim NIG(\alpha^X, \beta^X, \delta^X, -\frac{\delta^X \beta^X}{\gamma^X})$ (we work with centered NIG distributions).

3. Simulate $(\tilde{X}_{t+\Delta}, \tilde{T}_{t+\Delta})$ given $(\tilde{X}_t, \tilde{T}_t)$ using the below scheme

$$\begin{cases} X_{t+\Delta} &= \mu_X(t+\Delta) + e^{-\kappa_X \Delta}(X_t - \mu_X(t)) + \lambda\sigma_T \sqrt{\frac{1 - e^{-2\kappa_X \Delta}}{2\kappa_X}}N_1 + e^{-\kappa_X \Delta/2}Z^X \\ T_{t+\Delta} &= \mu_T(t+\Delta) + e^{-\kappa_T \Delta}(T_t - \mu_T(t)) + \sigma_T \sqrt{\frac{1 - e^{-2\kappa_T \Delta}}{2\kappa_T}}(\rho N_1 + \sqrt{1 - \rho^2}N_2), \end{cases}$$

with ρ defined as in Proposition 6.2.

Note that this is the exact scheme for T, and the only discretization error on X comes from the approximation of $\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dL_v^X$. When comparing the pricing by Monte-Carlo with the formulas using the Fourier transform as in Section 5, it is then worth to use the characteristic

function associated to this discretization scheme. This avoids to have a bias between both methods. Namely, we use for $t_0 < t$ such that $t - t_0$ is a multiple of the discretization step Δ

$$\hat{\psi}_{X_t}(u; t_0 - t) = e^{iu(\mu_X(t) + e^{-\kappa_X(t-t_0)}(X_t - \mu_X(t)) - \frac{1}{2}\lambda^2 \sigma_T^2 \frac{1 - e^{-2\kappa_X(t-t_0)}}{2\kappa_X} u^2} \times$$

$$\exp\left(\Delta \sum_{\ell=0}^{\frac{t-t_0}{\Delta} - 1}\left(ium^X e^{-\kappa_X(\ell+1/2)\Delta} + \delta^X \gamma^X - \delta^X \sqrt{(\alpha^X)^2 - (\beta^X + iue^{-\kappa_X(\ell+1/2)\Delta})^2}\right)\right).$$

instead of (4.1).

6.4 Proofs of the results of Section 5

6.4.1 Results on the dependence between X and T

From (1.2), we are interested in the law of $(\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dW_v^T, \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)} dW_v^T)$ that captures the dependence between $X_{t+\Delta}$ and $T_{t+\Delta}$ given X_t and T_t.

Proposition 6.2. *The random vector* $\begin{pmatrix} \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dW_v^T \\ \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)} dW_v^T \end{pmatrix}$ *is a centered Gaussian vector with covariance matrix*

$$(\Delta) := \begin{bmatrix} k_X^2(\Delta) & k_{XT}^2(\Delta) \\ k_{XT}^2(\Delta) & k_T^2(\Delta) \end{bmatrix} = \begin{bmatrix} \frac{1-e^{-2\kappa_X\Delta}}{2\kappa_X} & \frac{1-e^{-(\kappa_X+\kappa_T)\Delta}}{\kappa_X+\kappa_T} \\ \frac{1-e^{-(\kappa_X+\kappa_T)\Delta}}{\kappa_X+\kappa_T} & \frac{1-e^{-2\kappa_T\Delta}}{2\kappa_T} \end{bmatrix}.$$

It has the same law as

$$\begin{pmatrix} k_X(\Delta)\varrho \\ k_T(\Delta) \end{pmatrix} G + \begin{pmatrix} k_X(\Delta)\sqrt{1-\varrho^2} \\ 0 \end{pmatrix} G^{\perp}$$

where G and $G^{\perp} \sim \mathcal{N}(0,1)$ are independent and $\varrho = \frac{k_{XT}^2(\Delta)}{k_X(\Delta)k_T(\Delta)} \in [0,1]$.

Proof. The Brownian motion is a Gaussian process, which gives the Gaussian property. We use then the Itô isometry to get the covariance matrix.

$\square$

From Proposition 6.2, we can quickly get the following corollary.

Corollary 6.1. *Conditionally on $(\int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)} dW_v^T)$, $(\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dW_v^T)$ follows a Gaussian distribution with mean*

$$\mathbb{E}\left(\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dW_v^T \bigg| \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)} dW_v^T\right)$$

$$= \frac{2\kappa_T}{1-e^{-2\kappa_T\Delta}} \frac{1-e^{-(\kappa_X+\kappa_T)\Delta}}{\kappa_X+\kappa_T} \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)} dW_v^T$$

and variance

$$\frac{1 - e^{-2\kappa_X\Delta}}{2\kappa_X} - \left(\frac{1 - e^{-(\kappa_X+\kappa_T)\Delta}}{\kappa_X + \kappa_T}\right)^2 \frac{2\kappa_T}{(1 - e^{-2\kappa_T\Delta})}.$$

6.4.2 Identities on the normal distribution

We note Φ the cumulative distribution function of the normal distribution $\mathcal{N}(0, 1)$.

Lemma 6.1. *Let $G \sim \mathcal{N}(0, 1)$, $a \in \mathbb{R}$, $b > 0$. We have*

$$\mathbb{E}[(a + bG)^+] = a\Phi(a/b) + \frac{b}{\sqrt{2\pi}}e^{-a^2/(2b^2)}.$$

Proof. We have $\mathbb{E}[(a + bG)^+] = \int_{-a/b}^{\infty}(a + bx)\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx = a\Phi(a/b) + \frac{b}{\sqrt{2\pi}}e^{-a^2/(2b^2)}$. □

Lemma 6.2. *Let $a \in \mathbb{R}$, $b \in \mathbb{R}$ and $G \sim \mathcal{N}(0, 1)$. We have*

$$\mathbb{E}[bG(a + bG)^+] = b^2\Phi(a/|b|)$$

Proof. It is sufficient to prove the result for $b > 0$ since $bG \overset{law}{=} |b|G$. Let $d = a/b$. We have $\mathbb{E}[bG(a + bG)^+] = b^2\mathbb{E}[G(d + G)^+]$ and

$$\mathbb{E}[G(d + G)^+] = \int_{-d}^{\infty} x(d + x)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx = \Phi(d). \qquad \square$$

Lemma 6.3. *Let $a \in \mathbb{R}$, $b > 0$ and $G \sim \mathcal{N}(0, 1)$. We have*

$$\mathbb{E}[((a + bG)^+)^2] = (a^2 + b^2)\Phi(a/b) + \frac{ab}{\sqrt{2\pi}}e^{-\frac{a^2}{2b^2}}.$$

Proof. Let $d = a/b$. We have $\mathbb{E}[((a + bG)^+)^2] = b^2\mathbb{E}[((d + G)^+)^2]$ and

$$\mathbb{E}[((d + G)^+)^2] = \int_{-d}^{\infty}(d^2 + 2dx + x^2)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx = d^2\Phi(d) + \frac{2d}{\sqrt{2\pi}}e^{-\frac{d^2}{2}} - \frac{d}{\sqrt{2\pi}}e^{-\frac{d^2}{2}} + \Phi(d). \quad \square$$

Lemma 6.4. *Let $a \in \mathbb{R}$, $b \in \mathbb{R}$ and $G \sim \mathcal{N}(0, 1)$. We have*

$$\mathbb{E}[bG((a + bG)^+)^2] = |b|^3\left(\sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}(\frac{a}{|b|})^2} + 2\frac{a}{|b|}\Phi(\frac{a}{|b|})\right)$$

Proof. Since bG has the same law as $|b|G$, it is sufficient to consider the case $b > 0$. We have $\mathbb{E}[bG((a + bG)^+)^2] = b^3\mathbb{E}[G((d + G)^+)^2]$ with $d = \frac{a}{b}$. We then get by integration by parts and then Lemma 6.1

$$\mathbb{E}[G((d + G)^+)^2] = \int_{-d}^{\infty}(d + x)^2 x\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx = 2\int_{-d}^{\infty}(d + x)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}dx$$

$$= 2\mathbb{E}[(d + G)^+] = \sqrt{\frac{2}{\pi}}e^{-\frac{d^2}{2}} + 2d\Phi(d). \qquad \square$$

6.4.3 Computations with Fourier transform

In the following section, we will develop some conditional expectations for different derivatives. However, some derivatives do not admit explicit formulas and were computed through inverse Fourier methods.

We first use Carr Madan formula [48, Equations (5) and (6)] to compute $\mathbb{E}((S_{t+\Delta} - \bar{S})^+ \mid \mathcal{F}_t)$ for $t \geq 0, \Delta > 0$:

$$\mathbb{E}((S_{t+\Delta} - \bar{S})^+ \mid \mathcal{F}_t) = \frac{\exp(-\alpha \bar{X})}{\pi} \int_0^\infty e^{-i\bar{X}v} \frac{\psi_{X_t}(v - (\alpha + 1)i)}{\alpha^2 + \alpha - v^2 + i(2\alpha + 1)v} dv, \tag{6.3}$$

where ψ corresponds to the characteristic function as in Equation (4.1), $\alpha > 0$ and $\bar{X} := \ln(\bar{S})$. Note that from (4.1), we have $\mathbb{E}[S_{t+\Delta}^{1+\alpha} \mid \mathcal{F}_t] < \infty$ a.s. and is equal to $\psi_{X_t}(-(\alpha + 1)i; \Delta)$. In practice, we take $\alpha = 0.5$ for (6.3) as well as for Proposition 6.3 below.

Second, we apply Gil-Pelaez [78] inversion formula to compute $\mathbb{P}\left(S_{t+\Delta} \geq \bar{S} \mid \mathcal{F}_t\right)$.

$$\mathbb{P}_{\lambda=0}\left(S_{t+\Delta} \geq \bar{S} \mid \mathcal{F}_t\right) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathcal{R}\left(\frac{e^{-iv\bar{X}} \psi_{X_t}(v)}{iv}\right) dv \tag{6.4}$$

where $\mathcal{R}$ denotes the real part, ψ the characteristic function as in Equation (4.1) and $\bar{X} := \ln(\bar{S})$.

Third, we leverage again Carr Madan [48] approach to compute $\mathbb{E}(((S_{t+\Delta} - \bar{S})^+)^2 \mid \mathcal{F}_t)$.

Proposition 6.3. *Under Model* (ETM), *we have*

$$\mathbb{E}(((S_{t+\Delta} - \bar{S})^+)^2 \mid \mathcal{F}_t)$$
$$= \frac{\exp(-\alpha \bar{X})}{\pi} \int_0^\infty e^{-i\bar{X}v} \psi(v - (\alpha + 2)i; \Delta)\left(\frac{1}{\alpha + iv} - 2\frac{1}{\alpha + 1 + iv} + \frac{1}{\alpha + 2 + iv}\right) dv \tag{6.5}$$

where ψ corresponds is the characteristic function as in Equation (4.1), $\alpha > 0$ and $\bar{X} := \ln(\bar{S})$.

Proof. We prove the result for a random variable Y with distribution μ on $\mathbb{R}$ such that $\mathbb{E}[e^{(2+\alpha)Y}] < \infty$. We define

$$C(k) = \mathbb{E}\left[\left((e^Y - e^k)^+\right)^2\right] = \int_k^\infty (e^y - e^k)^2 \mu(dy)$$

and $c(k) = e^{\alpha k} C(k)$. The function c is nonnegative and integrable on $\mathbb{R}$ since

$$\int_{\mathbb{R}} c(k) dk \leq \int_{\mathbb{R}} e^{\alpha k} \mathbb{E}[e^{2Y} \mathbb{1}_{Y > k}] dk = \frac{\mathbb{E}[e^{(2+\alpha)Y}]}{\alpha} < \infty,$$

by Fubini's theorem. Following Carr Madan [48], as c is integrable, we can define its inverse Fourier transform $\tilde{\psi} : \mathbb{C} \longrightarrow \mathbb{C}$ such that:

$$
\begin{aligned}
\tilde{\psi}(v) &= \int_{-\infty}^{\infty} e^{ivk} c(k) dk \\
&= \int_{-\infty}^{\infty} e^{ivk} \int_{k}^{\infty} e^{-\alpha k} (e^k - e^y)^2 \mu(dy) dk \\
&= \int_{-\infty}^{\infty} \Big(\frac{1}{\alpha + iv} - 2 \frac{1}{\alpha + 1 + iv} + \frac{1}{\alpha + 2 + iv} \Big) e^{(\alpha + 2 + iv)y} \mu(dy) \\
&= \psi(v - (\alpha + 2)i) \Big(\frac{1}{\alpha + iv} - 2 \frac{1}{\alpha + 1 + iv} + \frac{1}{\alpha + 2 + iv} \Big),
\end{aligned}
$$

by using Fubini's theorem. We have $|\tilde{\psi}(v)| \leq \mathbb{E}[e^{(2+\alpha)Y}] \frac{2}{|\alpha+iv||\alpha+1+iv||\alpha+2+iv|}$, and thus $\tilde{\psi}(v)$ is integrable on $\mathbb{R}$ and bounded. We get then the claim by Fourier inversion and using that $\mathbb{E}[S_{t+\Delta}^{2+\alpha}|\mathcal{F}_t] = \psi_{X_t}(-(\alpha+2)i) < \infty$ a.s. $\qquad \square$

6.4.4 Results to calculate the average payoffs of derivatives

Futures

Proposition 6.4. *Under Model* (ETM)*, we have*

$$
\mathbb{E}(S_{t+\Delta} T_{t+\Delta} \mid \mathcal{F}_t) = \exp\left(\mu_X(t + \Delta) + e^{-\kappa_X \Delta}(X_t - \mu_X(t)) \right) \varphi(-i; \Delta) \times
$$

$$
\left((\mu_T(t + \Delta) + e^{-\kappa_T \Delta}(T_t - \mu_T(t))) e^{\frac{1}{2} k_X(\Delta)^2 \lambda^2 \sigma_T^2} + \lambda \sigma_T^2 k_{XT}^2(\Delta) e^{\frac{1}{2} \lambda^2 \sigma_T^2 k_X(\Delta)^2} \right)
$$

where φ is the characteristic function defined in Equation (3.5) *and $k_T(\cdot)$, $k_X(\cdot)$ and $k_{XT}(\cdot)$ are as in Equation* (5.3)*.*

Proof. From (1.2), we get

$$
\begin{aligned}
\mathbb{E}(e^{X_{t+\Delta}} T_{t+\Delta} \mid \mathcal{F}_t) &= \mathbb{E}\Bigg(\exp\Bigg(\mu_X(t + \Delta) + e^{-\kappa_X \Delta}(X_{t+\Delta} - \mu_X(t + \Delta)) + \lambda \sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T \\
&\quad + \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dL_u^X \Bigg) T_{t+\Delta} \Bigg| \mathcal{F}_t \Bigg) \\
&= \exp\Bigg(\mu_X(t + \Delta) + e^{-\kappa_X \Delta}(X_t - \mu_X(t)) \Bigg) \mathbb{E}(e^{\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dL_u^X}) \times \\
&\quad \mathbb{E}\Bigg(e^{\lambda \sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T} T_{t+\Delta} \Bigg| T_t \Bigg),
\end{aligned}
$$

since $\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dL_u^X$ is independent of $\mathcal{F}_t$ and T_t. The first term is deterministic, the second term is equal to $\varphi(-i; \Delta)$ by (3.4). We use (1.2) to write as follows the third term:

$$\mathbb{E}\left(e^{\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T} T_{t+\Delta} \,\Big|\, T_t \right) = (\mu_T(t+\Delta) + e^{-\kappa_T\Delta}(T_t - \mu_T(t)))\mathbb{E}\left(e^{\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T} \right)$$

$$+ \sigma_T \mathbb{E}\left(e^{\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T} \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-u)} dW_u^T \right).$$

From Proposition 6.2, we get $\mathbb{E}\left(e^{\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T} \right) = e^{\frac{1}{2}k_X(\Delta)^2\lambda^2\sigma_T^2}$ and

$$\mathbb{E}\left(e^{\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T} \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-u)} dW_u^T \right)$$

$$= \mathbb{E}\left(e^{\lambda\sigma_T\left(\frac{k_{XT}^2(\Delta)}{k_T(\Delta)}G + \frac{\sqrt{k_X(\Delta)^2 k_T(\Delta)^2 - k_{XT}^4(\Delta)}}{k_T(\Delta)}G^\perp\right)} k_T(\Delta)G \right)$$

$$= \mathbb{E}\left(e^{\lambda\sigma_T \frac{\sqrt{k_X(\Delta)^2 k_T(\Delta)^2 - k_{XT}^4(\Delta)}}{k_T(\Delta)}G^\perp} \right) k_T(\Delta)\mathbb{E}\left(Ge^{\lambda\sigma_T \frac{k_{XT}^2(\Delta)}{k_T(\Delta)}G} \right).$$

Since $\mathbb{E}[e^{xG}] = e^{x^2/2}$ and $\mathbb{E}[Ge^{xG}] = xe^{x^2/2}$, we get the claim. $\qquad\square$

Swaps

Proposition 6.5. *Under Model* (ETM) *and for* $\Delta > 0$ *, we have*

$$\mathbb{E}((S_{t+\Delta} - \bar{S})(\bar{T} - T_{t+\Delta}) \mid \mathcal{F}_t) = \bar{T}\psi_{X_t}(-i; \Delta) - \mathcal{F}(t, t+\Delta) - \bar{S}\bar{T} + \bar{S}(\mu_T(t) + e^{-\kappa_T\Delta}(T_t - \mu_T(t)))$$

where ψ *is the characteristic function defined in Equation* (4.1), $\mathcal{F}(t, t+\Delta)$ *as in Equation* (5.2) *and* $k_X(\cdot)$ *is as in Equation* (5.3).

Proof. Let first develop the formula:

$$\mathbb{E}((S_{t+\Delta} - \bar{S})(\bar{T} - T_{t+\Delta}) \mid \mathcal{F}_t) = \bar{T}\mathbb{E}(S_{t+\Delta} \mid \mathcal{F}_t) - \mathbb{E}(S_{t+\Delta}T_{t+\Delta} \mid \mathcal{F}_t) - \bar{S}\bar{T} + \bar{S}\mathbb{E}(T_{t+\Delta} \mid \mathcal{F}_t)$$

We have $\mathbb{E}(S_{t+\Delta}T_{t+\Delta} \mid \mathcal{F}_t)$ from Equation (5.2) and Proposition 6.4. The first term is equal to:

$$\bar{T}\mathbb{E}(S_{t+\Delta} \mid \mathcal{F}_t) = \bar{T}\psi_{X_t}(-i; \Delta)$$

and, by using the independence between L^X and W^T, we get the forth term:

$$\bar{S}E(T_{t+\Delta} \mid \mathcal{F}_t) = \bar{S}(\mu_T(t+\Delta) + e^{-\kappa_T\Delta}(T_t - \mu_T(t)))\square$$

Quantos

Proposition 6.6. *Under Model* (ETM) *and for* $\Delta > 0$ *, we have*

$$\mathbb{E}[(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t] = \left(\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t)\right) \times$$
$$\Phi\left(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)}\right)$$
$$+ \frac{\sigma_T k_T(\Delta)}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)}\right)^2\right)$$

Proof. Under Model (ETM), the distribution of $\bar{T} - T_{t+\Delta}$ given $\mathcal{F}_t$ is $\mathcal{N}(\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t)), \sigma_T^2 k_T(\Delta)^2)$. We then apply Lemma 6.1 to obtain the result. $\square$

Proposition 6.7. *Under Model* (ETM) *and for* $\Delta > 0$, *we can write the following Taylor expansion on* λ:

$$\mathbb{E}((S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t) = \mathbb{E}_{\lambda=0}((S_{t+\Delta} - \bar{S})^+ \mid \mathcal{F}_t) \times$$
$$\left(\left(\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t))\right) \times\right.$$
$$\Phi\left(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)}\right)$$
$$+ \frac{\sigma_T k_T(\Delta)}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)}\right)^2\right)\right)$$
$$- \left(\mathbb{E}_{\lambda=0}((S_{t+\Delta} - \bar{S})^+ \mid \mathcal{F}_t) + \bar{S}\mathbb{P}_{\lambda=0}\left(S_{t+\Delta} \geq \bar{S} \mid \mathcal{F}_t\right)\right) \times$$
$$\sigma_T^2 k_{XT}(\Delta)^2 \Phi\left(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)}\right)\right)\lambda + o(\lambda)$$

where $k_T(\cdot)$ *,* $k_X(\cdot)$ *and* $k_{XT}(\cdot)$ *are as in Equation* (5.3).

Proof. Under Model (ETM) and for $\Delta > 0$, we want to compute the first order Taylor expansion of $\mathbb{E}((S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t)$. For $\lambda = 0$,

$$\mathbb{E}((S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t) = \mathbb{E}((S_{t+\Delta} - \bar{S})^+ \mid \mathcal{F}_t)\mathbb{E}((\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t)$$

We have $\mathbb{E}((\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t)$ from Proposition 6.6 and $\mathbb{E}((S_{t+\Delta} - \bar{S})^+ \mid \mathcal{F}_t)$ from Equation (6.3).

Let now consider the derivative of $\mathbb{E}((S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t)$ in $\lambda = 0$:

$$\left.\frac{d}{d\lambda}\right|_{\lambda=0} \mathbb{E}((S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t)$$
$$= \mathbb{E}_{\lambda=0}\left(\mathbb{1}_{S_{t+\Delta} \geq \bar{S}}\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-s)}dW_s e^{X_{t+\Delta}}(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t\right)$$
$$= \mathbb{E}_{\lambda=0}\left(\mathbb{1}_{S_{t+\Delta} \geq \bar{S}}e^{X_{t+\Delta}} \mid \mathcal{F}_t\right) \mathbb{E}_{\lambda=0}\left(\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-s)}dW_s(\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t\right)$$

Let consider the first term,

$$\mathbb{E}_{\lambda=0} \left(\mathbb{1}_{S_{t+\Delta} \geq \bar{S}} S_{t+\Delta} \mid \mathcal{F}_t \right) = \mathbb{E}_{\lambda=0}((S_{t+\Delta} - \bar{S})^+ \mid \mathcal{F}_t) + \bar{S} \mathbb{P}_{\lambda=0} \left(S_{t+\Delta} \geq \bar{S} \mid \mathcal{F}_t \right)$$

The first element is computed with Equation (6.3) and the second with Equation (6.4).

We can now develop the second term by using Proposition 6.2 and (1.2):

$$\mathbb{E} \left(\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-s)} dW_s^T (\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t \right)$$

$$= \frac{k_{XT}(\Delta)^2}{k_T(\Delta)^2} \mathbb{E} \left(\sigma_T \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-s)} dW_s^T (\bar{T} - T_{t+\Delta})^+ \mid \mathcal{F}_t \right)$$

$$= \frac{k_{XT}(\Delta)^2}{k_T(\Delta)^2} \mathbb{E} \Bigg(\sigma_T \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-s)} dW_s^T$$

$$\times \left(\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t)) - \sigma_T \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-s)} dW_s^T \right)^+ \Bigg| \mathcal{F}_t \Bigg).$$

$$(6.6)$$

Now we apply Lemma 6.2 with $a = \bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t))$ and $b = -\sigma_T k_T(\Delta)$ to get

$$\mathbb{E} \left(\sigma_T \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-s)} dW_s^T (\bar{T} - T_{t+\Delta})^+ \Big| \mathcal{F}_t \right)$$

$$= -\sigma_T^2 k_T(\Delta)^2 \Phi \left(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)} \right)$$

$$(6.7)$$

$$\square$$

6.4.5 Results for static hedging portfolios

Results for $\mathcal{E}$-HDD

Proposition 6.8. *Under Model* (ETM), *we have*

$$\mathbb{E}[((\bar{T} - T_{t+\Delta})^+)^2 \mid \mathcal{F}_t] = \Big((\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t)))^2 + \sigma_T^2 k_T(\Delta)^2 \Big) \times$$

$$\Phi \Big(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)} \Big)$$

$$+ \frac{(\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t)))\sigma_T k_T(\Delta)}{\sqrt{2\pi}} \times$$

$$\exp \Big(-\frac{1}{2} \Big(\frac{\bar{T} - \mu_T(t + \Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)} \Big)^2 \Big).$$

Proof. Since $\bar{T} - T_{t+\Delta}$ follows a Gaussian distribution given $\mathcal{F}_t$, we have an explicit formula by using Lemma 6.3. $\square$

Proposition 6.9. *Under Model* (ETM), *we have*

$$\mathbb{E}(e^{X_{t+\Delta}}\mathbb{1}_{T_{t+\Delta}\leq u} \mid \mathcal{F}_t) = \psi_{X_t}(-i;\Delta)\Phi\Big(\frac{\tilde{u}(T_t)}{k_T(\Delta)} - \lambda\sigma_T\frac{k_{XT}^2(\Delta)}{k_T(\Delta)}\Big).$$

$$\mathbb{E}(e^{X_{t+\Delta}}\mathbb{1}_{u\leq T_{t+\Delta}} \mid \mathcal{F}_t) = \psi_{X_t}(-i;\Delta)\Phi\Big(\lambda\sigma_T\frac{k_{XT}^2(\Delta)}{k_T(\Delta)} - \frac{\tilde{u}(T_t)}{k_T(\Delta)}\Big).$$

where ψ is the characteristic function defined in Equation (4.1), *$k_T(\cdot)$, $k_X(\cdot)$ and $k_{XT}(\cdot)$ are as in Equation* (5.3) *and $\tilde{u}(T_t) = \frac{u-(\mu_T(t+\Delta)+e^{-\kappa_T\Delta}(T_t-\mu_T(t)))}{\sigma_T}$.*

Proof. From (1.2), we can write

$$\mathbb{E}(e^{X_{t+\Delta}}\mathbb{1}_{T_{t+\Delta}\leq u} \mid \mathcal{F}_t) = \exp\left(\mu_X(t) + e^{-\kappa_X\Delta}(X_t - \mu_X(t))\right)\mathbb{E}(e^{\int_t^{t+\Delta}e^{-\kappa_X(t+\Delta-u)}dL_u^X})\times$$

$$\mathbb{E}\left(e^{\lambda\sigma_T\int_t^{t+\Delta}e^{-\kappa_X(t+\Delta-u)}dW_u^T}\mathbb{1}_{T_{t+\Delta}\leq u}\Big|\mathcal{F}_t\right)$$

The second term is $\varphi(-i;\Delta)$, see (3.4). We now consider the last term:

$$\mathbb{E}\left(e^{\lambda\sigma_T\int_t^{t+\Delta}e^{-\kappa_X(t+\Delta-s)}dW_s^T}\mathbb{1}_{T_{t+\Delta}\leq u}\Big|\mathcal{F}_t\right) = \mathbb{E}\left(e^{\lambda\sigma_T\int_t^{t+\Delta}e^{-\kappa_X(t+\Delta-s)}dW_s^T}\mathbb{1}_{\int_t^{t+\Delta}e^{-\kappa_X(t+\Delta-s)}dW_s^T\leq\tilde{u}(T_t)}\Big|\mathcal{F}_t\right)$$

where $\tilde{u}(T_t) = \frac{u-(\mu_T(t+\Delta)+e^{-\kappa_T\Delta}(T_t-\mu_T(t)))}{\sigma_T}$. For $\tilde{u}\in\mathbb{R}$ and $G,G^\perp\sim\mathcal{N}(0,1)$ independent, we have:

$$\mathbb{E}\left(e^{\lambda\sigma_T(\frac{k_{XT}^2(\Delta)}{k_T(\Delta)}G+\frac{\sqrt{k_X(\Delta)^2k_T(\Delta)^2-k_{XT}^2(\Delta)}}{k_T(\Delta)}G^\perp)}\mathbb{1}_{k_T(\Delta)G\leq\tilde{u}}\right) = e^{\frac{\lambda^2\sigma_T^2}{2}\frac{k_X(\Delta)^2k_T(\Delta)^2-k_{XT}^4(\Delta)}{k_T^2(\Delta)}}\times$$

$$e^{\frac{(\lambda\sigma_T)^2}{2}(\frac{k_{XT}^2(\Delta)}{k_T(\Delta)})^2}\Phi\Big(\frac{\tilde{u}}{k_T(\Delta)} - \lambda\sigma_T\frac{k_{XT}^2(\Delta)}{k_T(\Delta)}\Big),$$

because $\mathbb{E}[e^{xG}\mathbb{1}_{G\leq a}] = e^{\frac{x^2}{2}}\Phi(a-x)$ for $x,a\in\mathbb{R}$. Since $\tilde{u}(T_t)$ is $\mathcal{F}_t$-measurable and the variables $\int_t^{t+\Delta}e^{-\kappa_X(t+\Delta-u)}dW_u^T$ and $\int_t^{t+\Delta}e^{-\kappa_T(t+\Delta-u)}dW_u^T$ are independent of $\mathcal{F}_t$, we get the claim by applying Proposition 6.2. $\qquad\square$

Proposition 6.10. *Under Model* (ETM), *we have*

$$\mathbb{E}(e^{2X_{t+\Delta}}\mathbb{1}_{T_{t+\Delta}\leq u} \mid \mathcal{F}_t) = \exp\left(2\mu_X(t+\Delta) + 2e^{-\kappa_X\Delta}(X_t - \mu_X(t))\right)\times$$

$$\exp\left(2m^X\frac{1-e^{-\kappa_X\Delta}}{\kappa_X} + \delta^X\gamma^X\Delta - \delta^X\int_0^\Delta\sqrt{(\alpha^X)^2-(\beta^X+2e^{-\kappa_X(\Delta-v)})^2}dv\right)\times$$

$$e^{4\frac{\lambda^2\sigma_T^2}{2}k_X(\Delta)^2}\Phi\Big(\frac{\tilde{u}(T_t)}{k_T(\Delta)} - 2\lambda\sigma_T\frac{k_{XT}^2(\Delta)}{k_T(\Delta)}\Big).$$

where $k_T(\cdot)$, $k_X(\cdot)$ and $k_{XT}(\cdot)$ are as in Equation (5.3) and $\tilde{u}(T_t) = \frac{u - (\mu_T(t+\Delta) + e^{-\kappa_T \Delta}(T_t - \mu_T(t)))}{\sigma_T}$.

Proof. From (1.2), we can write

$$\mathbb{E}(e^{2X_{t+\Delta}} \mathbb{1}_{T_{t+\Delta} \leq u} \mid \mathcal{F}_t) = \exp\left(2\mu_X(t+\Delta) + 2e^{-\kappa_X \Delta}(X_t - \mu_X(t))\right) \mathbb{E}(e^{2\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dL_u^X}) \times$$

$$\mathbb{E}\left(e^{2\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T} \mathbb{1}_{T_{t+\Delta} \leq u} \middle| \mathcal{F}_t\right)$$

The second term is $\varphi(-i; \Delta)$ by (3.4), and the last one is

$$\mathbb{E}\left(e^{2\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-s)} dW_s^T} \mathbb{1}_{T_{t+\Delta} \leq u} \middle| \mathcal{F}_t\right) = \mathbb{E}\left(e^{2\lambda\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-s)} dW_s^T} \mathbb{1}_{\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-s)} dW_s^T \leq \tilde{u}(T_t)} \middle| \mathcal{F}_t\right)$$

where $\tilde{u}(T_t) = \frac{u - (\mu_T(t+\Delta) + e^{-2\kappa_T \Delta}(T_t - \mu_T(t)))}{\sigma_T}$. We now calculate for $\tilde{u} \in \mathbb{R}$, and $G, G^\perp \sim \mathcal{N}(0,1)$ independent:

$$\mathbb{E}\left(e^{2\lambda\sigma_T(\frac{k_{XT}^2(\Delta)}{k_T(\Delta)} G + \frac{\sqrt{k_X(\Delta)^2 k_T(\Delta)^2 - k_{XT}^2(\Delta)}}{k_T(\Delta)} G^\perp)} \mathbb{1}_{k_T(\Delta) G \leq \tilde{u}}\right) = e^{\frac{4\lambda^2 \sigma_T^2}{2} \frac{k_X(\Delta)^2 k_T(\Delta)^2 - k_{XT}^4(\Delta)}{k_T^2(\Delta)}} \times$$

$$e^{\frac{(2\lambda\sigma_T)^2}{2}(\frac{k_{XT}^2(\Delta)}{k_T(\Delta)})^2} \Phi\left(\frac{\tilde{u}}{k_T(\Delta)} - 2\lambda\sigma_T \frac{k_{XT}^2(\Delta)}{k_T(\Delta)}\right).$$

Since $\tilde{u}(T_t)$ is $\mathcal{F}_t$-measurable and the variables $\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-u)} dW_u^T$ and $\int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-u)} dW_u^T$ are independent of $\mathcal{F}_t$, we get the claim by applying Proposition 6.2. $\qquad\square$

Proposition 6.11. *Under Model* (ETM)*, we have*

$$\mathbb{E}[S_{t+\Delta}((\bar{T} - T_{t+\Delta})^+)^2 | \mathcal{F}_t] = 2 \int_{T^0}^{\bar{T}} (\bar{T} - u) \mathbb{E}[S_{t+\Delta} \mathbb{1}_{u \leq T_{t+\Delta}} | \mathcal{F}_t] du,$$

with $T_0 = -\infty$.

We write this result with T^0, because for numerical purposes we use $T^0 = -273.15$ or $T^0 = -100$. Note that $\mathbb{E}[S_{t+\Delta} \mathbb{1}_{u \leq T_{t+\Delta}} | \mathcal{F}_t]$ can be calculated by using Proposition 6.9.

Proof. We have

$$((\bar{T} - T_{t+\Delta})^+)^2 = \int_{T^0}^{\bar{T}} \int_{T^0}^{\bar{T}} \mathbb{1}_{T_{t+\Delta} \leq u} \mathbb{1}_{T_{t+\Delta} \leq v} du dv$$

$$= 2 \int_{T^0}^{\bar{T}} \int_{T^0}^{\bar{T}} \mathbb{1}_{T_{t+\Delta} \leq u} \mathbb{1}_{T_{t+\Delta} \leq v} \mathbb{1}_{u \leq v} du dv = 2 \int_{T^0}^{\bar{T}} (\bar{T} - u) \mathbb{1}_{T_{t+\Delta} \leq u} du,$$

and therefore

$$\mathbb{E}[S_{t+\Delta}((\bar{T} - T_{t+\Delta})^+)^2 | \mathcal{F}_t] = 2 \int_{T^0}^{\bar{T}} (\bar{T} - u) \mathbb{E}[S_{t+\Delta} \mathbb{1}_{T_{t+\Delta} \leq u} | \mathcal{F}_t] du. \qquad\square$$

Results for two-sided quantos

Proposition 6.12. *Under Model* (ETM), *we have the following Taylor decomposition:*

$$\mathbb{E}[(S_{t+\Delta} - \bar{S})^+((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t] = \mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t]\mathbb{E}[((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$$

$$- \lambda\Big(\mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t] + \bar{S}\mathbb{P}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t]\Big) \times$$

$$\sigma_T^3 k_T(\Delta)k_{XT}^2(\Delta)\Big(\sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}(\frac{\bar{T}-\mu(t+\Delta)-e^{-\kappa_T\Delta}(T_t-\mu(t))}{\sigma_T k_T(\Delta)})^2}$$

$$+ 2\frac{\bar{T} - \mu(t+\Delta) - e^{-\kappa_T\Delta}(T_t - \mu(t))}{\sigma_T k_T(\Delta)}\times$$

$$\Phi\Big(\frac{\bar{T} - \mu(t+\Delta) - e^{-\kappa_T\Delta}(T_t - \mu(t))}{\sigma_T k_T(\Delta)}\Big)\Big) + o(\lambda)$$

where $k_T(\cdot)$, $k_{XT}(\cdot)$ *are as in Equation* (5.3).

Note that $\mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t]$ and $\mathbb{P}_{\lambda=0}\big(S_{t+\Delta} \geq \bar{S} \mid \mathcal{F}_t\big)$ are computed through Equation (6.3) and (6.4) respectively, and that $\mathbb{E}_{\lambda=0}[((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$ can be calculated by using Proposition 6.8.

Proof. As for $\mathbb{E}[(S_{t+\Delta} - \bar{S})^+]$ we will perform a Taylor decomposition of $\mathbb{E}[(S_{t+\Delta} - \bar{S})^+((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$. Let consider the 0 order term,

$$\mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t] = \mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t]\mathbb{E}[((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t].$$

Now, let us compute the derivative at $\lambda = 0$:

$$\frac{d}{d\lambda}\Big|_{\lambda=0} \mathbb{E}\left[(S_{t+\Delta} - \bar{S})^+((\bar{T} - T_{t+\Delta})^+)^2\Big|\mathcal{F}_t\right]$$

$$= \mathbb{E}\left[\sigma_T\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dW_v^T S_{t+\Delta}\mathbb{1}_{S_{t+\Delta}\geq\bar{S}}((\bar{T} - T_{t+\Delta})^+)^2\Big|\mathcal{F}_t\right]$$

$$= \mathbb{E}_{\lambda=0}\left(S_{t+\Delta}\mathbb{1}_{S_{t+\Delta}\geq\bar{S}} \mid \mathcal{F}_t\right)\mathbb{E}\left[\sigma_T\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dW_v^T((\bar{T} - T_{t+\Delta})^+)^2\Big|\mathcal{F}_t\right],$$

and $\mathbb{E}_{\lambda=0}[S_{t+\Delta}\mathbb{1}_{S_{t+\Delta}\geq\bar{S}}|\mathcal{F}_t] = \mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t] + \bar{S}\mathbb{P}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t]$.
By Proposition 6.2, we have

$$\mathbb{E}[\sigma_T\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dW_v^T((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$$

$$= \frac{k_{XT}^2(\Delta)}{k_T^2(\Delta)}\mathbb{E}\left[\sigma_T\int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)}dW_v^T((\bar{T} - T_{t+\Delta})^+)^2\Big|\mathcal{F}_t\right].$$

We can now use Lemma 6.4 with $a = \bar{T} - \mu(t + \Delta) - e^{-\kappa_T \Delta}(T_t - \mu(t))$ and $b = -\sigma_T k_T(\Delta)$ to get

$$\mathbb{E}\left[\sigma_T \int_t^{t+\Delta} e^{-\kappa_T(t+\Delta-v)} dW_v^T((\bar{T} - T_{t+\Delta})^+)^2 \Big| \mathcal{F}_t\right] = -\sigma_T^3 k_T(\Delta)^3 \left(\sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}(\frac{\bar{T}-\mu(t+\Delta)-e^{-\kappa_T\Delta}(T_t-\mu(t))}{\sigma_T k_T(\Delta)})^2}\right.$$

$$+ 2\frac{\bar{T} - \mu(t+\Delta) - e^{-\kappa_T\Delta}(T_t - \mu(t))}{\sigma_T k_T(\Delta)} \times$$

$$\left.\Phi\left(\frac{\bar{T} - \mu(t+\Delta) - e^{-\kappa_T\Delta}(T_t - \mu(t))}{\sigma_T k_T(\Delta)}\right)\right) \quad \square$$

Proposition 6.13. *Under Model* (ETM)*, we have the following Taylor expansion:*

$$\mathbb{E}[((S_{t+\Delta} - \bar{S})^+)^2(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] = \mathbb{E}_{\lambda=0}[((S_{t+\Delta} - \bar{S})^+)^2|\mathcal{F}_t]\mathbb{E}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]$$

$$- 2\Big(\mathbb{E}_{\lambda=0}[((S_{t+\Delta} - \bar{S})^+)^2|\mathcal{F}_t] + \bar{S}\mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t]\Big) \times$$

$$\sigma_T^2 k_{XT}(\Delta)^2 \Phi\Big(\frac{\bar{T} - \mu_T(t+\Delta) - e^{-\kappa_T\Delta}(T_t - \mu_T(t))}{\sigma_T k_T(\Delta)}\Big)\lambda + o(\lambda)$$

where $k_T(\cdot)$ and $k_{XT}(\cdot)$ are as in Equation (5.3)*.*

Note that $\mathbb{E}_{\lambda=0}[(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t]$ and $\mathbb{E}_{\lambda=0}[((S_{t+\Delta} - \bar{S})^+)^2|\mathcal{F}_t]$ can be computed through Equation (6.3) and (6.5) respectively, $\mathbb{E}_{\lambda=0}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]$ is given by Proposition 6.6 and $\mathbb{E}_{\lambda=0}[((\bar{T} - T_{t+\Delta})^+)^2|\mathcal{F}_t]$ can be calculated by using Proposition 6.8.

Proof. For $\lambda = 0$, we have

$$\mathbb{E}_{\lambda=0}[((S_{t+\Delta} - \bar{S})^+)^2(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] = \mathbb{E}_{\lambda=0}[((S_{t+\Delta} - \bar{S})^+)^2|\mathcal{F}_t]\mathbb{E}_{\lambda=0}[(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t].$$

Now, let us compute the derivative in $\lambda = 0$:

$$\frac{d}{d\lambda}\Big|_{\lambda=0}\mathbb{E}[((S_{t+\Delta} - \bar{S})^+)^2(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t] = \mathbb{E}_{\lambda=0}\Big[\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dW_v^T \times$$

$$2S_{t+\Delta}(S_{t+\Delta} - \bar{S})^+(\bar{T} - T_{t+\Delta})^+\Big|\mathcal{F}_t\Big]$$

$$= 2\mathbb{E}_{\lambda=0}[S_{t+\Delta}(S_{t+\Delta} - \bar{S})^+|\mathcal{F}_t] \times$$

$$\mathbb{E}\Big[\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dW_v^T(\bar{T} - T_{t+\Delta})^+\Big|\mathcal{F}_t\Big].$$

The calculation of $\mathbb{E}[\sigma_T \int_t^{t+\Delta} e^{-\kappa_X(t+\Delta-v)} dW_v^T(\bar{T} - T_{t+\Delta})^+|\mathcal{F}_t]$ has already been done in Equations (6.6) and (6.7). $\qquad\square$

Bibliography

[1] Enric Aguilar, Inge Auer, Manola Brunet, Thomas C. Peterson, and Jon Wieringa. "Guidelines on climate metadata and homogenization". In: *WMO Tech. Doc* 1186 (2003), p. 51.

[2] Yacine Aït-Sahalia, Robert Kimmel, et al. "Maximum likelihood estimation of stochastic volatility models". In: *Journal of financial economics* 83.2 (2007), pp. 413–452.

[3] Peter Alaton, Boualem Djehiche, and David Stillberger. "On modelling and pricing weather derivatives". In: *Applied mathematical finance* 9.1 (2002), pp. 1–20.

[4] Antonis Alexandridis and Achilleas D. Zapranis. *Weather derivatives: modeling and pricing weather-related risk*. Springer Science & Business Media, 2012.

[5] Antonis Alexandridis and Achilleas D. Zapranis. "Wind derivatives: Modeling and pricing". In: *Computational Economics* 41 (2013), pp. 299–326.

[6] Aurélien Alfonsi. *Affine diffusions and related processes: simulation, theory and applications*. Vol. 6. Bocconi & Springer Series. Springer, Cham; Bocconi University Press, Milan, 2015, pp. xiv+252. ISBN: 978-3-319-05220-5; 978-3-319-05221-2.

[7] Aurélien Alfonsi. "High order discretization schemes for the CIR process: application to affine term structure and Heston models". In: *Mathematics of Computation* 79.269 (2010), pp. 209–237.

[8] Aurélien Alfonsi and Nerea Vadillo. "A stochastic volatility model for the valuation of temperature derivatives". In: *arXiv preprint arXiv:2209.05918* (2022).

[9] Aurélien Alfonsi and Nerea Vadillo. "Risk valuation of quanto derivatives on temperature and electricity". In: *arXiv preprint arXiv:2310.07692* (2023).

[10] Robert Azencott, Peng Ren, and Ilya Timofeyev. "Realised volatility and parametric estimation of Heston SDEs". In: *Finance and Stochastics* 24 (2020), pp. 723–755.

[11] Martin T. Barlow. "A diffusion model for electricity prices". In: *Mathematical finance* 12.4 (2002), pp. 287–298.

[12] Ole E. Barndorff-Nielsen. "Normal inverse Gaussian distributions and stochastic volatility modelling". In: *Scandinavian Journal of statistics* 24.1 (1997), pp. 1–13.

[13] Pauline Barrieu and Olivier Scaillet. "A primer on weather derivatives". In: *Uncertainty and Environmental Decision Making*. Springer, 2009, pp. 155–175.

[14] Martin Baxter and Andrew Rennie. *Financial calculus: an introduction to derivative pricing*. Cambridge university press, 1996.

[15] Fred Espen Benth. "The stochastic volatility model of Barndorff-Nielsen and Shephard in commodity markets". In: *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 21.4 (2011), pp. 595–625.

[16] Fred Espen Benth and Jūratė Benth Šaltytė-Benth. "The volatility of temperature and pricing of weather derivatives". In: *Quantitative Finance* 7.5 (2007), pp. 553–561.

[17] Fred Espen Benth, Griselda Deelstra, and And Sinem Kozpınar. "Pricing energy quanto options in the framework of Markov-modulated additive processes". In: *IMA Journal of Management Mathematics* 34.1 (2023), pp. 187–220.

[18] Fred Espen Benth, Luca Di Persio, and Silvia Lavagnini. "Stochastic modeling of wind derivatives in energy markets". In: *Risks* 6.2 (2018), p. 56.

[19] Fred Espen Benth and Noor Ibrahim. "Stochastic modeling of photovoltaic power generation and electricity prices". In: *Journal of Energy Markets, Forthcoming* (2017).

[20] Fred Espen Benth, Jan Kallsen, and Thilo Meyer-Brandis. "A non-Gaussian Ornstein–Uhlenbeck process for electricity spot price modeling and derivatives pricing". In: *Applied Mathematical Finance* 14.2 (2007), pp. 153–169.

[21] Fred Espen Benth and Steen Koekebakker. "Stochastic modeling of financial electricity contracts". In: *Energy Economics* 30.3 (2008), pp. 1116–1157.

[22] Fred Espen Benth, Nina Lange, and Tor Age Myklebust. "Pricing and hedging quanto options in energy markets". In: *Journal of Energy Markets* 8.1 (2015).

[23] Fred Espen Benth and Thilo Meyer-Brandis. "The information premium for non-storable commodities". In: *Journal of Energy Markets* 2.3 (2009), pp. 111–140.

[24] Fred Espen Benth and Jūratė Šaltytė-Benth. *Modeling and pricing in financial markets for weather derivatives*. Vol. 17. World Scientific, 2012.

[25] Fred Espen Benth and Jūratė Šaltytė-Benth. "Stochastic modelling of temperature variations with a view towards weather derivatives". In: *Applied Mathematical Finance* 12.1 (2005), pp. 53–85.

[26] Fred Espen Benth and Jūratė Šaltytė-Benth. "The normal inverse Gaussian distribution and spot price modelling in energy markets". In: *International journal of theoretical and applied finance* 7.02 (2004), pp. 177–192.

[27] Fred Espen Benth and Jūratė Šaltytė-Benth. "Weather derivatives and stochastic modelling of temperature". In: *International Journal of Stochastic Analysis* 2011 (2011), pp. 1–21.

[28] Fred Espen Benth, Jūratė Šaltytė-Benth, and Steen Koekebakker. "Putting a price on temperature". In: *Scandinavian Journal of Statistics* 34.4 (2007), pp. 746–767.

[29]   Fred Espen Benth and Carlo Sgarra. "The risk premium and the Esscher transform in power markets". In: *Stochastic Analysis and Applications* 30.1 (2012), pp. 20–43.

[30]   Tesfahun Berhane, Nurilign Shibabaw, Gurju Awgichew, and Tesfaye Kebede. "Option pricing of weather derivatives based on a stochastic daily rainfall model with Analogue Year component". In: *Heliyon* 6.1 (2020).

[31]   Saptarshi Bhattacharya, Aparna Gupta, Koushik Kar, and Abena Owusu. "Hedging strategies for risk reduction through weather derivatives in renewable energy markets". In: *2015 International Conference on Renewable Energy Research and Applications (ICRERA)*. IEEE. 2015, pp. 1190–1195.

[32]   Tomas Björk and Camilla Landén. "On the term structure of futures and forward prices". In: *Mathematical Finance—Bachelier Congress 2000: Selected Papers from the First World Congress of the Bachelier Finance Society, Paris, June 29–July 1, 2000*. Springer. 2002, pp. 111–149.

[33]   Tim Bollerslev. "Generalized autoregressive conditional heteroskedasticity". In: *Journal of econometrics* 31.3 (1986), pp. 307–327.

[34]   Beáta Bolyog and Gyula Pap. "On conditional least squares estimation for affine diffusions based on continuous time observations". In: *Statistical Inference for Stochastic Processes* 22.1 (2019), pp. 41–75.

[35]   Bruno Paolo Bosco, Lucia P. Parisio, and Matteo M. Pelagatti. "Deregulated wholesale electricity prices in Italy: an empirical analysis". In: *International Advances in Economic Research* 13.4 (2007), pp. 415–432.

[36]   Colin F.H. Boyle, Jannik Haas, and Jordan D. Kern. "Development of an irradiance-based weather derivative to hedge cloud risk for solar energy systems". In: *Renewable Energy* 164 (2021), pp. 1230–1243.

[37]   Patrick L. Brockett, Mulong Wang, and Chuanhou Yang. "Weather derivatives and weather risk management". In: *Risk Management and Insurance Review* 8.1 (2005), pp. 127–140.

[38]   Dorje C. Brody, Joanna Syroka, and Mihail Zervos. "Dynamical pricing of weather derivatives". In: *Quantitative finance* 2.3 (2002), p. 189.

[39]   Nick Buckley, Alex Hamilton, Justyn Harding, Nicola Roche, Nick Ross, Emma Sands, Richard Skelding, Nick Watford, and Heidi Whitlow. "European weather derivatives". In: *General Insurance Convention*. 2002.

[40]   Rodrigo Caballero, Stephen Jewson, and Anders Brix. "Long memory in surface air temperature: detection, modeling, and application to weather derivative valuation". In: *Climate Research* 21.2 (2002), pp. 127–140.

[41]   Sean D. Campbell and Francis X. Diebold. "Weather forecasting for weather derivatives". In: *Journal of the American Statistical Association* 100.469 (2005), pp. 6–16.

[42] Melanie Cao and Jason Wei. *Pricing weather derivative: an equilibrium approach.* Tech. rep. Rotman School of Management, University of Toronto, 1999.

[43] Melanie Cao and Jason Wei. "Weather derivatives valuation and market price of weather risk". In: *Journal of Futures Markets: Futures, Options, and Other Derivative Products* 24.11 (2004), pp. 1065–1089.

[44] Massimiliano Caporin, Juliusz Preś, and Hipolit Torro. "Model based Monte Carlo pricing of energy and temperature quanto options". In: *Energy Economics* 34.5 (2012), pp. 1700–1712.

[45] René Carmona and Michael Coulon. "A survey of commodity markets and structural models for electricity prices". In: *Quantitative energy finance: modeling, pricing, and hedging in energy and commodity markets.* Springer, 2013, pp. 41–83.

[46] René Carmona and Pavel Diko. "Pricing precipitation based derivatives". In: *International Journal of Theoretical and Applied Finance* 8.07 (2005), pp. 959–988.

[47] Rene Carmona and Michael Ludkovski. "Spot convenience yield models for the energy markets". In: *Contemporary mathematics* 351 (2004), pp. 65–80.

[48] Peter Carr and Dilip Madan. "Option valuation using the fast Fourier transform". In: *Journal of computational finance* 2.4 (1999), pp. 61–73.

[49] Alvaro Cartea and Marcelo G. Figueroa. "Pricing in electricity markets: a mean reverting jump diffusion model with seasonality". In: *Applied Mathematical Finance* 12.4 (2005), pp. 313–335.

[50] Carolyn W. Chang, Jack S.K. Chang, and Kian Guan Lim. "Global warming, extreme weather events, and forecasting tropical cyclones: A market-based forward-looking approach". In: *ASTIN Bulletin: The Journal of the IAA* 42.1 (2012), pp. 77–101.

[51] CME Group. *CME Group Weather Suite Expanded.* 2023. URL: https://www.cmegroup.com/articles/2023/cme-group-weather-suite-expanded.html (visited on 07/23/2023).

[52] CME Group. *CME Weather Derivatives Establish New Records.* 2005. URL: http://investor.cmegroup.com/news-releases/news-release-details/cme-weather-derivatives-establish-new-records (visited on 07/23/2023).

[53] CME Group. *Managing Climate Risk with CME Group Weather Futures and Options.* 2021. URL: https://www.cmegroup.com/education/articles-and-reports/managing-climate-risk-with-cme-group-weather-futures-and-options.html (visited on 07/23/2023).

[54] Ana Cristina Costa and Amílcar Soares. "Homogenization of climate data: review and new perspectives using geostatistics". In: *Mathematical geosciences* 41 (2009), pp. 291–305.

[55] John C. Cox, Jonathan E. Ingersoll Jr., and Stephen A. Ross. "A theory of the term structure of interest rates". In: *Econometrica* 53.2 (1985), pp. 385–407.

[56] Laura Cucu, Rainer Döttling, Pascal Heider, and Samuel Maina. "Managing temperature-driven volume risks". In: *Journal of Energy Markets* 9.2 (2016).

[57] John Curtis, Muireann Á Lynch, and Laura Zubiate. "The impact of the North Atlantic Oscillation on electricity markets: A case study on Ireland". In: *Energy Economics* 58 (2016), pp. 186–198.

[58] Mark H.A. Davis. "Option pricing in incomplete markets". In: *Mathematics of derivative securities* 15 (1997), pp. 216–226.

[59] Freddy Delbaen and Jean Haezendonck. "A martingale approach to premium calculation principles in an arbitrage free market". In: *Insurance: Mathematics and Economics* 8.4 (1989), pp. 269–277.

[60] Shijie Deng. *Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes*. University of California Energy Institute Berkeley, 2000.

[61] Michel Denuit, Jan Dhaene, Marc Goovaerts, and Rob Kaas. *Actuarial theory for dependent risks: measures, orders and models*. John Wiley & Sons, 2006.

[62] Thomas Deschatre, Olivier Féron, and Pierre Gruet. "A survey of electricity spot and futures price models for risk management applications". In: *Energy Economics* 102 (2021), p. 105504.

[63] Gregor Dorfleitner and Maximilian Wimmer. "The pricing of temperature futures at the Chicago Mercantile Exchange". In: *Journal of Banking & Finance* 34.6 (2010), pp. 1360–1370.

[64] R.S. Elias, M.I.M. Wahab, and L. Fang. "A comparison of regime-switching temperature modeling approaches for applications in weather derivatives". In: *European Journal of Operational Research* 232.3 (2014), pp. 549–560.

[65] Caroline R. Ely, David J. Brayshaw, John Methven, James Cox, and Oliver Pearce. "Implications of the North Atlantic Oscillation for a UK–Norway renewable power system". In: *Energy Policy* 62 (2013), pp. 1420–1427.

[66] Kerry Emanuel, Sai Ravela, Emmanuel Vivant, and Camille Risi. "A statistical deterministic approach to hurricane risk assessment". In: *Bulletin of the American Meteorological Society* 87.3 (2006), pp. 299–314.

[67] Environmental Finance. *Annual Market Rankings 2022*. 2022. URL: https://www.environmental-finance.com/content/awards/annual-market-rankings-2022/categories/weather-markets-continue-to-heat-up.html (visited on 08/04/2023).

[68] Tobias Michael Erhardt, Claudia Czado, and Ulf Schepsmeier. "R-Vine Models for Spatial Time Series with an Application to Daily Mean Temperature". In: *Biometrics* 71.2 (2015), pp. 323–332.

[69] Alvaro Escribano, Ignacio J. Peña, and Pablo Villaplana. "Modelling electricity prices: International evidence". In: *Oxford bulletin of economics and statistics* 73.5 (2011), pp. 622–650.

[70] Hans Föllmer and Peter Leukert. "Efficient hedging: cost versus shortfall risk". In: *Finance and Stochastics* 4 (2000), pp. 117–146.

[71] Hans Föllmer and Peter Leukert. "Quantile hedging". In: *Finance and Stochastics* 3 (1999), pp. 251–273.

[72] Philip Hans Franses, Jack Neele, and Dick van Dijk. "Modeling asymmetric volatility in weekly Dutch temperature data". In: *Environmental Modelling & Software* 16.2 (2001), pp. 131–137.

[73] David García-León, Ana Casanueva, Gabriele Standardi, Annkatrin Burgstall, Andreas D. Flouris, and Lars Nybo. "Current and projected regional economic impacts of heatwaves in Europe". In: *Nature communications* 12.1 (2021), p. 5807.

[74] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. "Volatility is rough". In: *Quantitative finance* 18.6 (2018), pp. 933–949.

[75] Hélyette Geman and Andrea Roncoroni. "Understanding the fine structure of electricity prices". In: *The Journal of Business* 79.3 (2006), pp. 1225–1261.

[76] The World Bank GFDRR. *Weather Index-based Crop Insurance in Malawi. Facilitating Farmers' Access to Agricultural Credit.* 2012. URL: https://www.gfdrr.org/en/publication/weather-index-based-crop-insurance-malawi (visited on 07/23/2023).

[77] Rajna Gibson and Eduardo S. Schwartz. "Stochastic convenience yield and the pricing of oil contingent claims". In: *The Journal of Finance* 45.3 (1990), pp. 959–976.

[78] J. Gil-Pelaez. "Note on the inversion theorem". In: *Biometrika* 38.3-4 (1951), pp. 481–482.

[79] Marc Goovaerts, Florent Etienne De Vylder, and Jean Haezendonck. *Insurance Premiums. Theory and Applications.* North-Holland, 1985.

[80] Andreas Groll, Brenda López-Cabrera, and Thylo Meyer-Brandis. "A consistent two-factor model for pricing temperature derivatives". In: *Energy Economics* 55 (2016), pp. 112–126.

[81] Nalân Gülpınar and Ethem Çanakoḡlu. "Robust portfolio selection problem under temperature uncertainty". In: *European Journal of Operational Research* 256.2 (2017), pp. 500–523.

[82] David Heath, Eckhard Platen, and Martin Schweizer. "A comparison of two quadratic approaches to hedging in incomplete markets". In: *Mathematical finance* 11.4 (2001), pp. 385–413.

[83] Markus Hess. "Pricing energy, weather and emission derivatives under future information". PhD thesis. PhD thesis, Universitätsbibliothek Duisburg-Essen, 2012.

[84] Steven L. Heston. "A closed-form solution for options with stochastic volatility with applications to bond and currency options". In: *The review of financial studies* 6.2 (1993), pp. 327–343.

[85] Sondre Hølleland and Hans Arnfinn Karlsen. "Decline in temperature variability on Svalbard". In: *Journal of Climate* 33.19 (2020), pp. 8475–8486.

[86] Joshua B. Horton. "Parametric insurance as an alternative to liability for compensating climate harms". In: *Carbon & Climate Law Review* 12.4 (2018), pp. 285–296.

[87] John Hull and Alan White. "Pricing Interest-Rate-Derivative Securities". In: *The Review of Financial Studies* 3.4 (2015), pp. 573–592.

[88] John C. Hull. *Options futures and other derivatives*. Pearson Education India, 2003.

[89] Cody B. Hyndman. "Gaussian factor models—futures and forward prices". In: *IMA Journal of Management Mathematics* 18.4 (2007), pp. 353–369.

[90] Stephen Jewson and Anders Brix. *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations*. Cambridge University Press, 2005.

[91] Ho Ming Kang and Fadhilah Yusof. "Homogeneity tests on daily rainfall series". In: *Int. J. Contemp. Math. Sciences* 7.1 (2012), pp. 9–22.

[92] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*. Second. Vol. 113. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991.

[93] Dimitris Karlis. "An EM type algorithm for maximum likelihood estimation of the normal–inverse Gaussian distribution". In: *Statistics & probability letters* 57.1 (2002), pp. 43–52.

[94] B. Philipp Kellerhals. "Pricing electricity forwards under stochastic volatility". In: *Available at SSRN 274788* (2001).

[95] Lawrence A. Klimko and Paul I. Nelson. "On conditional least squares estimation for stochastic processes". In: *The Annals of statistics* (1978), pp. 629–642.

[96] Christopher R. Knittel and Michael R. Roberts. "An empirical examination of restructured electricity prices". In: *Energy Economics* 27.5 (2005), pp. 791–817.

[97] Rodwell Kufakunesu, Farai Julius Mhlanga, and Calisto Guambe. "On the sensitivity analysis of energy quanto options". In: *Stochastic Analysis and Applications* 40.6 (2022), pp. 1104–1125.

[98] Roger J.A. Laeven, Marc J. Goovaerts, et al. "Premium calculation and insurance pricing". In: *Encyclopedia of quantitative risk analysis and assessment* 3 (2008), pp. 1302–1314.

[99]   Shuying Lai, Jing Qiu, Yuechuan Tao, and Yinyan Liu. "Risk hedging strategies for electricity retailers using insurance and strangle weather derivatives". In: *International Journal of Electrical Power & Energy Systems* 134 (2022), p. 107372.

[100]   Damien Lamberton and Bernard Lapeyre. *Introduction to stochastic calculus applied to finance*. CRC press, 2011.

[101]   All Lari-Lavassani, A. Sadeghi, and Antony Ware. "Mean reverting models for energy option pricing". In: *Electronic Publications of the International Energy Credit Association* (2001).

[102]   Karl Larsson. "Parametric heat wave insurance". In: *Journal of Commodity Markets* (2023), p. 100345.

[103]   Yongheon Lee and Shmuel S. Oren. "An equilibrium pricing model for weather derivatives in a multi-commodity setting". In: *Energy Economics* 31.5 (2009), pp. 702–713.

[104]   Gunther Leobacher and Philip Ngare. "On modelling and pricing rainfall derivatives with seasonality". In: *Applied Mathematical Finance* 18.1 (2011), pp. 71–91.

[105]   Zenghu Li and Chunhua Ma. "Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model". In: *Stochastic Processes and their Applications* 125.8 (2015), pp. 3196–3233.

[106]   JoAnne Linnerooth-Bayer, Swenja Surminski, Laurens M. Bouwer, Ilan Noy, and Reinhard Mechler. "Insurance as a Response to Loss and Damage?" In: *Loss and damage from climate change: Concepts, methods and policy options* (2019), pp. 483–512.

[107]   Julio J. Lucia and Eduardo S. Schwartz. "Electricity prices and power derivatives: Evidence from the Nordic Power Exchange". In: *Review of derivatives research* 5.1 (2002), pp. 5–50.

[108]   Giovanni Masala, Marco Micocci, and Andrea Rizk. "Hedging wind power risk exposure through weather derivatives". In: *Energies* 15.4 (2022), p. 1343.

[109]   Takuji Matsumoto and Yuji Yamada. "Improving the Efficiency of Hedge Trading Using Higher-Order Standardized Weather Derivatives for Wind Power". In: *Energies* 16.7 (2023), p. 3112.

[110]   Takuji Matsumoto and Yuji Yamada. "Simultaneous hedging strategy for price and volume risks in electricity businesses using energy and weather derivatives". In: *Energy Economics* 95 (2021), p. 105101.

[111]   Reinhard Mechler, Elisa Calliari, Laurens M. Bouwer, Thomas Schinko, Swenja Surminski, JoAnne Linnerooth-Bayer, Jeroen Aerts, Wouter Botzen, Emily Boyd, Natalie Delia Deckard, et al. "Science for loss and damage. Findings and propositions". In: *Loss and damage from climate change: Concepts, methods and policy options* (2019), pp. 3–37.

[112]   Edward L. Melnick and Brian S. Everitt. *Encyclopedia of quantitative risk analysis and assessment*. Vol. 1. John Wiley & Sons, 2008.

[113] Xiaochun Meng. *Improving Probabilistic Forecasts by Using Intra-Day Data: An Application to Financial and Temperature Data*. 2018.

[114] Xiaochun Meng and James W. Taylor. "Comparing probabilistic forecasts of the daily minimum and maximum temperature". In: *International Journal of Forecasting* 38.1 (2022), pp. 267–281.

[115] Thilo Meyer-Brandis and Peter Tankov. "Multi-factor jump-diffusion models of electricity prices". In: *International Journal of Theoretical and Applied Finance* 11.05 (2008), pp. 503–528.

[116] José D. Morcillo, Fabiola Angulo, and Carlos J. Franco. "Analyzing the hydroelectricity variability on power markets from a system dynamics and dynamic systems perspective: Seasonality and ENSO phenomenon". In: *Energies* 13.9 (2020), p. 2381.

[117] Andreas Müller and Marcel Grandi. "Weather derivatives: a risk management tool for weather-sensitive industries". In: *The Geneva Papers on Risk and Insurance. Issues and Practice* 25.2 (2000), pp. 273–287.

[118] John A. Nelder and Roger Mead. "A simplex method for function minimization". In: *The computer journal* 7.4 (1965), pp. 308–313.

[119] Syoiti Ninomiya and Nicolas Victoir. "Weak approximation of stochastic differential equations and application to derivative pricing". In: *Appl. Math. Finance* 15.1-2 (2008), pp. 107–121.

[120] Martin Odening, Oliver Mußhoff, and Wei Xu. "Analysis of rainfall derivatives using daily precipitation models: Opportunities and pitfalls". In: *Agricultural Finance Review* 67.1 (2007), p. 135.

[121] Ostap Okhrin, Martin Odening, and Wei Xu. "Systemic weather risk and crop insurance: the case of China". In: *Journal of Risk and Insurance* 80.2 (2013), pp. 351–372.

[122] Yumi Oum and Shmuel S. Oren. "Optimal static hedging of volumetric risk in a competitive wholesale electricity market". In: *Decision Analysis* 7.1 (2010), pp. 107–122.

[123] Ludger Overbeck and Tobias Ryden. "Estimation in the Cox-Ingersoll-Ross model". In: *Econometric Theory* 13.3 (1997), pp. 430–461.

[124] Michal Pawlowski and Piotr Nowak. "Modelling spot prices on the Polish energy market". In: *Intelligent Systems' 2014*. Springer, 2015, pp. 781–792.

[125] Francisco Pérez-González and Hayong Yun. "Risk management and firm value: Evidence from weather derivatives". In: *The Journal of Finance* 68.5 (2013), pp. 2143–2176.

[126] Timothy J. Richards, Mark R. Manfredo, and Dwight R. Sanders. "Pricing weather derivatives". In: *American Journal of Agricultural Economics* 86.4 (2004), pp. 1005–1017.

[127] Paulina A. Rowińska, Almut E.D. Veraart, and Pierre Gruet. "A multi-factor approach to modelling the impact of wind energy on electricity spot prices". In: *Energy Economics* 104 (2021), p. 105640.

[128] Piergiacomo Sabino and Nicola Cufaro Petroni. "Fast simulation of tempered stable Ornstein–Uhlenbeck processes". In: *Computational Statistics* 37.5 (2022), pp. 2517–2551.

[129] Jūratė Šaltytė-Benth and Fred Espen Benth. "A critical view on temperature modelling for application in weather derivatives markets". In: *Energy Economics* 34.2 (2012), pp. 592–602.

[130] Frank Schiller, Gerold Seidler, and Maximilian Wimmer. "Temperature models for pricing weather derivatives". In: *Quantitative Finance* 12.3 (2012), pp. 489–500.

[131] Eduardo Schwartz and James E Smith. "Short-term variations and long-term dynamics in commodity prices". In: *Management Science* 46.7 (2000), pp. 893–911.

[132] Eduardo S. Schwartz. "The stochastic behavior of commodity prices: Implications for valuation and hedging". In: *The Journal of finance* 52.3 (1997), pp. 923–973.

[133] Andrea Stoppa and Ulrich Hess. "Design and use of weather derivatives in agricultural policies: the case of rainfall index insurance in Morocco". In: *International Conference "Agricultural Policy Reform and the WTO: Where are we heading", Capri (Italy)*. 2003.

[134] Swiss Re Institute. *The economics of climate change*. 2021. URL: https://www.swissre.com/institute/research/topics-and-risk-dialogues/climate-and-natural-catastrophe-risk/expertise-publication-economics-of-climate-change.html (visited on 07/31/2023).

[135] Che Mohd Imran Che Taib and Fred Espen Benth. "Pricing of temperature index insurance". In: *Review of development finance* 2.1 (2012), pp. 22–31.

[136] James W. Taylor and Roberto Buizza. "A comparison of temperature density forecasts from GARCH and atmospheric models". In: *Journal of forecasting* 23.5 (2004), pp. 337–355.

[137] The European Climate Adaptation Platform Climate-ADAPT. *Weather derivatives as risk management tool*. 2019. URL: https://climate-adapt.eea.europa.eu/en/metadata/adaptation-options/weather-derivatives-as-risk-management-tool/#websites (visited on 07/23/2023).

[138] The World Bank. *MN-Index-Based Livestock Insurance*. 2005. URL: https://projects.worldbank.org/en/projects-operations/project-detail/P088816?lang=en (visited on 07/23/2023).

[139] The World Bank. *Uruguay buys insurance against lack of rain and high oil prices*. 2018. URL: https://www.worldbank.org/en/results/2018/01/10/uruguay-insurance-against-rain-oil-prices (visited on 07/23/2023).

[140]   Richard S.J. Tol. "Autoregressive conditional heteroscedasticity in daily temperature measurements". In: *Environmetrics* 7.1 (1996), pp. 67–75.

[141]   Anders B. Trolle and Eduardo S. Schwartz. "Unspanned stochastic volatility and the pricing of commodity derivatives". In: *The Review of Financial Studies* 22.11 (2009), pp. 4423–4461.

[142]   Calum G. Turvey. "The pricing of degree-day weather options". In: *Agricultural Finance Review* 65.1 (2005), pp. 59–85.

[143]   Oldrich Vasicek. "An equilibrium characterization of the term structure". In: *Journal of financial economics* 5.2 (1977), pp. 177–188.

[144]   Dmitry V. Vedenov and Barry J. Barnett. "Efficiency of weather derivatives as primary crop insurance instruments". In: *Journal of Agricultural and Resource Economics* (2004), pp. 387–403.

[145]   Carlos Villa-Loaiza, Irvin Taype-Huaman, Julián Benavides-Franco, Guillermo Buenaventura-Vera, and Jaime Carabalí-Mosquera. "Does climate impact the relationship between the energy price and the stock market? The Colombian case". In: *Applied Energy* 336 (2023), p. 120800.

[146]   Daniel Weagley. "Financial sector stress and risk sharing: evidence from the weather derivatives market". In: *The Review of Financial Studies* 32.6 (2019), pp. 2456–2497.

[147]   Weather Risk Management Association. "Survey Results". In: *Weather Risk Management Association*. Washigton, DC, 2006.

[148]   Weather Risk Management Association. "Weather Risk Derivative Survey". In: *Weather Risk Management Association*. Washigton, DC, 2006.

[149]   Rafal Weron. "Electricity price forecasting: A review of the state-of-the-art with a look into the future". In: *International journal of forecasting* 30.4 (2014), pp. 1030–1081.

[150]   Rafal Weron and Adam Misiorek. "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models". In: *International journal of forecasting* 24.4 (2008), pp. 744–763.

[151]   www.artemis.bm. *CCRIF parametric payouts on Hurricane Irma reach $31.2m*. 2017. URL: https://www.artemis.bm/news/ccrif-parametric-payouts-on-hurricane-irma-reach-31-2m/ (visited on 07/23/2023).

[152]   www.artemis.bm. *Exchange traded weather derivatives planned for China*. 2021. URL: https://www.artemis.bm/news/exchange-traded-weather-derivatives-planned-for-china/ (visited on 07/23/2023).

[153]   Wei Xu, Martin Odening, and Oliver Mußhoff. "Indifference pricing of weather derivatives". In: *American Journal of Agricultural Economics* 90.4 (2008), pp. 979–993.

[154] Yuji Yamada. "Optimal hedging of prediction errors using prediction errors". In: *Asia-Pacific financial markets* 15 (2008), pp. 67–95.

[155] Yuji Yamada and Takuji Matsumoto. "Construction of Mixed Derivatives Strategy for Wind Power Producers". In: *Energies* 16.9 (2023), p. 3809.

[156] Peter Guangping Zhang. *Exotic options: a guide to second generation options*. World scientific, 1997.