



HAL
open science

Modélisation et calibration des prix spot électriques

O. Senhadji El Rhazi

► **To cite this version:**

O. Senhadji El Rhazi. Modélisation et calibration des prix spot électriques : Le modèle factoriel de Schwartz & Le modèle hyperbolique généralisé. Finance [q-fin.GN]. Paris 6, 2004. Français. NNT : . tel-01086146

HAL Id: tel-01086146

<https://theses.hal.science/tel-01086146>

Submitted on 3 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Modélisation et calibration des prix spot électriques :

*Le modèle factoriel de Schwartz & Le modèle
hyperbolique généralisé*

EDF - Optimisation Trading Commercialisation
Opération de marché - Pricing

Université Paris VI
Thèse de Master de Probabilités et Finance

présentée par
M. O. Senhadji El Rhazi

Soutenue le **20 Septembre 2004** devant le jury composé de :

Frédéric Noat	Directeur de thèse
Emmanuel Gobet	Examineur
Gilles Pages	Examineur
Nicole El Karoui	Présidente du jury

Table des matières

1	Introduction	4
1.1	Objets étudiés	4
1.2	Présentation du contexte	4
2	Le marché spot d'électricité	7
2.1	Généralités sur le spot	7
2.2	Structure des résidus	8
2.3	Approche non gaussienne	8
3	Le modèle factoriel de Schwartz	9
3.1	Dynamique des prix et primes de risque	9
3.1.1	Le modèle factoriel horaire	9
3.1.2	Analyse du prix spot	10
3.1.3	La prime de risque	13
3.1.4	Le modèle par produit	14
3.1.5	Calibration des modèles	15
3.2	Le prix spot un facteur désaisonnalisé	17
3.2.1	Solution du modèle	18
3.2.2	Modèle en temps discret	18
3.2.3	La vraisemblance	19
3.2.4	Estimation par moindres carrés	20
3.2.5	Analyse asymptotique	20
3.3	Le prix spot un facteur saisonnalisé	21
3.3.1	Solution du modèle	22
3.3.2	Modèle en temps discret	22
3.3.3	La vraisemblance	23
3.3.4	Incidence du risque	24
3.4	Fiabilité de l'estimateur log-vraisemblance	24
3.4.1	Description de la méthode d'estimation	24
3.4.2	Choix de la condition initiale	25
3.4.3	Performances de l'estimation	26
3.4.4	Structure des résidus	29
3.5	Calibration du prix spot	31
3.5.1	Problématique	31
3.5.2	Tendance constante	32
3.5.3	Tendance variable - risque neutre -	33
3.5.4	Tendance variable - prime de risque -	33
3.5.5	Analyse des résultats	34

4	Le modèle hyperbolique généralisé	34
4.1	Approche mathématique	34
4.1.1	Lois indéfiniment divisibles	35
4.1.2	Les processus de Lévy	36
4.2	Les processus hyperboliques généralisés	43
4.2.1	Quelques notions utiles	43
4.2.2	La distribution de loi \mathcal{HG}	44
4.2.3	Les mélanges de lois \mathcal{IGG}	45
4.2.4	Les processus subordonnés	46
4.2.5	Exemples de lois \mathcal{HG}	46
4.2.6	Propriétés de la loi \mathcal{HG}	47
4.2.7	La loi normal inverse gaussien	47
4.2.8	La loi variance gamma	50
4.3	Modélisation de la dynamique du sous-jacent	52
4.3.1	La prime de risque	52
4.3.2	L'approche martingale	54
4.3.3	Solution du modèle	56
4.4	Calibration du modèle	57
4.4.1	Estimation par moindres carrés	58
4.4.2	Estimation par maximum de vraisemblance	58
4.5	Validation et choix de modèles	60
4.5.1	Structure des résidus	60
4.5.2	Critères de choix de modèles	61
4.6	Évaluation des options européennes	61
4.6.1	Méthode de Fourier	62
4.6.2	Approche par Monte-Carlo	64
5	Conclusion	66

Remerciments

Je tiens à remercier **Beatrice Bigois** de m'avoir fait confiance et de m'avoir accueilli au sein de son équipe dynamique et chaleureuse.

Je remercie chaleureusement **Frédéric Noat** pour son encadrement exemplaire du présent stage. Sa disponibilité et son attention sont sans doute les atouts majeurs qui m'ont permis de mener à bien mon travail. Je me dois également de mentionner sa clarté et sa rigueur qui sont les caractéristiques de son incomparable encadrement.

Je tiens à remercier particulièrement **Guillaume Lehgaret, Virginie Gillet** et **Nathalie Delacroix-vaubois** qui m'ont soutenu et m'ont aidé tout au long de mon stage et sans qui mes projets n'auraient pas pu avancer. Merci à **Nadia Oudjane** de m'avoir donnée une occasion de travailler avec elle.

Je remercie également **Kaina**, du secrétariat, pour son aide au quotidien sur tous les détails administratifs et autres.

Je remercie mes professeurs de DEA, et tout particulièrement **Nicole El Karoui** qui fut pour moi une enseignante étonnante et motivante, et qui a fait tout pour avancer ma carrière professionnelle tout le long de mon DEA.

Enfin, merci aussi à l'Ecole Nationale des Ponts et chaussées et à l'université de paris VI, qui m'ont permis de réaliser ce stage de fin d'étude.

1 Introduction

1.1 Objets étudiés

Ce stage est une contribution à un approfondissement des études faites sur la modélisation et la calibration des prix spot et prix à terme d'électricité. Les travaux portent sur trois sujets principaux :

- La modélisation et la calibration des prix sous le modèle factoriel de Schwartz.
- La modélisation et l'étude des prix sous le modèle hyperbolique généralisé.
- La calibration et le pricing des prix sous le modèle hyperbolique généralisé.

Le rapport est structuré ainsi : dans cette partie introductive (1), on présente succinctement l'objectif ainsi que le contexte du stage; elle est l'occasion de présenter la division Optimisation Trading commercialisation dans laquelle se déroule le stage.

Le marché d'électricité (2), présente de manière générale les caractéristiques du marché d'électricité; On se fixera alors les objectifs de l'étude dans les parties (3 et 4).

Le modèle factoriel de Schwartz (3) correspond à la première modélisation du marché de l'électricité, cette partie reprend, synthétise et explique les différentes méthodes utilisées pour la calibration des prix d'électricité.

Le modèle hyperbolique généralisé (4) présente une extension du premier modèle, le but de cette partie est de comprendre la modélisation du marché de l'électricité et de fournir les fondements mathématiques et financiers pour la calibration et le pricing des prix sur ce marché.

Bonne lecture !

1.2 Présentation du contexte

La division OTC (Optimisation Trading, Commercialisation) est une des huit divisions de la branche Energies du groupe EDF. Elle compte 170 personnes en 2004. La division a été créée le 1er janvier 2003 pour adapter l'organisation des missions, jusqu'alors confiées à l'ancienne division SVO (Stratégie Valorisation Optimisation), aux évolutions du contexte économique, réglementaire et aux nouveaux enjeux de la Branche Energies. Cette organisation doit, en effet, permettre à la Branche Energies, responsable de ses

résultats et de ses performances :

✘ De renforcer sa relation commerciale avec l'agrégateur (nouvelle fonction de la Branche Commerce chargée de gérer l'approvisionnement en énergie électrique destinée aux clients finaux éligibles), dans le cadre de l'ouverture du marché de l'électricité français aux professionnels en juillet 2004 et aux clients résidentiels en 2007.

✘ D'assumer pleinement les arbitrages qui relèvent de la gestion des risques du producteur, face à l'essor continu des échanges sur les marchés de gros dans un environnement concurrentiel.

Dans ce contexte, la division OTC s'est fixé comme ambition, de garantir l'utilisation optimale des moyens de production physique ou contractuels de la Branche Energies, au service de la demande de ses clients et en tension avec les marchés de gros, dans le but de créer un maximum de valeur pour la Branche Energie et pour le groupe EDF, à travers une gestion des risques équilibrée.

Parallèlement, le directeur de la Branche Energies a créé la division Développement et Valorisation du Portefeuille d'actifs (DVP), chargée d'adapter, de développer et de valoriser le portefeuille d'actifs physiques et contractuels de la Branche Energies à moyen et long terme. Elle développe notamment une vision prospective des marchés énergétiques pour construire le futur mix de production d'EDF en France et en Europe dans une optique de développement durable et de respect des nouvelles contraintes environnementales.

Les missions de la division OTC :

OTC est une division opérationnelle, chargée de l'optimisation, du trading et de la commercialisation en gros de toutes les énergies (hors combustible nucléaire), en étroite coopération avec EDF Trading. Ses principales missions sont les suivantes :

✘ Optimiser l'exploitation du parc de production EDF existant, à tout les horizons de temps, en fonction des prévisions d'offre et de la demande, en relation étroite avec les divisions Production Nucléaire (DNP) et Energie Nouvelles Et Renouvelables, Thermiques et Hydrauliques (ENERTHY).

✘ Fournir à la Branche Commerce un approvisionnement fiable à un prix compétitif, source de rentabilité durable pour le Groupe, et élaborer avec elle des offres d'énergie attractives pour les marchés de gros européens.

✕ Maitriser les risques physiques et financiers du producteur liés aux aléas de l'équilibre offre/demande et des marchés, dans le cadre de la politique de gestion des risques de la Branche Energies.

✕ Assurer la gestion des flux financiers alimentant le chiffre d'affaires de la Branche Energies. OTC gère un flux de transactions annuelles de l'ordre de 15 milliards (dont près de 12 milliards de cessions à la Branche Commerce et près de 3 milliards de ventes d'énergie hors EDF).

Ces missions nécessitent une coopération étroite avec la Branche Commerce et toutes les divisions de la Branche Energies. La division DVP est un partenaire privilégié d'OTC, notamment en ce qui concerne l'évolution des relations entre la Branche Commerce et la Branche Energies liée à l'ouverture du marché, la relation contractuelle avec EDF Trading, la définition de la politique de risque de marché de la Branche Energies, la mise en oeuvre opérationnelle des évolutions réglementaires, l'adaptation du système d'information.

L'organisation de la division OTC :

La division OTC est organisée en quatre entités : Centre Opérationnel d'Optimisation de la Production (COOP), Opération de marchés, Risk Management et Back Office. La direction de la division s'appuie en outre sur deux fonctions communes aux divisions OTC et DVP : Contrôle de Gestion et Ressources Humaines. Le directeur de la division est assisté d'un directeur délégué et d'un chef de cabinet. Mon stage est effectué dans cette division OTC et précisément dans l'entité Opération de marché, qu'on expose ses missions dans la suite.

Les principales missions d'opérations de marchés :

Opérations de marché a en charge la gestion des relations commerciales avec le Branche Commerce. Elle est l'entité compétente sur toutes les questions relevant des marchés de gros européens. Les activités des 9 collaborateurs d'Opérations de marchés se répartissent entre :

✕ La gestion des relations avec les Branche Commerce :

- Les propositions d'offres de prix et, d'une manière générale, les transactions liées aux cessions d'énergie entre Branche Energies Commerce,
- La cotation de produits non standards, ces produits pouvant être des produits simples non cotés sur Platts (le cotation d'un ruban pour 2005 par exemple) ou des produits structurés (typiquement des clauses contractuelles d'effacement),
- Les échanges de données concernant les prévisions de consommation,

nécessaires à la gestion de l'équilibre offre-demande par le COOP et au cycle de gestion.

✗ Les analyses de marchés :

- Analyse des prix sur les marchés de gros européens, leur évolution, les corrélations entre les marchés, leur liquidité. . . ,
- Analyse des comportements des concurrents d'EDF sur les marchés de gros,
- Contribution aux analyses de débouchés sur le marché France.

Opérations de marchés est également en charge des achats de sécurité auprès d'EDF Trading, selon les modalités définies par Risk Management, et développe avec les filiales d'EDF des échanges sur les analyses de marchés ou la cotation de produit France.

2 Le marché spot d'électricité

Il est reconnu que les marchés ne suivent pas en général des bruits gaussiens (Black Scholes). La littérature a montré expérimentalement ce constat. D'une part les modèles de retour à la moyenne présentent un excès de Kurtosis et de Skewness par rapport aux modèles gaussiens (Mandelbrot (1963), Fama (1965)..). D'autre part, la volatilité implicite fait surgir des déformations qui illustrent l'insuffisance du modèle à reproduire la réalité. Pour palier à ces faiblesses, on introduit alors une classe plus exhaustive de lois qui sont les lois de Lévy hyperboliques généralisées.

2.1 Généralités sur le spot

Le marché de l'électricité est marqué par trois caractéristiques très particulières. Premièrement, **la forte saisonnalité des prix**, qui résulte fondamentalement du changement périodique de la demande sur le marché. Elle intègre la variabilité dans le temps de l'activité industrielle des acheteurs : jours fériés - ouvrables, week-end - semaine et jour - nuit. De plus, elle est modulée par une saisonnalité temporelle de la température : hiver - mi-saison - été. Deuxièmement, leurs **retour à la moyenne**, qui est une conséquence de l'équilibre offre-demande. Finalement, leurs variations brutales, qui sont dues au changement par palier (discontinu) des coûts fixes de la production d'électricité et le manque de liquidité. Cette spécificité, engendre des **pics de prix**, accentués par l'aspect non stockable du sous-jacent. On illustre dans les graphes de la Figure 1 les trois caractéristiques du marché spot d'électricité en France :

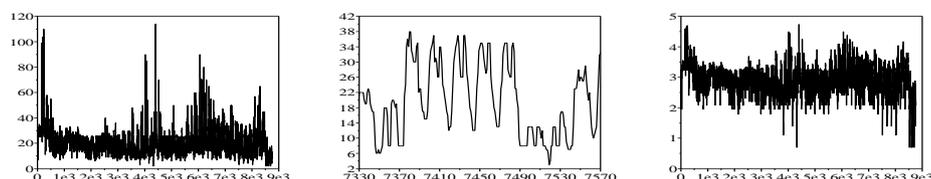


Figure 1: Prix spot Powernext heure par heure sur 1 an en Euros/MWh faisant apparaître les pics (à gauche), zoom sur une semaine pour voir la saisonnalité horohebdomadaire (au centre) et le logarithme du prix faisant apparaître le retour à la moyenne (à droite).

La problématique du marché de l'électricité est donc de trouver un modèle qui reproduit parfaitement ces trois aspects expérimentaux tout en étant théoriquement facile à utiliser.

2.2 Structure des résidus

Une première approche pour modéliser le prix spot est de postuler d'un modèle factoriel de Schwartz gaussien. L'étude dans la partie (3) résume les méthodes et les résultats dans le cas un facteur gaussien. Cette étude prouve que le modèle gaussien est inapproprié pour deux raisons. D'une part, les prix simulés avec les paramètres du calage ne permettent pas de reproduire les pics de prix. Ce défaut provient principalement du caractère gaussien du modèle. D'autre part, les résidus récupérés du calage présentent une autocorrélation et sont loin d'être gaussien. On a remarqué que les résidus sont leptokurtiques, i.e ils ont une queue de distribution épaisse avec un Kurtosis légèrement supérieur à 3.

2.3 Approche non gaussienne

Pour améliorer ce modèle, on peut proposer une extension à deux facteurs gaussien. Par ailleurs, on peut penser qu'une approche non gaussienne résoudre le problème. Il existe quatre grandes approches non gaussiennes : les modèles à saut, à changement de régime, à volatilité stochastique et à lois leptokurtiques. Cette dernière approche fera le cadre de notre étude. Les lois considérées doivent satisfaire les deux contraintes opérationnelles suivantes :

•⇨ Finitude des moments :

Il est impératif d'imposer l'existence des moments de tout ordre pour le pricing d'actifs conditionnels sur le spot ou sur le forward. Les lois dont un ou plusieurs moment sont infinis ne seront pas pris en compte.

•♦ **Stabilité par convolution :**

En simulation, il est préférable de connaître la loi des prix directement (horaire, journalier...), pour pouvoir passer d'un modèle à un autre sans beaucoup de calculs. On privilégiera alors les lois stables par convolution.

3 Le modèle factoriel de Schwartz

3.1 Dynamique des prix et primes de risque

3.1.1 Le modèle factoriel horaire

Plaçons-nous d'abord dans le cadre d'un marché risque-neutre pour une probabilité \mathbb{Q} . Sous cette probabilité, le processus $(F(t, T))_{t \geq 0}$ des prix à terme horaire à la date t pour une livraison d'un MWh d'électricité à la date T , est une martingale locale. Le modèle factoriel de Schwartz suppose que la courbe des prix suit la diffusion HJM suivante :

$$\frac{dF(t, T)}{F(t, T)} = \sum_{k=1}^N \sigma_{(t, T)}^k dW_t^k$$

Où, $(W_t^1 \dots W_t^N)$ désigne des mouvements browniens, $(\sigma_{(t, T)}^1 \dots \sigma_{(t, T)}^N)$ les volatilités à la date de cotation t pour une livraison unitaire (1 heure) de l'électricité en date T au prix $F(t, T)$. On s'intéresse particulièrement aux deux cas simples, **un facteur** et **deux facteurs**, dont les dynamiques sont respectivement les suivantes :

$$\text{Un facteur} \quad \frac{dF(t, T)}{F(t, T)} = \sigma_c e^{-a(T-t)} dW_t^1$$

$$\text{Deux facteurs} \quad \frac{dF(t, T)}{F(t, T)} = \sigma_c e^{-a(T-t)} dW_t^1 + \sigma_l dW_t^2$$

Il faut noter que dans le modèle un facteur, le prix à terme dépend uniquement des aléas à court terme avec une volatilité $\sigma_{(t, T)}^1 = \sigma_c e^{-a(T-t)}$. Par contre, dans le modèle deux facteurs, ce dernier, dépend de deux types d'aléas : les aléas court terme comme dans le modèle un facteur $\sigma_{(t, T)}^1 = \sigma_c e^{-a(T-t)}$ et les aléas long terme avec une volatilité $\sigma_{(t, T)}^2 = \sigma_l$.

Il est intéressant de voir la **volatilité globale instantanée** $\Sigma(T-t)$, diffère selon qu'on considère un modèle un ou deux facteurs. Dans le premier elle vaut $\Sigma(T-t) = \sigma_c e^{-a(T-t)}$, par contre dans le deuxième elle vaut

$\Sigma(T-t) = \sqrt{\sigma_c^2 e^{-2a(T-t)} + \sigma_l^2}$. Dans la Figure 2, on représente la différence entre ces deux grandeurs pour des valeurs typiques de la volatilité et du retour à la moyenne.

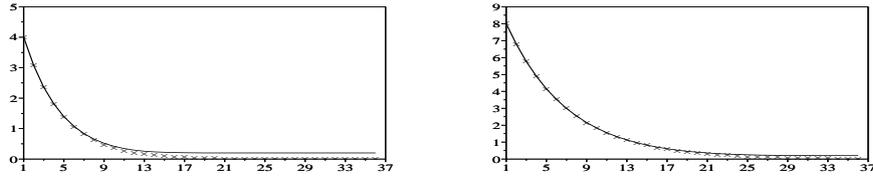


Figure 2: Evolution de la volatilité globale instantané $\Sigma(T-t)$ en fonction de l'échéance relative $T-t$ pour le modèle un facteur (en étoiles) et le modèle deux facteurs (trait continu), pour $(a, \sigma_c, \sigma_l) = (97, 400\%, 20\%)$ (à gauche) et $(a, \sigma_c, \sigma_l) = (97, 800\%, 20\%)$ (à droite).

Remarque :

1. On suppose que les deux sources de bruits, dans le modèle à deux facteurs, sont indépendantes :

$$\text{Cov}(dW_t^1, dW_t^2) = 0$$

2. Le modèle à deux facteurs intègre deux sources de bruits et donc le changement de probabilité se fera avec deux primes de risques.

3.1.2 Analyse du prix spot

Les produits d'électricité ne sont pas délivrables le même jour (la même heure) et leur prix n'est en général pas observable sur le marché. De ce fait la notion du prix spot est purement théorique. Mathématiquement il est préférable de considérer le processus logarithme du prix spot $X_t^v = \ln(\frac{S_t}{S_0})$. On rappelle la relation entre le prix spot et le prix à terme $S_T = \lim_{t \rightarrow T} F(t, T)$. Cette relation permet de formuler la dynamique des prix spot :

Le modèle un facteur :

Le prix spot, dans le modèle un facteur, vérifie l'équation différentielle stochastique (EDS) suivante :

$$dS_t = a(\theta_t - \ln S_t)S_t dt + \sigma_c S_t dW_t^1$$

$$\theta_t = \ln F(0, t) + \frac{1}{a} \frac{\partial \ln F(0, t)}{\partial t} + \frac{\sigma_c^2}{4a} (1 - e^{-2at})$$

En utilisant le lemme d'**Itô** pour le logarithme du prix spot, on trouve que ce dernier suit la diffusion d'Ornstein-Uhlenbeck suivante :

$$dX_t^v = a(m_t - X_t^v)dt + \sigma_c dW_t^1, \quad X_0^v = 0 \quad (1)$$

$$m_t = \ln F(0, t) + \frac{1}{a} \frac{\partial \ln F(0, t)}{\partial t} - \frac{\sigma_c^2}{4a} (1 + e^{-2at}) \quad (2)$$

Par ailleurs, on peut décomposer le prix en une partie déterministe et une partie stochastique. La partie déterministe notée $S_0 D_t$ représente la saisonnalité des prix spot. La partie stochastique, dont le logarithme est X_t^c , intègre la volatilité et le retour à la moyenne des prix spot. Il est aisé de montrer que celui-ci est solution de la diffusion d'Ornstein-Uhlenbeck suivante¹ :

$$dX_t^c = a(m - X_t^c)dt + \sigma_c dW_t^1, \quad X_0^c = 0$$

$$\ln S_0 D_t = \ln F(0, t) - \frac{\sigma_c^2}{4a} (1 - e^{-2at}) - m(1 - e^{-at})$$

Le prix S_t peut, alors, être exprimé selon deux expressions : $S_t = S_0 D_t e^{X_t^c} = S_0 e^{X_t^v}$. Pour la première approche, on utilise une méthode de désaisonnalisation du prix (méthode par classe, par processus moyennes mobiles...), puis on utilise l'historique du prix désaisonnalisé pour calibrer le processus $(X_t^c)_{t \geq 0}$ et quantifier les paramètres du modèle (m, a, σ) . Par contre, pour la deuxième approche, on utilise directement la courbe forward initiale $F(0, t)$ et l'historique du prix, pour calibrer le processus $(X_t^v)_{t \geq 0}$ et quantifier les paramètres (a, σ) . On peut donner une autre expression analytique à la partie Dt :

$$D_t = \frac{1}{S_0} e^{\mathbb{E}_Q[\ln S_t]}, \quad m = 0$$

Remarque :

La première approche dépend beaucoup de la méthode de désaisonnalisation. En plus toute erreur sur la saisonnalité se répercute sur l'estimation des paramètres du modèle et sur la structure des résidus.

La Figure 3 représente l'évolution du prix spot dans un modèle de Schwartz un facteur **variable** suivant différentes courbes forwards :

Les prix de l'électricité sont marqués par une forte saisonnalité de la demande et de la température. Pour cette raison les prix sont différents selon

¹ Injectons $\ln(S_t) = X_t^c + \ln(S_0 D_t)$ dans (1) ; on obtient :

$$dX_t^c + d(\ln(S_0 D_t)) = a(m_t - X_t^c - \ln(S_0 D_t))dt + \sigma_c dW_t^1$$

En prenant l'espérance de l'EDS ci-dessus, on trouve que $\ln(D_t)$ est solution de l'équation différentielle suivante :

$$d(\ln(S_0 D_t)) = a(m_t - \ln(S_0 D_t))dt$$

En réinjectant dans l'EDS vérifiée par X_t^c , on a le résultat.

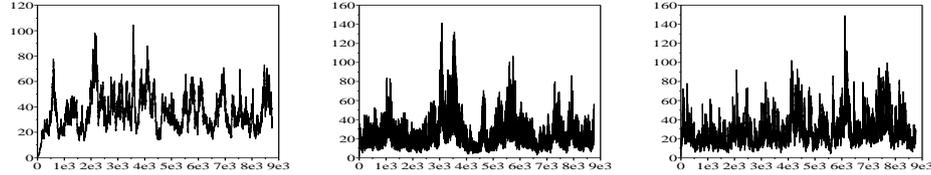


Figure 3: Evolution du prix spot S_t en fonction du temps pour une tendance : constante (à gauche), sinus-polynomiale (au centre) et celle d'EDF 2004-2005 (à droite).

les heures de la journée, les jours de la semaine et les saisons de l'année. On voit que la tendance polynomiale prend plus en compte ce phénomène que la tendance constante.

Le modèle deux facteur :

Le prix spot, dans le modèle deux facteurs, vérifie l'équation différentielle stochastique (EDS) suivante :

$$dS_t = a(\theta_t - \ln S_t - \frac{1}{2}\sigma_l^2 t + \sigma_l W_t^2)dt + \sigma_c dW_t^1 + \sigma_l dW_t^2$$

En utilisant le lemme d'**Itô** pour le logarithme du prix spot, on trouve que ce dernier suit la diffusion Markovienne suivante (voir [SS00]) :

$$dX_t^v = dc_t + dl_t$$

Avec,

$$dc_t = a(m_t - c_t)dt + \sigma_c dW_t^1 \quad \text{et} \quad dl_t = -\frac{1}{2}\sigma_l^2 dt + \sigma_l dW_t^2$$

L'estimation du modèle à deux facteurs peut se faire à l'aide du **filtre de Kalman**, qui se base sur le maximum de vraisemblance des résidus. Vue la forme de la volatilité globale, on peut penser qu'un découplage long-court terme permet d'estimer les paramètres (a, σ_c, σ_l) . En effet, pour estimer la composante long terme, on suppose que les rendements des logarithmes des prix à terme suivent une marche aléatoire, soit :

$$\text{Var}\left[\ln \frac{F(t + \delta, T)}{F(t, T)}\right] = \sigma_l^2 \delta$$

Par ailleurs, pour estimer la composante court terme, on dégrade le modèle deux facteurs en un modèle un facteur, puis on estime les paramètres (a, σ_c) . Dans la suite seul le modèle un facteur sera étudié.

Remarque :

L'estimation de la volatilité long terme se fait sur le produit calendar (durée de livraison 1 an), et on considère sa cotation loin de l'échéance.

3.1.3 La prime de risque

Dans l'estimation des paramètres, la probabilité risque-neutre \mathbb{Q} n'est pas utilisée et laisse place à la probabilité historique \mathbb{P} . Dans cet univers historique, il faut introduire le biais dû à la prime de risque ξ_t , que nous supposons déterministe. En étant risque neutre, on trouve que l'espérance d'un actif à terme sous \mathbb{P} ne représente pas la valeur du prix spot attendu à maturité. Il est en général accepté que les investisseurs aient de l'aversion pour le risque². Par ailleurs, vu la relation entre le prix forward et le spot, ce changement affecte aussi les prix spot.

Modèle factoriel :

La prime de risque permet le changement de probabilité unique $\mathbb{Q} \leftrightarrow \mathbb{P}$ qui rend les prix à terme martingales locales. Ce changement de probabilité est possible grâce au théorème de **Girsanov**. Le modèle factoriel devient alors :

$$\frac{dF(t, T)}{F(t, T)} = \sum_{k=1}^N \sigma_{(t, T)}^k (dW_t^k + \xi_t^k dt)$$

Les grandeurs $(\xi_t^1 \dots \xi_t^N)$ représentent les primes de risques spécifiques du prix aux sources de bruits $(W_t^1 \dots W_t^N)$, à la date t . Pour les deux cas simples un et deux facteurs, la dynamique des prix sous la probabilité historique \mathbb{P} est :

$$\text{Un facteur} \quad \frac{dF(t, T)}{F(t, T)} = \xi_t^c \sigma_c e^{-a(T-t)} dt + \sigma_c e^{-a(T-t)} dW_t^1$$

$$\text{Deux facteurs} \quad \frac{dF(t, T)}{F(t, T)} = \xi_t^c \sigma_c e^{-a(T-t)} dt + \sigma_c e^{-a(T-t)} dW_t^1 + \xi_t^l \sigma_l dt + \sigma_l dW_t^2$$

Dans le modèle à deux facteurs, il existe deux sources de bruits, auxquelles on attribue les primes de risques (ξ_t^c, ξ_t^l) .

Le prix spot :

²Les acheteurs sont quelques fois prêts à s'engager dans un contrat à terme à un prix supérieur à celui du prix spot espéré pour sécuriser leur fourniture, on parlera de prime négative. Au contraire, les vendeurs peuvent s'engager dans des contrats futurs à des prix inférieurs à ceux attendus par le marché pour pouvoir assurer un débouché à leurs actifs), on parlera de prime de risque positive.

Après avoir modélisé la dynamique du sous-jacent à terme, nous répercutons cette dynamique sur le sous-jacent au comptant. Plus précisément on remarque que ce changement se manifeste dans la tendance du logarithme du spot X_t . En effet, la nouvelle tendance s'exprime facilement en fonction de la tendance risque-neutre et la prime de risque comme suite :

$$m_t^{\mathbb{P}} = m_t^{\mathbb{Q}} + \frac{\sigma_c}{a} \xi_t \quad (3)$$

Dans le modèle à deux facteurs, le logarithme garde la même dynamique avec une nouvelle tendance :

$$m_t^{\mathbb{P}} = m_t^{\mathbb{Q}} + \frac{\sigma_c}{a} \xi_t^c + \frac{\sigma_l}{a} \xi_t^l \quad (4)$$

3.1.4 Le modèle par produit

L'avantage du modèle horaire est qu'il propose une modélisation à granularité très fine, ce qui permet de reconstruire n'importe quel prix pour des périodes de livraison non unitaire d'une heure d'un MW d'électricité. On note $F(t, T, \theta)$ le prix future côté à la date t pour une livraison unitaire entre T et $T + \theta$ (on note $\theta = Nh$ et $h = \frac{1}{24 \cdot 365}$), par absence d'opportunité d'arbitrage on a :

$$F(t, T, \theta) = \frac{1}{N} \sum_{i=0}^{N-1} F(t, T + ih)$$

On introduit les variables aléatoires suivantes :

$$\lambda_i^{t, T, \theta} = \frac{F(t, T + ih)}{F(t, T, \theta)}$$

Le coefficient $\lambda_i^{t, T, \theta}$ ainsi défini peut s'interpréter comme un facteur de forme (Le poids de l'heure i dans le prix moyen sur la période $[T, T + \theta]$). On supposera les facteurs de formes constants et égaux à 1 (voir [Ouj0903] et pour plus d'informations sur la pertinence de ces résultats voir [ND04]). On définit, pour la suite, la fonction $\Psi(\theta) = \frac{h}{\theta} \frac{1 - e^{-a\theta}}{1 - e^{-ah}}$. Le prix à terme pour un produit de livraison de durée θ suit, les dynamiques suivantes :

$$\begin{aligned} \text{Un facteur} \quad \frac{dF(t, T, \theta)}{F(t, T, \theta)} &= \sigma_c e^{-a(T-t)} \Psi(\theta) dW_t^1 \\ \text{Deux facteurs} \quad \frac{dF(t, T, \theta)}{F(t, T, \theta)} &= \sigma_c e^{-a(T-t)} \Psi(\theta) dW_t^1 + \sigma_l dW_t^2 \end{aligned}$$

Pour l'estimation d'un modèle "spot" ($T = t$) pour un produit de livraison de durée θ (ex : le DayAhead, $\theta = \frac{1}{365}$, $N = 24$, $h = \frac{1}{24 \cdot 365}$), on utilise le même mécanisme de calibration que dans le cas du modèle horaire en intégrant le facteur $\Psi(\theta)$. Ce facteur n'influence pas la méthode d'estimation

des paramètres. En effet, il suffit d'adopter le changement de variable : $\sigma_c^\theta = \sigma_c \Psi(\theta)$ et d'utiliser la courbe forward initiale : $F(0, t, \theta) = \frac{1}{\theta} \sum_{i=0}^{\theta-1} F(0, t+i)$. Ces derniers permettent de retrouver le cas du modèle horaire :

$$(a^h, \sigma_c^h, \xi_c^h) = (a^\theta, \frac{\sigma_c^\theta}{\Psi(\theta)}, \xi_c^\theta)$$

Les paramètres $(a^h, \sigma_c^h, \xi_c^h)$ étant ceux du modèle horaire, et $(a^\theta, \sigma_c^\theta, \xi_c^\theta)$ ceux du modèle produit de livraison θ . Ces différentes grandeurs sont compatibles et exprimées en unité annuelle.

Dans la suite on s'intéresse exclusivement au modèle un facteur horaire (Ou journalier avec le changement de variables ci-haut).

3.1.5 Calibration des modèles

Dans le cadre de notre étude, le marché de l'électricité, dans sa totalité, est modélisé par trois paramètres (a, σ_c, σ_l) ³. Il est clair que cette modélisation ne permet pas de représenter parfaitement la réalité du marché. De manière générale, plus le nombre de paramètres est élevé plus l'historique nécessaire à leurs calibration doit être long. Or le marché d'électricité est particulièrement **instationnaire**, l'utilisation d'un historique très long est inefficace. Par ailleurs, il existe deux approches possibles pour l'identification des modèles :

- Les **méthodes empiriques** utilisant des données de cotations historiques.

- Les **méthodes implicites** basées sur l'observation des prix des options.

Dans une perspective de pricing, il est clair que la deuxième approche est la plus appropriée. On rappelle quelques éléments de cette méthode de calibration : Le prix d'options, la formule fermée. On considère comme sous-jacent un produit à terme à la date T_0 pour une livraison unitaire à la date T_f : $F(T_0, T_f)$. Le prix d'une option, à la date t , d'échéance T_0 , dont le sous-jacent est ce forward et le pay-off est \mathcal{G} est donné par :

$$C(t, T_0, T_f, \mathcal{G}) = \mathbb{E}_Q[e^{-r(T_0-t)} \mathcal{G}(F(T_0, T_f)) | \mathcal{F}_t]$$

Pour une option européenne de sous-jacent le forward $F(T_0, T_f)$, le prix d'option admet la formule analytique suivante :

✗ Le prix d'un Call européen

$$C(t, T_0, T_f, K) = e^{-r(T_0-t)} (F(t, T_f) \mathcal{N}(d_1) - K \mathcal{N}(d_0))$$

³Cette modélisation suppose que $\rho = \text{Cov}(dW_t^1, dW_t^2)$.

✗ Le prix d'un Put européen

$$P(t, T_0, T_f, K) = e^{-r(T_0-t)}(F(t, T_f)(\mathcal{N}(d_1) - 1) - K(\mathcal{N}(d_0) - 1))$$

Où,

$$d_0 = d_1 - \sqrt{\Sigma(t, T_0, T_f)}$$

$$d_1 = \frac{\ln\left(\frac{F(t, T_f)}{K}\right) + \frac{1}{2}\Sigma(t, T_0, T_f)}{\sqrt{\Sigma(t, T_0, T_f)}}$$

$$\Sigma(t, T_0, T_f) = \frac{\sigma_c^2}{2a}(e^{2a(T_0-T_f)} - e^{2a(t-T_f)}) + \sigma_l^2(T_0 - t)$$

Supposons qu'on dispose que N cotations d'options \hat{C}_i qui proviennent du marché. On note θ le vecteur des variables (a, σ_c, σ_l) et par $C_i(\theta)$ le prix donné par la formule analytique. Pour calibrer le modèle factoriel (un ou deux facteurs), on utilise une méthode de moindres carrés :

$$\theta = \operatorname{argmin}_{\theta} \sum_{i=1}^N (\hat{C}_i - C_i(\theta))^2$$

Cette optimisation permet de calibrer les paramètres du modèle (a, σ_c, σ_l) . Malheureusement **le manque d'informations concernant le prix des options**, ne permet pas d'utiliser cette approche. Les méthodes statistiques restent le seul recours et elles seront exposées par la suite. Pour ces méthodes, il faut dégager une certaine **stationnarité** dans les données, faire des **tests d'adéquation des modèles**, notamment tester la normalité des résidus et leurs non autocorrélation. Dans la Figure 4 et 5, on représente la dépendance du prix des paramètres $(a, \sigma_c, \sigma_l, K)$.

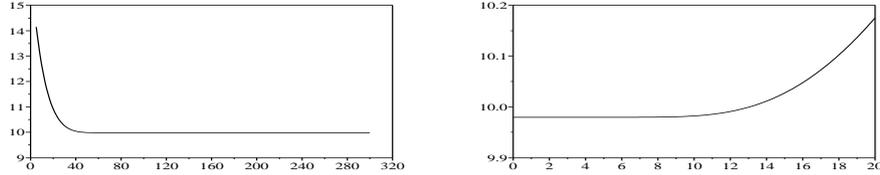


Figure 4: Evolution du prix du Call européen en fonction de a (à gauche) et de σ_c (à droite).

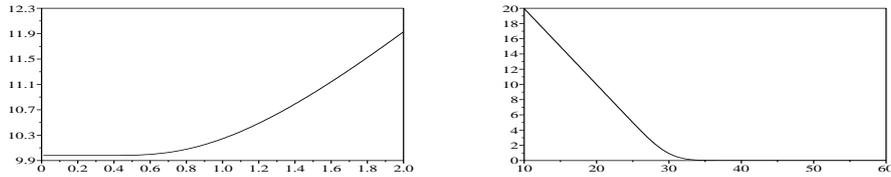


Figure 5: Evolution du prix du Call européen en fonction de σ_l (à gauche) et de K (à droite).

3.2 Le prix spot un facteur désaisonnalisé

Dans cette partie on considère que la tendance du logarithme du spot reste constante au cours du temps. On se place, dans un premier temps, sous la probabilité risque-neutre \mathbb{Q}^4 . Le processus $(X_t^c)_{t \geq 0}$ vérifie alors l'EDS suivante :

$$dX_t^c = a(m - X_t^c)dt + \sigma dW_t, \quad X_0^c = 0 \quad (5)$$

On admettra que $X_0^c = 0$, dans le cas contraire on considère $Y_t = X_t^c - X_0^c$. Cette EDS admet une unique solution. On peut démontrer que le processus

⁴Si on suppose que la prime de risque de risque est constante dans le temps, alors elle sera intégrée implicitement dans l'estimation du paramètre m .

$(X_t)_{t \geq 0}$ est un processus gaussien.

3.2.1 Solution du modèle

On peut expliciter la solution de l'EDS (5) en utilisant la formule d'Itô. En effet, posons $Z_t = (X_t^c - m)e^{at}$, alors Z_t vérifie l'EDS : $dZ_t = \sigma e^{at} dW_t$. L'expression de X_t^c est donnée par :

$$X_t^c = m(1 - e^{-at}) + \sigma \int_0^t e^{-a(t-s)} dW_s$$

La variable aléatoire X_t^c suit une loi normale $\mathcal{N}(M_{(m,a,t)}, V_{(a,\sigma,t)})$ de densité $f_{(m,a,\sigma,t)}$.

Où,

$$\begin{aligned} M_{(m,a,t)} &= m(1 - e^{-at}) \\ V_{(a,\sigma,t)} &= \frac{\sigma^2}{2a}(1 - e^{-2at}) \\ f_{(m,a,\sigma,t)}(x) &= \frac{e^{-\frac{(x-M_{(m,a,t)})^2}{2V_{(a,\sigma,t)}}}}{\sqrt{2\pi V_{(a,\sigma,t)}}} \end{aligned}$$

3.2.2 Modèle en temps discret

L'utilisation numérique du modèle ne peut se faire qu'en temps discret. En effet, en pratique, on dispose d'observations à des instants discrets (jour par jour, heure par heure...). On considère une discrétisation sur un intervalle $[0, T]$ avec un pas de temps h . Pour alléger les formulations on note : $X_{t_i}^c = X_{ih}^c = X_i^c$. La formule explicite de X_t^c donne le schéma suivant :

$$X_i^c - \phi_1 X_{i-1}^c - \phi_0 = \varepsilon_i \quad (6)$$

Où les ϕ_i ($i = 0, 1, 2$) sont des réels et $\varepsilon_i \sim BB(0, \sigma_\varepsilon^2)$ avec :

$$\begin{aligned} \phi_0 &= m(1 - e^{-ah}) \\ \phi_1 &= e^{-ah} \\ \sigma_\varepsilon &= \sqrt{\frac{1 - e^{-2ah}}{2a}} \sigma \\ \phi_2 &= \ln(\sigma_\varepsilon^2) \end{aligned}$$

Les variables aléatoires $\varepsilon_1, \dots, \varepsilon_n$ sont des bruits blancs gaussiens, c'est à dire de moyenne nulle, de variance constante et non autocorrélés. On remarque que X_t^c est un processus AR(1) (Processus Autorégressif) avec drift

constant. L'estimation des paramètres de ce processus peut se faire par moindres carrés ordinaire ou maximum de vraisemblance.

Remarque :

La condition de stationnarité du processus AR(1) est $|\phi_1| < 1$, cette condition est vérifiée dans la mesure où $a > 0$. Dans le cas limite $a = 0$, le processus de logarithme du prix spot est une marche aléatoire.

3.2.3 La vraisemblance

On dispose de n observations du logarithme du prix spot, sur un intervalle $[0, T]$, à des instants discrets de pas de temps h ($n = \frac{T}{h}$). Pour un échantillon $(x_1 \dots x_n)$ de loi X_t^c , la log-vraisemblance s'écrit :

$$\mathcal{L}(m, a, \sigma, x_1 \dots x_n) = \sum_{i=1}^n \ln f(\varepsilon_i)$$

La fonction $f(x)$ représente la densité de la loi normale $\mathcal{N}(0, \sigma_\varepsilon^2)$. Il s'agit maintenant de trouver le vecteur $(\hat{m}, \hat{a}, \hat{\sigma})$ qui maximise la log-vraisemblance, soit :

$$\mathcal{L}(\hat{m}, \hat{a}, \hat{\sigma}, x_1 \dots x_n) = \max_{(m, a, \sigma)} \sum_{i=1}^n \ln f(\varepsilon_i)$$

Il est préférable de maximiser \mathcal{L} par rapport aux paramètres (ϕ_0, ϕ_1, ϕ_2) . Ce changement de variables⁵ permet d'assurer la positivité de la volatilité σ et de garantir une expression simple de \mathcal{L} et de son gradient. En utilisant les calculs de la partie précédente, on peut écrire :

$$\mathcal{L}(\phi_0, \phi_1, \phi_2, x_1 \dots x_n) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \phi_2 - \frac{1}{2} \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \phi_0)^2] e^{-\phi_2}$$

En dérivant $\mathcal{L}(\phi_0, \phi_1, \phi_2, x_1 \dots x_n)$ par rapport à ϕ_0, ϕ_1 et ϕ_2 on trouve :

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \phi_0} = \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \phi_0)] e^{-\phi_2} \\ \frac{\partial \mathcal{L}}{\partial \phi_1} = \sum_{i=1}^n [x_{i-1} (x_i - \phi_1 x_{i-1} - \phi_0)] e^{-\phi_2} \\ \frac{\partial \mathcal{L}}{\partial \phi_2} = \frac{n}{2} \left(\frac{1}{n} \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \phi_0)^2] e^{-\phi_2} - 1 \right) \end{cases} \quad (7)$$

On montre que dans ce cas simple la maximisation de la vraisemblance revient à l'utilisation de la méthode des moindres carrés. Ce résultat particulier est en fait général pour tous les processus $ARMA(p, q)$.

⁵Ce changement est possible car il est difféomorphisme continue.

3.2.4 Estimation par moindres carrés

Les paramètres du processus X_t^c peuvent être estimés par la méthode des moindres carrés MCO. Cette méthode repose sur la minimisation de la moyenne quadratique des bruits ε_t sur toute la trajectoire $[0, T]$. On montre dans la suite l'équivalence entre cette méthode et celle du maximum de vraisemblance. En effet, Les paramètres $(\hat{\phi}_0, \hat{\phi}_1, \hat{\phi}_2)$, qui maximisent la log-vraisemblance, vérifient :

$$\frac{\partial \mathcal{L}}{\partial \phi_0} = 0, \quad \frac{\partial \mathcal{L}}{\partial \phi_1} = 0 \quad \frac{\partial \mathcal{L}}{\partial \phi_2} = 0.$$

Donc,

$$\begin{cases} \sum_{i=1}^n [(x_i - \hat{\phi}_1 x_{i-1} - \hat{\phi}_0)] = 0 \\ \sum_{i=1}^n [x_{i-1}(x_i - \hat{\phi}_1 x_{i-1} - \hat{\phi}_0)] = 0 \\ e^{-\hat{\phi}_2} = \frac{1}{n} \sum_{i=1}^n [(x_i - \hat{\phi}_1 x_{i-1} - \hat{\phi}_0)^2] \end{cases}$$

Enfin,

$$\begin{bmatrix} \hat{\phi}_0 \\ \hat{\phi}_1 \end{bmatrix} = (A' A)^{-1} A' B, \quad \text{où } A = \begin{bmatrix} 1 & x_0 \\ \vdots & \vdots \\ 1 & x_{n-1} \end{bmatrix} \text{ et } B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad (8)$$

et

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n} \sum_{i=1}^n [(x_i - \hat{\phi}_1 x_{i-1} - \hat{\phi}_0)^2]$$

3.2.5 Analyse asymptotique

L'estimateur du maximum de vraisemblance du paramètre vectorielle $\theta = (\phi_0, \phi_1, \phi_2)$ est un estimateur sans biais, convergent et asymptotiquement normal. On analysera dans ce qui suit ces deux caractéristiques de l'estimateur EMV.

D'abord commençons par écrire l'expression de la log-vraisemblance :

$$\mathcal{L}(\theta, x_1 \dots x_n) = \sum_{i=1}^n \ln f(\theta, \varepsilon_i)$$

Par la loi forte des grands nombres et sous réserve d'intégrabilité, on a :

$$\frac{1}{n} \sum_{i=1}^n \ln f(\theta, \varepsilon_i) \xrightarrow{p.s} \mathbb{E}[\ln f(\theta, \varepsilon_1)]$$

On note θ_0 la vraie valeur du paramètre inconnu. Ce paramètre est identifiable, i.e θ_0 est l'unique maximum global de la fonction $\theta \mapsto \mathbb{E}[\ln f(\theta, \varepsilon_1)]$. Donc, sous des conditions de régularité, il existe une suite d'estimateurs de maximum de vraisemblance $\hat{\theta}_n$ convergeant vers θ_0 en probabilité :

$$\hat{\theta}_n = \operatorname{argmax}_{\theta} \left[\frac{1}{n} \mathcal{L}(\theta, x_1 \dots x_n) \right]$$

Ensuite, rappelons la définition de la matrice d'information de Fisher $\mathcal{I}(\theta_0)$ pour une observation ε_i :

$$\mathcal{I}(\theta_0) = -\mathbb{E} \left[\frac{\partial^2 \ln f(\theta_0, \varepsilon_i)}{\partial \theta \partial \theta'} \right] = \mathbb{E} \left[\frac{\partial \ln f(\theta_0, \varepsilon_i)}{\partial \theta} \frac{\partial \ln f(\theta_0, \varepsilon_i)}{\partial \theta'} \right]$$

On estime la matrice de Fisher par :

$$\hat{\mathcal{I}}(\hat{\theta}_0) = \frac{1}{n} \sum_{i=1}^n \left[\frac{\partial \ln f(\hat{\theta}_n, \varepsilon_i)}{\partial \theta} \frac{\partial \ln f(\hat{\theta}_n, \varepsilon_i)}{\partial \theta'} \right]$$

Cette notion permet de comparer l'information contenue dans une statistique. En effet, si la log-vraisemblance est deux fois continûment différentiable dans un voisinage ouvert de θ_0 , et si la matrice d'information de Fisher pour les observations $\mathcal{I}(\theta_0)$ existe et est inversible, alors une suite $\hat{\theta}_n$ de maxima locaux convergeant est telle que :

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{Loi} \mathcal{N}(0, \mathcal{I}^{-1}(\theta_0))$$

Enfin, on montre que la variance du maximum de vraisemblance atteint asymptotiquement la borne de **Cramer-Rao** qui est l'inverse de la l'information de Fisher. Cette borne est la variance minimale de tout estimateur sans biais.

Dans le cadre de cette étude on se limite au calcul de la matrice de Fisher dans le cas de la tendance constante :

$$\mathcal{I}(\phi_0, \phi_1, \phi_2) = \begin{bmatrix} e^{-\phi_2} & e^{-\phi_2} \frac{\phi_0}{1-\phi_1} & 0 \\ e^{-\phi_2} \frac{\phi_0}{1-\phi_1} & e^{-\phi_2} \left(\frac{e^{\phi_2}}{1-\phi_1^2} + \frac{\phi_0^2}{(1-\phi_1)^2} \right) & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$$

3.3 Le prix spot un facteur saisonnalisé

On traite dans cette partie le cas général dans lequel la tendance dépend du temps. On se place dans un premier temps sous la probabilité risqué-neutre \mathbb{Q} . Le processus $(X_t^v)_{t \geq 0}$ vérifie alors l'EDS suivante :

$$dX_t^v = a(m_t - X_t^v)dt + \sigma dW_t, \quad X_0^v = 0 \quad (9)$$

Cette EDS admet une unique solution et on peut démontrer que le processus $(X_t^v)_{t \geq 0}$ reste toutefois un processus gaussien.

3.3.1 Solution du modèle

On peut expliciter la solution de l'EDS (9) en utilisant la formule d'Itô. En effet, posons $Z_t = (X_t^v - a \int_0^t m_s e^{-a(t-s)} ds) e^{at}$, alors Z_t vérifie l'EDS : $dZ_t = \sigma e^{at} dW_t$. L'expression de X_t^v est :

$$X_t^v = a \int_0^t m_s e^{-a(t-s)} ds + \sigma \int_0^t e^{-a(t-s)} dW_s$$

La variable aléatoire X_t^v suit une loi normale $\mathcal{N}(M_{(a,\sigma,t)}, V_{(a,\sigma,t)})$ de densité $f_{(a,\sigma,t)}$.

Où,

$$\begin{aligned} M_{(a,\sigma,t)} &= a \int_0^t m_s e^{-a(t-s)} ds \\ V_{(a,\sigma,t)} &= \frac{\sigma^2}{2a} (1 - e^{-2at}) \\ f_{(a,\sigma,t)}(x) &= \frac{e^{-\frac{(x-M_{(a,\sigma,t)})^2}{2V_{(a,\sigma,t)}}}}{\sqrt{2\pi V_{(a,\sigma,t)}}} \end{aligned}$$

On remarque que le cas de la tendance variable ne diffère du cas constant qu'en terme de moyenne, la variance reste la même.

3.3.2 Modèle en temps discret

On utilise les mêmes notations que dans le cas constant. On considère alors une discrétisation sur un intervalle $[0, T]$ avec un pas de temps h . Pour alléger les formulations on note : $\ln F(0, t_i) = \ln F(0, ih) = f_i$, le logarithme du forward à la date ih . La formule explicite de X_t^v donne le schéma suivant :

$$X_i^v - \phi_1 X_{i-1}^v - \varphi_0^i = \varepsilon_i \quad (10)$$

Où,

$$\begin{aligned} \varphi_0^i &= a \int_{(i-1)h}^{ih} m_s e^{-a(ih-s)} ds \\ \phi_1 &= e^{-ah} \\ \sigma_\varepsilon &= \sqrt{\frac{1 - e^{-2ah}}{2a}} \sigma \\ \phi_2 &= \ln(\sigma_\varepsilon^2) \end{aligned}$$

L'expression de φ_0^i peut être simplifiée et se formule comme suite :

$$\varphi_0^i = f_i - \phi_1 f_{i-1} - \frac{\sigma_\varepsilon^2}{2} \left(\frac{1 + \phi_1^{2i-1}}{1 + \phi_1} \right)$$

Les variables aléatoires $\varepsilon_1, \dots, \varepsilon_n$ sont des bruits blancs gaussiens $BB(0, \sigma_\varepsilon^2)$. On remarque que X_t^v est un processus AR(1) avec drift variable. Dans le cas où m_s est constant dans le temps, on retrouve les formules de la partie précédente. On peut supposer que m_s est constant par morceaux dans chaque intervalle de temps $[ih, (i+1)h]$. L'expression de φ_0^i devient alors

$$\varphi_0^i = m_i(1 - e^{-ah})$$

3.3.3 La vraisemblance

On suppose que l'on dispose de n observations du logarithme du prix spot, sur un intervalle de temps $[0, T]$, à des instants discrets de pas de temps h . Pour un échantillon $(x_1 \dots x_n)$ de X_t^v , la log-vraisemblance s'écrit :

$$\mathcal{L}(a, \sigma, x_1 \dots x_n) = \sum_{i=1}^n \ln f(\varepsilon_i)$$

Il s'agit maintenant d'optimiser cette fonction par rapport à (a, σ) sous la contrainte (2). Il est préférable de maximiser \mathcal{L} par rapport aux paramètres (ϕ_1, ϕ_2) . En utilisant les calculs de la partie précédente on peut écrire :

$$\mathcal{L}(\phi_1, \phi_2, x_1 \dots x_n) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \phi_2 - \frac{1}{2} \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \varphi_0^i)^2] e^{-\phi_2}$$

$$\varphi_0^i = f_i - \phi_1 f_{i-1} - \frac{e^{\phi_2}}{2} \left(\frac{1 + \phi_1^{2i-1}}{1 + \phi_1} \right)$$

En dérivant $\mathcal{L}(\phi_1, \phi_2, x_1 \dots x_n)$ par rapport à (ϕ_1, ϕ_2) , on trouve :

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \phi_1} = \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \varphi_0^i)(x_{i-1} - f_{i-1} - \frac{e^{\phi_2}}{2} g_i)] e^{-\phi_2} \\ \frac{\partial \mathcal{L}}{\partial \phi_2} = \frac{n}{2} \left(\frac{1}{n} \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \varphi_0^i)^2] e^{-\phi_2} - (x_i - \phi_1 x_{i-1} - \varphi_0^i) \frac{1 + \phi_1^{2i-1}}{1 + \phi_1} \right) - 1 \end{cases} \quad (11)$$

Avec,

$$g_i = \frac{(2i-1)(1 + \phi_1)\phi_1^{2i-2} - (1 + \phi_1^{2i-1})}{(1 + \phi_1)^2}$$

La résolution de ce système fait appel à des méthodes numériques, basées sur l'utilisation de méthodes itératives (Newton, Gradient conjugué, ...). Le problème avec ces méthodes est qu'on ne peut pas savoir si les estimations maximisent localement ou globalement la fonction de vraisemblance et par conséquent le choix des valeurs initiales pour les paramètres du modèle est important.

3.3.4 Incidence du risque

On se place dans cette partie sous la probabilité historique \mathbb{P} . On suppose que **la prime de risque est constante** sur l'intervalle $[0, T]$. L'estimation du processus $(X_t^v)_{t \geq 0}$ passe par l'expression de la vraisemblance et de son gradient pour les trois paramètres (ϕ_1, ϕ_2, ϕ_3) . On rappelle le schéma du modèle :

$$X_i^v - \phi_1 X_{i-1}^v - \varphi_0^i - \phi_3 = \varepsilon_i \quad (12)$$

Où,

$$\phi_3 = \frac{\sigma \xi (1 - e^{-ah})}{a}$$

L'expression de la vraisemblance et de son gradient deviennent :

$$\mathcal{L}(\phi_1, \phi_2, \phi_3, x_1 \dots x_n) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \phi_2 - \frac{1}{2} \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \varphi_0^i - \phi_3)^2] e^{-\phi_2}$$

En dérivant $\mathcal{L}(\phi_1, \phi_2, \phi_3, x_1 \dots x_n)$ par rapport à ϕ_1, ϕ_2 et ϕ_3 on trouve :

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \phi_1} = \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \varphi_0^i - \phi_3)(x_{i-1} - f_{i-1} - \frac{e^{-\phi_2}}{2} g_i)] e^{-\phi_2} \\ \frac{\partial \mathcal{L}}{\partial \phi_2} = \frac{n}{2} (\frac{1}{n} \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \varphi_0^i - \phi_3)^2] e^{-\phi_2} - (x_i - \phi_1 x_{i-1} - \varphi_0^i - \phi_3) \frac{1 + \phi_1^{2i-1}}{1 + \phi_1} - 1) \\ \frac{\partial \mathcal{L}}{\partial \phi_3} = \sum_{i=1}^n [(x_i - \phi_1 x_{i-1} - \varphi_0^i - \phi_3)] e^{-\phi_2} \end{cases}$$

On remarque que l'introduction de la prime de risque permet de corriger l'estimation des résidus. En effet, si on suppose à tort qu'on est sous une probabilité risque neutre, la calibration du modèle du spot donne des estimations erronées des paramètres (a, σ) et donc aussi de celle des résidus. Normalement cette erreur apparaîtra dans la structure des résidus (moyenne, variance, stationnarité, homoscédasticité...).

3.4 Fiabilité de l'estimateur log-vraisemblance

Dans cette partie on teste la fiabilité de l'estimateur log-vraisemblance décrit dans les deux parties précédentes. Cette étape est primordiale puisqu'elle permet de valider notre approche et nos calculs.

3.4.1 Description de la méthode d'estimation

On commence par fixer une période T et un pas de temps h (le nombre d'observations est alors $n = \frac{T}{h}$).

Modèle constant

Après le choix d'un vecteur (m, a, σ) , on simule un échantillon $(x_1 \dots x_n)$

suivant la loi de X_t . La simulation de X_i se fait par récurrence sur le modèle discret (19). Ensuite, on utilise l'échantillon $(x_1 \dots x_n)$ pour calibrer le modèle Ornstien-Ulhenbeck (5). Enfin, on compare les paramètres estimés $(\hat{m}, \hat{a}, \hat{\sigma})$ aux vraies valeurs (m, a, σ) .

Modèle variable

Il s'agit, d'abord, de choisir un vecteur (a, σ) . Ensuite, de simuler un échantillon $(x_1 \dots x_n)$ suivant la loi du processus (X_t) . En effet, on se donne une courbe forward initiale $F(0, t)$ et on simule X_i par récurrence suivant le modèle discret (10). Puis, on utilise l'échantillon $(x_1 \dots x_n)$, ainsi simulé, pour calibrer le modèle Ornstien-Ulhenbeck (9). Enfin, on compare les paramètres estimés $(\hat{a}, \hat{\sigma})$ aux vraies valeurs (a, σ) .

Modèle à risque

On se donne, d'abord, une courbe forward historique $F(0, t)$ et un vecteur (a, σ, ξ) . On récupère alors une courbe forward risque neutre $F^Q(0, t)$. Puis, on simule un échantillon $(x_1 \dots x_n)$ suivant la loi du processus (X_t) avec la courbe $F^Q(0, t)$ en utilisant le modèle risque neutre (9). Puis, on utilise l'échantillon $(x_1 \dots x_n)$, ainsi simulé, et la courbe forward historique pour calibrer le modèle Ornstien-Ulhenbeck (9). Enfin, on compare les paramètres estimés $(\hat{a}, \hat{\sigma}, \hat{\xi})$ aux vraies valeurs (a, σ, ξ) .

3.4.2 Choix de la condition initiale

Le choix de la condition initiale est comme pour tout algorithme itératif d'optimisation déterminant dans la convergence de la méthode. On approche au mieux ϕ_0, ϕ_1, ϕ_2 et ϕ_3 en utilisant des approximations naïves.

Modèle constante

Pour estimer m , on propose d'utiliser la moyenne empirique de l'échantillon $(x_1 \dots x_n)$. L'estimateur proposé est alors :

$$\hat{m} = \frac{1}{n} \sum_{i=1}^n x_i$$

Il s'agit maintenant de trouver une estimation à ϕ_1 . On suppose alors que $\varepsilon_1 = \varepsilon_2 = 0$, on obtient deux valeurs approximatives de ϕ_1 et leur moyenne donne :

$$\hat{\phi}_1 = \frac{1}{2} \left(\frac{x_2 - \hat{m}}{x_1 - \hat{m}} - \frac{x_1 - \hat{m}}{\hat{m}} \right)$$

Tout naturellement on peut estimer ϕ_0 et σ_ε on donne les valeurs suivantes :

$$\hat{\phi}_0 = \hat{m}(1 - \hat{\phi}_1)$$

$$\hat{\sigma}_\varepsilon = \sqrt{\frac{1}{n} \sum_{i=1}^n [(x_i - \hat{\phi}_1 x_{i-1} - \hat{\phi}_0)^2]}$$

On obtient alors les conditions initiales suivantes :

$$\boxed{\phi_{0_{init}} = \hat{\phi}_0, \phi_{1_{init}} = \hat{\phi}_1, \phi_{2_{init}} = \ln \hat{\sigma}_\varepsilon^2}$$

Modèle variable

Le choix de la condition initiale dans le cas de la tendance variable est plus compliqué que dans le cas de la tendance constante. On suppose que m_t est constant, puis on donne comme valeurs initiales à ϕ_1 et ϕ_2 les valeurs estimées par MCO, du modèle supposé avec tendance constante. On donne alors les conditions initiales suivantes :

$$\boxed{\phi_{1_{init}} = \phi_{1_{MCO}}, \phi_{2_{init}} = \phi_{2_{MCO}}}$$

Modèle à risque

Dans la partie précédente on a fixé le choix de la condition initiale pour (ϕ_1, ϕ_2) . Il s'agit maintenant de trouver une initialisation pour le paramètre ϕ_3 . On peut estimer ϕ_3 par :

$$\hat{\phi}_3 = \frac{1}{n} \sum_{i=1}^n [x_i - \hat{\phi}_1 x_{i-1} - \hat{\phi}_0]$$

On donne comme conditions initiales les valeurs suivantes :

$$\boxed{\phi_{1_{init}} = \phi_{1_{MCO}}, \phi_{2_{init}} = \phi_{2_{MCO}}, \phi_{3_{init}} = \hat{\phi}_3}$$

3.4.3 Performances de l'estimation

Comparaison MCO-EMV

Pour comparer les deux méthodes de calibration (dans le cas constant), on a réalisé plusieurs estimations suivant les deux estimateurs moindres carrés et maximum de vraisemblance. On a calculé, ensuite, à chaque fois, l'écart absolu entre les valeurs données par les deux méthodes. Le nombre de simulations étant $k = 100$, la durée d'observations $T = 1 \text{ an}$, le pas de temps $h = 1 \text{ heure}$. Les vraies valeurs des paramètres sont : $m = 35$, $a = 100$, $\sigma = 5$. L'erreur EMV-MCO est obtenue comme moyenne des valeurs absolues des écarts relatifs aux deux méthodes sur k simulations :

$\frac{1}{km} \sum_{i=1}^k m_{EMV}^i - m_{MCO}^i $	7.3310^{-10}
$\frac{1}{ka} \sum_{i=1}^k a_{EMV}^i - a_{MCO}^i $	1.1510^{-8}
$\frac{1}{k\sigma} \sum_{i=1}^k \sigma_{EMV}^i - \sigma_{MCO}^i $	4.7310^{-9}

Les résultats suivants prouvent que les deux méthodes donnent des résultats très proches. De plus, on admettra la convergence de la méthode de vraisemblance et on négligera l'erreur numérique due à l'optimisation par gradient conjugué. Dans la suite on retiendra les estimations par maximum de vraisemblance.

Précision de l'estimateur

Le but de cette partie est de donner les caractéristiques de l'estimateur de maximum de vraisemblance avec deux périodes d'observations $T = 38 \text{ jours}$ et $T = 1 \text{ an}$ et avec un pas de temps $h = 1 \text{ jour}$. Les vraies valeurs des paramètres sont :

$$m_0 = 10, \quad a_0 = 255, \quad \sigma_0 = 10, \quad \xi_0 = 40.$$

Ensuite, on utilise les trois modèles développés $k = 100$ fois pour simuler un échantillon de X_i et estimer les paramètres de cet échantillon. Enfin, on synthétise sous forme de tableaux les résultats obtenus⁶ :

38 jours Day-Ahead :

Variabes	Modèle constant	Modèle variable	Modèle à risque
m	6.0010^{-4} (1.1%)	-	-
a	2.3010^{-3} (10.4%)	1.6010^{-2} (21%)	4.5010^{-2} (20.4%)
σ	2.3010^{-2} (12.2%)	7.610^{-3} (12.3%)	1.5010^{-2} (14.2%)
ξ	-	-	1.2010^{-1} (29.%)

365 jours Day-Ahead :

Variabes	Modèle constant	Modèle variable	Modèle à risque
m	2.5010^{-5} (0.4%)	-	-
a	4.4010^{-4} (7%)	1.6010^{-2} (11%)	1.4010^{-2} (11%)
σ	4.6010^{-3} (3.6%)	5.1010^{-3} (4.4%)	2.6010^{-4} (4.7%)
ξ	-	-	1.0010^{-2} (7.6%)

Ces résultats montrent que les estimateurs \hat{m} , \hat{a} , $\hat{\sigma}$ et $\hat{\xi}$ sont sans biais et convergents. De plus, on a une quantification de la variance de ces estimateurs. On remarque particulièrement que l'écart-type relatif est inférieur à 11% pour tous les estimateurs pour une période $T = 1 \text{ an}$.

Biais dû au risque

La prime de risque représente un problème reconnu en calibration et pricing d'actifs. On teste dans la suite l'influence de cette grandeur sur la calibration d'un modèle variable -risque neutre-. Pour cela, on se donne une courbe

⁶La valeur correspond à l'erreur relative : $\frac{|v-\hat{v}|}{\hat{v}}$ et entre parenthèses l'écart-type relatif en pourcentage $\sigma_r(v) = \frac{\sigma(\hat{v})}{\hat{v}}\%$.

forward initiale risque neutre $F^Q(0, t)$ et une prime de risque ξ constante. On récupère alors une courbe historique $F(0, t) = F^Q(0, t)e^{\frac{\sigma}{a}\xi}$. La courbe $F^Q(0, t)$ est utilisée pour simuler un échantillon de X_t , par contre la courbe $F(0, t)$ est utilisée pour calibrer le modèle variable. Les vraies valeurs des paramètres sont :

$$a = 127.41, \sigma = 4.8, \xi = -73.92$$

La calibration sur le spot simulé donne alors les résultats suivants :

$$a = 8.18, \sigma = 4.46$$

Ces résultats sont à comparer avec les données de la partie "calibration du prix spot" entre l'estimation "tendance variable -risque neutre" et "tendance variable -prime de risque". On remarque que la valeur de a est très faible car le retour à la moyenne du processus mean-reverting n'as pas lieu, ce qui est due à un décalage permanent du processus du prix de sa moyenne, via la prime de risque.

Courbe du logarithme du prix spot

Les Figures 6 représentent l'évolution du logarithme du prix spot dans un modèle de Schwartz un facteur **constant** suivant différentes valeurs des paramètres (m, a, σ) :

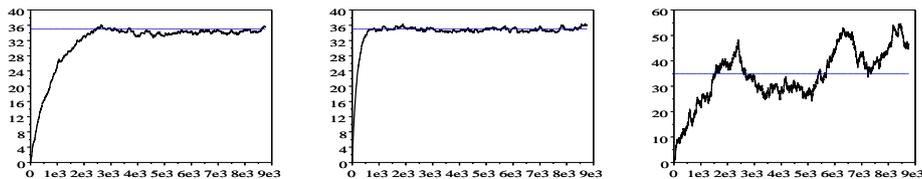


Figure 6: Evolution du processus X_t en fonction du temps pour (m, a, σ) égale à : $(35, 10, 5)$ à gauche, $(35, 50, 5)$ au centre et $(35, 10, 50)$ à droite.

Le paramètre σ influe beaucoup sur l'allure de la courbe du logarithme du spot et cela en rajoutant des fluctuations d'autant plus significatives que la volatilité est grande. Ensuite, le paramètre a représente bien une vitesse de retour à la moyenne, qui est d'autant plus grande que ce paramètre est grand. Enfin le paramètre m est une moyenne à long terme autour de laquelle la courbe du logarithme du prix fluctue.

3.4.4 Structure des résidus

La structure des résidus donne une information principale sur la qualité des estimations. Les résidus simulés dans notre cas sont de nature gaussiennes iid. On vérifie dans ce qui suit que les estimations n'ont pas changé cette structure et que les résidus estimés $\hat{\varepsilon}_i$ suivent bien un processus de bruit blanc.

Test d'autocorrélation

Si les résidus $(\varepsilon_t)_{t \geq 0}$ obéissent à un bruit blanc, il ne doit pas exister d'autocorrélation dans cette série temporelle. Il existe différents tests qui permettent de répondre à cette demande (Box-Pierce, Ljung-Box, Durbin-Watson...). On choisit d'appliquer le test de **Box-Pierce**, sa statistique de test s'écrit :

$$BP(K) = n \sum_{k=1}^K \rho_k^2(\hat{\varepsilon}_t)$$

Où, $\rho_k(\hat{\varepsilon}_t)$ est le coefficient d'autocorrélation d'ordre k . Le nombre maximal de retards retenu est $K = \frac{n}{3}$. On peut donner une estimation à ce coefficient :

$$\rho_k = \frac{\frac{1}{n} \sum_{i=1}^{n-k} (\varepsilon_i - \bar{\varepsilon})(\varepsilon_{i+k} - \bar{\varepsilon})}{\frac{1}{n} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2}$$

Sous l'hypothèse nulle H_0 d'absence d'autocorrélation entre les résidus, la statistique $BP(K)$ suit une loi de Khi-deux à $(K-1)$ degrés de liberté. En revanche si cette hypothèse n'est pas vraie, alors $BP(K) \rightarrow +\infty$ quand $K \rightarrow +\infty$. On rejette donc H_0 si on observe des valeurs anormalement grandes pour $BP(K)$, par exemple supérieures à un z donné. Il existe donc une probabilité α de rejeter à tort (risque de première espèce), qui dépend de z :

$$\mathbb{P}(BP(K) \geq z) = \mathbb{P}(\chi(K-1) \geq z) = \alpha$$

Dans le cadre de ce test on commence par calculer $\alpha_K = \mathbb{P}(\chi(K-1) \geq \widehat{BP}(K))$. On le compare ensuite à 5%, si ce quantile est plus faible on rejette H_0 .

Exemple : le modèle constant

On obtient pour une trajectoire du logarithme du prix simulé les valeurs suivantes :

Durée	Quantile	Référence	Moyenne	Variance
1 mois	87.5%	5%	$1.8e^{-8}$	1
1 an	100%	5%	$2.7e^{-10}$	1

Le test prouve que les résidus sont non autocorrélés. Cette conclusion apparaît aussi dans l'histogramme de ε_t et dans l'évolution de la fonction d'autocorrélation ρ_h (voir la Figure 7) :

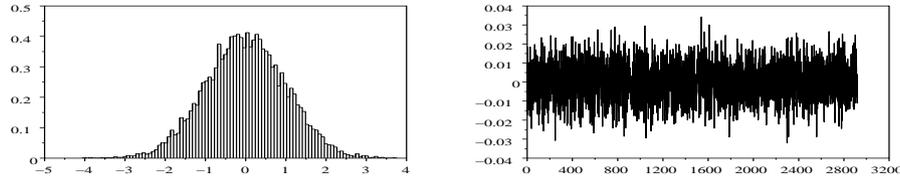


Figure 7: Histogramme des résidus ε_t (à gauche) et fonction d'autocorrélation ρ_h en fonction du retard du temps (à droite).

Test de normalité

Pour vérifier si le processus des résidus $(\varepsilon_t)_{t \geq 0}$ est un bruit blanc gaussien, plusieurs tests peuvent être utilisés, mais le test le plus courant est celui de **Jarque Bera**. Ce dernier est fondé sur la notion de skewness (moment d'ordre 3 et asymétrie) et de kurtosis (moment d'ordre 4 et queue de distribution). Soit U_k le moment empirique d'ordre k du processus $(\varepsilon_t)_{t \geq 0}$:

$$U_k = \frac{1}{n} \sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^k$$

Les coefficients de Skewness et de Kurtosis sont alors définis par

$$S = \frac{U_3}{U_2^3} \quad K = \frac{U_4}{U_2^2}$$

La statistique de Jarque Bera s'écrit :

$$JB(n) = \frac{n-1}{6} \left[S^2 + \frac{1}{4}(K-3)^2 \right]$$

Sous l'hypothèse nulle H_0 de normalité, la statistique $JB(n)$ suit une loi Khi-deux à deux degrés de liberté. Dans le cadre de ce test on commence par calculer $\alpha_K = \mathbb{P}(\chi(2) \geq \widehat{JB}(n))$. On le compare ensuite à 5%, si ce quantile est plus faible on rejette H_0 .

Exemple : le modèle constant

On obtient pour une trajectoire du logarithme du prix les valeurs suivantes :

Durée	Quantile	Référence	Skewness	Kurtosis
1 mois	41.3%	5%	$1.0e^{-3}$	3.36
1 an	94.6%	5%	$1.0e^{-4}$	2.96

3.5 Calibration du prix spot

3.5.1 Problématique

L'estimation sur le prix spot est primordiale dans la valorisation d'actifs électriques. On dispose chaque jour, à 11 : 00 *heure*, de 24 cotations pour les 24 livraisons unitaires d'une heure du jour suivant. Les 24 prix sont alors fixés d'avance et à la même heure. De ce fait on ne peut utiliser les modèles avec un pas horaire. Il existe deux solutions à ce problème :

- Utiliser un modèle avec un pas journalier Day-Ahead.
- Utiliser 1 modèle horaire pour chaque jour de la journée.

La première solution sera adoptée dans la suite. On considère deux périodes d'observations $T_1 = 38 \text{ jours}$ (01/04/2004 - 08/05/2004) et $T_2 = 1 \text{ an}$ (01/01/2002 - 31/12/2002), le pas de temps étant 1 *jour*. Ensuite, on récupère l'historique du prix spot sur le site de PowerNext. Puis on reconstruit la courbe forward initiale $F(0, t)$ avec l'outil VisualFuture d'EDF. Enfin, on utilise ces données pour calibrer un modèle factoriel au choix. Dans la suite on présente les résultats pour les différents modèles. On représente dans la Figure 8 le prix spot Day-Ahead, son logarithme et la courbe forward pour $[0, T_1]$:

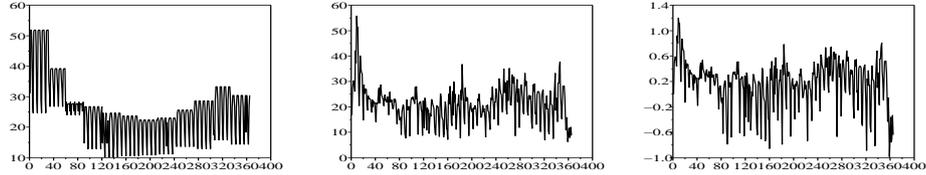


Figure 8: Evolution du forward (à gauche), du prix spot Day-Ahead (au centre) et de son logarithme (a droite) sur $[0, T_1]$.

3.5.2 Tendence constante

La calibration du prix spot sur un modèle un facteur gaussien, constant et risque-neutre donne les résultats suivants :

VARIABLES	Valeurs 38 jours	Valeurs 1 an
m	-1.10	$1.73e^{-1}$
a	443.50	231.70
σ	9.95	7.80

Le test d'absence d'autocorrélation de Box-Pierce donne les résultats suivants :

Grandeurs	Valeurs 38 jours	Valeurs 1 an
Probabilité de rejet	0	1
Moyenne résidus	$6.76e^{-10}$	$1.45e^{-10}$
Variance résidus	1.00	1.00

Le test de normalité de Jarque Bera donne les résultats suivants :

Grandeurs	Valeurs 38 jours	Valeurs 1 an
Probabilité de rejet	1	1
Skewness	1.92	0.23
Kurtosis	4.37	3.50

Ces résultats prouvent que le calage des prix sur un modèle un facteur gaussien constant génère des résidus qui sont non normaux et autocorrélés. L'hypothèse de bruit blanc sur les résidus est donc rejetée. Ces résultats correspondent aux résultats donnés par l'outil ED-Visual-Analysis d'EDF.

3.5.3 Tendance variable - risque neutre -

La calibration du prix spot sur un modèle un facteur gaussien, variable et risque-neutre donne les résultats suivants :

Variabes	Valeurs 38 jours	Valeurs 1 an
a	8.99	8.30
σ	3.70	4.42

Le test d'absence d'autocorrélation de Box-Pierce donne les résultats suivants :

Grandeurs	Valeurs 38jours	Valeurs 1 an
Probabilité de rejet	0.74	1
Moyenne résidus	$-2.34e^{-2}$	$-3.38e^{-2}$
Variance résidus	1	0.99

Le test de normalité de Jarque Bera donne les résultats suivants :

Grandeurs	Valeurs 38 jours	Valeurs 1 an
Probabilité de rejet	1	0.93
Skewness	0.14	$2.1e^{-2}$
Kurtosis	5.97	3.51

Ces résultats prouvent que le calage des prix sur un modèle un facteur gaussien variable génère des résidus qui sont autocorrélés. L'hypothèse de bruit blanc sur les résidus est donc rejetée. Cette calibration donne une vitesse de retour à la moyenne très faible due principalement à la prime de risque.

3.5.4 Tendance variable - prime de risque -

La calibration du prix spot sur un modèle un facteur gaussien, variable et prime de risque constante donne les résultats suivants :

Variabes	Valeurs 38 jours	Valeurs 1 an
a	601.46	127.41
σ	5.29	4.8
ξ	-367.44	-73.92

Le test d'absence d'autocorrélation de Box-Pierce donne les résultats suivants :

Grandeurs	Valeurs 38 jours	Valeurs 1 an
Probabilité de rejet	0	0
Moyenne résidus	$-1.44e^{-5}$	$-6.71e^{-8}$
Variance résidus	1	1

Le test de normalité de Jarque Bera donne les résultats suivants :

Grandeurs	Valeurs 38 jours	Valeurs 1 an
Probabilité de rejet	1	0.98
Skewness	3.31	$8.10e^{-2}$
Kurtosis	8.05	3.50

Ces résultats prouvent que le calage des prix sur un modèle un facteur gaussien variable avec une prime de risque génère des résidus qui sont non normaux et autocorrélés. L'hypothèse de bruit blanc sur les résidus est donc rejetée. Toutefois, on remarque que cette calibration donne une meilleure vitesse de retour à la moyenne.

3.5.5 Analyse des résultats

Les calibrations effectués sur données historiques donnent des résultats mitigés.

Premièrement, l'estimation sur le spot, sans désaisonnalisation, avec un modèle à tendance constante suppose une forme particulière à la courbe forward :

$$\ln\left(\frac{F(0, t)}{F(0, 0)}\right) = m(1 - e^{-at}) + \frac{\sigma^2}{4a}(1 - e^{-2at})$$

Dans le cas de la désaisonnalisation, la calibration dépend de la manière de déterminer la partie saisonnière D_t . En pratique, en cas de non désaisonnalisation, il faut absolument enlever les week-end sur l'échantillon du spot, travailler en Day-Ahead et considérer un pas de temps de cotation $h = \frac{1}{252}$. Deuxièmement, la calibration avec un modèle variable suppose que le marché est risque neutre. Cette hypothèse est très restrictive (voir la partie " Biais dû au risque "). Pour utiliser ce modèle il faut impérativement enlever la prime de risque à la courbe forward initiale. De plus le calage doit prendre en compte les week-end et un pas de temps de cotations $h = \frac{1}{365}$.

Troisièmement, la calibration d'un modèle à risque, représente un très bon candidat parmi les trois présentés dans cette étude. Cette calibration, intègre les deux points forts des deux calages précédents : la prime de risque comme le calage constant et la variabilité de la courbe forward comme le calage variable. Dans ce modèle, il faut considérer les week-ends et le pas de temps est doit être $h = \frac{1}{365}$.

4 Le modèle hyperbolique généralisé

4.1 Approche mathématique

Notations 1 On adopte, dans la suite, les notations suivantes :

- Toutes les lois sont des lois sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
- On se place sur un espace filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$.
- Si μ est une loi, on note $\hat{\mu}$ sa transformée de Fourier.

4.1.1 Lois indéfiniment divisibles

Les lois indéfiniment divisibles ont été étudiées au cours des années 1930. D'abord en 1932, par A.N. Kolmogorov en caractérisant les lois de variance finie par la forme du logarithme de leur fonction caractéristique. Ensuite en 1934, par P. Lévy en généralisant la formule donnée par A.N. Kolmogorov pour toutes les lois indéfiniment divisibles. La formule dite de Lévy-Khintchine est une formule équivalente à la formule donnée par P. Lévy en 1934 et qui caractérise les processus dit de Lévy.

Définitions et remarques

Définition 1 Une loi μ est dite **indéfiniment divisible** s'il existe une suite de probabilités $(\mu_n)_{n \in \mathbb{N}^*}$ telle que

$$\forall n \in \mathbb{N}^*, \quad \mu = \mu_1 \otimes \dots \otimes \mu_n$$

C'est-à-dire si pour tout $n \in \mathbb{N}^*$, μ est la loi de la somme de n variables aléatoires i.i.d. à valeurs dans \mathbb{R} .

Il est facile de voir que les lois normales, exponentielles et de Poisson sont des lois indéfiniment divisibles. Par ailleurs, une loi de Bernoulli de paramètre $p \in]0, 1[$ n'en est pas une.

Remarque :

Une variable aléatoire L est infiniment divisible si pour tout entier N , L peut se décomposer comme la somme de N v.a indépendantes et de même loi,

$$L = L_1 + L_2 + \dots + L_N, \quad i.i.d.$$

Définition 2 Une **mesure de Lévy** est une mesure positive ν sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\nu(0) = 0$ et $\int_{\mathbb{R}} (|x|^2 \wedge 1) \nu(dx) < +\infty$ ⁷.

Dans la littérature, on trouve d'autres caractérisations des mesures de Lévy, caractérisations équivalentes à la définition précédente. La mesure de Lévy est une notion indispensable pour exprimer la formule de Lévy-Khintchine.

Formule de Lévy-Khintchine

Théorème 1 Une loi μ est une loi indéfiniment divisible si et seulement si sa transformée de Fourier s'écrit sous la forme $\hat{\mu} = e^\psi$:

$$\forall u \in \mathbb{R}, \quad \psi(u) = imu - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}} (e^{iux} - 1 - iux\mathbf{1}_{\{|x|<1\}}) \nu(dx) \quad (13)$$

Avec $m \in \mathbb{R}$, $\sigma \geq 0$ et ν une mesure de Lévy. Réciproquement, pour ψ défini par (13), il existe une loi μ indéfiniment divisible telle que $\hat{\mu} = e^\psi$.

⁷ \wedge est l'opérateur inf : $x \wedge y = \inf(x, y)$.

Preuve :

La preuve de ce théorème consiste à montrer que toute loi indéfiniment divisible est limite presque sûre de loi gaussienne et Poisson composée (voir [App03] pp. 7).

On constate que l'écriture de $\hat{\mu}$, en fonction du triplet de Lévy-Khintchine (m, σ, ν) , est unique. Particulièrement si μ est une loi gaussienne de moyenne e et variance v alors $m = e$, $\sigma = \sqrt{v}$, $\nu = 0$.

Remarques :

La transformée de Fourier d'une loi indéfiniment divisible ne s'annule pas et admet un logarithme continu.

4.1.2 Les processus de Lévy

Les processus de Lévy ont été nommés par Paul Lévy, qui a établi leur relation avec les lois indéfiniment divisibles. Théoriquement, ils représentent une sous classe importante des processus Markovien. Ils sont un exemple de semi-martingale (soit une sur-martingale ou une sous-martingale) pour lesquelles les calculs stochastiques sont applicables. Expérimentalement, ils représentent des lois à queues épaisses qui prennent en compte la liquidité et le risque du marché.

Définition et propriétés

Définition 3 On appelle **processus de Lévy** un processus stochastique L_t à valeurs dans \mathbb{R} , adapté avec $L_0 = 0$ (presque sûrement), à accroissements indépendants et stationnaires dont les trajectoires sont continues en probabilité. Ce qui signifie que :

1. *Indépendance* : si $0 \leq s < t$, $L_t - L_s$ est indépendant de la tribu \mathcal{F}_s .
2. *Stationnarité* : $L_t - L_s$ a la même loi que L_{t-s} pour $0 \leq s < t$.
3. *Continuité* : L_t est continu en probabilité $\lim_{t \rightarrow s} L_t \stackrel{\mathbb{P}}{=} L_s$.

Les processus de Lévy sont une sous famille des processus P.A.I.S. On peut citer deux processus de Lévy bien connus, le mouvement brownien W_t et le processus de Poisson N_t .

Remarques :

La continuité en probabilité n'implique pas la continuité des trajectoires. Les seuls processus de Lévy dont les trajectoires sont continues sont les mouvements browniens.

Proposition 1 Soit L_t un processus de Lévy, il existe une unique (à indistingabilité près) modification \tilde{L}_t de L_t qui est càdlàg (continue à droite

ayant des limites à gauche). De plus \tilde{L}_t est un processus de Lévy.

Cette proposition nous permet de supposer dorénavant qu'un processus de Lévy est càdlàg. Les discontinuités d'un processus de Lévy sont donc par sauts pures.

Transformée de Fourier et Laplace

Proposition 2 Soit L_t un processus de Lévy à valeurs dans \mathbb{R} . Il existe une unique fonction $\psi : \mathbb{R} \rightarrow \mathbb{C}$ continue telle que $\psi(0) = 0$ et telle que

$$\forall t \geq 0, \quad \forall u \in \mathbb{R}, \quad \mathbb{E}[\exp(iuL_t)] = \exp(t\psi(u)) \quad (14)$$

La fonction ψ est nommée l'**exposant caractéristique** de la loi de L_1 .

Preuve :

On commence par démontrer que la loi de L_1 est indéfiniment divisible. En effet,

$$\forall n \in \mathbb{N}^*, \quad L_1 = (L_{\frac{1}{n}} - L_{\frac{n-1}{n}}) + \dots + (L_{\frac{1}{n}} - L_0)$$

La Formule de Lévy-Khintchine permet d'écrire : $\mathbb{E}[\exp(iuL_1)] = \exp(\psi(u))$. D'autres parts, on a :

$$\begin{aligned} \mathbb{E}[\exp(iuL_t)] &= \mathbb{E}[\exp(iu(L_t - L_s))] \cdot \mathbb{E}[\exp(iuL_s)] \\ &= \mathbb{E}[\exp(iuL_{t-s})] \cdot \mathbb{E}[\exp(iuL_s)] \end{aligned}$$

Cette stabilité par multiplication est seul vérifiée par les fonctions $t \mapsto \exp(ta)$, la condition sur a est obtenue pour $t = 1$ et permet de conclure que $a = \psi(u)$.

Proposition 3 Soit L_t un processus de Lévy. Pour toute fonction f suffisamment régulière telle que $|\operatorname{Re}(f)|$ borné,

$$\mathbb{E}[\exp(\int_0^t f(s)dL_s)] = \exp(\int_0^t \varphi \circ f(s)ds) \quad (15)$$

La fonction φ représente la fonction génératrice des cumulants de L_1 . Elle permet de formuler l'expression de la transformée de Laplace : $\mathbb{E}[\exp(uL_t)] = \exp(t\varphi(u))$.

Preuve :

En utilisant la formule (14) pour les fonction en escalier, sachant que toute fonction continue est limite uniforme d'une suite de ce type de fonctions, puis en passant à la limite.

Remarque :

1. Les deux transformées sont un résultat direct de la formule (13). En effet, on peut exprimer φ et ψ en fonction du triplet de Lévy (m, σ, ν) :

$$\psi(u) = imu - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}} (e^{iux} - 1 - iux\mathbb{1}_{\{|x|<1\}})\nu(dx)$$

$$\varphi(u) = mu + \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}} (e^{ux} - 1 - ux\mathbb{1}_{\{|x|<1\}})\nu(dx)$$

2. Dans cette représentation le couple (m, σ, ν) à une interprétation financière. En effet, en calculant les dérivés de la transformée de Laplace on obtient la moyenne et la variance de la variable L_1 , soit :

$$\mathbb{E}[L_1] = m + \int_{|x|\geq 1} x\nu(dx) \quad \text{et} \quad \text{Var}[L_1] = \sigma^2 + \int_{\mathbb{R}} x^2\nu(dx)$$

Caractère Markovien

Rappel :

Soit (E, \mathcal{E}) un espace mesurable. On appelle processus de Markov (homogène) à valeurs dans (E, \mathcal{E}) un terme $X = (\Omega, \mathcal{F}_t, \mathcal{F}, X_t, (\mathbb{P}_x)_{x \in E})$ où

- (i) $\forall x \in E$, \mathbb{P}_x est une probabilité sur (Ω, \mathcal{F}) ,
- (ii) $\forall A \in \mathcal{F}$, $x \mapsto \mathbb{P}_x(A)$ est \mathcal{E} -mesurable,
- (iii) $\forall t \geq 0$, X_t est une v.a \mathcal{F}_t -mesurable à valeurs (E, \mathcal{E}) ,
- (iv) $\forall x \in E$, $\mathbb{P}_x(X_0 = x) = 1$,
- (v) $\forall t, h \geq 0$, $\forall x \in E$, $\forall \Gamma \in \mathcal{E}$,

$$\mathbb{P}_x(X_{t+h} \in \Gamma | \mathcal{F}_t) = \mathbb{P}_{X_t}(X_h \in \Gamma), \quad \mathbb{P}_x \text{ p.s}$$

La fonction de transition de X est donnée par : $P_t f(x) = \mathbb{E}_x[f(X_t)]$. On dit que X admet la propriété de semi-groupe si et seulement si

$$P_{t+s} = P_s P_t = P_t P_s$$

Proposition 4 Soit L_t un processus de Lévy de loi μ_t . Alors L_t est un processus de Markov homogène dont le semi-groupe de transition P_t , semi-groupe d'opérateurs, est défini par :

$$P_t f(x) = \mathbb{E}_x(f(L_t)) = \int_{\mathbb{R}} f(x+y)\mu_t(dy)$$

Pour toute fonction f mesurable, bornée non négative, pour tout $t \geq 0$, pour tout $x \in \mathbb{R}$.

Les processus de Lévy sont des processus markoviens, qui admettant une caractérisation plus fine grâce à la formule (13).

Formule d'Itô discontinue

Nous allons maintenant introduire la formule d'Itô pour une semi-martingale (sous- ou sur- martingale) càdlàg L_t pour une fonction $f(t, x)$ suffisamment régulière.

Rappel :

Soient X_t un processus d'Itô et f de classe $C^{1,2}(\mathbb{R}, \mathbb{R})$. Alors $f(t, X_t)$ est un processus d'Itô et on a

$$f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s) dX_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s) d\langle X \rangle_s$$

Théorème 2 Soit L_t une semi-martingale càdlàg et $f(t, x)$ une fonction deux fois continûment différentiable, on a :

$$\begin{aligned} f(t, L_t) &= f(0, L_0) + \int_0^t \frac{\partial f}{\partial s}(s, L_{s-}) ds + \int_0^t \frac{\partial f}{\partial x}(s, L_{s-}) dL_s + \frac{\sigma^2}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, L_{s-}) ds \\ &\quad + \int_0^t \int_{\mathbb{R}} \{f(s, L_{s-} + y) - f(s, L_{s-}) - \frac{\partial f}{\partial x}(s, L_{s-}) y\} \mu(ds, dy) \end{aligned}$$

Où, $\mu(ds, dy)$ est la mesure de saut associée au processus et $L_{s-} = \lim_{t \rightarrow s, t < s} L_t$.

La mesure de saut compte le nombre de sauts de L_t en temps et a pour compensateur⁸ $\nu(dy)ds$. La formule d'Itô correspond à un processus de Lévy continu ($\nu = 0$). Pour la définition de l'intégrale sur un processus de Lévy voir [App03] pp. 54.

Décomposition de Lévy-Itô

Dans la suite, on apporte une analyse particulière à la discontinuité du processus de Lévy, grâce à la décomposition de Lévy-Itô on peut approcher un processus de Lévy par un processus de Poisson composé et un mouvement brownien avec un drift.

Rappel :

Un processus de Poisson N_t d'intensité λ est un processus de Lévy dont le triplet de Lévy est $(\lambda, 0, \delta_1)$ ⁹, il a une loi donnée par :

$$\mathbb{P}(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

⁸ Le compensateur d'une semi-martingale Z est l'unique processus à variation fini \bar{Z} tel que $Z - \bar{Z}$ est une martingale.

⁹ δ_1 est la fonction Dirac au point 1.

Soit $(Z_n, n \in \mathbb{N})$ une séquence de v.a i.i.d de loi μ_Z , le processus de Poisson composé Y_t (de Z_n et N_t) est défini par

$$Y_t = \sum_{i=1}^{N_t} Z_i$$

L'exposant caractéristique de Y_t et N_t est donné par :

$$\psi_{N_t}(u) = \lambda(e^{iu} - 1), \quad \psi_{Y_t}(u) = \int_{\mathbb{R}} (e^{iux} - 1) \lambda \mu_Z(dx)$$

Soit L_t un processus de Lévy, l'amplitude du saut à l'instant t est donnée par

$$\Delta L_t = L_t - L_{t-}$$

On se fixe $\Lambda \in \mathcal{B}(\mathbb{R})$ tel que $0 \notin \bar{\Lambda}$, et on définit les temps d'arrêt T_i^Λ , $i = 0 \dots n$ avec

$$T_0^\Lambda = 0 \\ T_i^\Lambda = \inf\{t > T_{i-1}^\Lambda, \Delta L_t \in \Lambda\}, \quad i = 1 \dots n$$

On montre que les variables aléatoires $T_i^\Lambda - T_{i-1}^\Lambda$, ont une loi exponentielle (voir [App03] pp. 44). De ce fait, le processus aléatoire N_t^Λ

$$N_t^\Lambda = \sum_{0 \leq s \leq t} \mathbf{1}_\Lambda(\Delta L_s) = \sum_{n=1}^{+\infty} \mathbf{1}_{\{T_n^\Lambda \leq t\}}, \quad \text{pour } t \geq 0$$

est bien un processus de Poisson d'intensité $\kappa = \nu(\Lambda) = \mathbb{E}[N_1^\Lambda]$. La mesure de saut $\mu(ds, dx)$ est, en fait, égale à la mesure $(t, \Lambda) \mapsto N_t^\Lambda$:

$$\mu(ds, dx) = N_{ds}^{dx} = \sum_{s \leq t \leq s+ds} \mathbf{1}_{[x, x+dx]}(\Delta L_t)$$

La décomposition de Lévy-Itô, du processus de Lévy L_t , à l'aide du développement fait ci-dessus et de la formule (13), se formule par l'égalité suivante :

$$L_t = mt + \sigma W_t + \int_{|x| < 1} x(N_t^{dx} - t\nu(dx)) + \int_{|x| \geq 1} x N_t^{dx}$$

Dans cette représentation on retrouve le triplet de Lévy (m, σ, ν) , les différents parties sont :

✗ Le terme de drift :

$$m = \mathbb{E}[L_1 - \int_{|x| \geq 1} x N_1^{dx}] = \mathbb{E}[L_1] - \int_{|x| \geq 1} x \nu(dx)$$

✗ Le terme continue martingale :

W_t est un mouvement brownien standard.

X Le terme discontinue martingale :

$\int_{|x|<1} x(N_t^{dx} - t\nu(dx))$ est la partie des sauts inférieurs à 1 de L_t compensée.

X Le terme discontinue de saut pure :

$\int_{|x|\geq 1} xN_t^{dx}$ est un processus de Poisson composé des sauts supérieurs ou égale à 1.

Le processus de Poisson composé $\int_{|x|\geq 1} xN_t^{dx}$ est la composition d'un processus de poisson $N_t \sim \mathcal{P}(\kappa)$ et d'une suite de v.a de loi $Z_i \sim \mu_Z$, avec :

$$\kappa = \int_{|x|\geq 1} \nu(dx), \quad \mu_Z = \frac{1}{\kappa} \nu \mathbf{1}_{\{|x|\geq 1\}}$$

Il faut noter que κ caractérise seulement le nombre des sauts et que la loi μ_Z caractérise leurs amplitudes.

Exemples ¹⁰

• La loi Normale Inverse Gaussienne \mathcal{NIG} :

$$\kappa = \int_{|x|\geq 1} \frac{\alpha\delta}{\pi|x|} e^{\beta x} K_1(\alpha|x|) dx, \quad \mu_Z(x) = \frac{1}{\kappa} \frac{\alpha\delta}{\pi|x|} e^{\beta x} K_1(\alpha|x|) \mathbf{1}_{\{|x|\geq 1\}}$$

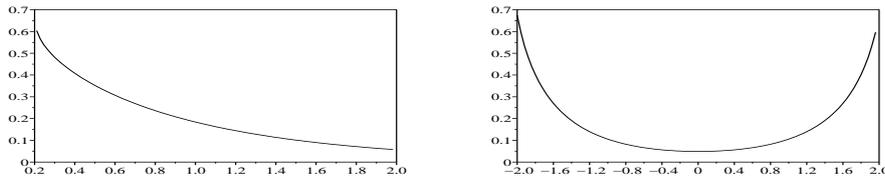


Figure 9: Variation de l'intensité de sauts pour la loi \mathcal{NIG} en fonction de α (à gauche), et β (à droite).

• La loi Variance Gamma \mathcal{VG} :

¹⁰Voir la suite pour la définition des processus \mathcal{NIG} et \mathcal{VG} .

$$\kappa = \int_{|x| \geq 1} \frac{\lambda}{|x|} e^{-\alpha|x| - \beta x} dx, \quad \mu_Z(x) = \frac{1}{\kappa} \frac{\lambda}{|x|} e^{-\alpha|x| - \beta x} \mathbf{1}_{\{|x| \geq 1\}}$$

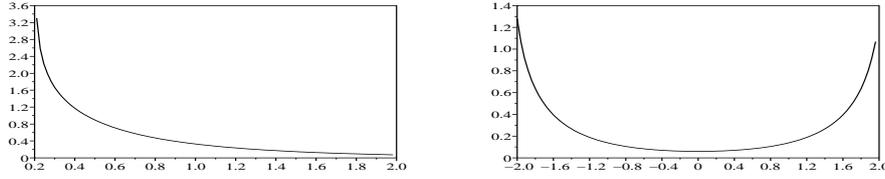


Figure 10: Variation de l'intensité de sauts pour la loi \mathcal{VG} en fonction de α (à gauche), et β (à droite).

On voit bien que les paramètres λ et δ sont proportionnels à l'intensité des sauts κ (le nombre de sauts supérieurs à 1). On se reporte à la Figure 9 et 10 pour remarquer que cette intensité augmente avec $|\beta|$ et diminue avec α . Cette conclusion est à rapprocher avec l'étude des résidus.

En général :

De manière plus générale, on peut se donner $0 < \varepsilon \ll 1$, on approche un processus, \mathcal{NIG} ou \mathcal{VG} , L_t par un processus de Poisson composé de saut supérieurs à ε :

$$L_t \cong \mu t + \sum_{i=1}^{N_t^\varepsilon} Z_i^\varepsilon, \quad N_t^\varepsilon \sim \mathcal{P}(\kappa_\varepsilon)$$

✘ La loi $\mathcal{NIG}(\alpha, \beta, \delta, \mu)$:

$$\kappa_\varepsilon = \int_{|x| \geq \varepsilon} \frac{\alpha \delta}{\pi |x|} e^{\beta x} K_1(\alpha |x|) dx, \quad \mu_{Z^\varepsilon}(x) = \frac{1}{\kappa_\varepsilon} \frac{\alpha \delta}{\pi |x|} e^{\beta x} K_1(\alpha |x|) \mathbf{1}_{\{|x| \geq \varepsilon\}}$$

✘ La loi $\mathcal{VG}(\lambda, \alpha, \beta, \mu)$:

$$\kappa_\varepsilon = \int_{|x| \geq \varepsilon} \frac{\lambda}{|x|} e^{-\alpha|x| - \beta x} dx, \quad \mu_{Z^\varepsilon}(x) = \frac{1}{\kappa_\varepsilon} \frac{\lambda}{|x|} e^{-\alpha|x| - \beta x} \mathbf{1}_{\{|x| \geq \varepsilon\}}$$

Remarque :

La mesure de Lévy de ces deux lois a une singularité en 0, de ce fait il présentent plusieurs sauts dans tout intervalle de temps fini.

4.2 Les processus hyperboliques généralisés

Notations 2 On adopte, dans la suite, les notations suivantes :

- Pour deux variables (α, β) , on note : $\gamma = \sqrt{\alpha^2 - \beta^2}$, $\gamma_x = \sqrt{\alpha^2 - (\beta + x)^2}$.

- Le Skewness et le Kurtosis d'une variable X sont notés, respectivement : $\mathbb{S}[X]$ et $\mathbb{K}[X]$.

Les processus hyperboliques généralisés \mathcal{HG} ont été étudiés par Eberlein-Prause (1998) et Prause (1999). Leurs densités sont connues, ce qui représente un avantage majeur pour la calibration. Ils sont caractérisés par cinq paramètres $(\lambda, \alpha, \beta, \delta, \mu)$, ces cinq degrés de libertés sont suffisants pour que le modèle de prix représente bien le marché. On peut les introduire par leurs densités de distribution pour la calibration, par mélange de lois pour les simulations ou par processus subordonné pour une interprétation financière.

4.2.1 Quelques notions utiles

Définition 4 On définit la fonction de Bessel K_ν modifiée de type 3 d'indice ν par :

$$K_\nu(x) = \frac{1}{2} \int_0^{+\infty} y^{\nu-1} e^{-\frac{1}{2}x(y+y^{-1})} dy$$

On peut donner une autre caractérisation à cette fonction par l'équation différentielle suivante :

$$x^2 y'' + xy' - (x^2 + \nu^2)y = 0$$

Cette fonction apparaît dans le développement de la densité de distribution des lois hyperboliques généralisées.

Propriétés :

1. La fonction $\nu \mapsto K_\nu$ est symétrique (paire)

$$K_\nu(x) = K_{-\nu}(x), \quad \forall x \in \mathbb{R}$$

2. La fonction $(x, \nu) \mapsto K_\nu(x)$ vérifie les deux relations de récurrence

$$2K_{\nu+1} = \frac{x}{\nu+1}(K_{\nu+2} - K_\nu) \quad \text{et} \quad K'_\nu(x) = -K_{\nu+1} + \frac{\nu}{x}K_\nu$$

3. Il existe une formule analytique de la fonction $x \mapsto K_{m+\frac{1}{2}}(x)$

$$K_{m+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}} e^{-x} \sum_{j=0}^m \frac{(m+j)!}{j!(m-j)!(2x)^j}, \quad \forall m \in \mathbb{Z}$$

Définition 5 Une variable aléatoire X est de **loi inverse gaussienne généralisée** $\mathcal{IGG}(\lambda, \delta, \gamma)$, si elle admet la densité suivante :

$$f_{\mathcal{IGG}}(x; \lambda, \delta, \gamma) = \left(\frac{\gamma}{\delta}\right)^\lambda \frac{1}{2K_\lambda(\delta\gamma)} x^{\lambda-1} e^{-\frac{1}{2}\left(\frac{\delta^2}{x} + \gamma^2 x\right)}, \quad \text{pour } x > 0$$

Avec,

$$\begin{cases} \delta \geq 0, & \gamma > 0 & \text{si } \lambda > 0 \\ \delta > 0, & \gamma > 0 & \text{si } \lambda = 0 \\ \delta > 0, & \gamma \geq 0 & \text{si } \lambda < 0 \end{cases}$$

La transformée de Laplace de la loi inverse gaussien généralisée $\mathcal{IGG}(\lambda, \delta, \gamma)$ est donnée par :

$$L_{\mathcal{IGG}}(x) = \frac{K_\lambda(\delta\gamma\sqrt{1 - \frac{2x}{\gamma^2}})}{K_\lambda(\delta\gamma)(1 - \frac{2x}{\gamma^2})^{\frac{\lambda}{2}}}$$

En dérivant cette fonction, on peut obtenir la moyenne et la variance d'une variable X de loi \mathcal{IGG} :

$$\mathbb{E}[X] = \frac{\delta K_{\lambda+1}(\delta\gamma)}{\gamma K_\lambda(\delta\gamma)}, \quad \text{Var}[X] = \left(\frac{\delta}{\gamma}\right)^2 \left[\frac{K_{\lambda+2}(\delta\gamma)}{K_\lambda(\delta\gamma)} - \left(\frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)}\right)^2 \right]$$

Remarques :

1. Pour $\lambda = -\frac{1}{2}$, on obtient une **loi gaussienne inverse** $\mathcal{IG}(\delta, \gamma)$ de densité :

$$f_{\mathcal{IG}}(x; \delta, \gamma) = \frac{\delta x^{-\frac{3}{2}}}{\sqrt{2\pi}} e^{-\frac{\gamma^2}{2x}(x - \frac{\delta}{\gamma})^2}, \quad \forall x > 0$$

2. Pour $\delta = 0$, on obtient une **loi gamma** $\Gamma(\lambda, \frac{\gamma^2}{2})$ dont la densité s'écrit :

$$f_\Gamma(x; \lambda, \frac{\gamma^2}{2}) = \frac{(\frac{\gamma^2}{2})^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\frac{\gamma^2}{2}x}, \quad \forall x > 0, \quad \Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$$

4.2.2 La distribution de loi \mathcal{HG}

Les distributions hyperboliques généralisées ont été introduites par Bandorf-Nielson en 1977. La densité d'une loi hyperbolique généralisée est donnée par :

$$f_{\mathcal{HG}}(x; \lambda, \alpha, \beta, \delta, \mu) = a(\lambda, \alpha, \beta, \delta) \sqrt{\delta^2 + (x - \mu)^2}^{\lambda - \frac{1}{2}} K_{\lambda - \frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2}) e^{\beta(x - \mu)}$$

Avec,

$$a(\lambda, \alpha, \beta, \delta) = \frac{\gamma^\lambda}{\sqrt{2\pi} \alpha^{\lambda - \frac{1}{2}} \delta^\lambda K_\lambda(\delta\gamma)}$$

$$\begin{cases} \delta \geq 0, & \alpha > 0, & \gamma > 0 & \text{si } \lambda > 0 \\ \delta > 0, & \alpha > 0, & \gamma > 0 & \text{si } \lambda = 0 \\ \delta > 0, & \alpha \geq 0, & \gamma \geq 0 & \text{si } \lambda < 0 \end{cases}$$

Les queues de distribution des lois hyperboliques sont en général plus épaisses que celles de la loi normale, on donne l'équivalence à l'infini :

$$f_{HG}(x) \sim \begin{cases} C|x|^{\lambda-1}exp((\beta - \alpha)x), & x \rightarrow +\infty \\ C|x|^{\lambda-1}exp((\beta + \alpha)x), & x \rightarrow -\infty \end{cases}$$

Avec,

$$C = \frac{\gamma^\lambda e^{-\beta\mu}}{2\alpha^\lambda \delta^\lambda K_\lambda(\delta\gamma)}$$

Les lois hyperboliques généralisées forment une large classe de lois. Cette diversité est due au nombre de paramètres qui les caractérisent : α détermine l'épaisseur de la queue de distribution (les pics), β détermine son degré d'asymétrie (Skewness), λ détermine la sous classe de cette loi, μ est un paramètre de position et δ est un paramètre d'échelle.

4.2.3 Les mélanges de lois \mathcal{IGG}

Les distributions hyperboliques généralisées $\mathcal{HG}(\lambda, \alpha, \beta, \delta, \mu)$ ¹¹ correspondent à un mélange en moyenne et variance d'une distribution gaussienne \mathcal{N} par une distribution inverse gaussienne généralisée $\mathcal{IGG}(\lambda, \delta, \gamma)$:

$$f_{HG}(x; \lambda, \alpha, \beta, \delta, \mu) = \int_0^\infty f_{\mathcal{N}(\mu+\beta y, y)}(x) f_{IGG}(y; \lambda, \delta, \gamma) dy$$

On peut voir cette v.a comme une gaussienne à moyenne et volatilité stochastique. En effet, si un processus L_t , dont L_1 suit la loi $\mathcal{HG}(\lambda, \alpha, \beta, \delta, \mu)$ alors il peut s'écrire sous la forme :

$$L_1 = \mu + \beta Y + \sqrt{Y}N, \quad \text{où } N \sim \mathcal{N}(0, 1), \quad \text{et } Y \sim \mathcal{IGG}(\lambda, \delta, \gamma)$$

Les variables aléatoires N et Y sont indépendantes. Comme conséquence, l'expression de la transformée de Laplace de la loi \mathcal{HG} en fonction de celle de la loi \mathcal{IGG} :

$$L_{HG}(x) = e^{\mu x} L_{IGG}\left(\frac{1}{2}x^2 + \beta x\right)$$

Cette représentation des lois hyperboliques en mélange de lois permet de simuler une variable de ce type en utilisant simplement une simulation d'une variable gaussienne et d'une variable \mathcal{IGG} . D'autres parts, on peut voir le modèle avec un bruit hyperbolique généralisé comme un modèle à volatilité stochastique. Dans le cas \mathcal{NIG} la volatilité suit une loi \mathcal{IG} par contre dans le cas \mathcal{VG} la volatilité suit une loi Γ . Un processus hyperbolique est centré si et seulement si $\mu = -\beta\mathbb{E}[Y]$ (par exemple $\beta = \mu = 0$).

¹¹La densité de distribution pour un processus L_t correspondent à la densité de son incrément L_1 .

4.2.4 Les processus subordonnés

Dans l'approche avec subordination le processus hyperbolique généralisé suit une loi gaussienne avec une modification du temps pour représenter les périodes d'activités sur le marché. Soit un processus de Lévy $(\tau_t)_{t \geq 0}$ telle que τ_1 est distribué suivant une loi inverse gaussienne généralisée $\mathcal{IGG}(\lambda, \delta, \gamma)$. Si on considère le processus

$$L_t = \mu t + \beta \tau_t + W_{\tau_t}$$

Où, (W_t) est un mouvement brownien standard, on montre que L_t est alors un processus hyperbolique généralisé $\mathcal{HG}_t(\lambda, \alpha, \beta, \delta, \mu)$. Le processus τ_t s'interprète comme un temps virtuel représentant les périodes d'agitation et de calme sur le marché :

• Une **période d'agitation** verra le temps s'écouler plus rapidement, soit

$$\tau_{t+dt} - \tau_t > dt \text{ et } \text{Var}[W_{\tau_{t+dt}} | \tau_{t+dt}] > \text{Var}[W_{\tau_t} | \tau_t] + \text{Var}[W_{dt}]$$

• Une **période de calme**, par contre, verra le temps s'écouler lentement, soit

$$\tau_{t+dt} - \tau_t \leq dt \text{ et } \text{Var}[W_{\tau_{t+dt}} | \tau_{t+dt}] \leq \text{Var}[W_{\tau_t} | \tau_t] + \text{Var}[W_{dt}]$$

Le processus L_t découle de la subordination du mouvement brownien W_t par le processus τ_t , le subordonateur dans ce cas est le temps. Un modèle de générateur de bruit L_t , peut se voir aussi comme un **modèle à volatilité stochastique** $\sigma_t = \sqrt{\tau_t}$.

4.2.5 Exemples de lois \mathcal{HG}

La classe des lois hyperboliques généralisées réunit plusieurs lois connues. Dans la suite, on cite quelques cas spécifiques :

1. La loi **Normale** $\mathcal{N}(\mu + \omega_0 \beta, \omega_0)$ correspond au cas limite :

$$(\alpha, \delta) \rightarrow \infty \text{ et } \frac{\delta}{\alpha} \rightarrow \omega_0$$

2. La loi **Hyperbolique** correspond à $\lambda = 1$, elle admet la densité suivante :

$$f_{HYP}(x; \alpha, \beta, \delta, \mu) = \frac{\gamma}{2\alpha\delta K_1(\delta\gamma)} e^{(-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu))}$$

3. La loi **Normale Inverse Gaussienne** correspond à $\lambda = -\frac{1}{2}$, elle admet la densité suivante :

$$f_{NIG}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha\delta}{\pi} e^{(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu))} \frac{K_1(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{\sqrt{\delta^2 + (x-\mu)^2}}$$

4. La loi **Variance Gamma** correspond à $\delta = 0$, elle admet la densité suivante :

$$f_{VG}(x; \lambda, \alpha, \beta, \mu) = \frac{\gamma^{2\lambda}}{\sqrt{\pi}\Gamma(\lambda)(2\alpha)^{\lambda-\frac{1}{2}}} |x - \mu|^{\lambda-\frac{1}{2}} K_{\lambda-\frac{1}{2}}(\alpha|x - \mu|) e^{\beta(x-\mu)}$$

4.2.6 Propriétés de la loi \mathcal{HG}

Un processus de Lévy hyperbolique généralisé $L_t(\lambda, \alpha, \beta, \delta, \mu)$, est un processus stationnaire issu de 0, ses accroissements sont i.i.d et dont l'incrément L_1 admet la densité $f(x; \lambda, \alpha, \beta, \delta, \mu)$. Cette caractérisation prouve que les accroissements de longueur 1 ont une loi hyperbolique généralisée mais ceux de longueurs différentes ne le sont pas en général. Seuls les lois \mathcal{NIG} et \mathcal{VG} , stable par convolution, ont des accroissement de même loi. Par ailleurs, la loi hyperbolique généralisée est stable par addition et multiplication, car $\tilde{L}_t = \sigma L_t + \xi t$ suit aussi la loi hyperbolique généralisée,

$$\tilde{L}_t \sim \mathcal{HG}_t(\lambda, \alpha/\sigma, \beta/\sigma, \sigma\delta, \sigma\mu + \xi)$$

La transformée de Laplace de la loi hyperbolique généralisée $\mathcal{HG}(\lambda, \alpha, \beta, \delta, \gamma)$ est donnée par :

$$L_{HG}(x) = \mathbb{E}[e^{xL_1}] = e^{\mu x} \frac{\gamma^\lambda K_\lambda(\delta\gamma x)}{\gamma^\lambda K_\lambda(\delta\gamma)}, \quad \text{pour tout } |\beta + x| < \alpha$$

En dérivant cette expression, on peut formuler la moyenne et la variance d'un processus hyperbolique généralisée $\mathcal{HG}(\lambda, \alpha, \beta, \delta, \mu)$:

$$\begin{aligned} \mathbb{E}[L_1] &= \mu + \frac{\delta\beta K_{\lambda+1}(\delta\gamma)}{\gamma K_\lambda(\delta\gamma)} \\ \text{Var}[L_1] &= \frac{\delta K_{\lambda+1}(\delta\gamma)}{\gamma K_\lambda(\delta\gamma)} + \frac{\beta^2 \delta^2}{\gamma^2} \left[\frac{K_{\lambda+2}(\delta\gamma)}{K_\lambda(\delta\gamma)} - \left(\frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)} \right)^2 \right] \end{aligned}$$

4.2.7 La loi normal inverse gaussien

Caractéristiques générales

La distribution de la loi \mathcal{NIG} à été introduite par Barndorff-Nielsen en 1977, elle correspond au cas $\lambda = -\frac{1}{2}$. La densité de cette loi est donnée par :

$$f_{NIG}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha\delta}{\pi} e^{(\delta\gamma + \beta(x-\mu))} \frac{K_1(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{\sqrt{\delta^2 + (x-\mu)^2}}$$

Un processus L_t de loi $\mathcal{NIG}_t(\alpha, \beta, \delta, \mu)$, admet la loi $\mathcal{NIG}(\alpha, \beta, t\delta, t\mu)$. Ce processus est purement de saut et n'a aucune composante brownienne, sa mesure de Lévy est donnée par :

$$\nu_{NIG}(dx) = \frac{\alpha\delta}{\pi|x|} e^{\beta x} K_1(\alpha|x|) dx$$

Cette loi infiniment divisible est stable par convolution i.e :

$$\mathcal{NIG}(\alpha, \beta, \delta_1, \mu_1) + \mathcal{NIG}(\alpha, \beta, \delta_2, \mu_2) = \mathcal{NIG}(\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2)$$

La transformée de Laplace de cette distribution est donnée par :

$$L_{\mathcal{NIG}}(x) = \mathbb{E}[e^{xL_t}] = e^{\mu tx + \delta t(\gamma - \gamma_x)}, \quad \text{pour } |\beta + x| < \alpha$$

$$\varphi_{\mathcal{NIG}}(x) = \mu x + \delta(\gamma - \gamma_x)$$

La moyenne, la variance, le Skewness et le Kurtosis de la loi \mathcal{NIG} sont :

$$\mathbb{E}[L_t] = \mu t + \frac{\delta\beta}{\gamma} t, \quad \text{Var}[L_t] = \frac{\delta\alpha^2}{\gamma^3} t$$

$$\mathbb{S}[L_t] = 3\beta\delta\frac{\alpha^2}{\gamma^5} t, \quad \mathbb{K}[L_t] = 3\left(1 + 4\left(\frac{\beta}{\alpha}\right)^2\right)\frac{\delta\alpha^4}{\gamma^7} t$$

Le processus L_t est centré réduit si et seulement si $\mu = -\frac{\delta\beta}{\gamma}$, et $\delta\alpha^2 = \gamma^3$. Dans la Figure 11, on compare la densité de la loi $\mathcal{NIG}(\alpha, \beta, \delta, \mu)$ avec celle de la loi normale $\mathcal{N}(0, 1)$, en montrant l'influence des paramètres β et μ :

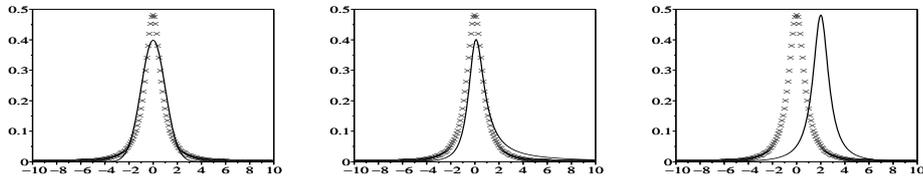


Figure 11: Densité de la loi normale (à gauche), de la loi $\mathcal{NIG}(0.33, 0.32, 0.8, 0)$ (au centre), de la loi $\mathcal{NIG}(0.33, 0, 0.8, 2)$ (à droite) et celle de $\mathcal{NIG}(0.33, 0, 0.8, 0)$ comme référence (en pointillés).

On voit particulièrement que la loi \mathcal{NIG} à une queue plus épaisse que la loi normale. Le paramètre μ est un paramètre de position et le paramètre β a une influence sur l'asymétrie de la queue de distribution. Dans la Figure 12, on représente plusieurs échantillons de cette loi, en variant les paramètres α et β :

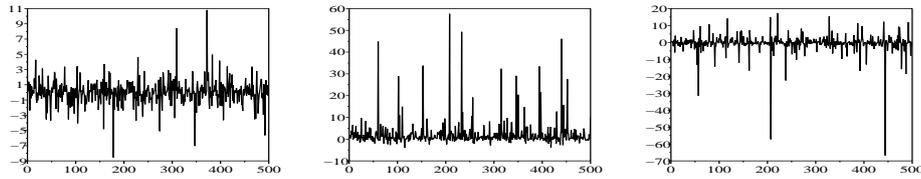


Figure 12: Echantillon de la loi $\mathcal{NIG}(0.33, 0, 1, 0)$ comme référence (à gauche), de la loi $\mathcal{NIG}(0.33, 0.3, 1, 0)$ (au centre) et celle de $\mathcal{NIG}(0.033, 0, 1, 0)$ (à droite).

Le paramètre α caractérise la volatilité de l'échantillon et le nombre de pics (plus α est petit, plus les pics sont marqués et nombreux) et le paramètre β influe sur leur signe, leur nombre et leur intensité (voir la partie Décomposition de Lévy pour une approche mathématique des sauts).

Simulation de loi \mathcal{NIG}

Soit V une variable suivant la loi $\chi^2(1)$, on pose l'équation : $V = \frac{(\gamma Z - \delta)^2}{Z}$, dont les deux racines sont :

$$Z_1 = \frac{\delta}{\gamma} + \frac{1}{2\gamma^2}(V + \sqrt{V^2 + 4\gamma\delta V}) \quad \text{et} \quad Z_2 = \frac{\delta}{\gamma} + \frac{1}{2\gamma^2}(V - \sqrt{V^2 + 4\gamma\delta V})$$

On remarque que $Z_1 Z_2 = \frac{\delta^2}{\gamma^2}$. On peut simuler une variable inverse gaussien $\mathcal{IG}(\delta, \gamma)$ en utilisant le pont de diffusion (voir [WR03]) par :

$$Z = Z_1 \quad \text{avec une probabilité} \quad \frac{\delta}{\delta + \gamma Z_1}$$

$$Z = Z_2 \quad \text{avec une probabilité} \quad \frac{\delta}{\delta + \gamma Z_2}$$

Alors pour simuler une v.a Normale Inverse Gaussien $L(\alpha, \beta, \mu, \delta)$, on utilise la représentation de la loi \mathcal{NIG} par mélange de lois :

$$L = \mu + \beta Z + \sqrt{Z}U, \quad U \sim \mathcal{N}(0, 1)$$

4.2.8 La loi variance gamma

Caractéristiques générales

La distribution de la loi \mathcal{VG} à été introduite par Madan et Seneta en 1980, elle correspond au cas $\delta = 0$. La densité de cette loi est donnée par :

$$f_{VG}(x; \lambda, \alpha, \beta, \mu) = \frac{\gamma^{2\lambda}}{\sqrt{\pi}\Gamma(\lambda)(2\alpha)^{\lambda-\frac{1}{2}}} |x - \mu|^{\lambda-\frac{1}{2}} K_{\lambda-\frac{1}{2}}(\alpha|x - \mu|) e^{\beta(x-\mu)}$$

Un processus L_t de loi $\mathcal{VG}_t(\lambda, \alpha, \beta, \mu)$, admet la loi $\mathcal{VG}(t\lambda, \alpha, \beta, t\mu)$. Ce processus est purement de saut et n'a aucune composante brownienne, sa mesure de Lévy est donnée par :

$$\nu_{VG}(dx) = \frac{\lambda}{|x|} e^{-\alpha|x| - \beta x} dx$$

Cette loi infiniment divisible est stable par convolution i.e :

$$\mathcal{VG}(\lambda_1, \alpha, \beta, \mu_1) + \mathcal{VG}(\lambda_2, \alpha, \beta, \mu_2) = \mathcal{VG}(\lambda_1 + \lambda_2, \alpha, \beta, \mu_1 + \mu_2)$$

La transformée de Laplace de cette distribution est donnée par :

$$L_{VG}(x) = \mathbb{E}[e^{xL_t}] = e^{\mu t x} \left(\frac{\gamma}{\gamma_x}\right)^{2\lambda t}, \quad \text{pour } |\beta + x| < \alpha$$

$$\varphi_{VG}(x) = \mu x + 2\lambda \ln\left(\frac{\gamma}{\gamma_x}\right)$$

La moyenne, la variance, le Skewness et le Kurtosis de la loi \mathcal{VG} sont :

$$\mathbb{E}[L_t] = \mu t + \frac{2\beta\lambda}{\gamma^2} t, \quad \text{Var}[L_t] = \frac{2\lambda t}{\gamma^2} \left(1 + 2\left(\frac{\beta}{\gamma}\right)^2\right)$$

$$\mathbb{S}[L_t] = \frac{4\lambda\beta t}{\gamma^4} \left(3 + 4\left(\frac{\beta}{\gamma}\right)^2\right), \quad \mathbb{K}[L_t] = \frac{12\lambda t}{\gamma^4} \left(1 + \frac{8\beta^2}{\gamma^2} + \frac{8\beta^4}{\gamma^4}\right) + \frac{12\lambda^2 t^2}{\gamma^4} \left(1 + \frac{4\beta^2}{\gamma^2} + \frac{4\beta^4}{\gamma^4}\right)$$

Le processus L_t est centré réduit si et seulement si $\mu = -\frac{2\lambda\beta}{\gamma}$, et $2\lambda(\alpha^2 + \beta^2) = \gamma^4$. Dans la Figure 13, on compare la densité de la loi $\mathcal{VG}(\lambda, \alpha, \beta, \mu)$ avec celle de la loi normal $\mathcal{N}(0, 1)$, en montrant l'influence des paramètres α et μ :

On voit particulièrement que la loi \mathcal{VG} à une queue plus épaisse que la loi normale. Le paramètre μ est un paramètre de position et le paramètre α a une influence sur l'épaisseur de la queue de distribution. Dans la Figure 14, on représente plusieurs échantillons de cette loi, en variant les paramètres λ et β : "

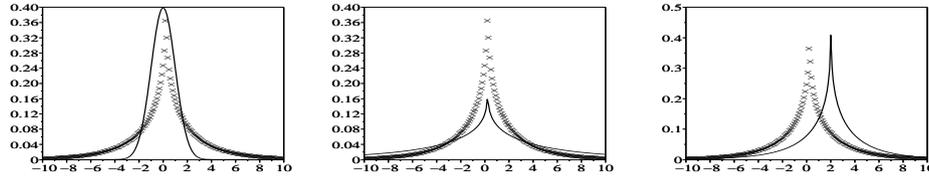


Figure 13: Densité de la loi normale (à gauche), de la loi $\mathcal{VG}(0.6, 0.13, 0, 0.01)$ (au centre), de la loi $\mathcal{VG}(0.6, 0.33, 0, 2.24)$ (à droite) et celle de $\mathcal{VG}(0.6, 0.13, 0, 0.23)$ comme référence (en pointillés).

Le paramètre λ caractérise la volatilité de l'échantillon et le nombre de pics (plus λ est petit, plus les pics sont marqués et moins nombreux) et le paramètre β influe sur leur signe, leur nombre et leur intensité (voir la partie Décomposition de Lévy pour une approche mathématique des sauts).

Simulation de loi \mathcal{VG}

Soit V une variable suivant la loi $\mathcal{T}(2)$, dont la densité g et la fonction de répartition G sont :

$$g(x) = \frac{1}{2\sqrt{2}}(1 + \frac{x^2}{2})^{-\frac{3}{2}}, \quad G(x) = \frac{1}{2}(1 + \frac{\frac{x}{\sqrt{2}}}{\sqrt{1 + \frac{x^2}{2}}})$$

On remarque qu'on peut simuler une variable Student V , par inversion de la fonction de répartition :

$$V \sim \frac{\sqrt{2}(U - \frac{1}{2})}{\sqrt{U(1-U)}}, \quad U \sim \mathcal{U}[0, 1]$$

Dans le cas $\lambda \geq 1$, on utilise une méthode de rejet sur la loi Student pour simuler une v.a de loi $\Gamma(\lambda, 1)$ (voir [Dev86]) :

$$f_{\Gamma}(x) \leq cg_{\lambda}(x), \quad c = e\sqrt{\frac{6}{\pi}}$$

Où, g_{λ} est la densité de loi de la variable $(\lambda - 1) + V\sqrt{\frac{3\lambda}{2} - \frac{3}{8}}$. Cette loi admet la densité de distribution suivante :

$$g_{\lambda}(x) = \frac{1}{2\sqrt{3\lambda - \frac{3}{4}}}(1 + \frac{(x - (\lambda - 1))^2}{3\lambda - \frac{3}{4}})^{-\frac{3}{2}}$$

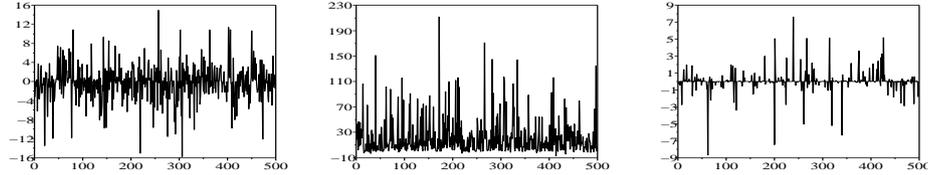


Figure 14: Echantillon de la loi $\mathcal{VG}(0.7, 0.33, 0, 0)$ comme référence (à gauche), de la loi $\mathcal{VG}(0.7, 0.33, 0.3, 0)$ (au centre) et celle de $\mathcal{VG}(0.07, 0.33, 0, 0)$ (à droite).

Par ailleurs, on montre que dans le cas $\lambda < 1$, la loi de $\Gamma(\lambda, 1)$ se déduit de celle de $\Gamma(\lambda + 1, 1)$ par :

$$\Gamma(\lambda, 1) \sim \Gamma(\lambda + 1, 1)U^{\frac{1}{\lambda}}$$

Finalement, pour générer une v.a Variance Gamma $L(\lambda, \alpha, \beta, \mu)$, on utilise la représentation de la loi \mathcal{VG} par mélange de lois :

$$L = \mu + \beta Z + \sqrt{Z}N, \quad \text{où } N \sim \mathcal{N}(0, 1) \text{ et } Z \sim \frac{2}{\gamma^2}\Gamma(\lambda, 1)$$

4.3 Modélisation de la dynamique du sous-jacent

Il existe deux approches possibles pour modéliser les prix d'un sous-jacent, en le considérant sous la probabilité historique ou risque-neutre. On adoptera la première approche dans une perspective de calibration et la deuxième pour des besoins de pricing. Vu le nombre d'acteurs sur le marché de l'électricité, on peut le supposer **viable** i.e, il n'existe pas d'opportunité d'arbitrage. Toutefois, vu la non-stockabilité du sous-jacent, les contrats forward, ne peuvent être couverts avec une méthode "Buy-and-Hold", donc le marché est **non complet**.

4.3.1 La prime de risque

Définition 6 Soit L_t un processus de Lévy défini sur l'espace $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}^+}, \mathbb{P})$.

On appelle **transformation d'Esscher**¹² le changement de probabilité de \mathbb{P} à \mathbb{Q}^ω avec la densité $Z_t^\omega = \frac{d\mathbb{Q}^\omega}{d\mathbb{P}}|_{\mathcal{F}_t}$ de la forme :

$$Z_t^\omega = e^{\omega L_t - \varphi(\omega)t}$$

¹²La transformation d'Esscher est un changement de probabilité risque-neutre.

Où φ est la fonction génératrice des moments de L_1 et ω une variable réelle.

On introduit la condition d'intégrabilité de la mesure de Lévy : Il existe une constante c telle que

$$\int_1^\infty e^{cx} \nu(dx)$$

Cette condition est suffisante pour l'existence du moment d'ordre k du prix spot (la valeur de k dépend évidemment de c). On voit particulièrement que si $c > 1$, alors Z_t^ω est bien définie ($\varphi(\omega) < \infty$). Il est facile de voir que le processus $(Z_t^\omega)_{t \geq 0}$ est une martingale de norme 1.

Proposition 5 Soit un processus de Lévy $(L_t)_{t \geq 0}$ et \tilde{S}_t son exponentielle¹³ de Lévy sous la probabilité \mathbb{P} . Soit ω_0 la solution de l'équation $\varphi(\omega + 1) - \varphi(\omega) = r$. Alors sous la probabilité \mathbb{Q}^{ω_0} , le processus \tilde{S}_t est une martingale.

De manière générale, on suppose qu'il existe, pour tout actif financier dont le générateur de bruit est L_t , une probabilité risque neutre \mathbb{Q}^ω sous laquelle cet actif est une martingale.

Preuve :

On suppose que un tel ω existe. On écrit l'espérance conditionnelle de S_t sachant la filtration \mathcal{F}_s avec $s < t$:

$$\begin{aligned} \mathbb{E}_\omega[\tilde{S}_t | \mathcal{F}_s] &= \tilde{S}_s \mathbb{E}\left[\frac{\tilde{S}_t}{\tilde{S}_s} \cdot \frac{Z_t^\omega}{Z_s^\omega} | \mathcal{F}_s\right] \\ &= \tilde{S}_s \mathbb{E}\left[e^{-r(t-s)} \cdot e^{(\omega+1)(L_t-L_s) - \varphi(\omega)(t-s)}\right] \\ &= \tilde{S}_s e^{(\varphi(\omega+1) - \varphi(\omega) - r)(t-s)} \\ &= \tilde{S}_s \end{aligned}$$

Dans ces calculs on utilise le fait que $(L_t - L_s)$ est indépendant de \mathcal{F}_s et que \tilde{S}_s est \mathcal{F}_s -mesurable. De plus on utilise (15) pour le calcul de la dernière espérance.

Remarque :

Le théorème de Girsanov est un cas particulier de cette transformation. En effet, on l'obtient en posant $L_t = W_t$, $\omega = \xi$ et $\varphi(u) = \frac{u^2}{2}$. On remarque que la variable ω , représente une mesure du risque inhérent au générateur de bruit L_t .

Proposition 6 Soit L_t un processus de Lévy de distribution hyperbolique généralisée sous la probabilité \mathbb{P} ,

$$L_t \stackrel{\mathbb{P}}{\sim} \mathcal{HG}_t(\lambda, \alpha, \beta, \delta, \mu)$$

¹³ On définit l'exponentielle de Lévy du processus L_t comme : $\tilde{S}_t = S_0 e^{L_t - rt}$.

Alors la distribution de loi de L_t , sous la probabilité \mathbb{Q}^ω , est aussi une distribution hyperbolique généralisée,

$$L_t \stackrel{\mathbb{Q}^\omega}{\sim} \mathcal{HG}_t(\lambda, \alpha, \beta + \omega, \delta, \mu)$$

Cette proposition permet de résumer le changement de probabilité $\mathbb{P} \leftrightarrow \mathbb{Q}^\omega$, dans le cas de processus hyperbolique généralisé, au changement des paramètres de la loi du générateur de bruit L_t .

4.3.2 L'approche martingale

On récapitule, dans cette partie, la méthode suivie pour modéliser les prix à terme et spot de l'électricité. On adopte la même démarche suivie dans [BEHN99], qui est motivé par la théorie des taux d'intérêts. Dans le marché de l'électricité les prix (sous la probabilité historique \mathbb{P}) présentent les caractéristiques suivantes : la saisonnalité, le retour à la moyenne et les pics de prix. Pour cette raison on propose le modèle suivant :

$$F(t, T) = \Lambda(t, T)e^{X_t^T}$$

Dans cette configuration, la saisonnalité est représentée par $\Lambda(t, T)$, une fonction déterministe périodique. Pour tenir compte du caractère discontinue des prix, on modélise la partie stochastique X_t^T par¹⁴ :

$$X_t^T = \int_0^t \sigma e^{-a(T-s)} dL_s$$

Ce processus intègre bien les pics des prix via la loi de L_t et le retour à la moyenne du prix spot :

$$dX_t^t = -aX_t^t + \sigma dL_t$$

Vu la non-complétude du marché et du modèle¹⁵, il existe plusieurs probabilités risque-neutre. Le prix des actifs est spécifié une fois cette probabilité est définie. On fait le choix de la **mesure martingale statistique**, ce choix a été déjà fait auparavant par Keller en 1997 et Eberlein, Keller et Prause en 1998. On choisit, alors, parmi les transformations possibles d'Esscher, celle qui représente au mieux le marché¹⁶. Soit ω la prime de risque liée au marché d'électricité, qu'on supposera constante. Sous la probabilité \mathbb{Q}^ω , les prix à termes sont des martingales ($\frac{\partial \mathbb{Q}^\omega}{\partial \mathbb{P}} = e^{\omega L_t - \varphi(\omega)t}$ la dérivée de Randon-Nikodym) :

$$\mathbb{E}_\omega[F(t, T)|\mathcal{F}_s] = F(s, T)$$

¹⁴Dans le modèle gaussien on a : $X_t^T = \int_0^t \sigma e^{-a(T-s)} dW_s$

¹⁵Voir [Sch03] pp. 46, pour une représentation martingale avec sauts du portefeuille de réplication.

¹⁶ ω est estimée sur données historiques.

En utilisant cette égalité et la condition initiale $\Lambda(0, T) = F(0, T)$, on trouve :

$$\Lambda(t, T) = F(0, T)e^{-\int_0^t [\varphi(\omega + \sigma e^{-a(T-s)}) - \varphi(\omega)] ds}$$

Finalement on peut exprimer le prix à terme, dans le modèle générale de Lévy, sous la probabilité historique \mathbb{P} , sous la forme :

$$F(t, T) = F(0, T)e^{-\int_0^t [\varphi(\omega + \sigma e^{-a(T-s)}) - \varphi(\omega)] ds + \int_0^t \sigma e^{-a(T-s)} dL_s}$$

On rappelle la relation entre le prix spot et le prix à terme $S_T = \lim_{t \rightarrow T} F(t, T)$. Cette relation permet d'exprimer le prix spot (sous \mathbb{P}) :

$$S_t = F(0, t)e^{-\int_0^t [\varphi(\omega + \sigma e^{-a(t-s)}) - \varphi(\omega)] ds + \int_0^t \sigma e^{-a(t-s)} dL_s}$$

Remarque :

Le paramètre ω est estimé sur données historiques, plus exactement en utilisant les cotations des prix forwards :

$$\omega = \operatorname{argmin}_\theta ||F(0, T) - F^\theta(t, T)||$$

Proposition 7 Soit L_t un générateur de bruit à une distribution hyperbolique généralisée $\mathcal{HG}_t(\lambda, \alpha, \beta + \omega, \delta, \mu)$, alors sous la probabilité risque neutre \mathbb{Q}^ω le prix à terme et sopt, s'écrivent :

$$F(t, T) = F(0, T)e^{-\int_0^t \varphi^*(\sigma e^{-a(T-s)}) ds + \int_0^t \sigma e^{-a(T-s)} dL_s^*}$$

$$S_t = F(0, T)e^{-\int_0^t \varphi^*(\sigma e^{-a(t-s)}) ds + \int_0^t \sigma e^{-a(t-s)} dL_s^*}$$

Avec,

$$L_t^* \sim \mathcal{HG}_t(\lambda, \alpha, \beta + \omega, \delta, \mu)$$

$$\varphi^*(x; \lambda, \alpha, \beta, \delta, \mu) = \varphi(x; \lambda, \alpha, \beta + \omega, \delta, \mu)$$

Sous la probabilité \mathbb{Q}^ω , le processus $(F(t, T))_{t \geq 0}$ des prix à terme horaire est, par construction, une martingale :

$$\mathbb{E}_\omega[F(t, T) | \mathcal{F}_s] = F(s, T)$$

Par ailleurs, on peut exprimer la relation d'absence d'opportunités d'arbitrage entre le spot et le forward :

$$\mathbb{E}_\omega[S_T | \mathcal{F}_t] = F(t, T)$$

Remarque :

Si on note μ^{L_t} la mesure de saut de L_t , alors le processus de prix forward vérifie l'EDS suivante :

$$\frac{dF(t, T)}{F(t, T)} = -(\varphi(\omega + \sigma e^{-a(T-t)}) - \varphi(\omega)) dt + \sigma e^{-a(T-t)} dL_t$$

$$+ \int_{\mathbb{R}} (e^{y + \sigma e^{-a(T-t)}} - 1 - \sigma e^{-a(T-t)} y) \mu^{L_t}(dy, dt)$$

Le seconde terme dans l'EDS, est maximal lorsque $t = T$, il représente une composante de saut pure pour le processus.

4.3.3 Solution du modèle

Le modèle saisonnalité

Mathématiquement il est préférable de considérer le processus $X_t^v = \ln\left(\frac{S_t}{F(0,t)}\right)$. Ce dernier suit la diffusion d'Ornstein-Uhlenbeck suivante :

$$dX_t^v = a(m_t - X_t^v)dt + \sigma dL_t, \quad X_0^v = 0 \quad (16)$$

$$m_t = -\frac{1}{a}[\varphi(\omega + \sigma e^{-at}) - \varphi(\omega)] - \int_0^t [\varphi(\omega + \sigma e^{-as}) - \varphi(\omega)]ds$$

On peut expliciter la solution de l'EDS (16) en utilisant la formule d'**Itô** pour les semi-martingales. En effet, posons $Z_t = (X_t^v - a \int_0^t m_s e^{-a(t-s)} ds) e^{at}$, alors Z_t vérifie l'EDS : $dZ_t = \sigma e^{at} dL_t$. L'expression de X_t est :

$$X_t^v = a \int_0^t m_s e^{-a(t-s)} ds + \sigma \int_0^t e^{-a(t-s)} dL_s$$

En pratique, on dispose d'observations à des instants discrets (jour par jour, heure par heure...). On considère une discrétisation sur un intervalle $[0, T]$ avec un pas de temps h . Pour alléger les formulations on note : $X_{t_i}^v = X_{ih}^v = X_i^v$. La formule explicite de X_t^v donne le schéma suivant :

$$X_i^v - \phi_1 X_{i-1}^v - \phi_0^i = \varepsilon_i \quad (17)$$

Où,

$$\varepsilon_i = \int_{(i-1)h}^{ih} \sigma e^{-a(ih-s)} dL_s, \quad \phi_1 = e^{-ah}$$

$$\phi_0^i = - \int_0^{ih} [\varphi(\omega + \sigma e^{-as}) - \varphi(\omega)] ds + \phi_1 \int_0^{(i-1)h} [\varphi(\omega + \sigma e^{-as}) - \varphi(\omega)] ds$$

Le modèle désaisonnalisé

On peut décomposer le prix en une partie déterministe et une partie stochastique. La partie déterministe notée $S_0 D_t$ représente la saisonnalité des prix spot. La partie stochastique, dont le logarithme est X_t^c , intègre la volatilité et le retour à la moyenne des prix spot. Il est aisé de montrer que celui-ci est solution de la diffusion d'Ornstein-Uhlenbeck suivante :

$$dX_t^c = a(m - X_t^c)dt + \sigma dL_t, \quad X_0^c = 0 \quad (18)$$

$$\ln D_t = \ln\left(\frac{F(0,t)}{F(0,0)}\right) - m(1 - e^{-at}) - \int_0^t [\varphi(\omega + \sigma e^{-as}) - \varphi(\omega)] ds$$

Le prix S_t peut, alors, être exprimé selon deux expressions : $S_t = S_0 D_t e^{X_t^c} = F(0,t) e^{X_t^v}$. On utilise une méthode de désaisonnalisation du prix (méthode par classe, par processus moyennes mobiles...), puis on utilise l'historique du prix désaisonnalisé pour calibrer le processus $(X_t^c)_{t \geq 0}$ et quantifier les

paramètres du modèle (m, a, σ) et les paramètres du bruit L_t . On peut expliciter la solution de l'EDS (18) en utilisant la formule d'**Itô** pour les semi-martingales. En effet, posons $Z_t = (X_t - a \int_0^t m e^{-a(t-s)} ds) e^{at}$, alors Z_t vérifie l'EDS : $dZ_t = \sigma e^{at} dL_t$. L'expression de X_t est :

$$X_t = a \int_0^t m e^{-a(t-s)} ds + \sigma \int_0^t e^{-a(t-s)} dL_s$$

En pratique, on dispose d'observations du prix désaisonnalisé à des instants discrets (jour par jour, heure par heure...). On considère une discrétisation sur un intervalle $[0, T]$ avec un pas de temps h . Pour alléger les formulations on note : $X_{t_i}^c = X_{ih}^c = X_i^c$. La formule explicite de X_t^c donne le schéma suivant :

$$X_i^c - \phi_1 X_{i-1}^c - \phi_0 = \varepsilon_i \quad (19)$$

Où,

$$\begin{aligned} \varepsilon_i &= \int_{(i-1)h}^{ih} \sigma e^{-a(ih-s)} dL_s, \quad \phi_1 = e^{-ah} \\ \phi_0 &= m(1 - e^{-ah}) \end{aligned}$$

Remarque :

On montre, à l'aide de la transformée de Laplace, que les résidus sont i.i.d (en particulier leurs loi est indépendante de i). On fera l'approximation suivante :

$$\varepsilon_i \sim \sigma L_h$$

On approche dans ce cas la loi des résidus ε_t , par une loi connue et celle de L_h . Cette approximation est vérifiée pour $ah \ll 1$.

On s'intéresse à la calibration de ce modèle discret pour deux sous-classes : Les processus - \mathcal{NIG} - Normal Inverse Gaussien et - \mathcal{VG} - Variance Gamma.

4.4 Calibration du modèle

Pour la calibration on choisira le modèle du prix désaisonnalisé. On suppose que l'on dispose de n observations $(x_1 \dots x_n)$ du prix désaisonnalisé sur un intervalle $[0, T]$ avec un pas de temps h . On commence par calibrer le processus d'Ornstein-Uhlenbeck à l'aide de la méthode des moindres carrés et estimer les paramètres a et σ . Ensuite les paramètres du bruit (centré réduit) seront estimés par méthode maximum de vraisemblance (dans le cas \mathcal{NIG} , on estime $(\alpha, \beta, \delta, \mu)$ et dans le cas \mathcal{VG} , on estime $(\lambda, \alpha, \beta, \mu)$. On rappelle le schéma du modèle :

$$\begin{aligned} X_i^c - \phi_1 X_{i-1}^c - \phi_0 &= \varepsilon_i \\ \phi_1 &= e^{-ah}, \quad \phi_0 = m(1 - e^{-ah}), \quad \varepsilon_i \sim BB(0, \sigma^2 h) \end{aligned}$$

4.4.1 Estimation par moindres carrés

Cette méthode repose sur la minimisation de la moyenne quadratique des bruits ε_t sur toute la trajectoire $[0, T]$. En effet, Les paramètres $(\widehat{\phi}_0, \widehat{\phi}_1, \sigma)$, vérifient :

$$\begin{bmatrix} \widehat{\phi}_0 \\ \widehat{\phi}_1 \end{bmatrix} = (A' A)^{-1} A' B, \quad \text{où } A = \begin{bmatrix} 1 & x_0 \\ \vdots & \vdots \\ 1 & x_{n-1} \end{bmatrix} \quad \text{et } B = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

et

$$\widehat{\sigma}^2 = \frac{1}{nh} \sum_{i=1}^n [(x_i - \widehat{\phi}_1 x_{i-1} - \widehat{\phi}_0)^2]$$

Cette méthode découplée permet de récupérer les paramètres du processus d'Ornstein-Uhlenbeck (a, σ) , et aussi les résidus centrés réduits qu'on notera $\eta_i = \frac{x_i - \widehat{\phi}_1 x_{i-1} - \widehat{\phi}_0}{\widehat{\sigma}}$. Il s'agit maintenant de calibrer les paramètres du bruit blanc η_t qui ne sont rien d'autres que $(\alpha, \beta, \delta h, \mu h)$ pour le \mathcal{NIG} et $(\lambda h, \alpha, \beta, \mu h)$ pour le \mathcal{VG} .

4.4.2 Estimation par maximum de vraisemblance

La vraisemblance du \mathcal{NIG}

Pour un échantillon $(\eta_1 \dots \eta_n)$ de η_t (on rappelle que η_i sont i.i.d), le logarithme de la vraisemblance s'écrit :

$$\mathcal{L}(\alpha, \beta, \delta, \mu) = \sum_{i=1}^n \ln f_{\mathcal{NIG}}(\eta_i, \alpha, \beta, \delta h, \mu h)$$

On note dans la suite :

$$s_i = \frac{\eta_i - \mu h}{\delta h}, \quad p_i = \sqrt{1 + s_i^2}, \quad R_1(x) = \frac{K_2(x)}{K_1(x)}$$

La log-vraisemblance se formule alors :

$$\mathcal{L}(\omega, \alpha, \beta, \delta, \mu) = -n \ln \pi + n \ln \alpha + n \delta h \gamma + \sum_{i=1}^n \left[\beta \delta h s_i + \ln \frac{K_1(\alpha \delta h p_i)}{p_i} \right]$$

Pour optimiser cette fonction on utilise une méthode de gradient conjugué. Pour ce besoin on calcul le gradient de $\mathcal{L}(\alpha, \beta, \delta, \mu)$ par rapport à α, β, δ et μ , on trouve :

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial \alpha} = n\left(\frac{2}{\alpha} + \frac{\delta \alpha h}{\gamma}\right) - \sum_{i=1}^n \delta h p_i R_1(\alpha \delta h p_i) \\ \frac{\partial \mathcal{L}}{\partial \beta} = -n \frac{\delta \beta h}{\gamma} + \sum_{i=1}^n \delta h s_i \\ \frac{\partial \mathcal{L}}{\partial \delta} = n\left(\frac{1}{\delta} + \gamma h\right) - \sum_{i=1}^n \frac{\alpha h}{p_i} R_1(\alpha \delta h p_i) \\ \frac{\partial \mathcal{L}}{\partial \mu} = -n \beta h + \sum_{i=1}^n \frac{\alpha h s_i}{p_i} R_1(\alpha \delta h p_i) \end{array} \right.$$

La vraisemblance du \mathcal{VG}

Pour un échantillon $(\eta_1 \dots \eta_n)$ de η_t (on rappelle que η_i sont i.i.d), le logarithme de la vraisemblance s'écrit :

$$\mathcal{L}(\lambda, \alpha, \beta, \mu) = \sum_{i=1}^n \ln f_{VG}(\eta_i, \lambda h, \alpha, \beta, \mu h)$$

On note dans la suite :

$$s_i = \eta_i - \mu h, \quad k_\lambda(x) = \frac{\partial \ln K_\lambda(x)}{\partial \lambda}, \quad R_\lambda(x) = \frac{K_{\lambda+1}(x)}{K_\lambda(x)}$$

La log-vraisemblance se formule alors :

$$\begin{aligned} \mathcal{L}(\omega, \alpha, \beta, \delta, \mu) &= 2n\lambda h \ln \gamma - \frac{n}{2} \ln \pi - n \ln \Gamma(\lambda h) - n\left(\lambda h - \frac{1}{2}\right) \ln(2\alpha) \\ &+ \sum_{i=1}^n \left[\left(\lambda h - \frac{1}{2}\right) \ln |s_i| + \ln K_{\lambda h - \frac{1}{2}}(\alpha |s_i|) + \beta s_i \right] \end{aligned}$$

Pour optimiser cette fonction on utilise une méthode de gradient conjugué. Pour ce besoin on calcul le gradient de $\mathcal{L}(\lambda, \alpha, \beta, \mu)$ par rapport à λ, α, β et μ , on trouve :

$$\left\{ \begin{array}{l} \frac{\partial \mathcal{L}}{\partial \lambda} = nh \left(\ln \frac{\gamma^2}{2\alpha} - \frac{\partial \ln \Gamma(x)}{\partial x}(\lambda h) \right) + \sum_{i=1}^n [h \ln |s_i| + h k_{\lambda h - \frac{1}{2}}(\alpha |s_i|)] \\ \frac{\partial \mathcal{L}}{\partial \alpha} = \frac{2nh\lambda\alpha}{\gamma^2} + \sum_{i=1}^n |s_i| R_{\lambda h - \frac{1}{2}}(\alpha |s_i|) \\ \frac{\partial \mathcal{L}}{\partial \beta} = -\frac{2nh\lambda\beta}{\gamma^2} + \sum_{i=1}^n s_i \\ \frac{\partial \mathcal{L}}{\partial \mu} = -nh\beta + \sum_{i=1}^n \left[\frac{\alpha h s_i}{|s_i|} R_{\lambda h - \frac{1}{2}}(\alpha |s_i|) - \frac{2\lambda h - 1}{s_i} h \right] \end{array} \right.$$

4.5 Validation et choix de modèles

4.5.1 Structure des résidus

Test d'autocorrélation

Si les résidus $(\eta_t)_{t \geq 0}$ obéissent à un bruit blanc, il ne doit pas exister d'autocorrélation dans cette série temporelle. Il existe différents tests qui permettent de répondre à cette demande (Box-Pierce, Ljung-Box, Durbin-Watson...). On choisit d'appliquer le test de **Box-Pierce**, sa statistique de test s'écrit :

$$BP(K) = n \sum_{k=1}^K \rho_k^2(\hat{\eta}_t)$$

Où, $\rho_k(\hat{\eta}_t)$ est le coefficient d'autocorrélation d'ordre k . Le nombre maximal de retards retenu est $K = \frac{n}{3}$. On peut donner une estimation à ce coefficient :

$$\rho_k = \frac{\frac{1}{n} \sum_{i=1}^{n-k} (\eta_i - \bar{\eta})(\eta_{i+k} - \bar{\eta})}{\frac{1}{n} \sum_{i=1}^n (\eta_i - \bar{\eta})^2}$$

Sous l'hypothèse nulle H_0 d'absence d'autocorrélation entre les résidus, la statistique $BP(K)$ suit une loi de Khi-deux à $(K-1)$ degrés de liberté. En revanche si cette hypothèse n'est pas vraie, alors $BP(K) \rightarrow +\infty$ quand $K \rightarrow +\infty$. On rejette donc H_0 si on observe des valeurs anormalement grandes pour $BP(K)$, par exemple supérieures à un z donné. Il existe donc une probabilité α de rejeter à tort (risque de première espèce), qui dépend de z :

$$\mathbb{P}(BP(K) \geq z) = \mathbb{P}(\chi(K-1) \geq z) = \alpha$$

Dans le cadre de ce test on commence par calculer $\alpha_K = \mathbb{P}(\chi(K-1) \geq \widehat{BP}(K))$. On le compare ensuite à 5%, si ce quantile est plus faible on rejette H_0 .

Test d'adéquation de loi

Dans cette partie on expose le test de **Kolmogorov-Smirnov**. On note F_n la fonction de répartition empirique des résidus $(\eta_t)_{t \geq 0}$, elle est donnée par :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{]-\infty, x](x_i)}$$

On peut tester l'hypothèse H_0 selon laquelle $(\eta_t)_{t \geq 0}$ ont la même loi, dont la fonction de répartition est F . En effet, on considère la variable aléatoire :

$$K_n = \sup_x |F_n(x) - F(x)|$$

On peut montrer que, sous H_0 , $\sqrt{n}K_n \xrightarrow{Loi} K$. Où K a une loi fixe (indépendante de F) définie par :

$$\mathbb{P}(K > k) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} \exp(-2j^2 k^2)$$

Le quantile $k_{1-\alpha}$ défini par $\mathbb{P}[K > k_{1-\alpha}] = \alpha$ aura α pour niveau asymptotique. En particulier, on a $k_{0.90} = 1.22$, $k_{0.95} = 1.36$ et $k_{0.99} = 1.63$. Dans le cadre de ce test, si $\sqrt{n}\widehat{K}_n < k_{0.95}$ on rejette H_0 .

4.5.2 Critères de choix de modèles

Il existe des critères de choix des modèles fondés sur le calcul de prévision que l'on cherche à minimiser η_t . Parmi ces critères, on choisit le critère d'erreur absolue moyenne *MAE* (Mean Absolute Error) et le critère d'erreur quadratique moyenne *RMSE* (Root Mean Squared Error) :

$$MAE = \frac{1}{n} \sum_{i=1}^n |\eta_i|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \eta_i^2}$$

On cherche à minimiser ces différents critères. Leur application nous permet de retenir un modèle parmi les modèles proposés.

4.6 Evaluation des options européennes

On considère comme sous-jacent un produit à terme à la date T_0 pour une livraison unitaire à la date T_f : $F(T_0, T_f)$. Le prix d'une option européenne, vu le caractère Markovien du prix du sous-jacent, à la date t , d'échéance T_0 , dont le sous-jacent est ce forward et le pay-off est Ψ est donné par¹⁷ :

$$\begin{aligned} C(t, T_0, T_f, \Psi) &= \mathbb{E}_\omega[e^{-r(T_0-t)} \Psi(F(T_0, T_f)) | \mathcal{F}_t] \\ &= \mathbb{E}_\omega[e^{-r(T_0-t)} \Psi(xe^{X_{T_0}})] |_{x=F(t, T_f)} \end{aligned}$$

$$X_{T_0} = - \int_t^{T_0} \varphi(\sigma e^{-a(T_f-s)}) ds + \int_t^{T_0} \sigma e^{-a(T_f-s)} dL_s, \quad x = F(t, T_f)$$

Dans cette partie, on présente les méthodes mathématiques pour le calcul de ce prix, par formule fermée et par Monte-Carlo. Les applications numériques sont faites pour les valeurs des paramètres suivants :

$$\lambda = 1, \quad \alpha = 0.3, \quad \beta = 0.001, \quad \delta = 0.3, \quad \mu = 10, \quad \omega = 0, \quad a = 0.19, \quad \sigma = 0.13$$

$$t = 0, \quad T_0 = 1, \quad T_f = 1 + \frac{30}{365}, \quad x = 100, \quad K = 100, \quad r = 0.25$$

¹⁷Formule facilement transportable pour le modèle par produit voir [?].

4.6.1 Méthode de Fourier

On introduit d'abord l'utilisation de la transformée de Fourier dans le pricing d'options standards (voir [CM99]). Ensuite, on l'applique pour le Call européen sur le forward, dans le cas d'un modèle gaussien, \mathcal{NIG} et \mathcal{VG} .

On note $S_T = xe^{X_T}$ le prix d'un actif à la maturité T et X_T son logarithme. La fonction caractéristique de X_T est :

$$\psi_T(u) = \mathbb{E}_\omega[\exp(iuX_T)]$$

Soit C_T le prix à maturité T d'un Call européen, de strike $\frac{K}{x} = \exp(k)$. On note q_T la densité du logarithme du prix X_T . La fonction caractéristique de S_T est alors :

$$\psi_T(u) = \int_{-\infty}^{+\infty} e^{ius} q_T(s) ds$$

Le prix du Call européen étant donné par $\mathbb{E}_\omega[e^{-r(T-t)}(S_T - K)_+]$, sa formulation en fonction de q_T donne :

$$C_T(k) = x \int_k^{+\infty} e^{-r(T-t)}(e^s - e^k) q_T(s) ds$$

Pour des raisons d'intégrabilité, on introduit un coefficient $\eta > 0$, et on considère le prix modifié c_T :

$$c_T(k) = e^{\eta k} C_T(k), \quad \text{Avec} \quad \mathbb{E}_\omega[S_T^{\eta+1}] < \infty \quad (20)$$

Au lieu de calculer la transformée de Fourier du prix $C_T(k)$, dont l'intégrabilité n'est pas assurée, on calcule celle de c_T , on la note $\tilde{\psi}$:

$$\tilde{\psi}_T(s) = \int_{-\infty}^{+\infty} e^{isk} c_T(k) dk$$

En utilisant la transformée inverse de Fourier de c_T et un calcul par Fubini de $\tilde{\psi}$, on trouve que le prix C_T est donné par :

$$C_T(k) = x \frac{\exp(-\eta k)}{\pi} \int_0^\infty e^{-isk} \tilde{\psi}_T(s) ds, \quad \text{Avec} \quad \tilde{\psi}_T(s) = \frac{e^{-r(T-t)} \psi_T(s - (\eta + 1)i)}{\eta^2 + \eta - s^2 + is(2\eta + 1)}$$

Pour une borne supérieure de calcul de l'intégral A^ε , on dispose de l'approximation suivante¹⁸ :

$$|C_T(k) - C_T^\varepsilon(k)| < \varepsilon, \quad \text{Avec} \quad C_T^\varepsilon(k) = x \frac{\exp(-\eta k)}{\pi} \int_0^{A^\varepsilon} e^{-isk} \tilde{\psi}_T(s) ds$$

Dans notre étude, on a

$$A^\varepsilon \approx x \frac{\exp(-\eta k)}{\pi} \frac{M}{\varepsilon}, \quad M = \exp(-r(T_0 - t) + \int_t^{T_0} [\varphi((\eta + 1)\sigma e^{-a(T_f - z)}) - (\eta + 1)\varphi(\sigma e^{-a(T_f - z)})] dz)$$

¹⁸ Approximation très grossière à améliorer numériquement.

La méthode des trapèzes, pour un pas d'intégration $h = \frac{A^\varepsilon}{N}$, permet de calculer le prix approché $C_T^\varepsilon(k)$, soit :

$$C_T^\varepsilon(k) \cong x \frac{\exp(-\eta k)}{\pi} \sum_{j=0}^{N-1} \Re \left\{ \frac{\tilde{\psi}_T(jh) + e^{-ikh} \tilde{\psi}_T(jh+h)}{2} e^{-ijkh} \right\} h$$

Application :

Dans les deux cas de modèles \mathcal{NIG} et \mathcal{VG} , on prend $\eta = 0.25$. La condition générale (20) sur η est vérifiée pour :

$$0 < \eta < \frac{\alpha - \beta - \omega}{\sigma e^{-a(T_f - T_0)}} - 1$$

✕ Le prix d'une option européenne -Modèle un facteur gaussien- :

La méthode décrite ci-haut est très générale, utilisable même dans le cas un facteur gaussien (voir la première partie pour la formule fermée par $B\&S$). On choisit de comparer cette méthode avec celle de $B\&S$ pour une option à la monnaie¹⁹ :

$\varepsilon = 10^{-2}, \quad h = 2.5 \cdot 10^{-4}$	$C_{FFT} = 3.637$	$C_{B\&S} = 3.624$
$\varepsilon = 10^{-3}, \quad h = 2.5 \cdot 10^{-4}$	$C_{FFT} = 3.625$	

✕ Le prix d'une option européenne \mathcal{NIG} et \mathcal{VG} :

Pour un Call européen de sous-jacent le forward $F(T_0, T_f)$, le prix d'option admet la formule analytique suivante :

$$\text{Modèle } \mathcal{NIG} \quad C(t, T_0, T_f, K) = \frac{x e^{-\eta k} e^{-r(T_0-t)}}{\pi} \int_0^\infty \frac{\exp(-isk + h_{\mathcal{NIG}}(s))}{\eta^2 + \eta - s^2 + is(2\eta + 1)} ds$$

$$\text{Modèle } \mathcal{VG} \quad C(t, T_0, T_f, K) = \frac{x e^{-\eta k} e^{-r(T_0-t)}}{\pi} \int_0^\infty \frac{\exp(-isk + h_{\mathcal{VG}}(s))}{\eta^2 + \eta - s^2 + is(2\eta + 1)} ds$$

Où,

$$h_\bullet(s) = \int_t^{T_0} [\varphi_\bullet((is + 1 + \eta)\sigma e^{-a(T_f-z)}) - (is + 1 + \eta)\varphi_\bullet(\sigma e^{-a(T_f-z)})] dz$$

Pour le prix d'un Put européen, on utilise la relation de **parité Call-Put** :

$$P(t, T_0, T_f, K) - C(t, T_0, T_f, K) = e^{-r(T_0-t)}(K - F(t, T_f))$$

La Figure 15, représente l'évolution du prix du Call en fonction du pas de discrétisation h avec une erreur $\varepsilon = 10^{-3}$ (les valeurs exactes : $\mathcal{NIG} = 2.457$, $\mathcal{VG} = 16.293$) :

¹⁹ L'option est d'échéance 1 an sur le forward 1 Month Ahead.

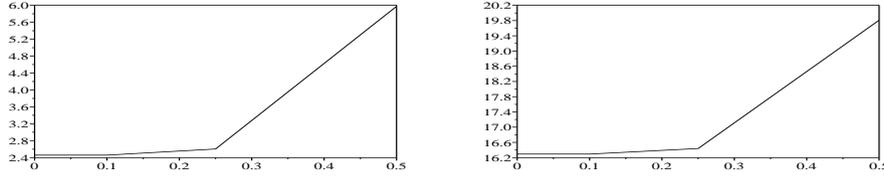


Figure 15: Evolution du prix du Call européen en fonction du pas de discrétisation h pour le modèle \mathcal{NIG} (à gauche), le modèle \mathcal{VG} (à droite).

4.6.2 Approche par Monte-Carlo

La base des méthodes de Monte-Carlo en finance est la capacité à simuler l'évolution des processus de prix. On s'intéresse donc aux méthodes de simulation d'une diffusion de la forme :

$$dX_t = b(t)dt + a(t)dL_t, \quad \text{Avec } X_{t_0} = 0 \quad (21)$$

Le générateur de bruit L_t est un processus \mathcal{NIG} ou \mathcal{VG} . On utilise un **Schéma d'Euler**²⁰ pour discrétiser l'EDS (21), sur une grille de $[t, T]$:

$$(t_0 \dots t_i \dots t_n), \quad \text{Avec } t_i = t + i \frac{T}{n}$$

On notera $\Delta^n t = \frac{T}{n}$ et $\Delta^n L_{i+1} = L_{t_{i+1}} - L_{t_i}$. L'idée de la discrétisation d'Euler est très simple. Pour tout $i = 1 \dots n$, on a :

$$\tilde{X}_{t_i} = \tilde{X}_{t_{i-1}} + b(t_{i-1})\Delta^n t + a(t_{i-1})\Delta^n L_i, \quad \text{Avec } \tilde{X}_{t_0} = 0$$

Pour les processus de Lévy \mathcal{NIG} et \mathcal{VG} la loi de l'incrément $\Delta^n L_i$ est connue. Par conséquent, on peut simuler la loi des \tilde{X}_T par récurrence. Le prix approché, avec un schéma à n pas de temps, de cette option européenne, est donné par Monte-Carlo en m simulations :

$$\tilde{C}(t, T_0, T_f, \Psi)_m^n = \frac{1}{m} \sum_{j=1}^m [e^{-r(T_0-t)} \Psi(xe^{\tilde{X}_{T_0}^{j,n}})]$$

$$\tilde{X}_{t_i}^{j,n} = \tilde{X}_{t_{i-1}}^{j,n} - \varphi(\sigma e^{-a(T_f-t_{i-1})})\Delta^n t + \sigma e^{-a(T_f-t_{i-1})}\Delta^n L_i^j, \quad \tilde{X}_{t_0} = 0$$

²⁰On peut améliorer le schéma d'Euler en considérant une méthode de trapèzes : $\bar{X}_{t_i} = \frac{\tilde{X}_{t_i} + \tilde{X}_{t_{i+1}}}{2}$.

Il existe, alors, deux sources d'erreurs dans cette méthode de calcul : erreurs de discrétisation et erreurs de Monte-Carlo. Les deux disparaissent en augmentant respectivement n et m (convergence faible en $\frac{1}{n}$ et $\frac{1}{\sqrt{m}}$). Par ailleurs, on peut améliorer la deuxième erreur en considérant des **méthodes de réduction de la variance** : Parité Call-Put et variables de contrôles :

$$\text{Var}[\Psi(xe^{X_{T_0}}) + \rho e^{X_{T_0}}] = \text{Var}[\Psi(xe^{X_{T_0}})] + 2\rho \text{Cov}[\Psi(xe^{X_{T_0}}), e^{X_{T_0}}] + \rho^2 \text{Var}[e^{X_{T_0}}]$$

Le candidat optimal ρ^* est calculé par Monte-Carlo indépendamment des simulations du prix, et il est donné par la formule suivante :

$$\rho^* = \mathbb{E}[\Psi(xe^{X_{T_0}})(1 - e^{X_{T_0}})](e^{\int_t^{T_0} [\varphi(2\sigma e^{-a(T_f-s)}) - 2\varphi(\sigma e^{-a(T_f-s)})] ds} - 1)^{-1}$$

Application²¹ :

✕ Le prix d'une option européenne \mathcal{NIG} :

$$\varphi_{\mathcal{NIG}}(x) = \mu x + \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + \omega + x)^2})$$

$$\Delta^n L_i \sim \mathcal{NIG}(\alpha, \beta, \delta \Delta^n t, \mu \Delta^n t), \quad \Delta^n t = \frac{T_0 - t}{n}$$

Le tableau ci-dessous résume les écart-types et les intervalles de confiance estimés avec et sans variable de contrôle pour le Call et le Put, le prix exact étant de 2.457, on prend comme référence $n = 10000$.

Nb de simulations m	Call Sans	Put Sans	Put Avec
10 000	0.080, [2.38, 2.69]	0.049, [2.39, 2.59]	0.041, [2.42, 2.58]
50 000	0.041, [2.39, 2.55]	0.021, [2.41, 2.49]	0.018, [2.42, 2.49]
500 000	0.013, [2.43, 2.48]	0.007, [2.44, 2.48]	0.006, [2.44, 2.47]

✕ Le prix d'une option européenne \mathcal{VG} :

$$\varphi_{\mathcal{VG}}(x) = \mu x + 2\lambda n \left(\frac{\sqrt{\alpha^2 - \beta^2}}{\sqrt{\alpha^2 - (\beta + \omega + x)^2}} \right)$$

$$\Delta^n L_i \sim \mathcal{VG}(\lambda \Delta^n t, \alpha, \beta, \mu \Delta^n t), \quad \Delta^n t = \frac{T_0 - t}{n}$$

Le tableau ci-dessous résume les écart-types et les intervalles de confiance estimés avec et sans variable de contrôle pour le Call et le Put, le prix exact étant de 16.293, on prend comme référence $n = 10000$.

Nb de simulations m	Call Sans	Put Sans	Put Avec
10 000	0.54, [15.84, 16.43]	0.17, [15.84, 16.43]	0.15, [15.84, 16.43]
50 000	0.19, [15.89, 16.65]	0.05, [16.19, 16.40]	0.04, [16.20, 16.39]
500 000	0.08, [16.07, 16.40]	0.02, [16.24, 16.34]	0.02, [16.24, 16.33]

²¹ Voir la partie précédente pour les valeurs numériques des paramètres utilisés.

5 Conclusion

Les données du prix spot sont récupérées sur le marché français d'électricité. Elles sont marquées par une saisonnalité quotidienne, hebdomadaire et annuelle. En dépit de cette saisonnalité les prix ont tendance à retourner à la moyenne. D'autre part, ils sont marqués par la présence de pics très importants. Ces trois caractéristiques ne sont pas totalement modélisées par un processus d'Ornstein-Uhlenbeck avec des accroissements gaussien (conclusion de la première partie).

Pour palier à ces faiblesses, nous avons proposé de modéliser les prix spot et forward électrique par un processus de Lévy de type hyperbolique généralisé (étude de la deuxième partie). On était spécialement attaché à travailler sur deux processus : Normal Inverse Gaussien et Variance Gamma. Ces processus présentent les degrés de libertés suffisantes pour calibrer les prix spots. Ils ont l'avantage d'être maniables mathématiquement. On a ainsi exposé le calcul d'option européenne, le calcul des prix d'option américaine se prolonge par simulation à l'aide de l'algorithme de Longstaff-Schwarz. On a utilisé une méthode de calibration par découplage. Elle est basée sur la méthode des moindres carrées et la méthode du maximum de vraisemblance.

Bibliographie

- [App03] D.Aalebaum, *Lectures on Lévy Processes in Euclidean Spaces And Groups*, Preprintreihe Mathematik, School of Computing and Mathematics, (2003).
- [BEHN99] F.E.Benth, L.Ekeland, R.Hauge, B.F.Nielsen, *On arbitrage-free pricing of forward contracts in energy markets*, Preprint, pp. 1-7, (2001).
- [Ben00] E.Benhamou, *Option pricing with Lévy Process*, Technical report, Goldman Sachs International, (2000).
- [CM99] P.Carr and D.B.Madan, *Option Valuation Using the Fast Fourier Transform*, J.Comp Finance, pp. 61-73, (1999).
- [CS00] L.Clewlow and C.Strickland, *Energy Derivatives. Pricing and Risk Management*, Lacima Publications, (2000).
- [Dev86] L.Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag, New York, pp. 401-412, (1986).
- [Ebexx] E.Eberlein, *Application of generalized hyperbolic Lévy motions to finance*, Technical report, Institut für Mathematische Stochastik, Preprint.
- [EH02] E.Eberlein and E.A.Hammerstein, *Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes*, Technical report, Department of Mathematical Stochastics University of Freiburg, (2002).
- [ES04] F.E.Benth and J.Saltyte-Benth, *The normal inverse Gaussian distribution and spot price modelling in energy markets*, Intern. J. Theor. Appl. Finance, 7 (2), pp. 177-192, (2004).
- [Gra01] J.Gran, *Electricity spot price dynamics*, Technical report, Stockholms Universitet Nationalekonomiska institutionen, (2001).
- [MCC02] D.B.Madan, P.Carr and E.C.Chang, *The variance gamma process and option pricing*, European Finance Review, pp. 5-13, 20-21, (1998).
- [ML02] V.Mignon et S.Lardic, *Econométrie des séries temporelles macroéconomiques et financières*, Economica, pp. 45, 274, 25-52, (2002).

-
- [Mon97] A.Monfort, *Cours de statistique mathématique, 3ème édition*, Economica, pp. 66, 109-115, (1997).
- [ND04] N.Oudjane et V.Duwing, *Etude des méthodes de calage du modèle à 2 facteurs*, Note HI-23/2003/018/A, Département SINETICS, EDF/R&D, Clamart, (20/04/2004).
- [Oud0603] N.Oudjane, *Modélisation des prix SPOT électriques par processus hyperbolique*, Note HI-23/02/022/A, Département SINETICS, EDF/R&D, Clamart, (20/06/2003).
- [Ouj0903] N.Oudjane, *Remarques sur le liens et différences entre le modèle factoriel horaire et le modèle par produit*, Note CR-I23/03/044, Département SINETICS, EDF/R&D, Clamart, (30/09/2003).
- [Pil03] B.Pillon, *Calage des paramètres factoriels*, Note HI-23/03/007, Département SINETICS, EDF/R&D, Clamart, (27/05/2003).
- [Pra99] K.Prause, *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*, Dissertation, Institut für Mathematische Stochastik, (1999).
- [Rai00] S.Raible, *Lévy Processes in Finance : Theory, Numerics, and Empirical Facts*, PhD thesis, Institut für Mathematische Stochastik, Universität Freiburg im Breisgau, (2000).
- [Sch03] W.Schoutens, *Lévy Processes in Finance : Pricing Financial Derivatives*, Wiley Publications, (2003).
- [SS00] E.Schwartz and J.E.Smith, *Short-Term Variations and Long-Term Dynamics in Commodity Prices*, Management Science, vol.46, no.7, (2000).
- [SS02] W.Schoutens and S.Symens, *The pricing of exotic options by Monte-carlo simulations in a Lévy market with stochastic volatility*, Technical report, pp.5-11, 20-21, (2002).
- [War03] X.Warin, *Pricing en dehors de la monnaie : processus de Lévy et valorisation*, Note HI-23/03/015/A, Département SINETICS, EDF/R&D, Clamart, (08/09/2003).
- [WR03] N.Webber and C.Ribeiro, *A Monte Carlo Methode for the Normal Inverse Gaussian Option Valuation Model using an Inverse Gaussian Bridge*, Technical report, City University Cass Business School, (2003).