

Analyse MCMC de certains modèles de diffusion  
avec application au marché européen du carbone

Jean-François Bégin  
Département de Mathématiques et Statistiques  
Université de Montréal  
Montréal, Canada

Août 2010

## Résumé

Cette analyse utilise une méthode d'estimation de paramètres dans un processus de diffusion proposée par Bjørn Eraker (2001) (voir la référence [5] dans la bibliographie pour plus de détails). En utilisant des séries chronologiques discrètes issues du marché du carbone européen (titres EUA et CER de la deuxième phase), nous utilisons les méthodes Monte Carlo par chaînes de Markov (MCMC) afin d'augmenter les données, ce qui fait disparaître le biais de discrétisation du processus de diffusion. De plus, nous utilisons les méthodes MCMC et le cadre de l'école statistique bayésienne afin de trouver des densités *a posteriori* empiriques pour chaque paramètre. Cette méthode est appliquée à trois modèles de la classe à élasticité constante de variance (en anglais, constant elasticity of variance) : processus de Orstein-Uhlenbeck, processus racine carrée et le processus d'Orstein-Uhlenbeck exponentiel. L'analyse compare les valeurs obtenues à l'aide de la méthode MCMC avec celles des estimateurs du maximum de vraisemblance (estimés par Godin (2010)).

MOTS-CLÉS : Analyse bayésienne ; processus de diffusion ; échantillonneur de Gibbs ; MCMC ; Marché du carbone européen ; EUA ; CER.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>6</b>
<b>2</b>	<b>Modèle bayésien</b>	<b>7</b>
2.1	Théorème de Bayes . . . . .	7
2.2	Grandes lignes de l'approche bayésienne . . . . .	7
2.3	Probabilités <i>a priori</i> et <i>a posteriori</i> . . . . .	8
<b>3</b>	<b>Méthodes Monte Carlo par chaînes de Markov (MCMC)</b>	<b>9</b>
3.1	Les chaînes de Markov, théorie . . . . .	9
3.2	Algorithme de Metropolis . . . . .	11
3.3	Algorithme de Metropolis-Hastings . . . . .	12
3.4	Échantillonnage de Gibbs . . . . .	13
3.5	Méthode d'acceptation-rejet . . . . .	13
3.6	Algorithme AR-MH (acceptation-rejet Metropolis-Hastings) . . . . .	14
<b>4</b>	<b>Processus de diffusion</b>	<b>16</b>
4.1	Introduction aux équations différentielles stochastiques . . . . .	16
4.2	Mouvement brownien ou processus de Wiener . . . . .	17
4.3	Classe de modèles à élasticité constante de variance . . . . .	18
4.3.1	Processus d'Orstein-Uhlenbeck . . . . .	19
4.3.2	Processus racine carrée . . . . .	19
4.3.3	Processus d'Orstein-Uhlenbeck exponentiel . . . . .	19
<b>5</b>	<b>Marché européen des certificats d'émission de carbone</b>	<b>20</b>
5.1	Historique . . . . .	20
5.2	Fonctionnement du marché européen . . . . .	21
<b>6</b>	<b>Méthodologie</b>	<b>22</b>
6.1	Définitions des éléments clés . . . . .	22
6.2	Algorithme . . . . .	23
<b>7</b>	<b>Applications et résultats</b>	<b>26</b>
7.1	<i>European Union Allowance</i> (EUA) . . . . .	28
7.1.1	Processus d'Orstein-Uhlenbeck . . . . .	28

7.1.2	Processus racine carrée . . . . .	28
7.1.3	Processus d'Orstein-Uhlenbeck exponentiel . . . . .	28
7.2	<i>Certified Emission Reductions</i> (CER) . . . . .	28
7.2.1	Processus d'Orstein-Uhlenbeck . . . . .	34
7.2.2	Processus racine carrée . . . . .	34
7.2.3	Processus d'Orstein-Uhlenbeck exponentiel . . . . .	34
7.3	Comparaisons avec l'expérience de Godin (2010) . . . . .	34
7.4	Comparaisons des modèles . . . . .	43
<b>8</b>	<b>Conclusion</b>	<b>46</b>
<b>A</b>	<b>Code <i>MATLAB</i> utilisé pour la simulation</b>	<b>47</b>
<b>B</b>	<b>Utilisation de la régression linéaire dans l'estimation des paramètres</b>	<b>49</b>

# Table des figures

7.1	Densité empirique <i>a posteriori</i> de $\sigma$ pour le processus d'Orstein-Uhlenbeck des EUA . . . . .	29
7.2	Densité empirique <i>a posteriori</i> de $\kappa_S$ pour le processus d'Orstein-Uhlenbeck des EUA . . . . .	29
7.3	Densité empirique <i>a posteriori</i> de $\theta_S$ pour le processus d'Orstein-Uhlenbeck des EUA . . . . .	30
7.4	Densité empirique <i>a posteriori</i> de $\sigma$ pour le processus racine carrée des EUA . . . . .	31
7.5	Densité empirique <i>a posteriori</i> de $\kappa_S$ pour le processus racine carrée des EUA . . . . .	31
7.6	Densité empirique <i>a posteriori</i> de $\theta_S$ pour le processus racine carrée des EUA . . . . .	32
7.7	Densité empirique <i>a posteriori</i> de $\sigma$ pour le processus d'Orstein-Uhlenbeck exponentiel des EUA . . . . .	33
7.8	Densité empirique <i>a posteriori</i> de $\kappa_S$ pour le processus d'Orstein-Uhlenbeck exponentiel des EUA . . . . .	33
7.9	Densité empirique <i>a posteriori</i> de $\theta_S$ pour le processus d'Orstein-Uhlenbeck exponentiel des EUA . . . . .	34
7.10	Densité empirique <i>a posteriori</i> de $\sigma$ pour le processus d'Orstein-Uhlenbeck des CER . . . . .	35
7.11	Densité empirique <i>a posteriori</i> de $\kappa_S$ pour le processus d'Orstein-Uhlenbeck des CER . . . . .	36
7.12	Densité empirique <i>a posteriori</i> de $\theta_S$ pour le processus d'Orstein-Uhlenbeck des CER . . . . .	36
7.13	Densité empirique <i>a posteriori</i> de $\sigma$ pour le processus racine carrée des CER . . . . .	37
7.14	Densité empirique <i>a posteriori</i> de $\kappa_S$ pour le processus racine carrée des CER . . . . .	38
7.15	Densité empirique <i>a posteriori</i> de $\theta_S$ pour le processus racine carrée des CER . . . . .	38
7.16	Densité empirique <i>a posteriori</i> de $\sigma$ pour le processus d'Orstein-Uhlenbeck exponentiel des CER . . . . .	39
7.17	Densité empirique <i>a posteriori</i> de $\kappa_S$ pour le processus d'Orstein-Uhlenbeck exponentiel des CER . . . . .	40

7.18	Densité empirique <i>a posteriori</i> de $\theta_S$ pour le processus d'Orstein-Uhlenbeck exponentiel des CER . . . . .	40
7.19	Un exemple de chemin emprunté par les processus d'Orstein-Uhlenbeck des EUA . . . . .	42
7.20	Un exemple de chemin emprunté par les processus racine carrée des EUA . . . . .	42
7.21	Un exemple de chemin emprunté par les processus d'Orstein-Uhlenbeck exponentiel des EUA . . . . .	43
7.22	Un exemple de chemin emprunté pour chaque processus des EUA à l'aide des estimations de notre analyse . . . . .	45

# Liste des tableaux

7.1	Statistiques descriptives des prix des EUA . . . . .	26
7.2	Autocorrélations de la série chronologique des prix des EUA .	27
7.3	Statistiques descriptives des prix des CER . . . . .	27
7.4	Autocorrélations de la série chronologique des prix des CER .	27
7.5	Estimation de moyenne et d'écart type pour les densités <i>a posteriori</i> des paramètres avec le processus d'Orstein-Uhlenbeck des EUA . . . . .	28
7.6	Estimation de moyenne et d'écart type pour les densités <i>a posteriori</i> des paramètres avec le processus racine carrée des EUA . . . . .	30
7.7	Estimation de moyenne et d'écart type pour les densités <i>a posteriori</i> des paramètres avec le processus d'Orstein-Uhlenbeck exponentiel des EUA . . . . .	32
7.8	Estimation de moyenne et d'écart type pour les densités <i>a posteriori</i> des paramètres avec le processus d'Orstein-Uhlenbeck des CER . . . . .	35
7.9	Estimation de moyenne et d'écart type pour les densités <i>a posteriori</i> des paramètres avec le processus racine carrée des CER . . . . .	37
7.10	Estimation de moyenne et d'écart type pour les densités <i>a posteriori</i> des paramètres avec le processus d'Orstein-Uhlenbeck exponentiel des CER . . . . .	39
7.11	Les valeurs de $K^*$ pour les trois modèles utilisés pour les données EUA . . . . .	44

# Chapitre 1

## Introduction

Le réchauffement climatique est une réalité d'aujourd'hui. Selon certains, c'est même le plus grand défi de notre génération. Il est maintenant très accepté que l'Homme est responsable ; du moins, c'est ce que conclut le GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat). Principalement, le dioxyde de carbone qui est émis essentiellement par nos moyens de transport est responsable. Ce réchauffement planétaire est susceptible d'affecter plusieurs pays de différentes manières : sécheresse, inondations, tempêtes, ouragans, etc.

Afin de réduire les émissions de gaz à effet de serre, et plus principalement les émissions de dioxyde de carbone, l'Union européenne a mis sur pied un marché du carbone. Dans cette analyse nous traiterons deux titres, l'EUA et le CER, transigés sur le marché du carbone européen, l'EU ETS. L'un des plus importants défis liés à ce marché est la modélisation mathématique des titres. De plus, il est important de noter que c'est un marché en pleine croissance (le marché double à chaque année).

L'objectif ici sera d'appliquer méthodes de Monte Carlo par chaînes de Markov à des modèles de diffusion de la classe à élasticité constante de variance (CEV) afin de modéliser les prix des titres EUA et CER. Pour ce faire, nous nous inspirerons du cadre présenté par Eraker (2001).

C'est dans le cadre d'un stage d'été en recherche du CRSNG co-supervisé par Madame Mylène Bédard et Monsieur Manuel Morales que nous vous présentons ce rapport. Il sera divisé en plusieurs sections. Les premiers chapitres traiteront des notions de base : l'approche bayésienne, les méthodes Monte Carlo par chaînes de Markov, les processus de diffusion et le marché du carbone européen. Ensuite, nous expliquerons la méthodologie utilisée pour produire l'expérience. Puis, nous montrerons les résultats obtenus à l'aide de nos simulations. Finalement, nous concluons l'analyse.



## Chapitre 2

# Modèle bayésien

L'essence de ce modèle et de cette école statistique nous provient du révérend Thomas Bayes. Toute la méthode de l'approche bayésienne découle essentiellement du fameux théorème de Bayes. Dans cette section, nous ferons une brève introduction à l'approche bayésienne : théorème de Bayes dans sa forme la plus simple, densité *a priori* et *a posteriori*.

### 2.1 Théorème de Bayes

Le théorème de Bayes (ou originalement appelé *de probabilité des causes*) est un résultat de base en théorie des probabilités. Étant donné deux événements  $A$  et  $B$ , il nous permet de trouver la probabilité de  $A$  sachant  $B$  si l'on connaît la probabilité de  $A$ , de  $B$  et de  $B$  sachant  $A$ . Dans sa forme la plus simple, on a que

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}. \quad (2.1)$$

### 2.2 Grandes lignes de l'approche bayésienne

En premier lieu, cette approche considère les paramètres comme des variables aléatoires. Ainsi, ces paramètres ont toutes une distribution dite *a priori* de densité  $\pi(\theta)$ . Ensuite, on peut utiliser les règles de probabilité (i.e. le théorème de Bayes vu en (2.1)) pour déduire une distribution *a posteriori* à partir de la distribution *a priori* et des données observées. Finalement, il est important de comprendre que la distribution *a priori* est subjective. On devrait la choisir avant d'observer les données. Chaque personne est libre de choisir sa propre densité *a priori* en fonction de ses croyances personnelles. Alors, on doit interpréter les probabilités dérivées à partir de cette approche comme un degré de croyances (ou de l'anglais *degree of belief*). Pour plus de détails sur l'approche bayésienne, consulter la référence [2].

## 2.3 Probabilités *a priori* et *a posteriori*

D'après ce qui a été décrit plus haut, on peut trouver finalement la distribution du paramètre qu'on veut estimer sachant nos données (ou encore  $\pi(\theta|x)$ ) en connaissant la distribution des données sachant le paramètre (ou  $f(x|\theta)$ ) et la distribution de notre paramètre (ou densité *a priori* ou  $\pi(\theta)$ ). Ainsi,

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int_{\Theta} f(x|\theta)\pi(\theta) d\theta}. \quad (2.2)$$

Cette dernière est la fameuse densité *a posteriori*. Elle représente ce qu'on sait par rapport au paramètre en considérant les données observées; c'est aussi la mise à jour de  $\pi(\theta)$  après l'observation de notre échantillon.

## Chapitre 3

# Méthodes Monte Carlo par chaînes de Markov (MCMC)

Dans ce chapitre, nous nous intéresserons aux méthodes MCMC, ou *Markov chain Monte Carlo*. Cette famille de méthodes de simulation d'une distribution arbitraire  $f$  (dans notre cas, la distribution *a posteriori*) a pour but de reproduire une chaîne de Markov avec une distribution stationnaire  $f$ . Ceci sera éclairci dans la page suivante. Alors, pour débiter, nous définirons ce qu'est une chaîne de Markov et nous donnerons ses principales propriétés. Ensuite, nous expliquerons quelques algorithmes MCMC utiles dans l'application qui nous intéresse.

### 3.1 Les chaînes de Markov, théorie

Une chaîne de Markov est une suite de variables aléatoires. Dénoteons  $X_t$  la valeur d'une variable aléatoire au temps  $t$ . Ces variables aléatoires peuvent être vues comme évoluant dans le temps, avec une probabilité de transition dépendante de l'état présent dans la chaîne. On peut définir un noyau de transition (ou *transition kernel*) comme une fonction qui détermine les transitions. Cette fonction est notée  $K$  et est définie sur  $\chi \times B(\chi)$  où  $\chi$  est l'espace des états (donc tous les états possibles). Si  $\chi$  est discret, alors le noyau de transition est simplement une matrice où chaque élément est

$$\begin{aligned} P_{ij} &= P(X_{n+1} = j | X_n = i, \dots, X_0 = i_0) \\ &= P(X_n = j | X_{n-1} = i_{n-1}) \end{aligned} \tag{3.1}$$

$\forall j, i_{n-1}, \dots, i_0 \in \chi$ . Si toutefois  $\chi$  est continu (comme ce sera le cas dans notre application) alors le noyau de transition est défini comme  $K(x, x')$ , où  $x$  est l'ensemble de l'état présent et  $x'$  est l'ensemble de l'état vers lequel on transite.

Alors, étant donné un noyau de transition  $K$ , une séquence  $\{X_n\}_{n \geq 0}$  de variables aléatoires est une chaîne de Markov si pour tout  $k$ , on a que

$$\begin{aligned} P(X_{k+1} \in A | x_0, x_1, \dots, x_k) &= P(X_{k+1} \in A | x_k) \\ &= \int_A K(x_k, dx) \end{aligned} \quad (3.2)$$

Cette propriété est appelée la propriété de Markov. En résumé, l'état suivant dépend uniquement de l'état présent et non de l'histoire de la chaîne.

Une chaîne de Markov peut avoir une distribution stationnaire. Intuitivement, la distribution stationnaire est telle que si on part la chaîne dans la distribution stationnaire, alors, à l'état prochain, la distribution de la chaîne sera encore la distribution stationnaire. Dans le cas des MCMC, on veut que cette distribution stationnaire soit la distribution *a posteriori*. Toutefois, pour que la distribution stationnaire soit unique et vrai peut importe le point de départ de la chaîne, cette dernière doit respecter quelques propriétés importantes.

### 1. Irréductibilité

Premièrement, on dit que  $i$  communique avec  $j$  si  $\exists n > 0$  tel que la probabilité d'aller à  $j$  à partir de  $i$  en  $n$  pas n'est pas nulle. De plus, on dit que  $C$  est une classe d'équivalence si  $\forall i, j \in C$ ,  $i$  communique avec  $j$  et  $j$  communique avec  $i$ . Alors, tous les états de cette classe sont du même type. Finalement, si une chaîne de Markov ne comporte qu'une seule classe d'équivalence, alors on dit que la chaîne est irréductible. Cette propriété se résume ainsi :

$$\exists n > 0 \text{ t.q. } P(X_n = j | X_0 = i) > 0 \quad \forall i, j. \quad (3.3)$$

### 2. Apériodicité

La période, notée  $d_i$ , d'un état  $i$  est le plus grand nombre qui divise tout  $n$  tel que la probabilité d'aller de  $i$  à  $i$  en  $n$  pas est strictement positive. En d'autres mots,

$$d_i = \text{PGCD}\{n : P_{ii}^{(n)} > 0, n \geq 1\} \quad (3.4)$$

où le *PGCD* est le plus grand commun diviseur. Si  $d_i = 1$ , on assume que l'état est apériodique. On peut aussi voir que si la probabilité d'aller à  $i$  à partir de  $i$  est non nulle, alors  $i$  est apériodique. Alors, si tous les états de la chaîne sont apériodiques, la chaîne est donc apériodique.

Avec ces deux propriétés, notre chaîne de Markov est assurée de converger vers une loi stationnaire. À ceci, on peut ajouter une dernière propriété : la réversibilité. Toutefois, elle n'est pas nécessaire ici. C'est seulement une propriété qui nous aide à construire une chaîne de Markov avec la loi stationnaire voulue,  $f$ .

### 3. Réversibilité

Une chaîne de Markov est réversible par rapport à  $f$  si la probabilité d'aller de  $i$  à  $j$  en  $n$  pas est la même que d'aller de  $j$  à  $i$  en  $n$  pas encore. Si on considère le cas discret, on a que :

$$f_i P_{ij}^{(n)} = f_j P_{ji}^{(n)} \quad (3.5)$$

où  $f_i$  est la valeur de la distribution stationnaire à l'état  $i$ .

Alors, ces propriétés nous permettent de construire notre chaîne de Markov dont la loi stationnaire converge vers la distribution que nous souhaitons (ici, la distribution *a posteriori*). Pour plus de détails et une introduction plus théorique, voir Robert et al. (2004). Ainsi, nous appliquerons les méthodes dites de Monte Carlo à ces chaînes de Markov pour construire nos algorithmes.

## 3.2 Algorithme de Metropolis

Nous commencerons notre visite des algorithmes MCMC par celui de Metropolis, même si ce dernier n'est pas utilisé directement dans l'article d'Eraker. La première version de cet algorithme a été décrite dans Metropolis et al. (1953) qui considérait le cas particulier de la distribution de Boltzmann, une distribution très utile en physique statistique. Depuis là, plusieurs autres applications ont été trouvées à cet algorithme (dont la finance). Notre but est de générer une chaîne de Markov qui converge vers la distribution souhaitée, disons ici  $f(x)$ . Pour ce faire, nous avons besoin d'une densité proposée arbitraire notée  $q$ . Cette loi instrumentale  $q$  doit être simulable rapidement (i.e. loi uniforme, normale, etc.) et doit être symétrique (i.e.  $q(\tilde{x}|x) = q(x|\tilde{x})$ ). L'algorithme va comme suit :

1. Initialiser  $x^{(0)}$ , soit le premier élément de la chaîne.
2. Poser  $i \leftarrow 1$ .
3. Simuler un candidat

$$\tilde{x} \sim q(\cdot|x^{(i-1)}) \quad (3.6)$$

qui dépend de la valeur précédente  $x^{(i-1)}$ .

4. Calculer la valeur

$$\alpha = \min\left\{1, \frac{f(\tilde{x})}{f(x^{(i-1)})}\right\}. \quad (3.7)$$

5. Accepter  $\tilde{x}$  avec la probabilité  $\alpha$  telle que

$$x^{(i)} = \begin{cases} \tilde{x} & \text{avec la probabilité } \alpha \\ x^{(i-1)} & \text{sinon} \end{cases}. \quad (3.8)$$

En pratique, on peut utiliser une loi uniforme sur  $[0, 1]$  pour effectuer cette étape.

6. Changer la valeur de  $i$  à  $i \leftarrow i + 1$  et aller en 3.

Étant donné une période de *burn-in* suffisante, la chaîne devrait tendre vers la distribution stationnaire (ou dans notre cas, la distribution  $f$  que nous voulons simuler). La période de *burn-in* dépend de la valeur initiale choisie et du comportement de la loi instrumentale  $q$ . Toutefois, cette méthode n'est pas sans défaut : elle ne génère pas d'échantillons indépendants et identiquement distribués parce que la probabilité d'acceptation dépend de la valeur précédente dans la chaîne.

### 3.3 Algorithme de Metropolis-Hastings

En 1970, quelques années après les travaux de Metropolis et ses collaborateurs, Hastings (1970) s'est penché sur le sujet. Il a étendu l'application de l'algorithme de Metropolis à des cas plus généraux. La grande différence entre l'algorithme de Metropolis et celui dit de Metropolis-Hastings est essentiellement la relaxation de l'hypothèse qui exige une distribution instrumentale symétrique (i.e.  $q(\tilde{x}|x) = q(x|\tilde{x})$ ). L'algorithme, qui ressemble beaucoup à celui décrit un peu plus haut, va comme suit :

1. Initialiser  $x^{(0)}$ , soit le premier élément de la chaîne.
2. Poser  $i \leftarrow 1$ .
3. Simuler un candidat

$$\tilde{x} \sim q(\cdot|x^{(i-1)}) \quad (3.9)$$

qui dépend de la valeur précédente  $x^{(i-1)}$ .

4. Calculer la valeur

$$\alpha = \min\left\{1, \frac{f(\tilde{x})}{f(x^{(i-1)})} \frac{q(x^{(i-1)}|\tilde{x})}{q(\tilde{x}|x^{(i-1)})}\right\}. \quad (3.10)$$

5. Accepter  $\tilde{x}$  avec la probabilité  $\alpha$  telle que

$$x^{(i)} = \begin{cases} \tilde{x} & \text{avec la probabilité } \alpha \\ x^{(i-1)} & \text{sinon} \end{cases}. \quad (3.11)$$

En pratique, on peut utiliser une loi uniforme sur  $[0, 1]$  pour effectuer cette étape.

6. Changer la valeur de  $i$  à  $i \leftarrow i + 1$  et aller en 3.

Essentiellement, la différence pratique entre les deux algorithmes se trouve dans le calcul de  $\alpha$ .

Il est à noter qu'il existe plusieurs autres algorithmes qui découlent de dernier. On peut penser à une variante du Metropolis-Hastings qui utilise une distribution instrumentale indépendante de  $x^{(i-1)}$  (i.e.  $q(\tilde{x})$ ). On nomme cette variante le *Independence Chain Metropolis-Hastings*.

### 3.4 Échantillonnage de Gibbs

On voit l'échantillonnage de Gibbs comme étant un cas particulier de l'algorithme de Metropolis-Hastings. On se sert de cet algorithme quand on veut échantillonner des valeurs d'une fonction multidimensionnelle avec un nombre fixe de variables (i.e.  $f(\mathbf{x})$  où  $\mathbf{x} = x_1, x_2, \dots, x_N$ ). La densité conjointe n'est pas nécessairement connue; toutefois, toutes les distributions conditionnelles doivent être connues. C'est dans cette optique que cette méthode semble être une bonne méthode pour faire l'échantillonnage d'une distribution *a posteriori* dans l'approche bayésienne. L'algorithme va comme suit :

1. Initialiser  $\mathbf{x}^{(0)} = x_1^{(0)}, x_2^{(0)}, \dots, x_N^{(0)}$ , soit le premier vecteur d'éléments de la chaîne.
2. Poser  $i \leftarrow 0$ .
3. Simuler

$$\begin{aligned} x_1^{(i+1)} &\sim f(\cdot | x_2^{(i)}, x_3^{(i)}, \dots, x_N^{(i)}) \\ x_2^{(i+1)} &\sim f(\cdot | x_1^{(i+1)}, x_3^{(i)}, \dots, x_N^{(i)}) \\ &\vdots \\ x_N^{(i+1)} &\sim f(\cdot | x_1^{(i+1)}, x_2^{(i+1)}, \dots, x_{N-1}^{(i+1)}) \end{aligned} \tag{3.12}$$

4. Changer la valeur de  $i$  à  $i \leftarrow i + 1$  et aller en 3.

Ici, tous les candidats simulés sont acceptés contrairement aux deux algorithmes traités précédemment (i.e.  $\alpha = 1$ ). La raison pour laquelle  $\alpha = 1$  est qu'on génère des valeurs des vraies distributions, alors nous n'avons pas besoin de corriger avec une certaine probabilité d'acceptation.

Comme on peut le voir, cette méthode est très facile à implémenter. Toutefois, cette méthode a quelques problèmes connus dans certains cas particuliers. Premièrement, dans une distribution à grande dimension, si deux variables sont parfaitement corrélées, alors l'échantillonneur *restera pris* et ne pourra pas changer les valeurs associées à ces variables. De plus, si une fonction est définie de telle sorte que tous les états ont une probabilité associée très près de zéro et qu'un seul point de l'espace a une très grande probabilité, alors l'échantillonneur prendra trop de temps pour converger. Ceci occasionnera des erreurs dans la simulation (voir les références [19] et [3]).

### 3.5 Méthode d'acceptation-rejet

En premier lieu, l'algorithme d'acceptation-rejet (ou *accept-reject*) est une technique de simulation. Toutefois, ce n'est pas une méthode MCMC car

elle n'implique pas la construction d'une chaîne de Markov. On doit premièrement choisir une densité instrumentale notée  $q$ . Cette densité doit respecter cette propriété :

$$f(x) \leq Mq(x), \forall x, M \geq 1. \quad (3.13)$$

Le terme  $Mq(x)$  est appelé l'enveloppe ; normalement, l'échantillonnage de cette distribution est facile. Le  $M$  sera déterminé en fonction que l'équation (3.13) soit toujours valide. Ensuite, il nous reste qu'à appliquer cet algorithme plutôt simple :

1. Poser  $i \leftarrow 1$ .

2. Simuler le candidat

$$\tilde{x} \sim q(\cdot). \quad (3.14)$$

3. Calculer la valeur

$$\alpha = \min\left\{1, \frac{f(\tilde{x})}{Mq(\tilde{x})}\right\} \quad (3.15)$$

avec  $M$  bien choisi.

4. Accepter  $\tilde{x}$  avec la probabilité  $\alpha$  tel que  $x^{(i)} = \tilde{x}$ . S'il y a rejet, on retourne à l'étape 2.

5. Changer la valeur de  $i$  à  $i \leftarrow i + 1$  et aller en 2.

Nous ne nous servons pas de cet algorithme directement ; toutefois, il est important de saisir le concept pour pouvoir comprendre la prochaine méthode de simulation.

### 3.6 Algorithme AR-MH (acceptation-rejet Metropolis-Hastings)

Le dernier algorithme que nous allons traiter dans le cadre de notre analyse est celui d'acceptation-rejet Metropolis-Hastings. C'est un algorithme hybride entre la méthode d'acceptation-rejet et le Metropolis-Hastings décrit plus haut. Nous l'utiliserons durant notre analyse. L'algorithme va comme suit :

1. Initialiser  $x^{(0)}$ , soit le premier élément de la chaîne.

2. Poser  $i \leftarrow 1$ .

3. Simuler le candidat

$$\tilde{x} \sim q(\cdot). \quad (3.16)$$

indépendant de l'état actuel de la chaîne de Markov.

4. Calculer la valeur

$$\alpha = \min\left\{1, \frac{f(\tilde{x})}{Mq(\tilde{x})}\right\} \quad (3.17)$$

avec  $M$  bien choisi.



5. Accepter  $\tilde{x}$  avec la probabilité  $\alpha$ . S'il y a rejet, on retourne à l'étape 3.

6. Calculer la valeur

$$\alpha' = \begin{cases} 1 & \text{si } f(x^{(i-1)}) < Mq(x^{(i-1)}) \\ \frac{f(x^{(i-1)})}{Mq(x^{(i-1)})} & \text{si } f(x^{(i-1)}) > Mq(x^{(i-1)}) \\ & \text{et } f(\tilde{x}) < Mq(\tilde{x}) \\ \min\left\{1, \frac{f(\tilde{x})}{f(x^{(i-1)})} \frac{q(x^{(i-1)})}{q(\tilde{x})}\right\} & \text{si } f(x^{(i-1)}) > Mq(x^{(i-1)}) \\ & \text{et } f(\tilde{x}) > Mq(\tilde{x}) \end{cases} . \quad (3.18)$$

7. Accepter  $\tilde{x}$  avec la probabilité  $\alpha'$  tel que

$$x^{(i)} = \begin{cases} \tilde{x} & \text{avec la probabilité } \alpha' \\ x^{(i-1)} & \text{sinon} \end{cases} . \quad (3.19)$$

8. Changer la valeur de  $i$  à  $i \leftarrow i + 1$  et aller en 3.

Comme on peut le remarquer, cet algorithme est un cas particulier d'un échantillonneur indépendant (*Independence Chain Metropolis-Hastings*).

Ceci fait très bien le tour de toutes les techniques MCMC que nous utiliserons durant notre travail.

# Chapitre 4

## Processus de diffusion

Ce chapitre traitera des processus de diffusion et des équations stochastiques différentielles. Essentiellement, nous ferons une introduction des concepts de base de cette théorie. Nous vous présenterons ensuite les trois modèles importants de notre analyse.

### 4.1 Introduction aux équations différentielles stochastiques

Les équations différentielles stochastiques sont des équations différentielles qui contiennent un ou plusieurs termes qui est un processus stochastique. La solution de ces équations est un processus stochastique aussi. Nous nous intéresserons ici plus particulièrement au processus de diffusion. Normalement, les EDS (ou équations différentielles stochastiques) incorporent du bruit blanc qui ici peut être vue comme la dérivée d'un mouvement brownien (sera introduit plus tard). Nos modèles ont cette caractéristique. En mathématiques financières, la notation est un peu différente si on la compare à celle vu en physique. Une équation typique utilisée, qui ressemblera à ce que nous verrons plus tard dans ce document est

$$dY_t = \mu(Y_t : \theta) dt + \sigma(Y_t : \theta) dW_t \quad (4.1)$$

où  $W_t$  est le mouvement brownien (dit standard) en  $d$  dimensions. On pourrait aussi exprimer la même équation sous forme d'intégrale :

$$Y_{t+s} - Y_t = \int_t^{t+s} \mu(Y_u : \theta) du + \int_t^{t+s} \sigma(Y_u : \theta) dW_u. \quad (4.2)$$

Pour s'assurer de l'existence et de l'unicité de la solution de l'équation différentielle stochastique, un théorème existe. On le connaît sous le nom de théorème général de l'existence et de l'unicité. Supposons les équations suivantes :

$$dY_t = \mu(Y_t : \theta) dt + \sigma(Y_t : \theta) dW_t \quad (4.3)$$

avec  $Y_0$  connu. Pour que l'existence et l'unicité soient respectées, on doit respecter ces trois contraintes (voir aussi la référence [10]) :

1. Les fonctions  $\mu$  et  $\sigma$  sont des fonctions mesurables qui vérifient la condition de Lipschitz, soit  $\forall(x_1, \theta), (x_2, \theta) \in \mathbb{R}^d \times \Theta$  :

$$\|\mu(x_1 : \theta) - \mu(x_2 : \theta)\| \leq C\|x_1 - x_2\|, \quad (4.4)$$

$$\|\sigma(x_1 : \theta) - \sigma(x_2 : \theta)\| \leq C\|x_1 - x_2\|, \quad (4.5)$$

avec  $C$  une constante positive quelconque.

2. Les fonctions  $\mu$  et  $\sigma$  sont bornées linéairement telles que

$$\|\mu(x_1 : \theta)\|^2 + \|\sigma(x_1 : \theta)\|^2 \leq C^2 (1 + \|x_1\|^2). \quad (4.6)$$

avec  $C$  une constante positive quelconque. Il est aussi à noter que  $\|\cdot\|$  symbolise la norme euclidienne.

3. La valeur initiale  $Y_0$  appartient à  $L^2(\Omega, \mathfrak{F}, P)$  et est indépendante de la  $\sigma$ -algèbre  $\sigma(W_t, t \in [0, T])$ , alors, il existe une solution pour tout  $t \in [0, T]$  appartenant à  $L^2(\Omega, \mathfrak{F}, P)$ , continu et unique sur l'intervalle.

Dans son article, Eraker fait part des deux premières conditions. Il ajoute aussi à ces deux conditions deux autres conditions.

1.  $\sigma(x : \theta) \sigma(x : \theta)'$  doit être défini positif pour tout  $x \in \mathbb{R}^d$ . Il est à noter que  $'$  définit la transposée de la matrice.
2. On a un échantillon d'observations  $Y_{t,j}$  pour  $t = 0, 1, 2, \dots, T$ , et  $j = 1, 2, \dots, d$ .

Ces conditions sont citées pour d'autres soucis que ceux mentionnés précédemment. La première condition est nécessaire pour avoir une fonction de vraisemblance bien définie. La deuxième condition ajoutée est présente pour la nous permettre d'utiliser des données observables et non-observables.

## 4.2 Mouvement brownien ou processus de Wiener

Le mouvement brownien est un processus stochastique. Un processus stochastique, c'est un processus aléatoire qui est fonction du temps. En gros, le mot *stochastique* est, comme nous l'explique N.N. Taleb dans son livre *Foolled by randomness* (voir référence [18]), une appellation huppée pour dire aléatoire (*Stochastic is a fancy Greek name for random*). Le mouvement brownien doit son nom au botaniste écossais Robert Brown ; ce dernier observait le pollen au microscope et a constaté la présence de très petites particules bougeant dans tous les sens. C'est de là que vient le nom.

Une manière d'expliquer et de définir ce mouvement est le suivant : supposons une marche aléatoire symétrique telle que

$$X_i = \begin{cases} 1 & \text{avec la probabilité } \frac{1}{2} \\ -1 & \text{avec la probabilité } \frac{1}{2} \end{cases}. \quad (4.7)$$

Ensuite, accélérons ce processus en prenant des pas de plus en plus petits (i.e.  $\Delta X \rightarrow 0$ ) et en prenant des intervalles de temps de plus en plus petits (i.e.  $\Delta t \rightarrow 0$ ). En passant à la limite, on obtient le mouvement brownien. Norbert Wiener a uniformisé cette théorie avec la rigueur mathématique nécessaire; c'est pour cela qu'on entend parfois parler du processus de Wiener lorsque certains parlent du mouvement brownien. Supposons que  $W_t$  est un processus de Wiener standard au temps  $t$ ; alors,  $W_t$  répond à ces caractéristiques :

1.  $Z_0 = 0$ ;
2.  $Z_{t+s} - Z_t \sim \mathcal{N}(0, s)$ ;
3.  $Z_{t+s} - Z_t$  est indépendant de  $Z_t - Z_{t-r}$  si  $r, s > 0$ ;
4.  $Z_t$  est continue;
5.  $Z_{t+s} - Z_t$  ne dépend pas de  $t$ .

Les processus de diffusion et les équations différentielles stochastiques utilisés dans notre analyse contiennent des mouvements browniens, alors ceci sera sans doute utile.

### 4.3 Classe de modèles à élasticité constante de variance

La classe des modèles *CEV* est un modèle à un facteur qui prend la forme suivante :

$$dS_t = (\theta_S + \kappa_S S_t)dt + \sigma S_t^\beta dW_t \quad (4.8)$$

où  $S_t$  est le facteur du modèle, soit le prix du sous-jacent dans notre cas,  $W_t$  est le processus de Wiener,  $\beta$  est l'élasticité de la variance. La propriété de réversion à la moyenne peut être présente (si  $\theta_r$  et  $\kappa_r$  sont bien choisis). Il est à noter que cette dernière écriture du modèle *CEV* est très générale et englobe plusieurs modèles particuliers de la classe *CEV* : si  $\beta = 0$ , on retrouve le processus d'Orstein-Uhlenbeck (ou plus connu sous le nom de modèle de Vasicek en finance) et si  $\beta = 0.5$ , on retrouve le processus racine carrée (ou modèle le Cox, Ingersoll et Ross pour les taux d'intérêt). Nous traiterons ces deux modèles dans les prochaines pages. Cette classe de modèles est facile à implémenter. Toutefois, Eraker (voir référence [5]) note dans son article que cette classe de modèle est incapable de reproduire la valeur élevée du coefficient d'aplatissement observé. Cette observation était liée à une analyse menée sur les taux d'intérêt. Dans une analyse portant cette fois sur les prix des certificats d'émission en question (voir référence [7]), l'auteur considère que le processus d'Orstein-Uhlenbeck est le modèle qui donne les meilleurs résultats. La méthodologie de cet auteur est cependant très différente de la nôtre; il utilise les maximums de vraisemblance afin de déterminer les valeurs des paramètres. Plus de détails sur ces modèles sont disponibles dans la référence [14].

### 4.3.1 Processus d'Orstein-Uhlenbeck

Ce modèle à un facteur de la classe à élasticité constante de variance prend cette forme

$$dS_t = (\theta_S + \kappa_S S_t)dt + \sigma dW_t. \quad (4.9)$$

Dans le cas où  $\theta_S > 0$  et  $\kappa_S < 0$ , nous avons la propriété de réversion à la moyenne. : si le prix du sous-jacent devient trop élevé, la moyenne du processus devient négative et si le prix devient trop bas, alors la moyenne du processus devient positive. Un problème de ce modèle est que les prix des sous-jacents peuvent devenir négatifs (il n'y a pas de contrainte pour empêcher ceci) ; nous ne souhaitons pas cela.

### 4.3.2 Processus racine carrée

Ce modèle, toujours dans la classe à élasticité constante de variance, prend la forme suivante

$$dS_t = (\theta_S + \kappa_S S_t)dt + \sigma\sqrt{S_t}dW_t. \quad (4.10)$$

Ce modèle permet aussi le retour à la moyenne. De plus, le problème majeur du processus d'Orstein-Uhlenbeck, soit le fait que les prix peuvent devenir négatifs, n'est plus un problème ici. La dérive du processus est maintenant proportionnelle à la racine carrée du prix du sous-jacent.

### 4.3.3 Processus d'Orstein-Uhlenbeck exponentiel

Ce modèle est dans la classe CEV. Il est très semblable au processus d'Orstein-Uhlenbeck standard. Toutefois, il est exponentiel. Il prend la forme suivante :

$$d(\ln S_t) = (\theta_S + \kappa_S (\ln S_t))dt + \sigma dW_t. \quad (4.11)$$

La différence par rapport aux autres modèles, c'est qu'ici on considère le logarithme du prix et non le prix en tant que tel. Ce modèle permet encore le retour à la moyenne.

Voilà ce qui conclut notre brève introduction aux modèles de diffusion que nous allons utiliser.

## Chapitre 5

# Marché européen des certificats d'émission de carbone

Dans ce chapitre, nous présenterons un bref historique du marché européen des certificats d'émission de carbone. De plus, nous introduirons le fonctionnement du marché en question.

### 5.1 Historique

Vers la fin des années 1980, plusieurs scientifiques et experts se sont mobilisés pour analyser la situation du réchauffement climatique et des gaz à effet de serre. Leurs conclusions furent très alarmantes : l'état de la planète se détériore à vue d'oeil. De plus, en 1995, le Groupe d'experts intergouvernemental sur l'évolution du climat écrit dans un rapport que l'étude des preuves suggère une influence détectable de l'activité humaine sur le climat planétaire. Ce rapport est le coup d'envoi de toutes les démarches internationales qui ont précédé la signature en 1997 du plus important traité mondial visant la réduction des émissions des gaz à effet de serre : le protocole de Kyoto. C'est dans le cadre de ce protocole que l'Union européenne s'est engagée à réduire ses émissions de gaz à effet de serre entre les années 2008 et 2012 de 8 % par rapport au niveau des années 1990. Pour ce faire, ils ont décidé d'implanter un marché de certificats d'émission de dioxyde de carbone (il est à noter que le dioxyde de carbone est le principal gaz à effet de serre produit par l'activité humaine). Les consultations ont débuté en mars 2000 ; en 2003, les directives sur le marché des émissions européennes (*European Emissions Trading Directive*) ont été entérinées. Deux ans plus tard, en 2005, le marché européen du carbone, le *European Union Emission Trading Scheme* (EU ETS) a vu le jour. C'est aujourd'hui le plus grand marché de ce type au monde.

## 5.2 Fonctionnement du marché européen

Le marché du carbone est basé sur le mécanisme du *Cap and Trade* (voir référence [6]) : on met un plafond d'émissions possibles sur le territoire européen et on laisse les entreprises faire le travail (par l'intermédiaire de bourses électroniques organisées). Les entreprises doivent acheter (ou vendre) des quotas supplémentaires pour ne pas dépasser leur quota permis. Ce quota se compte en terme de EUA (*European Union Allowance*), ou permis d'émission de dioxyde de carbone. Chaque EUA détenue par une compagnie donne le droit d'émettre une tonne de  $CO_2$ . Si la compagnie a émis plus de dioxyde de carbone qu'elle ne détient de EUA, elle aura une pénalité à payer. Le EU ETS divise ces permis d'émission en trois catégories selon le moment où l'entreprise a le droit d'émettre sa tonne de  $CO_2$  : la première phase (2005-2007), la deuxième phase (2008-2012) et la troisième phase (2013-2020). Si, par exemple, une compagnie a un EUA pour la première phase, alors elle peut émettre une tonne de dioxyde de carbone entre 2005 et 2007. Toutefois, en 2008, l'EUA expire. L'entreprise aura donc besoin d'un EUA pour la deuxième période dans ce cas. De plus, les EUA d'une phase sont émis, vendus ou donnés au début de la phase en question.

Malheureusement, ce ne sont pas toutes les entreprises européennes qui sont soumises à la réglementation du EU ETS ; en 2007, environ 50% des émissions de la zone euro étaient légiférées.

Afin d'harmoniser le fonctionnement du marché du carbone européen avec les directives du protocole de Kyoto, l'UE permet l'émission de certificats alternatifs, ou CER (*certified emission reductions*). Ces certificats permettent aussi d'émettre une tonne de dioxyde de carbone. Pour obtenir un tel certificat, une entreprise doit réaliser un projet qui permet la réduction d'émissions des gaz à effet de serre à l'étranger.

Il existe plusieurs bourses électroniques européennes qui s'occupent des échanges des titres CER et EUA tel que Bluenext, European Energy Exchange, Nord Pool, European Climate Exchange, etc. De plus, certaines bourses transigent des contrats à terme (pour la phase actuelle ou pour les phases futures) et des produits dérivés des titres EUA et CER.

# Chapitre 6

## Méthodologie

Dans ce chapitre, nous décrivons en profondeur l'algorithme utilisé. L'algorithme implémenté ici est très semblable à celui qu'Eraker a utilisé pour les taux d'intérêt des bons du Trésor américain (voir Eraker (2001)). Toutefois, dans notre analyse, nous utiliserons un coefficient d'élasticité de la variance,  $\beta$ , fixé à l'avance (c'est-à-dire 0 et 0,5 pour les processus de Orstein-Uhlenbeck et racine carrée respectivement). Le fait de fixer le coefficient d'élasticité de la variance n'a pas vraiment d'effet sur l'expérience, selon Eraker. De plus, nous définirons une densité *a priori*, chose qui n'a pas été faite explicitement par Eraker.

### 6.1 Définitions des éléments clés

Pour débiter, nous allons définir une matrice, nommée  $\hat{Y}$ , qui contient les informations connues,  $S_i$  (soient les prix des sous-jacents) et des informations inconnues,  $\hat{S}_j$  (qui seront entre les données connues). On simulera ces dernières à l'aide des méthodes MCMC ; le but ici est d'augmenter les données hebdomadaire (soit les  $S_j$ ) afin d'avoir, entre chaque donnée connue, une, trois ou sept données simulées. Comme il a été mentionné un peu plus tôt dans cette analyse, le fait d'estimer les paramètres d'un processus à l'aide de données très discontinue amène un certain biais (dit de discrétisation). Maintenant, en augmentant les données (c.-à.-d. en ajoutant des points entre chaque point connu de données), nous allons tenté de diminuer le biais de discrétisation à un niveau le plus minimal possible. La matrice  $\hat{Y}$  prend donc la forme suivante

$$\hat{Y} = \begin{bmatrix} S_{1,t_0} & \hat{S}_{1,t_1} & \dots & S_{1,t_m} & \hat{S}_{1,t_{m+1}} & \dots & S_{1,t_n} \end{bmatrix}. \quad (6.1)$$

Cette matrice consiste en une matrice  $1 \times (n+1)$  où  $n+1$  est le nombre total de données (observées et non observées). Ce nombre dépend de  $m$ . Il est à noter que  $m-1$  est défini comme le nombre de points ajouté entre deux



données observées et  $\frac{1}{m} = \Delta t$ . Dans cette analyse,  $\Delta t = 1$  est équivalent à une semaine. Dorénavant, nous noterons  $\hat{Y}_i$  la colonne  $i$  de la matrice  $\hat{Y}$ .

Nous savons que la densité de l'observation  $i$  sachant les paramètres du processus,  $\theta = (\theta_S, \kappa_S, \sigma)$ , est

$$p(\hat{Y}_i|\theta) = \frac{1}{\sigma \hat{Y}_{i-1}^\beta} \exp\left(-\frac{1}{2} \frac{(\hat{Y}_i - \hat{Y}_{i-1} - (\theta_S + \kappa_S \hat{Y}_{i-1}) \Delta t)^2}{\sigma^2 \hat{Y}_{i-1}^{2\beta} \Delta t}\right) \quad (6.2)$$

pour des raisons évidentes liées aux propriétés du mouvement brownien standard (c'est-à-dire  $W_s - W_{s-t} \sim \mathcal{N}(0, t)$ ). Il est à noter que nous utiliserons  $p$  pour les densités normalisées et  $\pi$  pour les formes non normalisées. De plus, nous poserons que la fonction *a priori* des paramètres du processus est non informative. Elle prendra la forme

$$p(\theta) \propto 1. \quad (6.3)$$

Cette densité *a priori* nous permet, comme son nom nous le dit, de n'introduire aucune information supplémentaire. Ici, on ne connaît pas les densités « naturelles » des paramètres du modèle (voir le chapitre 4.1 pour plus de détails sur ces fameux paramètres). Donc, en utilisant les bases de la méthodologie bayésienne (voir chapitre 2), on a que

$$\pi(\theta|\hat{Y}) \propto \prod_{i=1}^n \frac{1}{\sigma \hat{Y}_{i-1}^\beta} \exp\left(-\frac{1}{2} \frac{(\hat{Y}_i - \hat{Y}_{i-1} - (\theta_S + \kappa_S \hat{Y}_{i-1}) \Delta t)^2}{\sigma^2 \hat{Y}_{i-1}^{2\beta} \Delta t}\right). \quad (6.4)$$

## 6.2 Algorithme

L'algorithme que nous utiliserons n'est pas très compliqué. Il consiste en 6 étapes qui forment un échantillonneur de Gibbs (décrit au chapitre 3). Certaines étapes de cet échantillonneur sont directes; d'autres nécessitent de passer par des algorithmes MCMC plus complexes tel que le AR-MH (voir la section 3.6). Nous allons maintenant décrire chacune des étapes en détails.

1. On initialise toutes les données inconnues de la matrice  $\hat{Y}$  et tous les paramètres du processus. Les valeurs initiales des paramètres n'ont pas vraiment d'importance; les MCMC finiront par converger vers les bonnes valeurs. Pour les  $\hat{S}_i$ , on les initialise en utilisant une simple interpolation linéaire; ici encore, ça n'a pas beaucoup d'importance. De plus, on pose  $h \leftarrow 1$ . La variable  $h$  comptera le nombre d'itérations qu'on fera.
2. Ensuite, pour tous les  $i$ , on utilise l'algorithme hybride d'acceptation-rejet Metropolis-Hastings pour trouver les  $\hat{Y}_i^{(h)} | \hat{Y}_{\setminus i}^{(h-1)}, \theta$  où  $\hat{Y}_{\setminus i}^{(h-1)}$

signifie la matrice  $\hat{Y}$  sans la colonne  $i$  à l'itération  $h - 1$ . La densité de cette dernière est assez compliquée :

$$p\left(\hat{Y}_i^{(h)} | \hat{Y}_{\setminus i}^{(h-1)}, \theta\right) \propto \frac{1}{\sigma \hat{Y}_{i-1}^{(h)\beta}} \frac{1}{\sigma \hat{Y}_i^{(h)\beta}} \exp\left(-\frac{1}{2} \frac{\left(\hat{Y}_i^{(h)} - \hat{Y}_{i-1}^{(h)} - \left(\theta_S + \kappa_S \hat{Y}_{i-1}^{(h)}\right) \Delta t\right)^2}{\sigma^2 \hat{Y}_{i-1}^{(h)2\beta} \Delta t} - \frac{1}{2} \frac{\left(\hat{Y}_{i+1}^{(h-1)} - \hat{Y}_i^{(h)} - \left(\theta_S + \kappa_S \hat{Y}_i^{(h)}\right) \Delta t\right)^2}{\sigma^2 \hat{Y}_i^{(h)2\beta} \Delta t}\right). \quad (6.5)$$

Il est très difficile de tirer un échantillon directement à partir de cette densité ; c'est pourquoi on utilise l'algorithme de AR-MH (selon Eraker (2001), la convergence serait plus rapide avec ce genre d'algorithme). On utilise une normale comme loi instrumentale :

$$q \sim \mathcal{N}\left(\frac{1}{2}(\hat{Y}_{i-1}^{(h)} + \hat{Y}_{i+1}^{(h-1)}), \frac{1}{2}(\sigma \hat{Y}_{i-1}^{(h)\beta})^2 \Delta t\right). \quad (6.6)$$

La simulation d'un échantillon pour chaque colonne de la matrice  $\hat{Y}$  devient alors très facile. Le choix de paramètres de la loi instrumentale est un peu naturel : la moyenne de la normale est la moyenne des deux points qui se trouve à côté du point qu'on cherche à générer et la variance correspond à la variance du processus (estimée au point précédent).

3. Puis, on va chercher à estimer des valeurs moyennes pour les paramètres de notre processus à l'aide d'une régression linéaire. Ici, on doit définir le vecteur  $y$  et une matrice de design  $X$ . Les éléments du vecteur  $y$  sont définis comme une empilade de

$$y_i = \begin{bmatrix} \hat{Y}_i - \hat{Y}_{i-1} \\ \sqrt{\Delta t} \hat{Y}_i^\beta \end{bmatrix} \quad (6.7)$$

pour chaque  $i$ . Pour la matrice de design  $X$ , les éléments

$$X_i = \begin{bmatrix} \frac{\Delta t}{\sqrt{\Delta t} \hat{Y}_i^\beta} & \frac{\hat{Y}_i \Delta t}{\sqrt{\Delta t} \hat{Y}_i^\beta} \end{bmatrix} \quad (6.8)$$

sont empilés. Le but de cette étape est de faire une régression linéaire du vecteur  $y$  sur la matrice de design  $X$ . Nous allons évaluer deux choses : les estimateurs  $(\bar{\theta}_S, \bar{\kappa}_S)$  et la moyenne du carré des résidus, notée  $\bar{s}^2$ . Comme dans une régression, on évalue les estimateurs par

$$\bar{\theta} = (\bar{\theta}_S, \bar{\kappa}_S)' = (X'X)^{-1}(X'y) \quad (6.9)$$

où le  $X'$  signifie la matrice transposée de  $X$ . La moyenne du carré des résidus se calcule ainsi :

$$\bar{s}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - X_i \bar{\theta})^2. \quad (6.10)$$

Des informations supplémentaires concernant les arguments qui nous poussent à faire cette étape sont disponibles dans l'annexe A de cette analyse.

4. Ensuite, on tire  $(\theta_S, \kappa_S)^{(h)}$ , les valeurs de  $\theta_S$  et  $\kappa_S$  à l'itération  $h$  à partir d'une densité normale :

$$(\theta_S, \kappa_S) | \sigma, \hat{Y} \sim \mathcal{N}(\bar{\theta}, \sigma^2 (X'X)^{-1}). \quad (6.11)$$

Le terme de variance de la normale multivariée est un résultat assez standard en régression linéaire. La loi des paramètres est alors une normale qui est très facile à simuler.

5. Puis alors, on tire  $\sigma^{(h)}$ , la valeur de  $\sigma$  à l'itération  $h$  à partir d'une densité chi-deux :

$$\sigma^{-2} | \hat{Y} \sim \frac{\chi_{(n-2)}^2}{(n-k)\bar{s}^2}. \quad (6.12)$$

Ici, nous n'utilisons pas la méthode Eraker qui semble avoir des lacunes : manque de définition dans son article et manque de précision sur les densités *a priori* utilisées. Voir la référence [9] pour plus d'informations sur les hypothèses qui nous permettent d'utiliser la distribution chi-deux.

6. Finalement, on augmente la valeur de  $h$  et on retourne à l'étape 2.

En suivant ces 6 étapes, nous obtenons des densités empiriques pour chaque paramètre du modèle. De plus, nous obtenons des densités empiriques pour chaque donnée non observée de la matrice  $Y$ . L'analyse des prix des certificats d'émission (EUA et CER) sera faite au chapitre suivant.

## Chapitre 7

# Applications et résultats

Dans ce chapitre, nous ferons part des résultats de l'analyse MCMC pour les certificats d'émission de type EUA et CER. Les données utilisées pour l'analyse proviennent de la bourse Bluenext (<http://www.bluenext.fr/statistics/downloads.html>). Pour les EUA, nous avons sélectionné les données pour la deuxième phase d'émission, soit 26 février 2008 au 20 juillet 2010. Nous travaillerons avec les prix de fermeture à la fin de chaque semaine (du vendredi). Pour les CER, nous travaillerons aussi avec les données de la deuxième phase. Toutefois, les données sur la bourse Bluenext pour les CER sont disponibles uniquement à partir du 12 août 2008. Nous prendrons encore les prix de fermeture à la fin de la semaine pour notre analyse. Comme Eraker (voir la référence [5]), le but ici est d'augmenter la quantité de données à l'aide de simulation MCMC ; c'est pourquoi nous n'utiliserons pas toutes les données disponibles. Toutefois, il est à noter que les statistiques descriptives calculées ci-dessous proviennent de l'ensemble des données disponibles sur la bourse Bluenext.

Les tables 7.1 et 7.3 contiennent les statistiques descriptives pour les deux séries chronologiques de cette analyse. En moyenne, les CER valent moins que les EUA. De plus, la volatilité associée au EUA est plus élevée que celle des CER. Les tables 7.2 et 7.4 nous montre que les séries utilisées sont hautement autocorrélées.

	$S_i$	$S_i - S_{i-1}$
Moyenne	16,5858	-0,0108
Médiane	14,5500	0,0100
Écart type	4,9403	0,4073
Coefficient d'asymétrie	0,8549	-0,4937
Coefficient d'aplatissement	2,3849	4,5007

TABLE 7.1 – Statistiques descriptives des prix des EUA

lag	$S_i$	$S_i - S_{i-1}$
1	0,9991	0,0689
2	0,9982	-0,1178
3	0,9973	0,0486
4	0,9964	0,0956
5	0,9954	0,0139

TABLE 7.2 – Autocorrélations de la série chronologique des prix des EUA

	$S_i$	$S_i - S_{i-1}$
Moyenne	13,0646	-0,0157
Médiane	12,5800	0,0000
Écart type	2,5629	0,3196
Coefficient d'asymétrie	1,4893	-0,5571
Coefficient d'aplatissement	6,061	4,8069

TABLE 7.3 – Statistiques descriptives des prix des CER

lag	$S_i$	$S_i - S_{i-1}$
1	0,9983	0,0432
2	0,9966	-0,0836
3	0,9950	0,0512
4	0,9932	0,1464
5	0,9914	-0,0015

TABLE 7.4 – Autocorrélations de la série chronologique des prix des CER

		$\sigma$	$\kappa_S$	$\theta_S$
$\Delta t = 1$ ( $m = 1$ )	Moyenne	0,8671	-0,0171	0,2296
	Écart type	0,0569	0,0160	0,2766
$\Delta t = \frac{1}{2}$ ( $m = 2$ )	Moyenne	0,8673	-0,0173	0,2326
	Écart type	0,0567	0,0160	0,2765
$\Delta t = \frac{1}{4}$ ( $m = 4$ )	Moyenne	0,8677	-0,0173	0,2331
	Écart type	0,0562	0,0160	0,2771
$\Delta t = \frac{1}{8}$ ( $m = 8$ )	Moyenne	0,8681	-0,0175	0,2357
	Écart type	0,0552	0,0160	0,2763

TABLE 7.5 – Estimation de moyenne et d'écart type pour les densités *a posteriori* des paramètres avec le processus d'Orstein-Uhlenbeck des EUA

## 7.1 *European Union Allowance (EUA)*

Nous nous pencherons maintenant sur la série des EUA. Pour chaque processus utilisé, nous donnerons les moyennes et les écarts types des distributions *a posteriori* de chaque paramètre.

### 7.1.1 Processus d'Orstein-Uhlenbeck

Pour ce processus utilisé avec les données des EUA, quelques informations relatives aux distributions *a posteriori* sont données dans la table 7.5. Les figures 7.1, 7.2 et 7.3 correspondent aux densités *a posteriori* des paramètres  $\sigma$ ,  $\kappa_S$  et  $\theta_S$  respectivement.

### 7.1.2 Processus racine carrée

Les résultats pour le processus racine carrée sont résumés dans la table 7.6. Les densités *a posteriori* des paramètres  $\sigma$ ,  $\kappa_S$  et  $\theta_S$  sont visibles dans les figures 7.4, 7.5 et 7.6 respectivement.

### 7.1.3 Processus d'Orstein-Uhlenbeck exponentiel

Pour le processus d'Orstein-Uhlenbeck exponentiel utilisé avec les données des EUA, des informations relatives aux distributions *a posteriori* sont données dans la table 7.7. Les figures 7.7, 7.8 et 7.9 correspondent aux densités *a posteriori* des paramètres  $\sigma$ ,  $\kappa_S$  et  $\theta_S$  respectivement.

## 7.2 *Certified Emission Reductions (CER)*

Maintenant, nous ferons la même chose. Toutefois, cette fois nous utiliserons la série chronologique contenant les prix de CER.

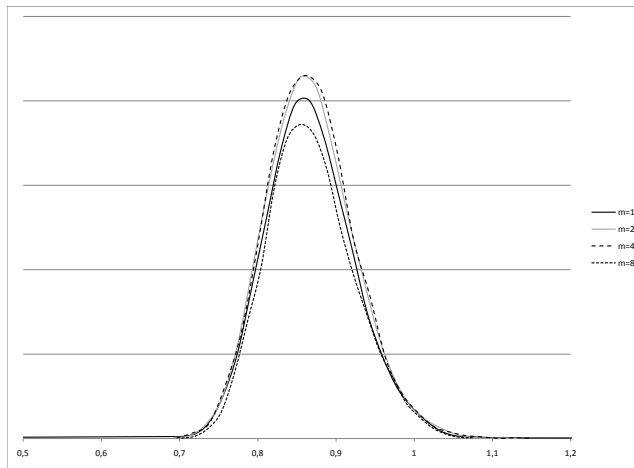


FIGURE 7.1 – Densité empirique *a posteriori* de  $\sigma$  pour le processus d'Orstein-Uhlenbeck des EUA

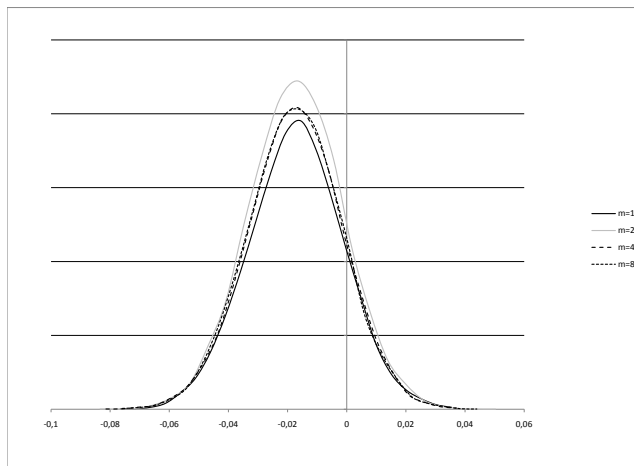


FIGURE 7.2 – Densité empirique *a posteriori* de  $\kappa_S$  pour le processus d'Orstein-Uhlenbeck des EUA

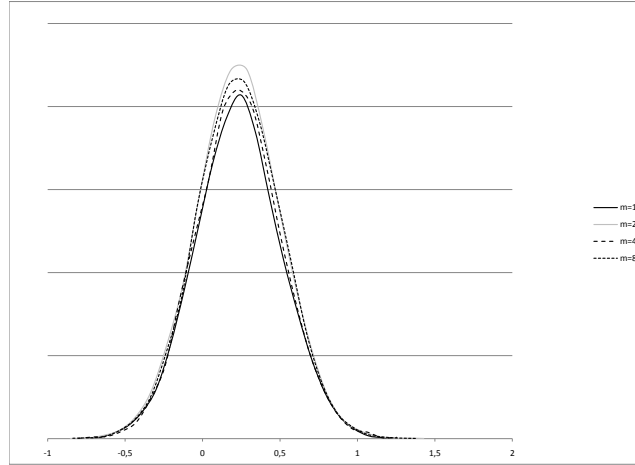


FIGURE 7.3 – Densité empirique *a posteriori* de  $\theta_S$  pour le processus d'Orstein-Uhlenbeck des EUA

		$\sigma$	$\kappa_S$	$\theta_S$
$\Delta t = 1$ ( $m = 1$ )	Moyenne	0,2164	-0,0247	0,3552
	Écart type	0,0141	0,0176	0,2807
$\Delta t = \frac{1}{2}$ ( $m = 2$ )	Moyenne	0,2171	-0,0239	0,3436
	Écart type	0,0141	0,0179	0,2856
$\Delta t = \frac{1}{4}$ ( $m = 4$ )	Moyenne	0,2164	-0,0235	0,3369
	Écart type	0,0137	0,0177	0,2826
$\Delta t = \frac{1}{8}$ ( $m = 8$ )	Moyenne	0,2150	-0,0235	0,3362
	Écart type	0,0133	0,0176	0,2812

TABLE 7.6 – Estimation de moyenne et d'écart type pour les densités *a posteriori* des paramètres avec le processus racine carrée des EUA



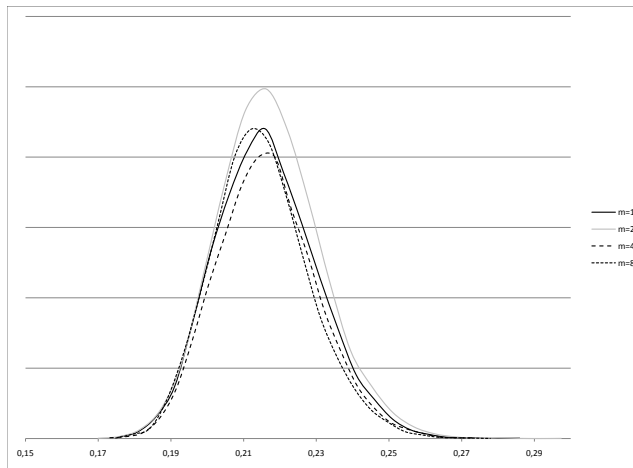


FIGURE 7.4 – Densité empirique *a posteriori* de  $\sigma$  pour le processus racine carrée des EUA

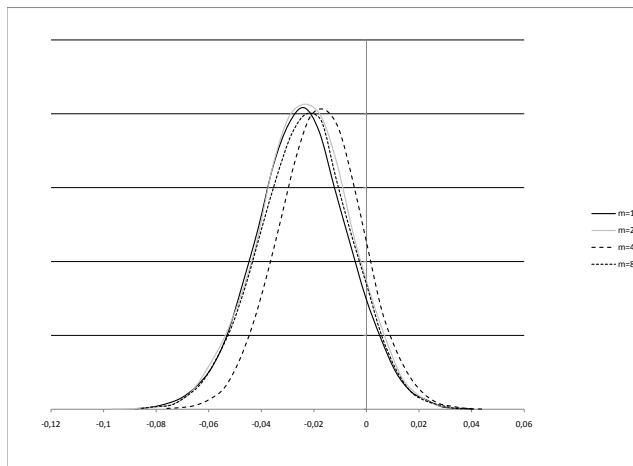


FIGURE 7.5 – Densité empirique *a posteriori* de  $\kappa_S$  pour le processus racine carrée des EUA

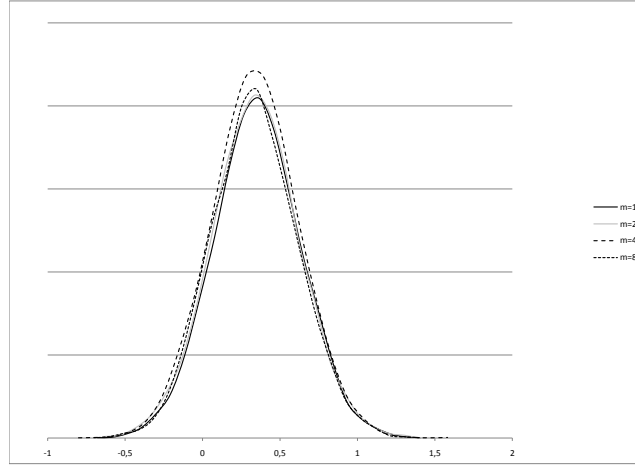


FIGURE 7.6 – Densité empirique *a posteriori* de  $\theta_S$  pour le processus racine carrée des EUA

		$\sigma$	$\kappa_S$	$\theta_S$
$\Delta t = 1$ ( $m = 1$ )	Moyenne	0,0574	-0,0235	0,0619
	Écart type	0,0037	0,0188	0,0523
$\Delta t = \frac{1}{2}$ ( $m = 2$ )	Moyenne	0,0574	-0,0240	0,0633
	Écart type	0,0037	0,0189	0,0525
$\Delta t = \frac{1}{4}$ ( $m = 4$ )	Moyenne	0,0572	-0,0237	0,0623
	Écart type	0,0036	0,0187	0,0519
$\Delta t = \frac{1}{8}$ ( $m = 8$ )	Moyenne	0,0573	-0,0240	0,0632
	Écart type	0,0573	0,0188	0,0522

TABLE 7.7 – Estimation de moyenne et d'écart type pour les densités *a posteriori* des paramètres avec le processus d'Orstein-Uhlenbeck exponentiel des EUA

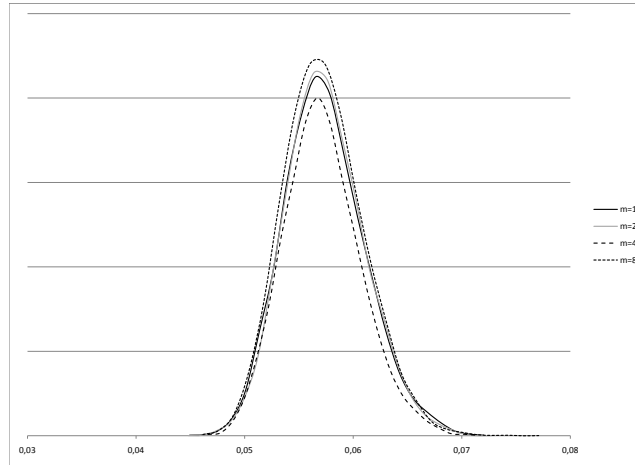


FIGURE 7.7 – Densité empirique *a posteriori* de  $\sigma$  pour le processus d'Orstein-Uhlenbeck exponentiel des EUA

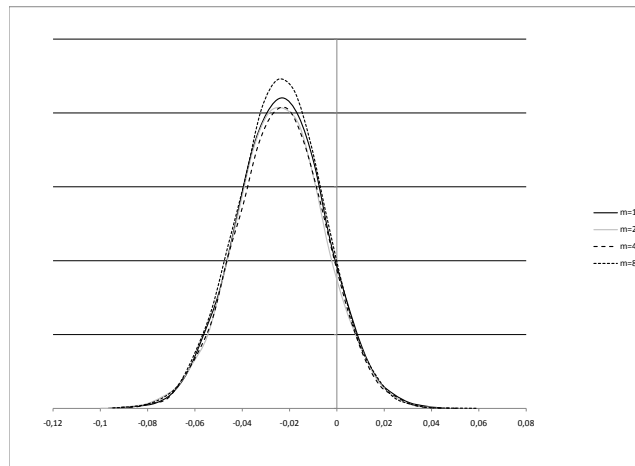


FIGURE 7.8 – Densité empirique *a posteriori* de  $\kappa_S$  pour le processus d'Orstein-Uhlenbeck exponentiel des EUA

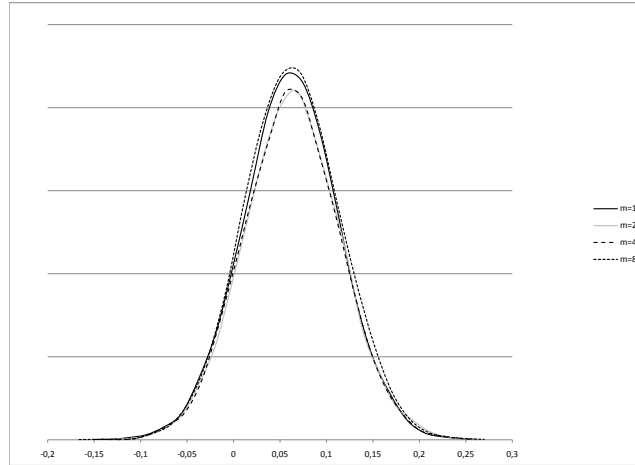


FIGURE 7.9 – Densité empirique *a posteriori* de  $\theta_S$  pour le processus d’Orstein-Uhlenbeck exponentiel des EUA

### 7.2.1 Processus d’Orstein-Uhlenbeck

Pour le processus d’Orstein-Uhlenbeck, le résumé concernant les paramètres se situe à la table 7.8. Les graphiques des densités empiriques *a posteriori* sont aux figures 7.10, 7.11 et 7.12.

### 7.2.2 Processus racine carrée

À la table 7.9, on peut voir le résumé des paramètres pour le processus racine carrée associé à la série des prix des CER. Pour les graphiques, consultez les figures 7.13, 7.14 et 7.15.

### 7.2.3 Processus d’Orstein-Uhlenbeck exponentiel

Pour le processus d’Orstein-Uhlenbeck exponentiel, le résumé concernant les paramètres se situe à la table 7.10. Les graphiques des densités empiriques *a posteriori* sont aux figures 7.16, 7.17 et 7.18.

## 7.3 Comparaisons avec l’expérience de Godin (2010)

Dans Godin (2010), l’auteur fournit des estimateurs pour les paramètres des trois modèles étudiés ici ; pour ce faire, il utilise la méthode du maximum de vraisemblance. Les séries chronologiques utilisées par Godin sont les

		$\sigma$	$\kappa_S$	$\theta_S$
$\Delta t = 1$ ( $m = 1$ )	Moyenne	0,6899	-0,0697	0,8344
	Écart type	0,0502	0,0268	0,3572
$\Delta t = \frac{1}{2}$ ( $m = 2$ )	Moyenne	0,6893	-0,0699	0,8361
	Écart type	0,0487	0,0269	0,3585
$\Delta t = \frac{1}{4}$ ( $m = 4$ )	Moyenne	0,6891	-0,0698	0,8316
	Écart type	0,0483	0,0272	0,3622
$\Delta t = \frac{1}{8}$ ( $m = 8$ )	Moyenne	0,6867	-0,0699	0,8349
	Écart type	0,0462	0,0272	0,3626

TABLE 7.8 – Estimation de moyenne et d'écart type pour les densités *a posteriori* des paramètres avec le processus d'Orstein-Uhlenbeck des CER.

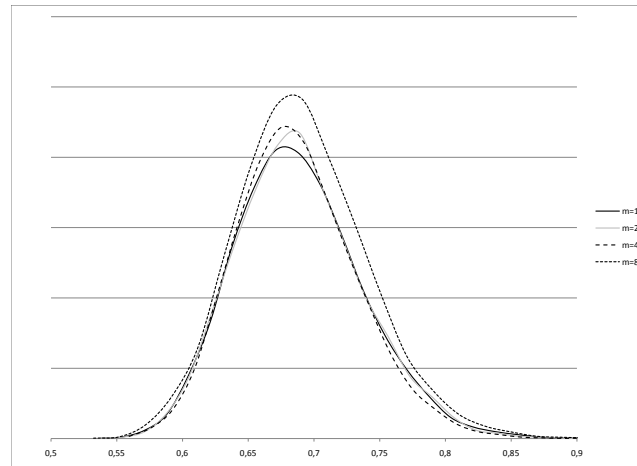


FIGURE 7.10 – Densité empirique *a posteriori* de  $\sigma$  pour le processus d'Orstein-Uhlenbeck des CER.

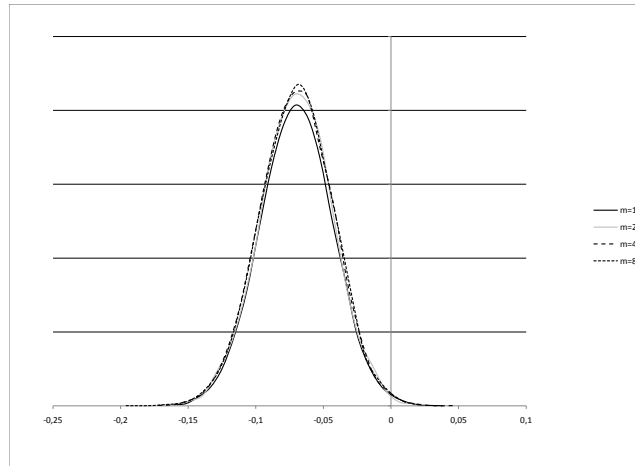


FIGURE 7.11 – Densité empirique *a posteriori* de  $\kappa_S$  pour le processus d'Orstein-Uhlenbeck des CER

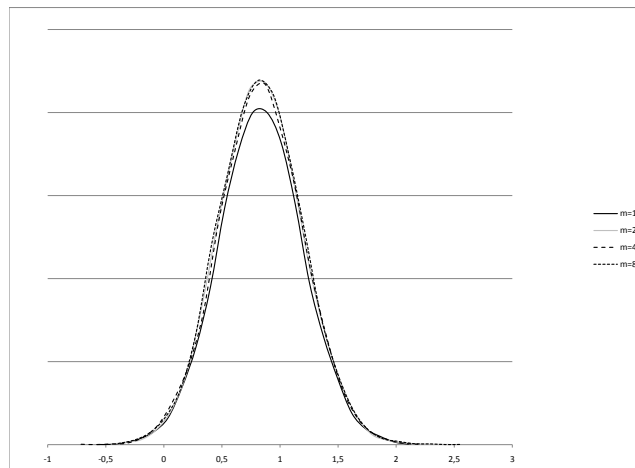


FIGURE 7.12 – Densité empirique *a posteriori* de  $\theta_S$  pour le processus d'Orstein-Uhlenbeck des CER

		$\sigma$	$\kappa_S$	$\theta_S$
$\Delta t = 1$ ( $m = 1$ )	Moyenne	0,1885	-0,0788	0,9533
	Écart type	0,0137	0,0293	0,3783
$\Delta t = \frac{1}{2}$ ( $m = 2$ )	Moyenne	0,1903	-0,0798	0,9644
	Écart type	0,0135	0,0299	0,3783
$\Delta t = \frac{1}{4}$ ( $m = 4$ )	Moyenne	0,1903	-0,0799	0,9649
	Écart type	0,0133	0,0298	0,3827
$\Delta t = \frac{1}{8}$ ( $m = 8$ )	Moyenne	0,1896	-0,0796	0,9611
	Écart type	0,0132	0,0300	0,3861

TABLE 7.9 – Estimation de moyenne et d'écart type pour les densités *a posteriori* des paramètres avec le processus racine carrée des CER

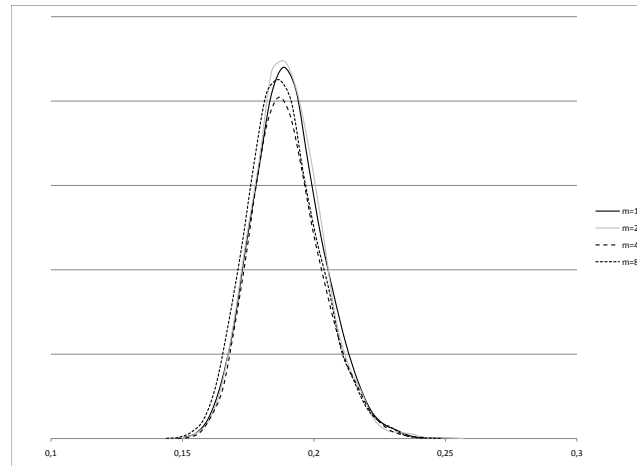


FIGURE 7.13 – Densité empirique *a posteriori* de  $\sigma$  pour le processus racine carrée des CER

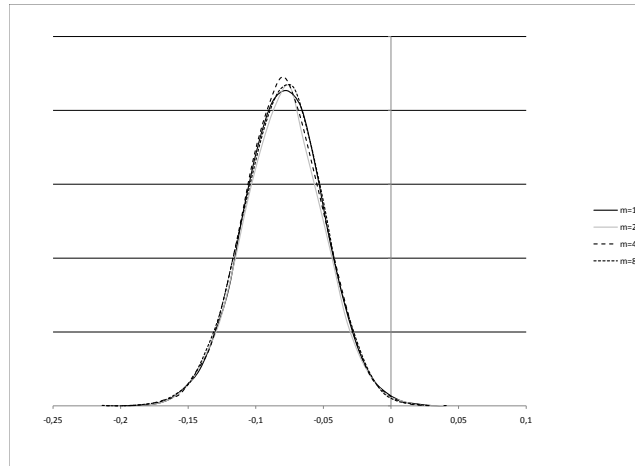


FIGURE 7.14 – Densité empirique *a posteriori* de  $\kappa_S$  pour le processus racine carrée des CER

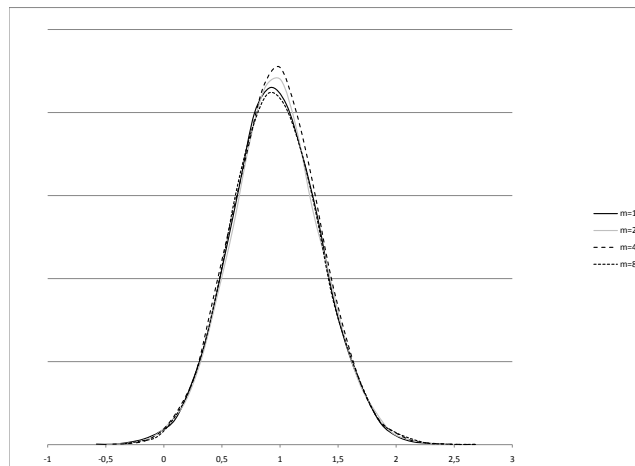


FIGURE 7.15 – Densité empirique *a posteriori* de  $\theta_S$  pour le processus racine carrée des CER



		$\sigma$	$\kappa_S$	$\theta_S$
$\Delta t = 1$ ( $m = 1$ )	Moyenne	0,0541	-0,0744	0,1851
	Écart type	0,0040	0,0299	0,0766
$\Delta t = \frac{1}{2}$ ( $m = 2$ )	Moyenne	0,0542	-0,0740	0,1840
	Écart type	0,0038	0,0308	0,0789
$\Delta t = \frac{1}{4}$ ( $m = 4$ )	Moyenne	0,0541	-0,0739	0,1835
	Écart type	0,0038	0,0307	0,0786
$\Delta t = \frac{1}{8}$ ( $m = 8$ )	Moyenne	0,0541	-0,0745	0,1850
	Écart type	0,0038	0,0308	0,0788

TABLE 7.10 – Estimation de moyenne et d'écart type pour les densités *a posteriori* des paramètres avec le processus d'Orstein-Uhlenbeck exponentiel des CER

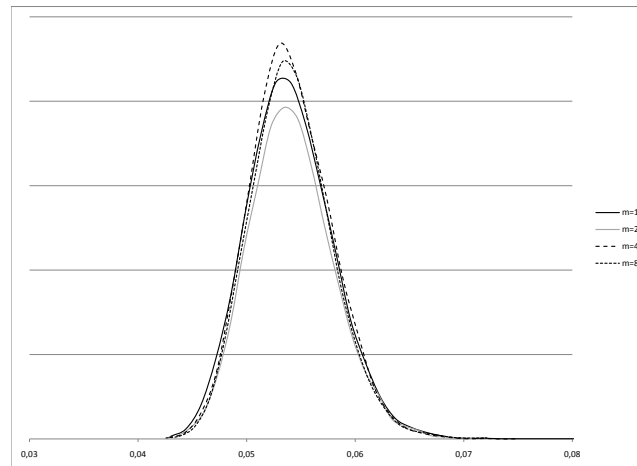


FIGURE 7.16 – Densité empirique *a posteriori* de  $\sigma$  pour le processus d'Orstein-Uhlenbeck exponentiel des CER

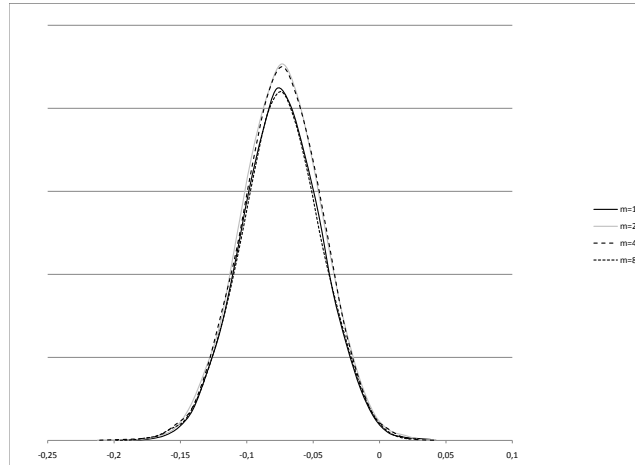


FIGURE 7.17 – Densité empirique *a posteriori* de  $\kappa_S$  pour le processus d'Orstein-Uhlenbeck exponentiel des CER

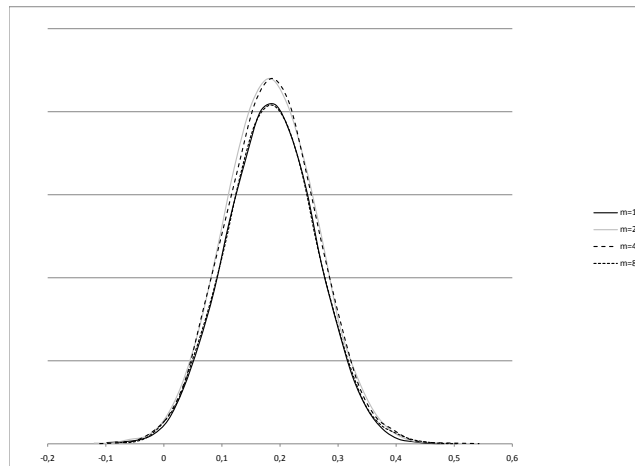


FIGURE 7.18 – Densité empirique *a posteriori* de  $\theta_S$  pour le processus d'Orstein-Uhlenbeck exponentiel des CER

mêmes que nous, toutefois nous avons utilisé plus de données que ce dernier (il y avait plus de données disponibles au moment de faire cette analyse). Alors, le processus de Orstein-Uhlenbeck obtenu par Godin pour les EUA est

$$dS_t = (0,1850 - 0,0153S_t)dt + 0,9439dZ. \quad (7.1)$$

Celui obtenu dans notre analyse en prenant  $m = 8$  est

$$dS_t = (0,2357 - 0,0175S_t)dt + 0,8681dZ. \quad (7.2)$$

En utilisant le même mouvement brownien pour les deux processus (le même chemin), on peut observer la différence visuelle entre les deux processus sur un an à la figure 7.19. D'après l'inspection, les deux processus sont assez proches. De plus, les estimateurs des paramètres sont très semblables. Pour le processus racine carrée des EUA, Godin a obtenu

$$dS_t = (0,2877 - 0,0213S_t)dt + 0,2318\sqrt{S_t}dZ \quad (7.3)$$

et nous avons obtenu

$$dS_t = (0,3362 - 0,0235S_t)dt + 0,2150\sqrt{S_t}dZ. \quad (7.4)$$

Encore une fois, les estimateurs des paramètres ont des valeurs numériques très semblables. Un exemple de chemin emprunté par ces processus est disponible à la figure 7.20. Finalement, le processus de Orstein-Uhlenbeck exponentiel pour les EUA de Godin va comme suit

$$d(\ln S_t) = (0,0519 - 0,0202(\ln S_t))dt + 0,0600dZ. \quad (7.5)$$

Le nôtre prend cette forme

$$d(\ln S_t) = (0,0632 - 0,0240(\ln S_t))dt + 0,0573dZ. \quad (7.6)$$

Les valeurs numériques sont quasi identiques. Un exemple de chemin emprunté par les processus de Orstein-Uhlenbeck exponentiel est à la figure 7.21.

Pour les modèles utilisés avec les données des CER, la conclusion est la même : les valeurs numériques des estimateurs sont très semblables de celle de Godin, et ce, même si les nombres de données utilisées diffèrent dans les deux expériences.

Alors, on peut assumer que si nous avons sélectionné les mêmes données que celles de la référence [7], les estimateurs des paramètres de notre analyse seraient encore plus près des estimateurs de Godin ; et ce, en utilisant cinq fois moins de données (et en pratiquant l'augmentation de données à l'aide des MCMC).

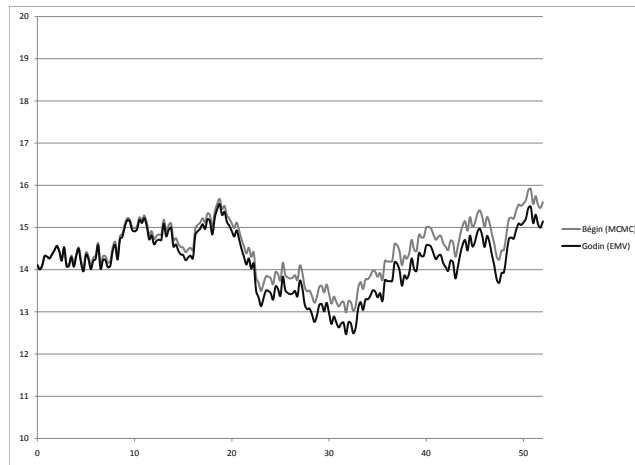


FIGURE 7.19 – Un exemple de chemin emprunté par les processus d’Orstein-Uhlenbeck des EUA

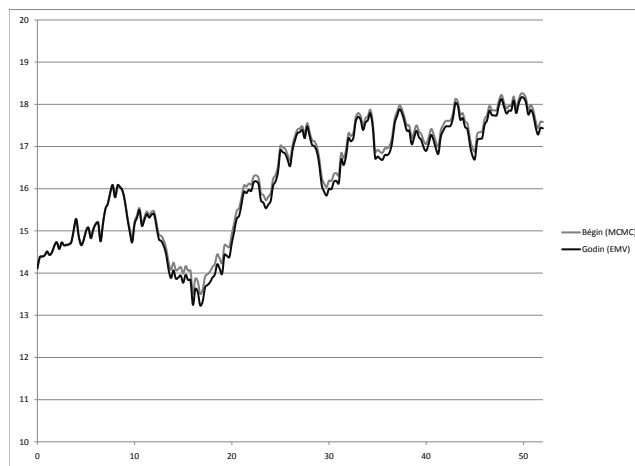


FIGURE 7.20 – Un exemple de chemin emprunté par les processus racine carrée des EUA

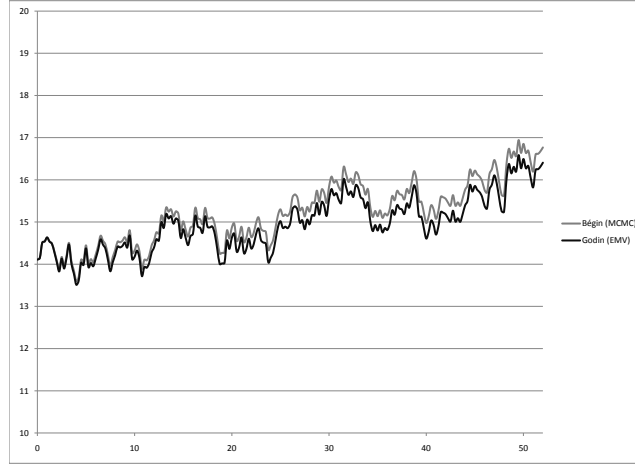


FIGURE 7.21 – Un exemple de chemin emprunté par les processus d’Orstein-Uhlenbeck exponentiel des EUA

## 7.4 Comparaisons des modèles

Dans les graphiques des sections 7.1 et 7.2 de ce présent chapitre, nous avons pu observer plusieurs graphiques de densités *a posteriori* des paramètres de nos modèles. En pratiquant une inspection visuelle de ces graphiques, on peut voir qu’à partir de  $m = 4$  (donc pour 4 et 8), on ne voit pas beaucoup de différences entre les distributions des paramètres en général. Donc, nous avons des raisons de croire que l’augmentation de données de  $m = 4$  à 8 n’apporte pas de meilleurs résultats selon les évidences graphiques présentées dans les pages précédentes. Cette conclusion est la même qu’Eraker obtenait dans son article (voir la référence [5]).

Pour comparer les trois modèles entre eux, nous utiliserons les densités obtenues avec  $m = 8$  malgré le fait énoncé plus haut. Une version approximative de la cote de Bayes sera utilisée ici afin de déterminer lequel des trois modèles est le meilleur. La cote de Bayes (*Bayes factor*),  $K$ , est équivalente à l’expression suivante dans notre analyse :

$$K = \frac{\int \int \int p(\theta_S | M_1) p(\sigma | M_1) p(\kappa_S | M_1) p(D | \theta_S, \kappa_S, \sigma, M_1) d\theta_S d\kappa_S d\sigma}{\int \int \int p(\theta_S | M_2) p(\sigma | M_2) p(\kappa_S | M_2) p(D | \theta_S, \kappa_S, \sigma, M_2) d\theta_S d\kappa_S d\sigma} \quad (7.7)$$

où  $\theta_S$ ,  $\sigma$  et  $\kappa_S$  sont les paramètres du modèle,  $M_1$  est le premier modèle,  $M_2$  est le deuxième modèle (qu’on veut comparer à  $M_1$ ) et  $D$  l’ensemble des données. Nous utiliserons une approximation de cette valeur dans notre analyse. Pour ce faire, nous poserons trois partitions telles que le domaine

$M_1 \setminus M_2$	O-U <sup>1</sup>	RC <sup>2</sup>	O-U exponentiel
O-U	1	-	-
RC	$2,4256 \times 10^{14}$	1	-
O-U exponentiel	$\approx 0$	$\approx 0$	1

TABLE 7.11 – Les valeurs de  $K^*$  pour les trois modèles utilisés pour les données EUA

de  $p(\theta_S|M_1)$  est divisé en  $t = \{t_0, t_1, \dots, t_{n_t}\}$ , le domaine de  $p(\kappa_S|M_1)$  est divisé en  $k = \{k_0, k_1, \dots, k_{n_k}\}$  et le domaine de  $p(\sigma|M_1)$  est divisé en  $s = \{s_0, s_1, \dots, s_{n_s}\}$ . Nous ferons la même chose pour le deuxième modèle  $M_2$ . Alors, l'approximation de  $K$  prendra la forme suivante :

$$\begin{aligned}
K^* = & \left( \sum_i \sum_j \sum_m p(t_m \leq \theta_S < t_{m+1}|M_1) p(s_j \leq \sigma < s_{j+1}|M_1) \right. \\
& \left. p(k_i \leq \kappa_S < k_{i+1}|M_1) p(D|s_j, t_m, k_i, M_1) \Delta s_j \Delta t_m \Delta k_i \right) \\
& \left( \sum_i \sum_j \sum_m p(t_m \leq \theta_S < t_{m+1}|M_2) p(s_j \leq \sigma < s_{j+1}|M_2) \right. \\
& \left. p(k_i \leq \kappa_S < k_{i+1}|M_2) p(D|s_j, t_m, k_i, M_2) \Delta s_j \Delta t_m \Delta k_i \right)^{-1} \quad (7.8)
\end{aligned}$$

On sait que  $K^*$  tend vers  $K$  (voir la théorie des séries de Riemann).

Pour comparer les modèles, des interprétations des valeurs de  $K$  existent (voir l'échelle de Jeffreys (1935)). Par exemple, une valeur de  $K$  entre 10 et 30 suppose que le modèle  $M_1$  est fortement meilleur que le modèle  $M_2$ . Les résultats sont disponibles dans la table 7.11. D'après ces derniers, le processus racine carrée semble le meilleur d'une manière très évidente. De plus, d'après les valeurs des  $K^*$ , on aurait tendance à supporter les processus d'Orstein-Uhlenbeck et racine carrée vis-à-vis le processus d'Orstein-Uhlenbeck exponentiel. En bref, le processus racine carrée semble le meilleur des trois, d'après les approximations des cotes de Bayes.

La figure 7.22 illustre un exemple de chemin emprunté par les trois processus considérés avec les données connues. Afin de ne pas alourdir le texte, les cotes de Bayes et les graphiques pour les données CER ne sont pas données. Les résultats sont identiques aux EUA pour ce qui a trait à la comparaison entre les modèles.

---

1. O-U signifie processus d'Orstein-Uhlenbeck.  
2. RC signifie procesus racine carrée.

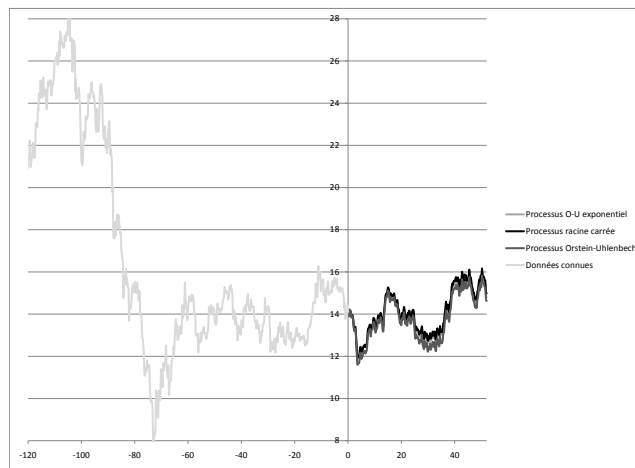


FIGURE 7.22 – Un exemple de chemin emprunté pour chaque processus des EUA à l'aide des estimations de notre analyse

## Chapitre 8

# Conclusion

Nous avons appliqué la méthode MCMC de Bjørn Eraker afin d'augmenter les données du marché du carbone européen et de trouver les densités *a posteriori* des paramètres d'un modèle de diffusion choisi dans la classe à élasticité constante de variance. En utilisant cinq fois moins de données que le nombre disponible dans le marché (soit une donnée par semaine au lieu d'une donnée par jour), on obtient des résultats très près d'une méthode qui utiliserait toutes les données disponibles. Donc, on peut assumer ici, à l'aide d'évidences visuelles et de comparaisons informelles, que l'augmentation de données a un effet positif sur les résultats. Toutefois, cet effet bénéfique devient de moins en moins visible si on augmente le nombre de données. De plus, nous avons comparé trois modèles de la classe CEV. Le processus racine carrée semble le plus approprié selon les résultats de la cote de Bayes. Le processus d'Orstein-Uhlenbeck nous donne de très mauvais résultats comparés aux deux autres modèles étudiés.



## Annexe A

# Utilisation de la régression linéaire dans l'estimation des paramètres

Dans son article, Eraker utilise une régression linéaire pour estimer les paramètres  $\theta_S$ ,  $\kappa_S$  et  $\sigma$  de la classe de modèles à élasticité constante de variance. Cette technique est justifiée puisque  $\Delta t \rightarrow 0$  et la densité normale est une loi d'échelle (c'est-à-dire, qu'en pratique, cette loi se voit aussi paramétrisée en fonction du carré de l'échelle,  $\sigma^2$ , qui correspond à la variance de la distribution). Alors, nous avons que

$$dS_t = (\theta_S + \kappa_S S_t) dt + \sigma S_t^\beta dW_t \quad (\text{B.1})$$

tel que défini pour la classe *CEV*. De plus, on sait que  $dW_t$ , le processus de Wiener, correspond à une loi normale. Si on discrétise l'équation (B.1) et qu'on assume que la différence de temps entre deux observations tend vers zéro (c'est-à-dire  $\Delta t \rightarrow 0$ ), on a

$$\Delta S_t = (\theta_S + \kappa_S S_t) \Delta t + \sigma S_t^\beta \Delta W_t. \quad (\text{B.2})$$

D'après ce que nous avons vu à la section 4.2,  $\Delta W_t \sim \mathcal{N}(0, \Delta t)$ . En utilisant l'argument de loi d'échelle, on peut faire les changements suivants

$$\frac{S_t - S_{t-1}}{\sqrt{\Delta t} S_t^\beta} = \theta_S \left( \frac{\Delta t}{\sqrt{\Delta t} S_t^\beta} \right) + \kappa_S \left( \frac{S_t \Delta t}{\sqrt{\Delta t} S_t^\beta} \right) + \epsilon_i \quad (\text{B.3})$$

où  $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ . Ainsi, on peut estimer les paramètres  $\theta_S$  et  $\kappa_S$  en faisant une régression linéaire. On définit premièrement une matrice de design  $X$  et un vecteur  $y$  tels que les éléments

$$y_i = \left[ \frac{S_t - S_{t-1}}{\sqrt{\Delta t} S_t^\beta} \right] \quad (\text{B.4})$$

sont empilés de manière à former le vecteur  $y$  et les éléments

$$X_i = \left[ \frac{\Delta t}{\sqrt{\Delta t S_t^\beta}} \quad \frac{S_t \Delta t}{\sqrt{\Delta t S_t^\beta}} \right] \quad (\text{B.5})$$

sont empilés de manière à former la matrice  $X$ . L'estimation des paramètres est alors donnée par

$$(\theta_S, \kappa_S)' = (X'X)^{-1}(X'y). \quad (\text{B.6})$$

# Bibliographie

- [1] Béliveau, A. (2008). *Reconstruction d'images par la statistiques ; projet de recherche CRSNG*.
- [2] Boldstad W. (2004). *Introduction to Bayesian Statistics*. Wiley.
- [3] Chib, S. (2010). The Gibbs Sampling Algorithm. <http://fedc.wiwi.hu-berlin.de/xplore/ebooks/html/csa/node28.html>.
- [4] Durrett, R. (2004). *Essentials of Stochastic Processes*. Springer texts in statistics.
- [5] Eraker, B. (2001). MCMC Analysis of Diffusion Models With Application to Finance. *Journal of Business & Economic Statistics*. Volume 19, No 2, 177-191.
- [6] Gilbertson, T. (2009). Carbon Trading : How it works and why it fails. *Critical currents*. No 7.
- [7] Godin, F. (2010). *Les produits dérivés des marchés européens du carbone*.
- [8] Hastings, W.K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*. Volume 57, No. 1, 97-109.
- [9] Hirano, K. (2010). Lecture Note 10 : Bayesian Estimation and MCMC. <http://www.u.arizona.edu/~hirano/696E/ln10.pdf>.
- [10] Janssen, J., Manca, R., Volpe, R. (2008). *Mathematical Finance : Deterministic and Stochastic Models*. Wiley.
- [11] Jeffreys, H. (1935). Some tests of significance treated by the theory of probability. *Proc. Cambridge Philosophical Soc.*. Volume 31, 203-222
- [12] Le Goff, P. (2010). Le marché européen du carbone : lancement réussi ; le marché canadien : en construction. <http://www2.parl.gc.ca/Content/LOP/ResearchPublications/prb0551-f.html>.
- [13] Martin, G. (2009). *Notes de cours, MAT2717*.
- [14] Mazet, V. (2010). Introduction aux méthodes de Monte Carlo par chaînes de Markov. [http://lsiit-miv.u-strasbg.fr/lsiit/perso/mazet/publis/mazet\\_mtde03.pdf](http://lsiit-miv.u-strasbg.fr/lsiit/perso/mazet/publis/mazet_mtde03.pdf).
- [15] McDonald, R.L. (2006). *Derivatives Markets, second edition*, Pearson Education.

- [16] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E. (1953). Equation of state calculation by fast computing machines. *Journal of Chemical Physics*. Volume 21, No 6, 1087–1092.
- [17] Rice, J. (2007). *Mathematical Statistics and Data Analysis*. Duxbury.
- [18] Robert, C., Casella, G. (2004). *Monte Carlo Statistical Methods, 2nd Edition*. Springer texts in statistics.
- [19] Taleb, N.N. (2005). *Foiled by randomness : the hidden role of chance in life and in the markets*, Random House Trade Paperback.
- [20] Walsh, B. (2010). Markov Chains Monte Carlo and Gibbs Sampling, Lecture Notes for EEB581. [http://membres-timc.imag.fr/Olivier.Francois/mcmc\\_gibbs\\_sampling.pdf](http://membres-timc.imag.fr/Olivier.Francois/mcmc_gibbs_sampling.pdf).