Application de la statistique des diffusions à un modèle de taux d'intérêt *

ÉRIC FOURNIÉ et DENIS TALAY **

1. Introduction

Motivés par la mise au point d'un algorithme d'identification des paramètres du modèle d'évolution des taux d'intérêts proposé par J. C. Cox, J. E. Ingersoll et S. A. Ross dans [6] (modèle dit CIR), nous nous sommes proposés d'évaluer les performances d'estimateurs paramétriques issus de la statistique des processus de diffusion : il semblerait que jusqu'à présent on ait plutôt cherché à estimer tous les paramètres simultanément, y compris la volatilité, à partir du prix d'obligations à court terme, sans utiliser la loi du taux en tant que processus : méthode GMM (Generalized Method of Moments) (Longstaff [21]), Gibbons and Ramaswamy [11] par exemple), moindres carrés pour minimiser l'écart entre les prix théoriques déduits du modèle et les prix du marché (Brown and Schaefer [5], Brown and Dybvig [4]). Ces diverses procédures conduisent à des ordres de grandeur très variables, et surtout à des erreurs d'estimation difficiles à décrire en fonction du nombre de relevés, donc incontrôlables (voir aussi les travaux de R. Merton [22] et S. Heston [14], et pour l'estimation paramétrique de modèles en temps discret par la méthode GMM, D. Duffie et K. Singleton [8]).

Nous avons donc voulu étudier dans quelle mesure la statistique des processus pouvait apporter des améliorations. Malheureusement, notre conclusion sera très nuancée : nous verrons que les estimateurs naturels (maximum de vraisemblance, méthode des moments), appliqués à des relevés de taux au jour le jour, ne pourraient être réellement précis que si les paramètres du modèle restaient constants sur de longs intervalles de temps, ce qui n'est pas réaliste d'un point de vue économique. Par contre, ils permettent d'obtenir des ordres de grandeur fiables ainsi que des intervalles de confiance, et par ailleurs l'estimation de la volatilité à partir de la variation quadratique donne des résultats satisfaisants; ceci nous

* Cette étude a été financée par un contrat ARCAS/INRIA; nous sommes redevables à Eric Briys et Michel Crouhy des relevés journaliers de différents taux et de nombreux fructueux commentaires.
** INRIA, Sophia Antipolis, Valbonne, F-06565

a permis de mettre en œuvre plus efficacement un algorithme d'optimisation comparable à ceux cités ci-dessus (cet aspect ne sera pas détaillé ici).

Plus précisément : soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé, $(\mathcal{F}_t)_{t \geq 0}$ une filtration de $\mathcal{F}$ et (W_t) un $(\mathcal{F}_t) - P$ processus de Wiener standard à valeurs dans $\mathbb{R}$. Soit $(X_t)_{t \geq 0}$ le processus réel défini sur $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ représentant le taux d'intérêt. On le suppose solution de l'équation différentielle stochastique suivante :

$$dX_t = c(\theta - X_t) + dt + \sigma\sqrt{X_t}dW_t, \; X_0 > 0 \tag{1}$$

où c, θ et σ sont des paramètres réels représentant respectivement le coefficient de rappel, la moyenne et la volatilité du taux d'intérêt.

Nous nous proposons de construire les estimateurs du maximum de vraisemblance des paramètres et ceux issus de la méthode des moments. Nous établirons leur normalité asymptotique et nous comparerons leurs vitesses de convergence, puis nous montrerons leur comportement sur des données simulées, ainsi que sur des données issues de relevés journaliers. Nous conclurons par la mise en œuvre d'estimateurs non paramétriques.

En ce qui concerne les algorithmes de simulation numérique de trajectoires de processus de diffusion, nous renvoyons à Talay [24]. D'autres aspects statistiques appliqués à des modèles financiers seront développés dans la thèse en cours d'Eric Fournié.

D'un point de vue mathématique, pour la partie « normalité asymptotique », nous avons adapté les preuves utilisées par Le Breton et Musiela [18] pour les systèmes bilinéaires, en détaillant quelques points de démonstration. Par ailleurs, nous avons essayé de faire une présentation la plus auto-suffisante possible.

2. Quelques propriétés du modèle

La proposition suivante se montre aisément comme son analogue pour les processus de Bessel (cf. [15], chap. IV, exemple 8.2.).

Proposition 2.1

Si $c > 0$, $\theta > 0$, $\sigma > 0$ et $2c\theta \geq \sigma^2$ alors l'équation (1) a une solution globale unique au sens fort strictement positive P – p. s.

Par ailleurs :

Proposition 2.2

Sous les hypothèses précédentes, le processus (X_t) solution de (1) est ergodique, et son unique mesure invariante admet une densité par rapport à la mesure de Lebesgue. C'est celle d'une loi $\Gamma\left(\frac{2c\theta}{\sigma^2}, \frac{2c}{\sigma^2}\right)$.

Preuve. — Soit $\mathcal{L}$ l'opérateur infinitésimal de la diffusion (X_t) :

$$\mathcal{L} = c(\theta - x)\frac{\partial}{\partial x} + \frac{\sigma^2 x}{2}\frac{\partial^2}{\partial x^2}$$

Montrons que le processus est récurrent positif relativement à un intervalle borné de $\mathbb{R}_*^+$.

Il suffit ([13], chap. 3, th. 7.1) de montrer l'existence d'un intervalle ouvert U de complémentaire borné et d'une fonction $V(x)$ non négative et deux fois continûment différentiable sur U tels que :

$$\mathcal{L}V(x) \leq -1, \quad \forall x \in U$$

En prenant $V(x) = x^2$, on obtient :

$$\mathcal{L}V(x) = -2cx^2 + (2c\theta + \sigma^2)x$$

Le trinôme $\mathcal{L}V(x) + 1$ admet une racine strictement positive, α, et est négatif pour $x > \alpha$; il suffit par conséquent de prendre $U = [\alpha, \infty[$ de complémentaire $\bar{U} =]0, \alpha[$ sur R_*^+, pour obtenir la récurrence positive relativement à $\bar{U}$. Le processus étant strictement positif, les lemmes 2.1. et 3.1. de [13], chapitre 4, montrent que la récurrence positive relativement à un domaine borné de $\mathbb{R}_*^+$ entraîne la récurrence positive par rapport à n'importe quel domaine non vide de $\mathbb{R}_*^+$.

Donc ([13], chap. 4, th. 4.1 et 5.1) le processus admet une unique mesure de probabilité invariante.

On vérifie aussi que cette mesure invariante admet une densité par rapport à la mesure de Lebesgue. En outre, W. Feller dans [9] montre que, pour toute densité $p_0(x)$ l'équation de Fokker-Planck :

$$\frac{\partial}{\partial t}p(t, x) = \mathcal{L}^* p(t, x), \quad p(0, x) = p_0(x) \tag{2}$$

a une unique solution positive telle que $\int_0^\infty p(t, x)dx < \infty$; la fonction $p(t, x)$ est donc la densité de la loi de X_t. La limite, quand $t \to \infty$, de $p(t, x)$ est la densité d'une loi $\Gamma\left(\frac{2c\theta}{\sigma^2}, \frac{2c}{\sigma^2}\right)$.

Nous noterons μ l'unique loi invariante de (X_t) et $\mathbb{E}_\mu[X]$ l'espérance d'une v.a. de loi μ.

Donnons quelques résultats qui nous seront utiles dans la suite.

$$\mathbb{E}_\mu[X] = \theta$$

$$\mathbb{E}_\mu[X^2] = \frac{\sigma^2\theta + 2c\theta^2}{2c}$$

$$\mathbb{E}_\mu[(X - E[X])^2] = \frac{\sigma^2\theta}{2c}$$

$$\mathbb{E}_\mu[X^3] = \frac{\theta}{4c^2}(4c^2 + 6c\theta\sigma^2 + 2\sigma^4)$$

$$\mathbb{E}_\mu\left[\frac{1}{X}\right] = \frac{2c}{2c\theta - \sigma^2}$$

Notons que la loi limite dépend des paramètres θ et $\frac{c}{\sigma^2}$: il n'est donc possible d'estimer séparément c et σ^2 en utilisant seulement des statistiques de cette loi.

3. Estimation paramétrique

Dans ce paragraphe, nous construisons des estimateurs des paramètres c, θ et σ. Nous commencerons par rappeler comment on peut estimer σ à partir de la variation quadratique. Nous estimerons c et θ par la méthode du maximum de vraisemblance puis par celle des moments. Nous étudierons les propriétés asymptotiques des estimateurs et nous comparerons, en un certain sens, leurs vitesses de convergence.

On suppose que les observations ont lieu en temps continu. Soit $C(\mathbb{R}^+, \mathbb{R}_*^+)$ l'espace des fonctions de $\mathbb{R}^+$ continues, strictement positives, que l'on munit de sa tribu borélienne $\mathcal{B} = \sigma(\pi_s, s \geq 0)$ engendrée par les fonctions coordonnées π_s, $s \geq 0$. Soit $Q_{c,\theta}^{\sigma,\nu}$ la loi induite par le processus (X_t) de loi initiale ν sur $(C(\mathbb{R}^+, \mathbb{R}_*^+), \mathcal{B})$, et notons ${}_TQ_{c,\theta}^{\sigma,\nu}$ sa restriction à $(C([0, T], \mathbb{R}_*^+), \mathcal{B}_T)$, où $\mathcal{B}_T$ est la trace de $\mathcal{B}$ sur $C([0, T], \mathbb{R}_*^+)$.

Nous considérerons donc le modèle statistique :

$$\left\{ C([0, T], \mathbb{R}_*^+), \mathcal{B}_T, \left({}_TQ_{c,\theta}^{\sigma,\nu}, \sigma > 0, c > 0, \theta > \frac{\sigma^2}{2c}\right) \right\}$$

3.1. *Estimation paramétrique pour le coefficient de diffusion*

On identifie le coefficient de diffusion σ en utilisant la relation entre variation quadratique et processus croissant. Cette technique est bien connue (voir, par exemple, Delebecque et Quadrat [7], Le Breton et Musiela [18]).

Appelons $\langle\pi\rangle_t$ la variation quadratique du processus canonique et considérons une suite de partitions $(0 = t_0^n < t_1^n < \ldots < t_N^n = T)$ de $[0, T]$ telles que

$\max_i |t_{i+1}^n - t_i^n|$ est une suite décroissante vers 0; alors (cf. [20], chapitre 4 par exemple) :

$$\langle\pi\rangle_T = \lim_{n\to\infty} \sum_{i=0}^{n-1} (\pi_{t_{i+1}^n} - \pi_{t_i^n})^2 = \int_0^T (\sigma\sqrt{\pi_s})^2\, ds,\ {}_TQ_{c,\theta}^{\sigma,\nu} - \text{p.s.}$$

D'autre part :

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} \pi_{t_i^n}(t_{i+1}^n - t_i^n) = \int_0^T \pi_s ds$$

Par conséquent, on obtient un estimateur $\hat{\sigma}_{N,T}$ de σ en choisissant une partition de pas T/N constant et en posant :

$$\hat{\sigma}_{N,T} = \sqrt{\frac{N\Sigma_{i=0}^{N-1}(\pi_{(i+1)T/N} - \pi_{iT/N})^2}{T\Sigma_{i=0}^{N-1}\pi_{iT/N}}}$$

Comme :

$$\lim_{N\to\infty} \hat{\sigma}_{N,T} = \sigma,\ {}_TQ_{c,\theta}^{\sigma,\nu} - \text{p. s.}$$

cet estimateur est fortement consistant.

On notera que la durée de l'observation n'a aucune influence sur la convergence : seule compte la fréquence des relevés. Nous verrons plus loin, à partir d'essais numériques, que cet estimateur donne des résultats satisfaisants pour le modèle CIR avec des relevés journaliers. Comme σ a été estimé en temps fini, indépendamment des autres paramètres, nous pouvons le supposer désormais connu.

3.2. *Estimateur du maximum de vraisemblance*

Nous allons à présent construire l'estimateur du maximum de vraisemblance de (c, θ) et étudier ses propriétés. Le résultat principal portera sur la consistance forte et la normalité asymptotique de l'estimateur.

Effectuons d'abord un changement de temps. Nous posons :

$$x_t = X_{kt},\ \forall t > 0$$
$$k = \frac{4}{\sigma^2}$$

La différentielle stochastique de (x_t) s'écrit :

$$dx_t = C(\theta - x_t)\, dt + 2\sqrt{x_t}\, dB_t \tag{3}$$

où l'on a posé $C = kc$ et $B_t = \dfrac{W_{kt}}{\sqrt{k}}$; (B_t) est un Brownien standard à valeurs réelles.

On notera $P^{\nu}_{C,\theta}$ la mesure induite par le processus (x_t) sur $(C(\mathbb{R}^+, \mathbb{R}^+_*), \mathcal{B})$, et ${}_TP^{\nu}_{C,\theta}$ sa restriction à $(C([0, T], \mathbb{R}^+_*), \mathcal{B}_T)$.

Après le changement de temps ci-dessus, le modèle statistique est devenu :

$$\{(C([0, T], \mathbb{R}^+_*), \mathcal{B}_T, ({}_TP^{\nu}_{C,\theta}, \theta >,0,C > 2/\theta)\}$$

Soit (Y_t) le processus réel Bes(2, 4) solution forte de l'EDS suivante :

$$dY_t = 4dt + 2\sqrt{Y_t}dB_t \tag{4}$$

On notera μ^{ν} la mesure induite sur $(C(\mathbb{R}^+, \mathbb{R}^+_*), \mathcal{B})$ par le processus (Y_t) de loi initiale ν et ${}_T\mu^{\nu}$ sa restriction à $(C([0, T], \mathbb{R}^+_*), \mathcal{B}_T)$.

Soit le processus $(Z^{C,\theta}_t)_{0\leq t\leq T}$ défini par :

$$Z^{C,\theta}_t = \exp\left\{\int_0^t (2\sqrt{\pi_s})^{-2}(C(\theta - \pi_s) - 4)d\pi_s - \frac{1}{2}\int_0^t (2\sqrt{\pi_s})^{-2}([C(\theta - \pi_s)]^2 - 16)ds\right\}$$

La proposition suivante montre que le modèle statistique est dominé par la mesure ${}_T\mu^{\nu}$ et justifie par conséquent l'utilisation de la méthode du maximum de vraisemblance.

Proposition 3.3.

Les mesures ${}_TP^{\nu}_{C,\theta}$ et ${}_T\mu^{\nu}$ sont équivalentes et la dérivée de Radon-Nykodym de ${}_TP^{\nu}_{C,\theta}$ par rapport à ${}_T\mu^{\nu}$ est donnée par :

$$\frac{d_TP^{\nu}_{C,\theta}}{d_T\mu^{\nu}} = Z^{C,\theta}_T, \ {}_T\mu^{\nu} - \text{p. s.},$$

les intégrales stochastiques étant définies par rapport à ${}_T\mu^{\nu}$.

La preuve de cette proposition est une application des théorèmes 7.18 et 7.19 de [20] (à condition de lire dans l'énoncé du théorème 7.18 que seule la finitude presque sûre de la fonction α_t est requise, ce qui est suffisant).

Nous pouvons donc appliquer la méthode du maximum de vraisemblance, et en posant :

$$I_T = \int_0^T \frac{d\pi_t}{\pi_t} = \log\left(\frac{\pi_T}{\pi_0}\right) + 2J_T,$$

$$J_T = \int_0^T \frac{dt}{\pi_t},$$

$$K_T = \int_0^T \pi_t dt,$$

on obtient pour expression de la log-vraisemblance :

$$\begin{aligned} L_T(C, \theta) &= \log \left(Z_T^{C,\theta}\right) \\ &= \frac{1}{4}\left[-\frac{J_T}{2}(C^2\theta^2 - 16) + TC^2\theta - \frac{K_T}{2}C^2 + I_T(C\theta - 4) \right. \\ &\quad \left. - (\pi_T - \pi_0)C \right] \end{aligned}$$

$L_T(C, \theta)$ admet des dérivées partielles de tout ordre. Le gradient de $L_T(C, \theta)$, noté $L_T^{(1)}(C, \theta)$, et son hessien, noté $L_T^{(2)}(C, \theta)$, sont donnés respectivement par :

$$L_T^{(1)}(C, \theta) = \frac{1}{4}\begin{pmatrix} -1 & \theta \\ 0 & C \end{pmatrix}\left[\begin{pmatrix} \pi_T - \pi_0 \\ I_T \end{pmatrix} - C\theta\begin{pmatrix} T \\ J_T \end{pmatrix} + C\begin{pmatrix} K_T \\ T \end{pmatrix}\right] \tag{5}$$

$$\begin{aligned} L_T^{(2)}(C, \theta) = \frac{1}{4}\Bigg\{ &- \begin{pmatrix} -1 & \theta \\ 0 & C \end{pmatrix}\begin{pmatrix} K_T & T \\ T & J_T \end{pmatrix}\begin{pmatrix} -1 & 0 \\ \theta & C \end{pmatrix} \\ &+ \begin{pmatrix} 0 & I_T - C\theta J_T + CT \\ I_T - C\theta J_T + CT & 0 \end{pmatrix}\Bigg\} \end{aligned} \tag{6}$$

L'estimateur du maximum de vraisemblance $\begin{pmatrix} \hat{C}_T \\ \hat{\theta}_T \end{pmatrix}$ de $\begin{pmatrix} C \\ \theta \end{pmatrix}$ est solution de l'équation de vraisemblance :

$$L_T^{(1)}(\hat{C}_T, \hat{\theta}_T) = 0 \tag{7}$$

Ce système a une unique solution acceptable (la solution $\left(\hat{C}_T = 0, \hat{\theta}_T = \dfrac{\pi_T - \pi_0}{I_T}\right)$ est rejetée car elle ne vérifie pas l'hypothèse $C > 2/\theta$) donnée par :

$$\begin{cases} \hat{C}_T = \dfrac{I_T T - J_T(\pi_T - \pi_0)}{J_T K_T - T^2} \\ \hat{\theta} = \dfrac{I_T K_T - T(\pi_T - \pi_0)}{I_T T - J_T(\pi_T - \pi_0)} \end{cases} \tag{8}$$

Les propriétés asymptotiques de cet estimateur sont données dans le théorème suivant.

Théorème 3.4.

L'estimateur $(\hat{C}_T, \hat{\theta}_T)$ est fortement consistant et asymptotiquement normal, *i. e.* sous la loi $P^{\nu}_{C,\theta}$:

$$\lim_{T\to\infty} \begin{pmatrix} \hat{C}_T \\ \hat{\theta}_T \end{pmatrix} \overset{p.s.}{=} \begin{pmatrix} C \\ \theta \end{pmatrix}$$

et

$$\lim_{T\to\infty} T^{1/2}\left[\begin{pmatrix} \hat{C}_T \\ \hat{\theta}_T \end{pmatrix} - \begin{pmatrix} C \\ \theta \end{pmatrix}\right] \overset{D}{=} N(0, 4\Gamma^{-1}_{C,\theta})$$

où l'on a posé :

$$\Gamma_{C,\theta} = M\Lambda_{C,\theta}M'$$

avec

$$\Lambda_{C,\theta} = \begin{pmatrix} \mathbf{E}_\mu[X] & 1 \\ 1 & \mathbf{E}_\mu\left[\frac{1}{X}\right] \end{pmatrix} \quad \text{et} \quad M = \begin{pmatrix} -1 & \theta \\ 0 & C \end{pmatrix}$$

Preuve. — L'ergodicité du processus (X_t) entraîne, sous la loi $P^{\nu}_{C,\theta}$:

$$\lim_{T\to\infty} T^{-1}J_T \overset{p.s.}{=} \mathbf{E}_\mu[1/X] < \infty, \tag{9}$$

$$\lim_{T\to\infty} T^{-1}K_T \overset{p.s.}{=} \mathbf{E}_\mu[X] < \infty, \tag{10}$$

$$\lim_{T\to\infty} T^{-1}I_T \overset{p.s.}{=} 2\mathbf{E}_\mu[1/X]. \tag{11}$$

Les expressions de ces espérances, données au paragraphe précédent, et (8) montrent que l'estimateur est fortement consistant.

Montrons maintenant qu'il suffit d'établir les égalités suivantes pour prouver la normalité asymptotique :

$$\lim_{T\to\infty} T^{-1}L^{(2)}_T(C, \theta) = -\frac{1}{4}\Gamma_{C,\theta}, \text{ inversible}, \quad P^{\nu}_{C,\theta} - \text{p. s.} \tag{12}$$

$$\lim_{T\to\infty} T^{-1/2}L^{(1)}_T(C, \theta) \overset{D}{=} N\left(0, \frac{1}{4}\Gamma_{C,\theta}\right) \tag{13}$$

Effectuons un développement limité coordonnée par coordonnée de $L^{(1)}_T(\hat{C}_T, \hat{\theta}_T)$; en notant u^i la i-ème composante d'un vecteur u, on obtient :

$$0 = \frac{1}{\sqrt{T}}\left[L^{(1)}_T(\hat{C}_T, \hat{\theta}_T)\right]^i = \frac{1}{\sqrt{T}}\left[L^{(1)}_T(C, \theta)\right]^i$$
$$+ \frac{1}{\sqrt{T}}\left[L^{(2)}_T(\tilde{C}^i_T, \tilde{\theta}^i_T)\begin{pmatrix} \hat{C}_T - C \\ \hat{\theta}_T - \theta \end{pmatrix}\right]^i \tag{14}$$

où $(\tilde{C}_T^i, \tilde{\theta}_T^i)$ appartient au segment de droite reliant $(\hat{C}_T, \hat{\theta}_T)$ à (C, θ) et donc vérifie puisque l'estimateur est fortement consistant :

$$\lim_{T\to\infty} (\tilde{C}_T^i, \tilde{\theta}_T^i) \overset{\text{p.s.}}{=} (C, \theta), \ i = 1,2$$

Ceci entraîne :

$$\lim_{T\to\infty} T^{-1}[L_T^{(2)}(\tilde{C}_T^i, \tilde{\theta}_T^i) - L_T^{(2)}(C, \theta)] \overset{\text{p.s.}}{=} ((0)), \ i = 1,2 \tag{15}$$

Comme $\frac{1}{4}\Gamma_{C,\theta}$ est inversible, en passant à la limite dans l'équation (14) et en appliquant les relations (7), (15), (12) et (13), nous remarquons que nous sommes dans la situation suivante : soit (X_n), (Z_n) des suites de v. a. à valeurs dans $\mathbb{R}^d$, et (Y_n) une suite de v. a. de $M(d, d)$ telles que :

$$\lim_n X_n \overset{D}{=} X, \text{ X gaussien,}$$

$$\lim_n Y_n \overset{\text{p.s.}}{=} Y, \text{ Y déterministe, inversible,}$$

$$X_n = Y_n Z_n, \ \forall n.$$

Sous ces hypothèses, $\lim_n (X_n, Y_n) \overset{D}{=} (X, Y)$.

Comme $\exists n_0$, $\forall n > n_0$, Y_n est inversible, on a aussi $\lim_{n, n>n_0} Y_n^{-1} \overset{\text{p.s.}}{=} Y^{-1}$, et $\lim_{n, n>n_0} (X_n, Y_n^{-1}) \overset{D}{=} (X, Y^{-1})$.

La fonction $f : \mathbb{R}^d \times M(d, d) \to \mathbb{R}^d$, $(x, M) \to f(x, M) = Mx$ étant continue, on en déduit $\lim_{n, n>n_0} Z_n \overset{D}{=} Y^{-1}X$. Le vecteur X étant gaussien, $Y^{-1}X$ est aussi un vecteur gaussien.

Il nous reste donc seulement à prouver les égalités (12) et (13).

Pour ce faire, énonçons maintenant un lemme dont on peut trouver une preuve dans, par exemple, [20], chapitre 17 :

Lemme 3.5.

Soit $W = (W_t, \mathcal{F}_t)$, $t > 0$ un processus de Wiener standard et $f = (f_t)$, $t > 0$ un processus $(\mathcal{F}_t)$-adapté vérifiant :

$$\int_0^T f_t^2 dt < \infty, \text{ p. s., } \forall T > 0$$

et

$$\int_0^\infty f_t^2 dt = \infty, \text{ p. s.}$$

alors

$$\lim_{t\to\infty} \frac{\int_0^t f_u dW_u}{\int_0^t f_u^2 du} = 0, \text{ p. s.}$$

Sous la loi $P^{\nu}_{C,\theta}$:

$$I_T = C\theta J_T - CT + 2\int_0^T \frac{1}{\sqrt{\pi_t}} dW_t, \tag{16}$$

$$\pi_T - \pi_0 = C\theta T - CK_T + 2\int_0^T \sqrt{\pi_t}\, dW_t. \tag{17}$$

En utilisant (9), (10), (11), (16), (17) et le lemme précédent, on en déduit :

$$\lim_{T\to\infty} T^{-1}(I_T - C\theta J_T + CT) = 0, \ P^{\nu}_{C,\theta} - \text{p. s.} \tag{18}$$

$$\lim_{T\to\infty} T^{-1}(\pi_T - \pi_0 - C\theta T + CK_T) = 0, \ P^{\nu}_{C,\theta} - \text{p. s.} \tag{18}$$

Ceci implique (12). La convergence en loi (13) est une conséquence du théorème suivant dû à Yu. A. Kutoyants [17] :

Théorème 3.6.

Soit $W = (W_t, \mathfrak{F}_t)$, $t > 0$ un processus de Wiener standard m-dimensionnel, $W_t = (W_t^1, \ldots, W_t^m)$ et $F = (F_t)$, $t > 0$, $F_t = ((f_t^{ij}))_{n,m}$, un processus matriciel $(\mathfrak{F}_t)$-adapté vérifiant :

$$\forall i, j = 1, \ldots, n, \ f^{ij} \in M_2[0, T], \ \forall T > 0$$

$M_2[0, T]$ étant le sous-espace de $\mathfrak{L}_2[0, T]$ des processus (X_t), $t > 0$ vérifiant :

$$\mathbf{E}\left[\int_0^T X_t^2 dt\right] < \infty, \ \forall T > 0$$

et, si l'on pose $f_t^k = (f_t^{k1}, \ldots, f_t^{km})$, $k = 1, \ldots, n$,

$$\lim_{T\to\infty} T^{-1} \int_0^T (f_t^i, f_t^j) dt \overset{P}{=} c_{ij} < \infty, i, j = 1, \ldots, n,$$

alors :

$$\lim_{T\to\infty} T^{-1/2} \int_0^T F_s dW_s \overset{D}{=} N((0, \ldots, 0), \ C = ((c_{ij}))_{n,n}).$$

En effet, les égalités (9), (10) et ce théorème montrent que :

$$\lim_{T\to\infty} T^{-1/2} \int_0^T \begin{pmatrix} \sqrt{\pi_t} \\ 1/\sqrt{\pi_t} \end{pmatrix} dW_t \overset{D}{=} N(0, \Lambda_{C,\theta}) \qquad (20) \qquad \square$$

Montrons à présent l'efficacité asymptotique de l'estimateur $(\hat{C}_T, \hat{\theta}_T)$.

Proposition 3.7.

L'estimateur $(\hat{C}_T, \hat{\theta}_T)$ est asymptotiquement efficace.

Preuve. — Calculons M_T la borne de Cramer-Rao du modèle. Posons $f(x) = C(\theta - x)$, et définissons :

$$M_T(C, \theta) = {}_T\mathbb{E}^{\nu}_{C,\theta}\left[\int_0^T \frac{1}{4\pi_t} \nabla f(\pi_t)(\nabla f(\pi_t))' dt\right]$$

où $(\nabla f(x))' = \left(\frac{\partial}{\partial C} f(x), \frac{\partial}{\partial \theta} f(x)\right)$, et ${}_T\mathbb{E}^{\nu}_{C,\theta}$ est l'espérance calculée sous la loi ${}_TP^{\nu}_{C,\theta}$. On sait (cf. Kutoyants [16], th. 3.4.6, par exemple) que tout estimateur sans biais $(\hat{C}_T, \hat{\theta}_T)$ de (C, θ) vérifie :

$${}_T\mathbb{E}^{\nu}_{C,\theta}[(\hat{C}_T - C, \hat{\theta}_T - \theta)'(\hat{C}_T - C, \hat{\theta}_T - \theta)] - M_T^{-1}(C, \theta)$$

est une matrice définie non négative $\left(M_T^{-1}(C, \theta)\right.$ est dite « borne de Cramer-Rao », et plus traditionnellement notée $\left.I_T^{-1}(C, \theta)\right)$.

En posant :

$$H_T^{(i)} = {}_T\mathbb{E}^{\nu}_{C,\theta}\left[\int_0^T \pi_t^i dt\right], \quad i = 1, -1$$

on obtient :

$$M_T(C, \theta) = \frac{1}{4}\begin{pmatrix} \theta^2 H_T^{(-1)} - 2\theta T + H_T^{(1)} & C\theta H_T^{(1)} - CT \\ C\theta H_T^{(-1)} - CT & C^2 H_T^{(-1)} \end{pmatrix}$$

Les égalités (9) et (10) entraînent :

$$\lim_{T\to\infty} T^{-1} M_T(C, \theta) = \frac{1}{4}\Gamma_{C,\theta} \qquad \square$$

En utilisant l'estimateur de la variance, $\hat{K}_T = \frac{2\hat{\theta}_T}{\hat{C}_T}$, on peut obtenir la loi asymptotique de $T^{1/2}((\hat{C}_T - C), (\hat{\theta}_T - \theta), (\hat{K}_{T-K}))$. Ceci nous permettra de comparer ces estimateurs avec ceux issus de la méthode des moments.

Théorème 3.8.

L'estimateur $(\hat{C}_T, \hat{\theta}_T, \hat{K}_T)$ est fortement consistant et asymptotiquement normal, *i. e.* sous la loi $P^{v}_{C,\theta}$:

$$\lim_{T\to\infty} \begin{pmatrix} \hat{C}_T \\ \hat{\theta}_T \\ \hat{K}_T \end{pmatrix} \overset{p.s.}{=} \begin{pmatrix} C \\ \theta \\ K \end{pmatrix}$$

et

$$\lim_{T\to\infty} T^{1/2} \left[\begin{pmatrix} \hat{C}_T \\ \hat{\theta}_T \\ \hat{K}_T \end{pmatrix} - \begin{pmatrix} C \\ \theta \\ K \end{pmatrix} \right] \overset{D}{=} N(0, \Sigma_{C,\theta})$$

où :

$$\Sigma_{C,\theta} = \begin{pmatrix} 2C & \dfrac{-4}{C} & \dfrac{-4}{C^2}(C\theta + 2) \\ \dfrac{-4}{C} & \dfrac{4\theta}{C^2} & \dfrac{16\theta}{C^3} \\ \dfrac{-4}{C^2}(C\theta + 2) & \dfrac{16\theta}{C^3} & \dfrac{8\theta}{C^4}(C\theta + 6) \end{pmatrix} \tag{21}$$

Preuve. — Remarquons :

$$T^{1/2}(\hat{K}_T - K) = 2\,\frac{CT^{1/2}(\hat{\theta}_T - \theta) - \theta T^{1/2}(\hat{C}_T - C)}{C(\hat{C}_T - C) + C^2}$$

On définit la famille de fonctions $(h_t)_{t>0}$ et la fonction h par :

$$\begin{aligned} h_t \quad &: \mathbb{R}^2 \to \mathbb{R}^3 \\ h_t(x, y) &= \left(x, y, 2\,\frac{Cy - \theta x}{(C/\sqrt{t})x + C^2} \right) \\ h \quad &: \mathbb{R}^2 \to \mathbb{R}^3 \\ h(x, y) &= \left(x, y, -\frac{2\theta}{C^2}x + \frac{2}{C}y \right) \end{aligned}$$

Notons que $h_t \to h$, $t \to \infty$ uniformément sur tout compact, et que h est continue. D'autre part, en appelant Q_t la loi du couple $(t^{1/2}(\hat{C}_t - C), t^{1/2}(\hat{\theta}_T - \theta))$, le théorème précédent montre que $Q_t \to Q$ étroitement quand $t \to \infty$, où Q est une loi gaussienne en dimension 2, loi d'un couple que nous noterons (X, Y). Ces conditions sur h_t, h, Q_t, Q suffisent à assurer la conver-

gence étroite de $Q_t h_t^{-1}$ vers Qh^{-1} quand $t \to \infty$, où Qh^{-1} est la loi de la variable $h(X, Y) = (X, Y, aX + bY)$ avec $a = -\dfrac{2\theta}{C^2}$ et $b = \dfrac{2}{C}$ ([2], chap. 1, th. 5.5). C'est une loi normale car toute combinaison linéaire de ses coordonnées suit une loi normale unidimensionnelle. Elle est bien évidemment dégénérée. Le calcul de la matrice de covariance, à l'aide des expressions en fonction de C, θ et σ des statistiques de la loi invariante données au paragraphe 2, termine la preuve. □

Pour terminer, on revient aux paramètres c, θ et σ : soit $\hat{c}_T = \dfrac{\sigma^2 \hat{C}_{T'}}{4}$; on obtient le corollaire suivant :

Corollaire 3.9.

L'estimateur $(\hat{c}_T, \hat{\theta}_T, \hat{K}_T)$ est fortement consistant et asymptotiquement normal. De plus, l'estimateur $(\hat{c}_T, \hat{\theta}_T)$ est asymptotiquement efficace.

La preuve est une conséquence directe des résultats précédents, et du fait que $C = \dfrac{4C}{\sigma^2}$.

3.3. *Estimateurs des moments*

Ce paragraphe est consacré à la construction des estimateurs issus de la méthode des moments et à l'étude de leurs propriétés. Le résultat principal sera un théorème de consistance forte et de normalité asymptotique. Soit $\tilde{\theta}_T$ et $\tilde{K}_T$ des estimateurs de la moyenne θ et de la variance $K = \dfrac{2\theta}{C}$:

$$\tilde{\theta}_T = T^{-1} \int_0^T \pi_t dt,$$

$$\tilde{K}_T = T^{-1} \int_0^T (\pi_t - \tilde{\theta}_T)^2 dt$$

$$= T^{-1} \int_0^T \pi_t^2 dt - \tilde{\theta}_T^2.$$

Théorème 3.10.

L'estimateur $(\tilde{\theta}_T, \tilde{K}_T)$ est fortement consistant et asymptotiquement normal, *i. e.* sous la loi $P^{\nu}_{C,\theta}$:

$$\lim_{T \to \infty} \begin{pmatrix} \tilde{\theta}_T \\ \tilde{K}_T \end{pmatrix} \overset{\text{p.s.}}{=} \begin{pmatrix} \theta \\ K \end{pmatrix}$$

et

$$\lim_{T\to\infty} T^{1/2}\left[\begin{pmatrix}\tilde{\theta}_T\\ \tilde{K}_T\end{pmatrix} - \begin{pmatrix}\theta\\ K\end{pmatrix}\right] \overset{D}{=} N(0, \Gamma_{C,\theta})$$

où l'on a posé :

$$\Gamma_{C,\theta} = M_{C,\theta}\Lambda_{C,\theta}M'_{C,\theta}$$

avec

$$\Lambda_{C,\theta} = \begin{pmatrix} E_\mu[X] & E_\mu[X^2]\\ E_\mu[X^2] & E_\mu[X^3]\end{pmatrix} \quad \text{et} \quad M_{C,\theta} = \frac{2}{C}\begin{pmatrix} 1 & 0\\ \dfrac{2 - C\theta}{C} & 1\end{pmatrix}.$$

Preuve. — La consistance forte de l'estimateur est une conséquence immédiate de l'ergodicité de (x_t) puisque, sous la loi $P^{\nu}_{C,\theta}$:

$$\lim_{T\to\infty} T^{-1}\int_0^T \pi_t^i dt \overset{p.s.}{=} E_\mu[X^i] < \infty, \ \forall i \geq -1 \tag{22}$$

L'équation différentielle stochastique :

$$d\pi_t = C(\theta - \pi_t)dt + 2\sqrt{\pi_t}dW_t$$

et la formule d'Itô permettent d'obtenir la représentation suivante :

$$\begin{aligned} T^{1/2}\left[\begin{pmatrix}\tilde{\theta}_T\\ \tilde{K}_T\end{pmatrix} - \begin{pmatrix}\theta\\ K\end{pmatrix}\right] &= \begin{pmatrix} C^{-1}T^{-1/2}(\pi_0 - \pi_T)\\ 2^{-1}C^{-1}T^{-1/2}[(\pi_0^2 - \pi_T^2) - \alpha_T(\pi_0 - \pi_T)]\end{pmatrix}\\ &\quad + \frac{2}{C}\begin{pmatrix} 1 & 0\\ -\alpha_T & 1\end{pmatrix}\begin{pmatrix} T^{-1/2}\displaystyle\int_0^T \pi_t^{1/2}\,dW_t\\ T^{-1/2}\displaystyle\int_0^T \pi_t^{3/2}\,dW_t\end{pmatrix}\\ &= U_T + V_T \end{aligned} \tag{23}$$

où $\alpha_T = \dfrac{C\theta - 2}{C} + (\tilde{\theta}_T - \theta)$.

Il suffit donc de montrer, sous $P^{\nu}_{C,\theta}$:

$$\lim_{T\to\infty} E^{\nu}_{C,\theta}\,|\,U_T\,| = 0, \tag{24}$$

$$\lim_{T\to\infty} V_T \overset{D}{=} N(0, \Gamma_{C,\theta}). \tag{25}$$

Or, sous $P^{\nu}_{C,\theta}$, pour $s < t$, la loi conditionnelle de $\pi(t)$ sachant $\pi(s)$ a pour moyenne :

$$\pi(s) \exp(-C(t-s)) + \theta(1 - \exp(-C(t-s)))$$

et pour variance :

$$\frac{4}{C}\pi(s)(\exp(-C(t-s)) - \exp(-2C(t-s))) + \frac{2\theta}{C}(1 - \exp(-C(t-s)))$$

On en déduit :

$$\lim_{T\to\infty} \mathbb{E}^{\nu}_{C,\theta} \left| T^{-1/2}(\pi^i_0 - \pi^i_T) \right| = 0, \ i = 1,2$$

L'estimateur étant fortement consistant, (24) est vérifiée.

Par ailleurs, les mêmes arguments que ceux utilisés dans le théorème précédent et les égalités (22) montrent que :

$$\lim_{T\to\infty} T^{-1/2} \int_0^T \begin{pmatrix} \pi_t^{1/2} \\ \pi_t^{3/2} \end{pmatrix} dW_t \overset{D}{=} N(0, \Lambda_{C,\theta})$$

Ceci implique (25).

En utilisant l'estimateur fortement consistant de C, $\tilde{C}_T = \dfrac{2\tilde{\theta}_T}{\tilde{K}_T}$, on obtient la loi asymptotique de $T^{1/2}((\tilde{C}_T - C), (\tilde{\theta}_T - \theta), (\tilde{K}_T - K))$.

Théorème 3.11.

L'estimateur $(\tilde{C}_T, \tilde{\theta}_T, \tilde{K}_T)$ est fortement consistant et asymptotiquement normal, *i. e.* sous la loi $P^{\nu}_{C,\theta}$:

$$\lim_{T\to\infty} \begin{pmatrix} \tilde{C}_T \\ \tilde{\theta}_T \\ \tilde{K}_T \end{pmatrix} \overset{p.s.}{=} \begin{pmatrix} C \\ \theta \\ K \end{pmatrix}$$

et

$$\lim_{T\to\infty} T^{1/2} \left[\begin{pmatrix} \tilde{C}_T \\ \tilde{\theta}_T \\ \tilde{K}_T \end{pmatrix} - \begin{pmatrix} C \\ \theta \\ K \end{pmatrix} \right] \overset{D}{=} N(0, \Sigma'_{C,\theta})$$

où :

$$\Sigma'_{C,\theta} = \begin{pmatrix} 2C + \dfrac{8}{\theta} & -\dfrac{4}{C} & -\dfrac{4}{C^2}(C\theta + 6) \\ -\dfrac{4}{C} & \dfrac{4\theta}{C^2} & \dfrac{16\theta}{C^3} \\ -\dfrac{4}{C^2}(C\theta + 6) & \dfrac{16\theta}{C^3} & \dfrac{8\theta}{C^4}(C\theta + 10) \end{pmatrix} \tag{26}$$

La preuve est identique à celle du théorème analogue obtenu au paragraphe précédent.

De même, on obtient aisément le corollaire suivant, où $\tilde{c}_T = \dfrac{\sigma^2 \tilde{C}_T}{4}$.

Corollaire 3.12.

L'estimateur $(\tilde{c}_T, \hat{\theta}_T, \tilde{K}_T)$ est fortement consistant et asymptotiquement normal.

Enfin, on peut remarquer que $\sqrt{T}\,(\tilde{\theta}_T - \theta)$ et $\sqrt{T}(\tilde{\theta}_T - \theta)$ ont même variance asymptotique. On dit que les estimateurs $\tilde{\theta}_T$ et $\tilde{\theta}_T$ et θ ont même risque quadratique asymptotique.

4. Essais numériques pour les estimateurs paramétriques

Nous présentons dans ce paragraphe les résultats d'essais numériques effectués sur les estimateurs du maximum de vraisemblance et sur les estimateurs issus de la méthode des moments.

Nous travaillons avec les versions discrétisées des estimateurs en temps continu des paramètres de la dérive. Soit $\{(t_0, t_1, \ldots, t_N),\ t_0 = 0,\ t_N = T,\ t_0 < t_1 < \ldots < t_N\}$ une discrétisation du temps. Posons $\delta_N = \max_{i=1,\ldots,N} |t_i - t_{i-1}|$, et soit $\hat{\alpha}_{N,T} = (\hat{c}_{N,T}, \hat{\theta}_{N,T})$ l'estimateur discrétisé de $\alpha = (c, \theta)$. On peut montrer de la même manière que Le Breton dans [19] que lorsque $\delta_N \to 0$, $\hat{\alpha}_{N,T} \to \hat{\alpha}_T$, et $\delta_N^{-1/2}(\hat{\alpha}_{N,T} - \hat{\alpha}_T)$ est bornée en probabilité.

Nous avons procédé de la manière suivante. En utilisant le schéma de Milshtein, qui est robuste trajectoriellement alors que le schéma d'Euler ne l'est pas (cf. Talay [24]), nous avons simulé des trajectoires du processus (X_t) avec différentes valeurs pour c, θ et σ jusqu'à un temps T dix fois plus grand que la durée des relevés de taux d'intérêt dont nous disposions : T = 110 ans. Une durée de relevés de 110 ans n'est évidemment pas réaliste du point de vue économique car les paramètres du modèle n'ont aucune raison de rester stables aussi longtemps, mais nous verrons que nos estimateurs convergent très lentement; l'inégalité de Cramer-Rao (cf. les résultats du paragraphe 3.2) indique qu'aucun estimateur sans biais ne peut converger très vite. En effet, la borne inférieure de la variance de l'erreur d'estimation commise au bout d'un temps T assez grand se déduit de l'approximation suivante :

$$M_T^{-1}(c, \theta) \simeq \frac{4}{T}\Gamma_{C,\theta}^{-1} = \frac{1}{T}\begin{pmatrix} 2c & -\sigma^2/c \\ -\sigma^2/c & \theta\sigma^2/c^2 \end{pmatrix}$$

Par conséquent, pour un temp T assez grand, pour tout estimateur, la trace de la matrice de covariance du couple $(\hat{c}_T - c,\ \hat{\theta}_T - \theta)$ est bornée inférieurement par $\dfrac{1}{T}\,\dfrac{2c^3 + \theta\sigma^2}{c^2}$. Nos simulations correspondant à des valeurs de para-

mètres $c = 0{,}75$, $\theta = 0{,}1$ et $\sigma = 0{,}105$, pour la durée d'observation $T = 110$ ans, cette trace vaut environ $1{,}4.10^{-2}$; pour $T = 11$ ans, elle vaut environ 0,14 (donc toute estimation fondée sur l'observation pendant 11 ans du modèle correspondant à ces paramètres est nécessairement grossière; néanmoins les estimateurs étudiés précédemment fournissent un ordre de grandeur acceptable, à partir duquel on peut initialiser des algorithmes d'optimisation fondés sur l'observation de la gamme des taux).

En parallèle à cette application des estimateurs à des trajectoires simulées, nous présentons les résultats obtenus en les mettant en œuvre sur une suite de relevés journaliers du taux au jour le jour pendant 11 années (marché monétaire français, de janvier 1978 à janvier 1989).

Commentaires sur les graphes

- L'unité de temps en abscisse est l'année.
- Le pas de discrétisation de l'EDS pour la simulation est choisi égal au dixième de jour, tandis que le pas d'échantillonnage des données vaut 1 jour. On espère ainsi obtenir une erreur due à la simulation négligeable. De fait, on a vérifié que le comportement des estimateurs ne changeait pas notoirement en raffinant plus la discrétisation en temps de l'équation différentielle stochastique.

Figures 1 et 5 : L'estimateur du maximum de vraisemblance du paramètre c semble converger plus vite que celui issu de la méthode des moments. Ceci est conforme avec nos résultats précédents. En effet, le risque quadratique asymptotique de $\hat{c}_T$ est plus petit que celui de $\hat{c}_T$.

Figures 3 et 7 : Les deux estimateurs de θ semblent converger à la même vitesse, conformément au fait qu'ils ont le même risque quadratique asymptotique. Avec les valeurs choisies pour les paramètres, nous obtenons un risque quadratique asymptotique de $5{,}4.10^{-6}$.

Figures 3 et 1, 3 et 5, 7 et 1, 7 et 5 : Les estimateurs du maximum de vraisemblance et de la méthode des moments du paramètre θ convergent plus vite que ceux du paramètre c.

Figures 1 et 2, 3 et 4, 5 et 6, 7 et 8 : Le comportement des estimateurs des paramètres de la dérive sur trajectoire simulée montre qu'une durée d'observation de 11 ans pour le problème réel ne peut permettre une très bonne identification des paramètres du modèle, notamment c.

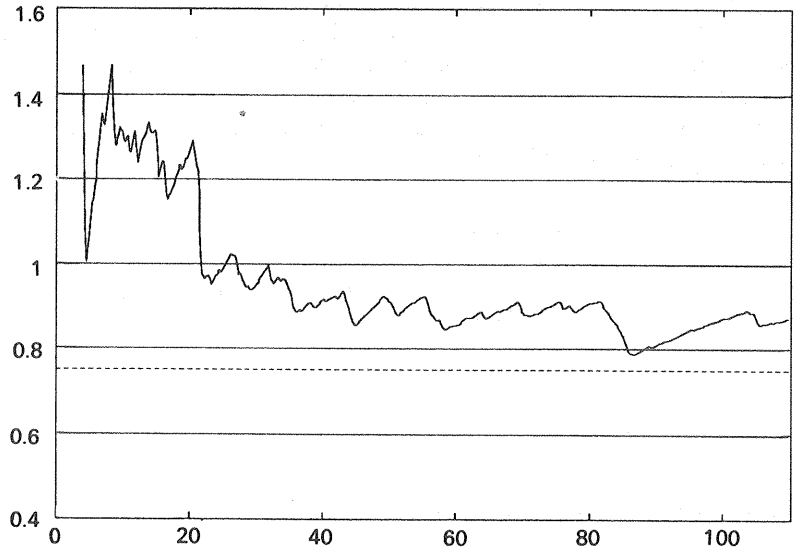


Fig. 1. — Estimateur $\hat{c}_t$ sur trajectoire simulée.

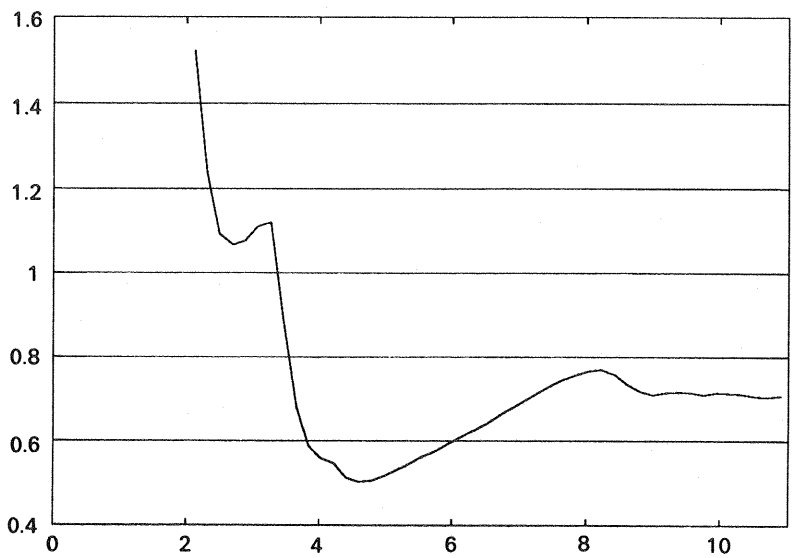


Fig. 2. — Estimateur $\hat{c}_t$ sur trajectoire réelle.

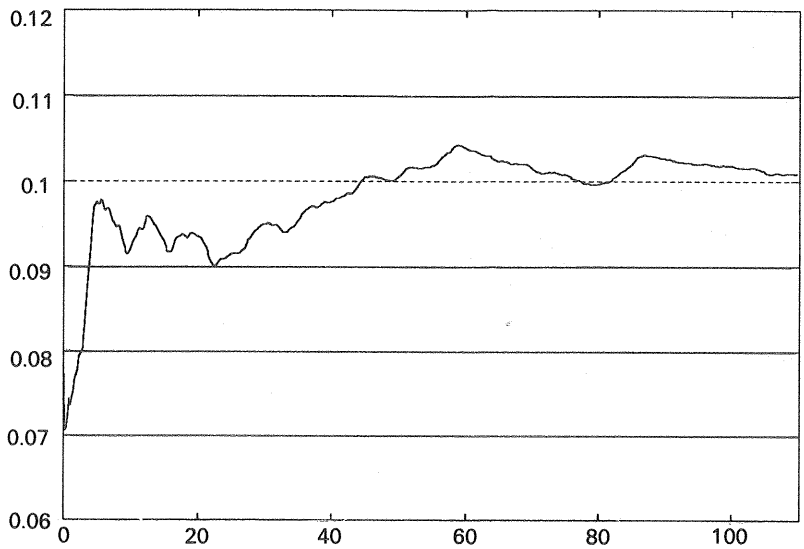


Fig. 3. — Estimateur $\hat{\theta}_t$ sur trajectoire simulée.

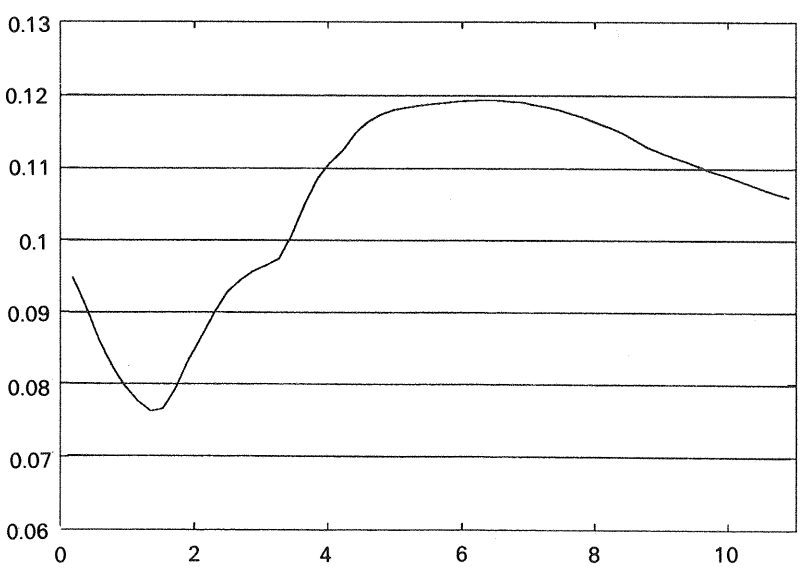


Fig. 4. — Estimateur $\hat{\theta}_t$ sur trajectoire réelle.

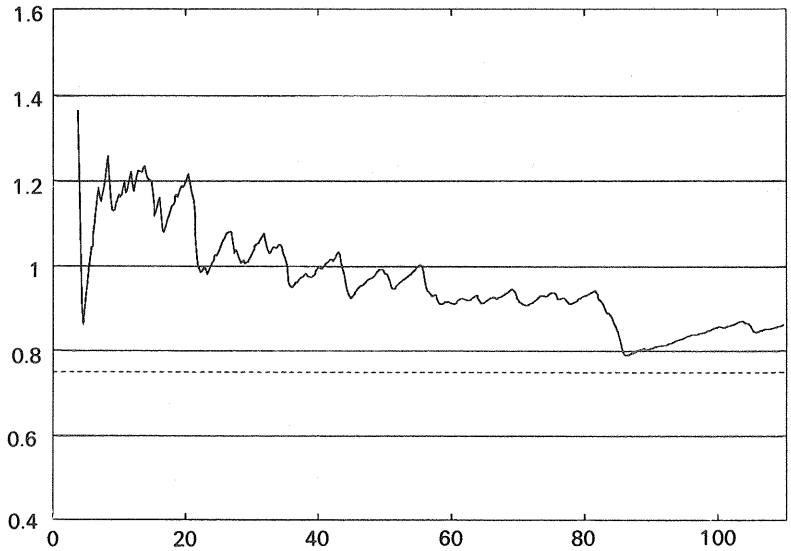


Fig. 5. — Estimateur $\hat{c}_t$ sur trajectoire simulée.

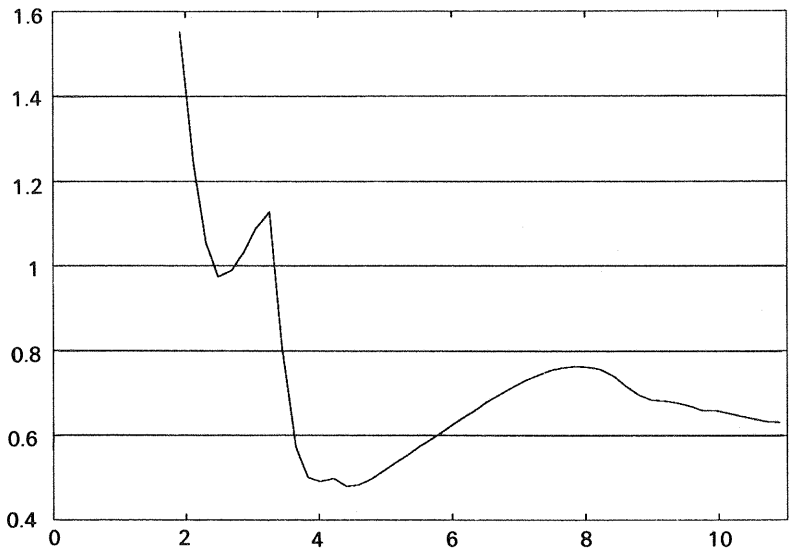


Fig. 6. — Estimateur $\hat{c}_t$ sur trajectoire réelle.

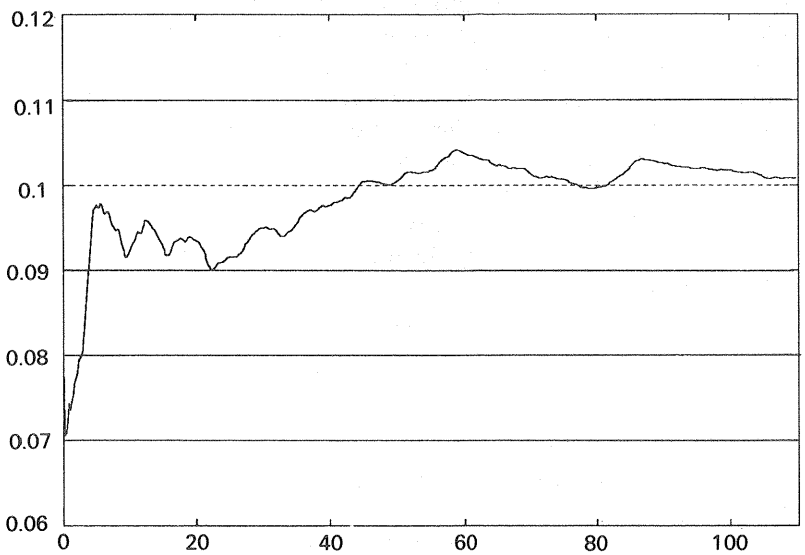


Fig. 7. — Estimateur $\tilde{\theta}_t$ sur trajectoire simulée.

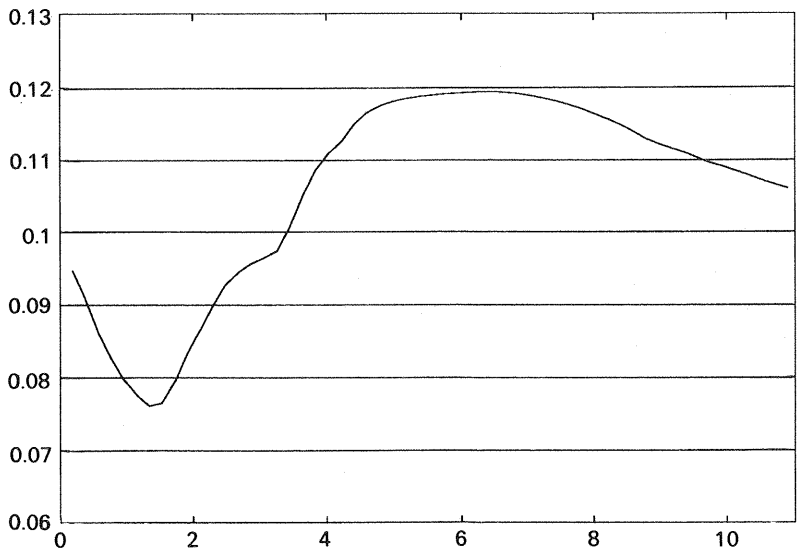


Fig. 8. — Estimateur $\tilde{\theta}_t$ sur trajectoire réelle.

Nous présentons aussi les résultats de l'estimation du paramètre de volatilité σ. L'estimateur $\hat{\sigma}_{N,T}$ dont nous avons donné l'expression au paragraphe 3.1 ne convient pas pour ce problème, car nous ne disposons que de relevés de taux à intervalle de temps fixé insuffisamment faible. Nous avons préféré un autre estimateur de σ, dont nous pouvons montrer la consistance, lorsque le temps d'observation t tend vers l'infini (la convergence est une simple application de la loi des grands nombres).

Soi Δ un pas de discrétisation fixé. Définissons le processus $(Y_t)_{0 \leq t \leq T}$ par $Y_t = 2\sqrt{X_t}$, $0 \leq t \leq T$. A l'aide de la formule d'Ito, nous obtenons facilement l'estimateur suivant :

$$\tilde{\sigma}_{\Delta,t} = \sqrt{\frac{1}{[t/\Delta]\Delta} \Sigma_{i=0}^{[t/\Delta]-1} (Y_{(i+1)\Delta} - Y_{i\Delta})^2}$$

où $[x]$ est la partie entière de x.

Nous avons testé cet estimateur sur des simulations avec les mêmes valeurs de paramètres que précédemment, c'est-à-dire $c = 0{,}75$, $\theta = 0{,}1$ et $\sigma = 0{,}105$ et $\Delta = 1$ jour, ainsi que sur les relevés de taux du marché. Les figures 9 et 10 montrent le bon comportement de cet estimateur. Remarquons que celui-ci suppose que le coefficient de volatilité σ reste constant au cours du temps. Ceci n'est pas réaliste du point de vue économique : il est probable que les paramètres du modèle subissent des chocs de temps à autre. Il nous a semblé intéressant de considérer, l'estimateur suivant :

$$\tilde{\sigma}^*_{\Delta,n,t} = \sqrt{\frac{1}{n\Delta} \Sigma_{i=0}^{n-1} (Y_{([t/\Delta]-i-1)\Delta} - Y_{([t/\Delta]-i)\Delta})^2}$$

$\tilde{\sigma}^*_{\Delta,n,t}$ exploite les n derniers relevés avant t. Nous avons cherché une valeur de n raisonnable; 90 jours semblait assurer une bonne convergence de $\tilde{\sigma}^*_{\Delta,n,t}$ vers σ. Les figures 11 et 12 montrent le comportement de cet estimateur.

• L'unité de temps en abscisse est l'année.

Figures 9 et 10 : La figure 9 montre que l'estimateur de σ converge rapidement.

Figures 11 et 12 : Le paramètre de volatilité σ estimé varie dans le temps (fig. 12). Il devrait avoir un comportement plus stable si les données provenaient d'un modèle CIR (fig. 11). Néanmoins, les estimations paraissent centrées autour d'une moyenne de la volatilité.

5. Estimateurs non paramétriques

Dans le but de tester la validité du modèle CIR, nous avons utilisé également des estimateurs non paramétriques d'une diffusion, décrits dans le rapport de Delebecque et Quadrat [7] et l'article de D. Florens-Zmirou [10], qui permettent

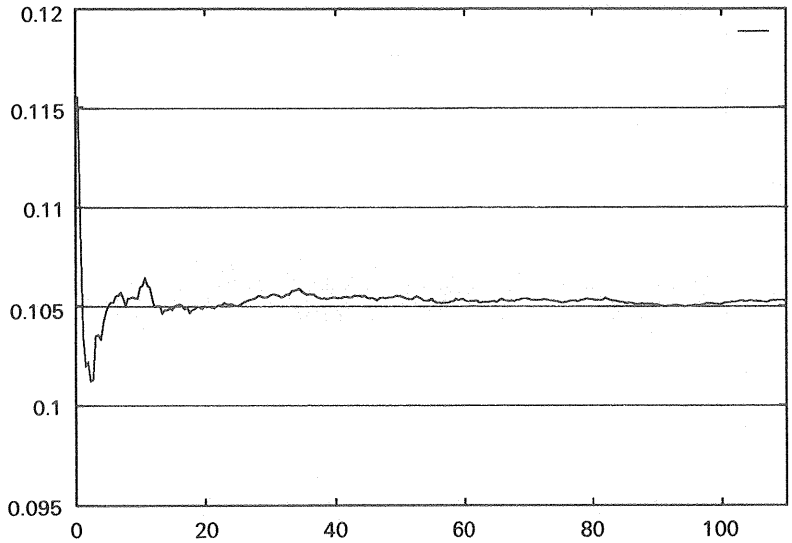


Fig. 9. — $\hat{\sigma}_{\Delta,t}$ sur trajectoire simulée.

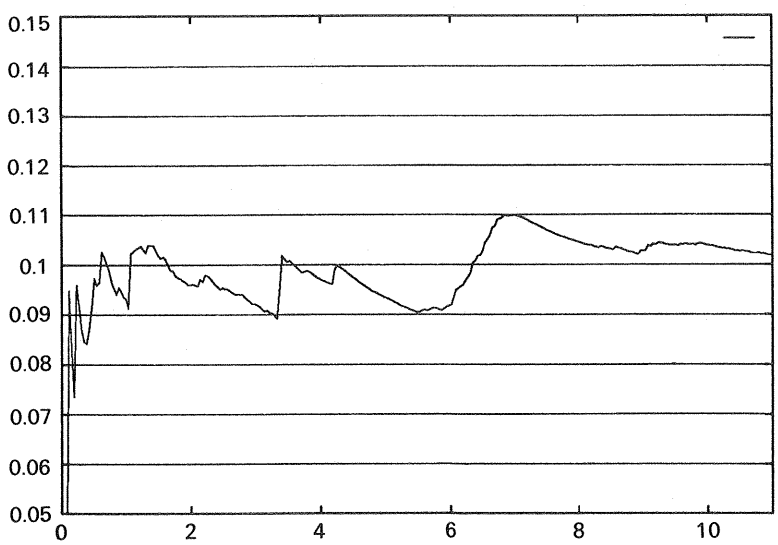


Fig. 10. — $\hat{\sigma}_{\Delta,t}$ sur trajectoire réelle.

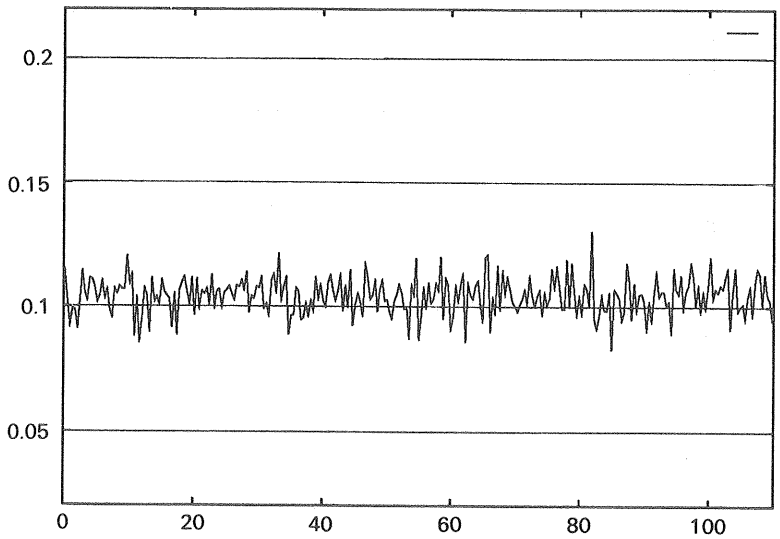


Fig. 11. — $\hat{\sigma}^*_{\Delta,90,t}$ sur trajectoire simulée.

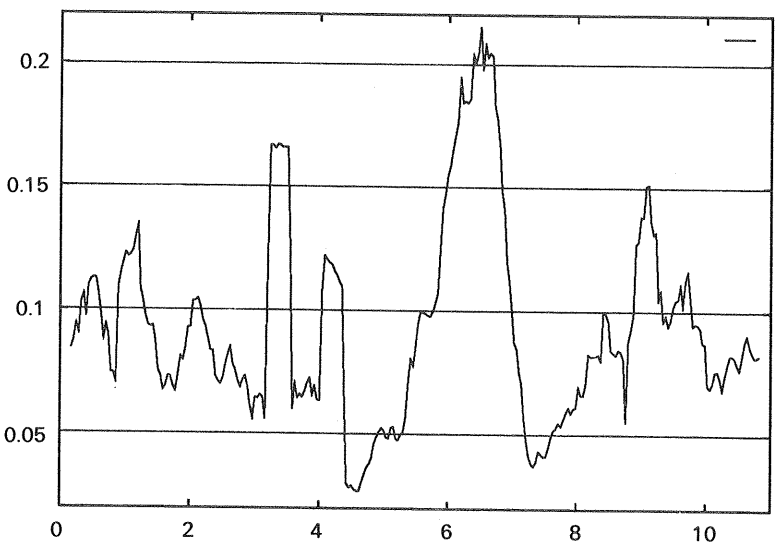


Fig. 12. — $\hat{\sigma}^*_{\Delta,90,t}$ sur trajectoire réelle.

d'identifier les fonctions de dérive et de diffusion à l'intérieur d'une famille de fonctions donnée (en pratique, pour des modèles très généraux, on cherche à identifier une approximation linéaire par morceaux des fonctions de dérive et de diffusion).

Considérons le processus ergodique (X_t) à valeur dans un espace de dimension fini vérifiant :

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

Pour simplifier, nous allons prendre une fonction de dérive $b(x)$ constante par morceaux sur une partition de l'espace d'état.

Choisissons une partition $(A_i)_{i=1\ldots n}$ de l'espace d'état et posons :

$$b(x) = b(\theta_1, \ldots, \theta_n, x) = \sum_{i=1}^{n} \theta_i \chi_{[x \in A_i]}$$

où $\chi_{[x \in A]} = 1$, si $x \in A$, $= 0$, sinon. On est ainsi ramené à un problème d'estimation paramétrique très simple pour le paramètre vectoriel $(\theta_1, \ldots, \theta_n)$.

L'estimateur du maximum de vraisemblance de θ_i est :

$$\hat{\theta}_i^T = \frac{\displaystyle\int_0^T \frac{Z_s^i}{\sigma^2(X_s)} dX_s}{\displaystyle\int_0^T \frac{Z_s^i}{\sigma^2(X_s)} ds}$$

où $Z_s^i = \chi_{[X_s \in A_i]}$.

On montre que, pour tout x, $b(\hat{\theta}_1^T, \ldots, \hat{\theta}_n^T, x)$ converge vers $b(\theta_1, \ldots, \theta_n, x)$ p. s. Delebecque et Quadrat donne un résultat de vitesse de convergence[1].

Pour l'estimateur non paramétrique de la fonction de diffusion, on utilise le résultat suivant. Considérons la diffusion :

$$X_t = X_0 + \int_0^t \sigma(X_s)dW_s, \quad 0 \leq t \leq 1$$

où (W_t) est un brownien unidimensionnel, et $\sigma(x)$ une fonction $\mathbb{R}$ dans $\mathbb{R}$ de classe C^1, localement lipschitzienne et strictement positive.

Soit enfin :

$$Z_n^i(x) = \chi_{[|X_{i/n} - x| < h_n]}, \quad i = 1, \ldots, n$$

et

$$S_n(x) = n \frac{\displaystyle\sum_{i=0}^{n-1} Z_n^i(x)(X_{(i+1)/n} - X_{i/n})^2}{\displaystyle\sum_{i=0}^{n-1} Z_n^i(x)}$$

où $(h_n)_{n \geq 0}$ est une suite tendant vers 0, telle que $(nh_n^2)^{-1} \log(n)$ tend vers l'infini; sur l'événement $[T(x) < 1]$, où $T(x)$ est le premier temps d'atteinte de x, cet estimateur converge en probabilité vers $\sigma^2(x)$.

On montre facilement que cet estimateur reste valable si l'on considère à la place de (X_t) un processus (Y_t) d'EDS :

$$Y_t = Y_0 + \int_0^T b(Y_s)d_s + \int_0^T \sigma(Y_s)dW_s$$

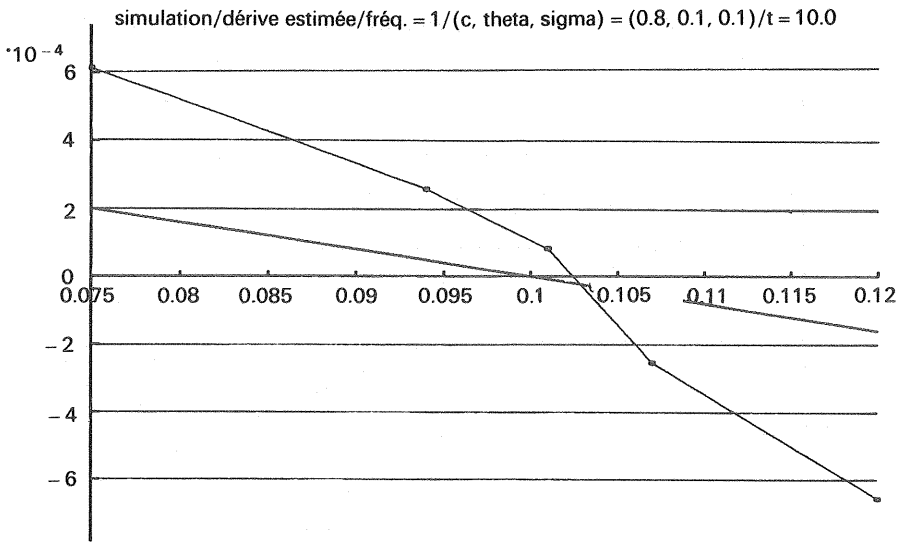
où la fonction de dérive $b(x)$ est lipschitzienne.

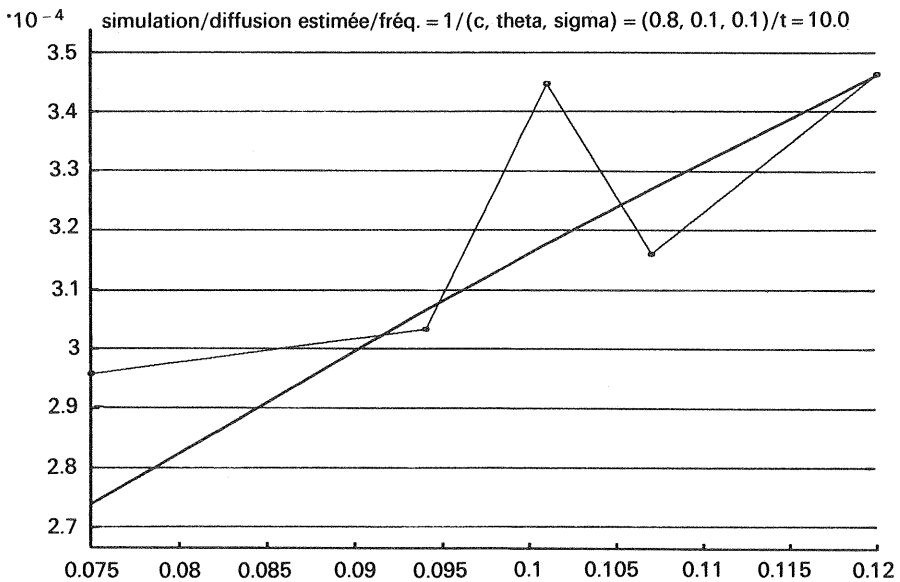
Les huit courbes qui suivent montrent le comportement de ces estimateurs : nous avons simulé une trajectoire d'une diffusion de type CIR, en discrétisant cette équation avec un pas égal à $\frac{\delta}{k}$, δ étant l'intervalle de temps entre deux relevés. On fait varier le temps d'observation ($t = 10, 50, 100$), ainsi que le pas de discrétisation. Ceci permet de mesurer la sensibilité des estimateurs à la qualité de la simulation.

On a choisi pour valeurs numériques : $c = 0{,}8$, $\theta = 0{,}1$, $\sigma = 0{,}1$. Les courbes en gras correspondent respectivement aux fonctions $c(\theta - x)$ et $\sigma\sqrt{x}$.

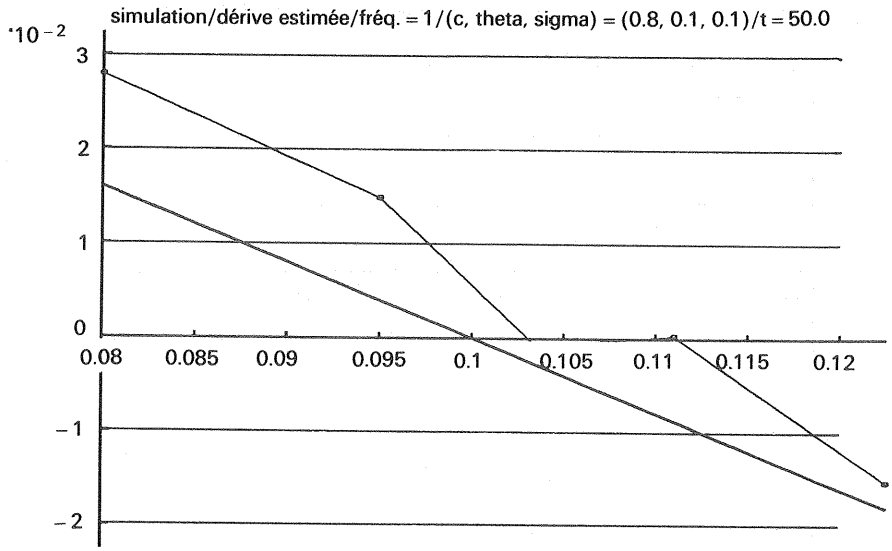
Pour le tracé des courbes, on a approché les fonctions constantes par morceaux estimées par des fonctions continues linéaires par morceaux.

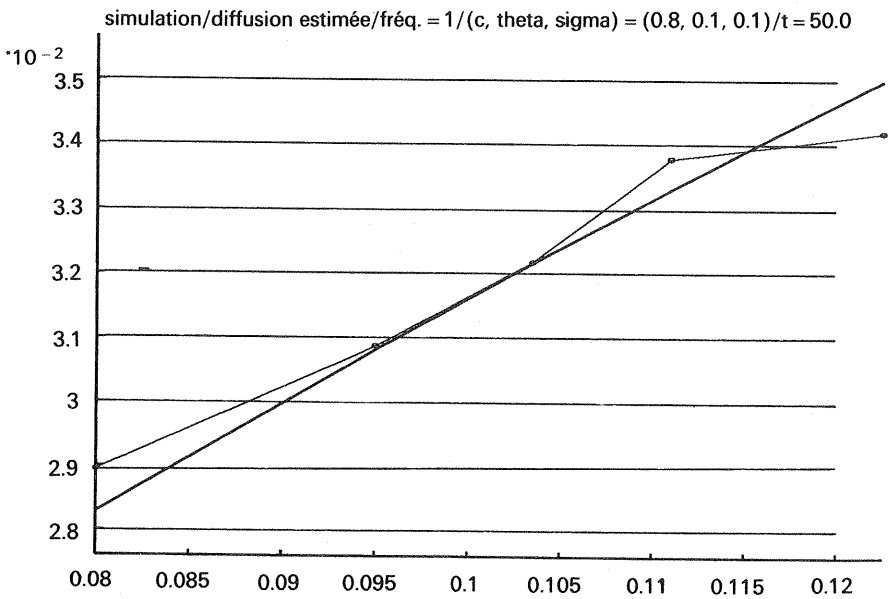
Estimation de la dérive et la diffusion sur 10 ans.



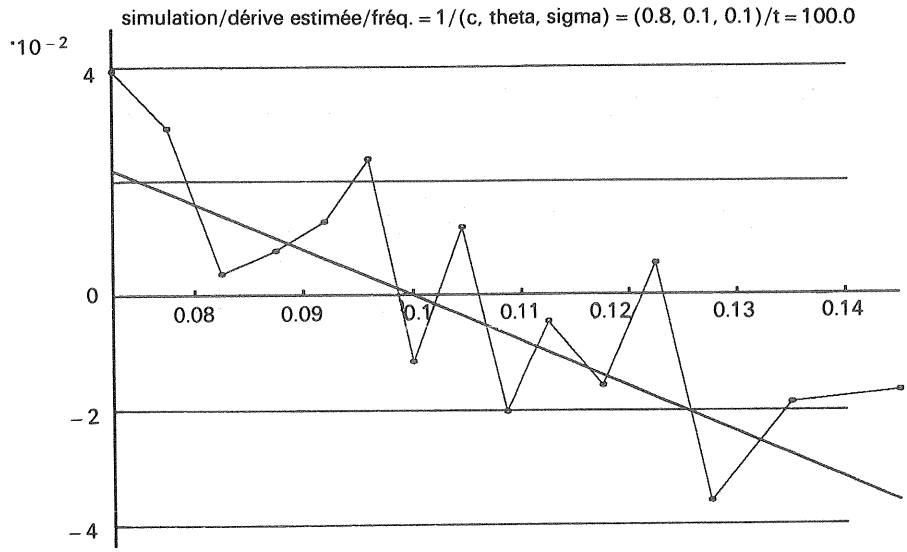


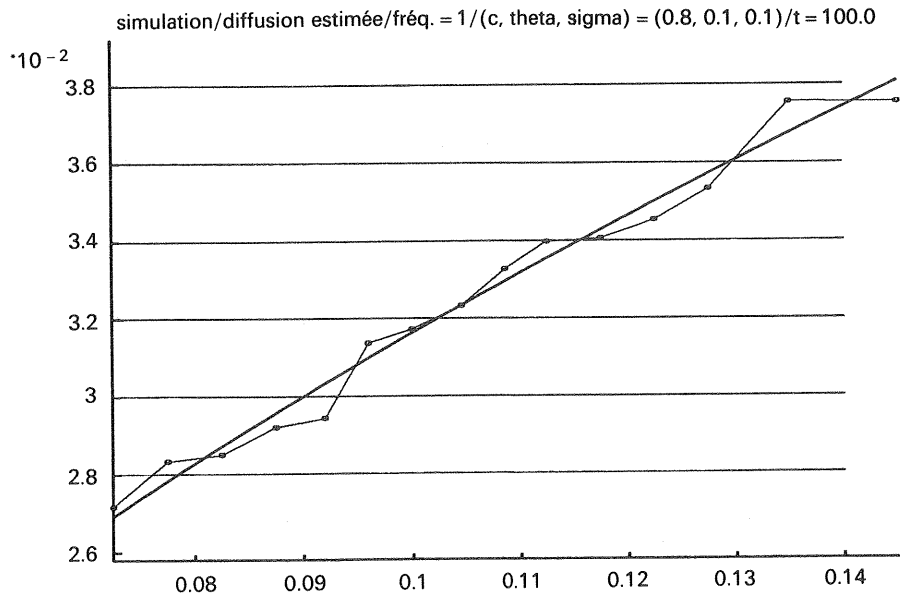
Estimation de la dérive et la diffusion sur 50 ans.



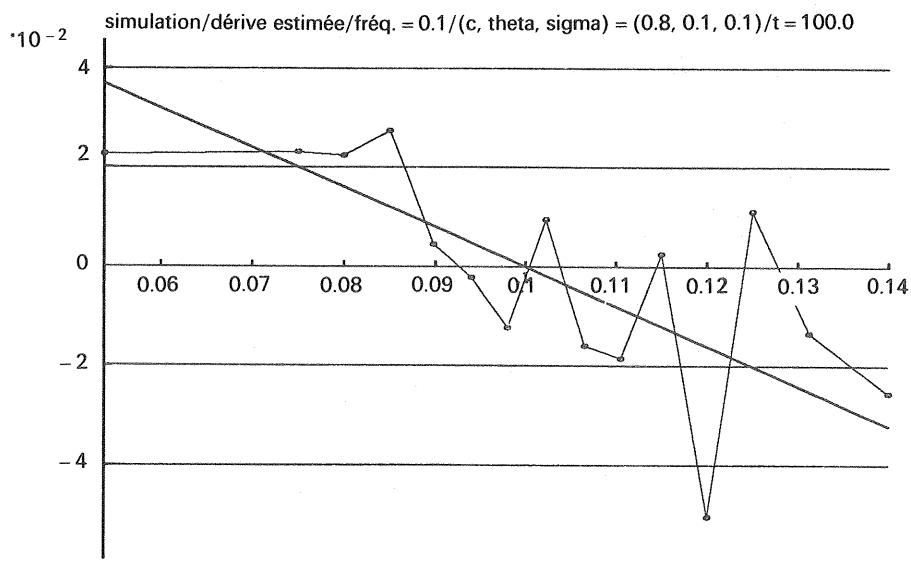


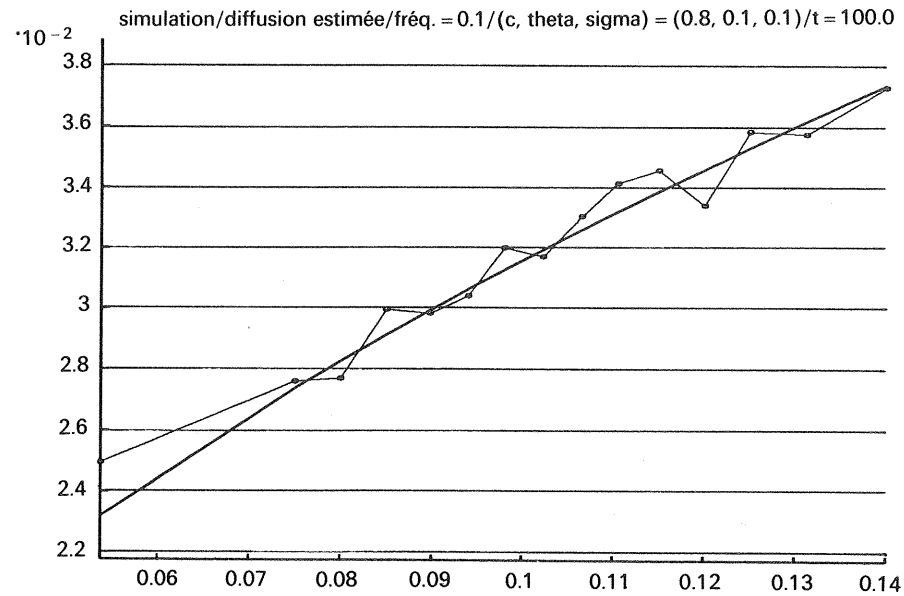
Estimation de la dérive et la diffusion sur 100 ans pour un pas de discrétisation égal à 1 jour.





Estimation de la dérive et la diffusion sur 100 ans pour un pas de discrétisation égal au dixième jour.



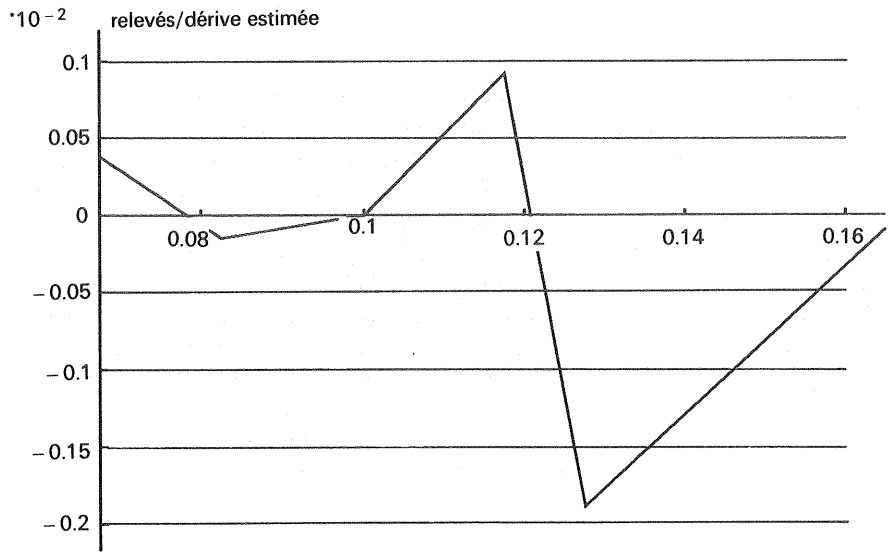


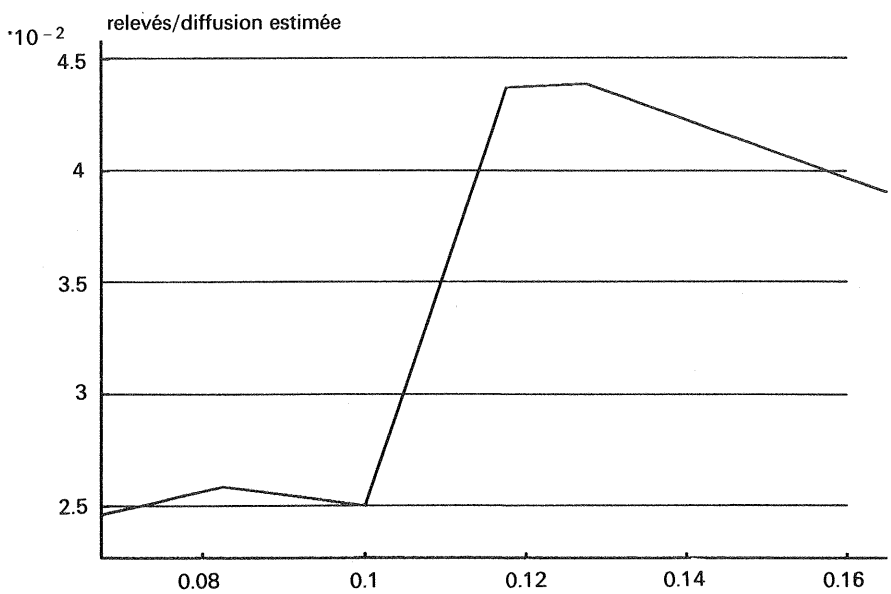
On remarquera que l'estimateur de la diffusion donne des résultats excellents à partir de $t = 50$, et acceptables à partir de $t = 10$. La sensibilité au pas de discrétisation en temps de l'EDS est faible.

Par contre, l'estimation de la dérive est plus délicate.

A titre indicatif, à partir des relevés dont nous disposions, les estimateurs précédents conduisent aux 2 courbes ci-dessous.

Estimation de la dérive et la diffusion sur 11 ans (données réelles).





Les essais numériques présentés ci-dessus montrent que l'estimation de la dérive ne doit pas être prise en compte. Celle de la diffusion, si tant est qu'elle est convenable, ne plaide pas pour une fonction de type $\sigma\sqrt{x}$.

Ceci ne remet pas nécessairement en cause le modèle CIR sur les intervalles de temps où les paramètres sont constants; de tels intervalles sont sûrement relativement courts, notamment en ce qui concerne la volatilité σ.

6. Conclusion

Nous avons vu que la statistique des processus appliquée au modèle CIR ne peut donner de réponses précises en un temps raisonnable au problème d'estimation paramétrique. En effet, la borne de Cramer-Rao montre clairement qu'un temps d'observation très long est nécessaire, et ce quelle que soit la méthode d'estimation utilisée, pour obtenir des résultats précis.

Néanmoins, ces méthodes statistiques nous permettent d'estimer précisément en temps fini le paramètre de volatilité et d'établir notamment que la volatilité semble dépendre du temps, et nous donnent des ordres de grandeurs ainsi que des intervalles de confiance pour les autres paramètres. Nous sommes ainsi capables de mettre en œuvre plus efficacement des méthodes plus robustes (adéquation entre des prix obligataires théoriques déduits du modèle et les prix du marché).

Nous conclurons en faisant remarquer que les méthodes non paramétriques existantes ne permettent pas d'accepter, ni de rejeter le modèle CIR. Il faudrait essayer de développer des tests de modèle efficaces.

NOTE

([1]) En supposant la dérive continue et, en posant :

$$\tau_i^{\mathrm{H}} = \mathit{inf}\left\{ t : \int_0^t \frac{Z_s^i}{\sigma^2(X_s)}\, ds \geq \mathrm{H} \right\},$$

et

$$\hat{\theta}_i^{\mathrm{H}} = \frac{1}{\mathrm{H}} \int_0^{\tau_i^{\mathrm{H}}} \frac{Z_s^i}{\sigma^2(X_s)}\, dX_s,$$

ils obtiennent le théorème d'approximation suivant :

Si

$$\int_0^\infty \frac{Z_s^i}{\sigma^2(X_s)}\, ds = \infty \text{ p. s.},$$

alors

$$\forall i = 1 \ldots n,\ \inf_{A_i} b(x) - |\eta| \leq \hat{\theta}_i^{\mathrm{H}} \leq \sup_{A_i} b(x) + |\eta|$$

où

$$\eta = \frac{1}{\mathrm{H}} \int_0^{\tau_i^{\mathrm{H}}} \frac{Z_s^i}{\sigma^2(X_s)}\, dW_s, \text{ et l'on sait que } \eta \simeq \mathrm{N}\left(0, \frac{1}{\mathrm{H}}\right).$$

RÉFÉRENCES

[1] Basawa I. V. and Prakasa Rao L. S. (1980), Statistical inference for stochastic processes, Academic Press.

[2] Billingsley P. (1968), Convergence of probability measures, John Wiley and sons.

[3] Brown B. M. and Hewitt H. I. (1975), Asymptotic likelihood theory for diffusion process, *J. Appl. Prob.,* 12, 228-238.

[4] Brown S. and Dybvig P. (1986), The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *Journal of Finance,* XLI-3, 1986.

[5] Brown R. H. and Schaefer S. M. (1988), Testing the CIR Model on British Government Index-linked Securities (Working Paper), London Business School, IFA 109-88.

[6] Cox J. C., Ingersoll J. E. and Ross S. A. (1985), A Theory of the Term Structure of the Interest Rates, Econometrica, vol. 53.

[7] Delebecque F. and Quadrat J. P. (1975), Identification d'une diffusion stochastique, Rapport INRIA 121.

[8] Duffie D. and Singleton K. (1990), Simulated Moments Estimation of Markov Models of Asset Prices, Stanford University, Graduate School of Business.

[9] Feller W. (1951), Two Singular Diffusion Problems, Annals of Mathematics.

[10] Florens-Zmirou D. (1989), Estimation de la variance d'une diffusion à partir d'une observation discrétisée, note au CRAS, t. 309, série I.

[11] Gibbons M. R. and Ramaswamy K. (1990), The Term Structure of the Interest Rates: Empirical Evidence (Working Paper), Graduate School of Business, Stanford University.

[12] Gihman I. I. and Skorohod A. V. (1972), Stochastic differential equations, Springer-Verlag.

[13] Has'Minskii R. Z. (1980), Stochastic stability of differential equation, Sijthoff and Norrdhoff.

[14] Heston S. (1990), Testing Continuous-time Models of the Term Structure of Interest Rates, Yale University, School of Organization and Management.

[15] Ikeda N. and Watanabe S. (1981), Stochastic differential equations and Diffusion processes, North-Holland Mathematical Library.

[16] Kutoyants Yu. A. (1984), Parameter estimation for stochastic processes, Heldermann Verlag, Berlin.

[17] Kutoyants Yu. A. (1975), On a Hypotheses Testing Problem and Asymptotic Normality of Stochastic Integral, Theory-Probability Appl. 20, 376-384.

[18] Le Breton A. and Musiela M. (1984), A Study of a One-Dimensionnal Bilinear Differential Model for Stochastic Processes, Prob. and Math. Stat., Vol. I, Fasc. 1, p. 91-107.

[19] Le Breton A. (1976), On continuous and discrete sampling for parameter estimation in diffusion type processes, Mathematical Programming Study 5, p. 124-144, North-Holland Publishing Company.

[20] Lipster R. S. and Shiryayev A. N. (1978), Statistics of Random Processes I, II, Springer-Verlag.

[21] Longstaff F. A. (1989), A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates, *Journal of Financial Economics,* 23-2, pp. 193-224.

[22] Merton R. (1980), On Estimating the Expected Return on the Market, *Journal of Financial Economics,* 8, 323-361.

[23] Rogers L. C. G., Williams D. (1987), Diffusion, Markov processes and Martingales, John Wiley and sons.

[24] Talay D. : Simulation and Numerical Analysis of Stochastic Differential Systems (rapport de recherche INRIA 1313 (1990), Effective Stochastic Analysis, P. Kree and W. Wedig (eds.), à paraître.