Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 347 (2009) 429–434
COMPTES RENDUS MATHEMATIQUE

Statistique/Systèmes dynamiques

Estimation de facteurs de Bayes entre modèles dynamiques non linéaires à espace d'état

Jean-Pierre Vila, Issa Saley

UMR Analyse des systèmes et biométrie, INRA-SupAgro, 2, place Pierre-Viala, 34060 Montpellier, France

Reçu le 15 septembre 2008 ; accepté après révision le 6 février 2009

Disponible sur Internet le 14 mars 2009

Présenté par Paul Deheuvels

Résumé

Les modèles non linéaires à espace d'état sont utilisés de façon croissante pour représenter de nombreux systèmes dynamiques stochastiques et pour les contrôler. De nouveaux outils de filtrage particulaire sont maintenant disponibles pour l'identification de ces modèles. Il n'en va pas de même pour le problème de leur sélection statistique car les vraisemblances associées sont le plus souvent non accessibles et d'estimation difficile. Ceci exclut a priori les critères classiques de comparaison de modèles de type Akaïke et compromet l'utilisation des méthodes performantes basées sur l'estimation d'un facteur de Bayes par simulations MCMC.

Cette Note propose un estimateur convergent non paramétrique d'un facteur de Bayes pour ces modèles, comme application directe de ces nouveaux filtres particulaires. ***Pour citer cet article : J.-P. Vila, I. Saley, C. R. Acad. Sci. Paris, Ser. I 347 (2009).***

Abstract

Bayes factor estimation for nonlinear dynamic state space models. The use of nonlinear state space models in the study and control of stochastic dynamic systems is regularly growing. With the new generation of particle filters, efficient filtering methods are now available for the identification of these models. However their statistical selection is still an open problem because of the frequent nonaccessibility of the related likelihoods and the intricate estimation of the latter. This rules out all the usual model comparison information criteria as Akaïke's and unfavour also the efficient methods relying on Bayes factor estimation by MCMC simulations.

This Note shows how a convergent nonparametric Bayes factor estimator can be built and used advantageously, as direct application of these new particle filters themselves. ***To cite this article: J.-P. Vila, I. Saley, C. R. Acad. Sci. Paris, Ser. I 347 (2009).***

Abridged English version

The stochastic state space models are tools especially well adapted to the representation of many dynamic systems with hidden structure and indirectly observable, of the following form:

Adresses e-mail : vila@supagro.inra.fr (J.-P. Vila), saley@supagro.inra.fr (I. Saley).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2009.02.017

$$\begin{cases} x_{t+1} = f_t(x_t, \theta_x, \varepsilon_t), \\ y_{t+1} = h_t(x_{t+1}, \theta_y, \eta_t) \end{cases} \tag{1}$$

in which t is the time index, f_t and h_t are Borelian functions, $x_t \in \mathbb{R}^d$ is a vector of unobserved state variables, $y_t \in \mathbb{R}^q$ a vector of observed variables, $\theta = (\theta_x^T, \theta_y^T)^T \in \Theta \subset \mathbb{R}^p$ a vector of unknown and constant parameters, ε_t, and η_t are vectors of independent white noises such that the probability density $p_t(y_t|x_t)$ exists and is bounded.

When the functions f_t and h_t are known and linear with respect to the state variables and the noises, the identification of such models can be performed by the Kalman filter, from prior densities p_0^x and p_0^θ for x and θ respectively. When f_t and h_t are known but nonlinear, the estimation of θ is a difficult enterprise, even with the most renowned particle filtering methods (Doucet et al. [7], Del Moral [5]). Recently a new method has been proposed, the resampling particle convolution filtering method, based on kernel estimators (Rossi [18], Rossi and Vila [20]), particularly suitable for parameter estimations (Rossi and Vila [19]). This method allows convergent estimates of the conditional densities of the state variables and parameter vectors, to be built at each time t. In this new approach the model (1) is first completed with the additional state equation $\theta_{t+1} = \theta_t$. Then at each time step t, n particles $(\tilde{x}_t^i, \tilde{\theta}_t^i, \tilde{y}_t^i), i = 1, \ldots, n$, are simulated from the model so enlarged. These particles and the current observation y_t are then used to estimate by kernel estimators the probability densities of x_t and θ_t conditional on the observations up to time t. Under weak assumptions the recursive use of these density estimates in the successive generations of sets of n particles, ensures at each time t the almost sure L_1-convergence of these estimates to their true counterparts as n tends to infinity. Moreover the almost sure convergence as t and n tend to infinity of the estimate $\hat{\theta}_t^n$ of the conditional expectation of θ_t to the true unknown parameter value θ is also ensured (as also that of the estimate $\hat{x}_t^n$ of the conditional expectation of x_t to its true counterpart as n tends to infinity).

However, the major issue which has to be considered before that of the parameter estimation itself, is that of the choice of the best state space model within a set of candidate models, given a time series of data observations $Y = y_1, y_2, \ldots, y_T$. This dynamic model selection problem is a particular case of the general problem of probabilistic model selection, for which one of the most efficient criteria with respect to the present-day computing facility, relies upon the estimation of the so-called Bayes Factor (Jeffreys [14], Kass and Raftery [15]). Given a data set of observations and two models M_1 and M_2, their Bayes factor is defined as the ratio of their respective marginal likelihoods, for which Kass and Raftery have proposed an interpretative scale for model selection help [15]. The exact computation of a Bayes factor is often rather intricate if not impossible. Actually, one has to resort to estimations by Monte Carlo Markov Chain (MCMC) simulations (Han and Carlin [12]) or even Reversible Jump MCMC (Green [11]). The most efficient among the MCMC-based Bayes factor estimation methods, embed the issue into the more general problem of the estimation of the ratio of the normalizing constants of two probability distributions (Meng and Wong [16], Chen and Shao [3], Gelman and Meng [9], Bartolucci et al. [1]): as a matter of fact the marginal likelihood of a model M (not to be confused with the likelihood function of the model) appears to be the normalizing constant of the model parameter posterior distribution. Applied to Bayes factor estimation these MCMC-based methods typically need exact numerical evaluations of the usual likelihood functions of the models, computed on MCMC outputs of the parameter posterior densities (these outputs can be obtained for example from the Metropolis-Hastings algorithm [13], the Gibbs sampler [8] or some various hybrid algorithms [23]). For linear Gaussian state space models the likelihood function is computable through the Kalman filter (Tanizaki [22]). However, in the case of general nonlinear state space models the likelihood functions are usually not reachable. These MCMC-based methods are then not relevant anymore for a Bayes factor estimation in this context, or need as will be seen, some heavy adjustments often providing uncertain results.

To contribute to a solution to this estimation problem, this Note shows how a convergent kernel-based estimator of the marginal likelihood of such a model, can be easily built from the resampling particle convolution filter mentioned above. This nonparametric estimator uses a classic decomposition of a model marginal likelihood. This decomposition was considered by several authors such as Dawid [4], Kass and Raftery [15], Gneiting and Raftery [10], to point out connections of the marginal (or integrated) likelihood, with related concepts such as prequential scores [4], the Bayes information criterion and other scoring rules to assess the quality of probabilistic forecasts.

Given two competing state space models of type (1), a convergent nonparametric estimator of their Bayes factor follows therefore immediately. Simulation studies conducted for state space models of different complexities, all confirm the practical interest of such an estimator.

1. Rappels préliminaires

Soit deux modèles probabilistes M_1 et M_2, et Θ_1 et Θ_2 les espaces de paramètres associés. On notera θ sans indice un élément de l'espace Θ_i, $i = 1, 2$, quand le modèle référé sera sans ambiguité.

Etant donné un échantillon d'observations $Y = y_1, \ldots, y_T$, le facteur de Bayes entre les deux modèles M_1 et M_2 est défini par

$$B_{12} = \frac{p_1(Y)}{p_2(Y)} \quad \text{avec } p_i(Y) = \int_{\Theta_i} p_i(Y|\theta) p_i(\theta)\, d\theta, \ i = 1, 2. \tag{2}$$

$p_i(Y)$, vraisemblance marginale du modèle M_i, est donc la constante de normalisation de $p_i(\theta, Y) = p_i(Y|\theta)p_i(\theta)$, densité a posteriori de θ, $p_i(\theta|Y)$, non normalisée.

Kass et Raftery [15] ont proposé une échelle d'interprétation des valeurs possibles de ce rapport, comme aide à la sélection entre M_1 et M_2.

Remarque 1. Cette définition du facteur de Bayes correspond à la définition plus courante mais moins commode dans notre approche, $B_{12} = \frac{p(M_1|Y)}{p(M_2|Y)} / \frac{p(M_1)}{p(M_2)}$ où $p(M_i)$ et $p(M_i|Y)$ sont respectivement les probabilités a priori et a posteriori du modèle M_i, $i = 1, 2$.

2. Estimation non paramétrique d'un facteur de Bayes entre deux modèles à espace d'état

Elle repose sur des estimations indépendantes des deux termes du rapport B_{12}.

2.1. Estimation non paramétrique d'une vraisemblance marginale $p(Y)$ d'un modèle $M(\theta)$ de type (1)

Notation : $y_{1:t} = y_1, \ldots, y_t$.

Remarquons que $p(Y) = p(y_{1:T}) = p(y_1) \prod_{t=1}^{t=T-1} p(y_{t+1}|y_{1:t})$.

Une estimation non paramétrique de $p(y_{t+1}|y_{1:t})$ s'obtient facilement par $p_{t+1}^n(y_{t+1}|y_{1:t}) = \frac{1}{n} \sum_{i=1}^n K_{\Delta_n}^y (y_{t+1} - \tilde{y}_{t+1}^i)$, avec

- $K_{\Delta_n}^y(z) = \Pi_{j=1}^q \frac{1}{\delta_{n,j}} K^y(\frac{z_j}{\delta_{n,j}})$, où $K^y(\cdot)$ est un noyau positif de Parzen–Rosenblatt [2] et z_j la jième composante du vecteur z de dimension q. $\delta_{n,j}$ est le paramètre de fenêtre du noyau. Le choix de sa valeur résulte le plus souvent de règles empiriques. Dans le cas présent un choix pertinent dérivé des recommandations de Silverman [21] est :

  $$\delta_{n,j} = 1,06 \times \min\left(\sqrt{\mathrm{var}(\tilde{y}_{t,j})}, \frac{\mathrm{iqr}(\tilde{y}_{t,j})}{1.34}\right) \times n^{-1/(4+q)},$$

  où $\tilde{y}_{t,j}$ est le vecteur des n particules $(\tilde{y}_{t,j}^i,\ i = 1, \ldots, n)$ et $\mathrm{iqr}(\tilde{y}_{t,j})$ est l'intervalle inter-quartile des $(\tilde{y}_{t,j}^i)$. On notera $\Delta_n^q = \prod_{j=1}^q \delta_{n,j}$.
- $\tilde{y}_{t+1}^i$, $i = 1, \ldots, n$, images par le sous-modèle d'état f_t puis par le sous-modèle d'observation h_t du modèle (1), des n particules $(\tilde{x}_t^i, \tilde{\theta}_t^i)$, $i = 1, \ldots, n$, générées à l'instant t selon la densité conditionnelle de (x_t, θ_t) estimée par noyau

  $$p_t^n(x, \theta|y_{1:t}) = \frac{\sum_{i=1}^n K_{\Delta_n}^y(\tilde{y}_t^i - y_t) \times K_{\Delta_n}^x(\tilde{x}_t^i - x) \times K_{\Delta_n}^\theta(\tilde{\theta}_t^i - \theta)}{\sum_{i=1}^n K_{\Delta_n}^y(\tilde{y}_t^i - y_t)} \quad \text{(avec notations appropriées).}$$

  D'après le théorème 3.1 de Rossi et Vila [20], $\lim_{n\to\infty} \| p_t^n(x, \theta|y_{1:t}) - p_t(x, \theta|y_{1:t}) \|_{L1} = 0$ *p.s.*, sous les simples conditions $\lim_{n\to\infty} \frac{n\Delta_n^{q+d+p}}{\log n} = \infty$, $\lim_{n\to\infty}(\delta_{n,j}) = 0$ pour $K^y(\cdot)$, $K^x(\cdot)$, $K^\theta(\cdot)$ et $\Delta_n^q = \mathrm{o}(n^{-\alpha/2})$, $0 < \alpha < 1$.

Proposition 2.1. *Soit $p^n(Y) = p_1^n(y_1) \prod_{t=1}^{t=T-1} p_{t+1}^n(y_{t+1}|y_{1:t})$. Sous les conditions précédentes sur l'évolution des paramètres de fenêtre de $K^y(\cdot)$, $K^x(\cdot)$ et $K^\theta(\cdot)$ quand n tend vers l'infini, $p^n(Y)$ tend vers $p(Y)$ p.s.*

Démonstration. Cette convergence est la conséquence de la convergence p.s. de $p^n_{t+1}(y_{t+1}|y_{1:t})$ vers $p(y_{t+1}|y_{1:t})$, qui elle même s'obtient en utilisant l'approche de Devroye pour l'étude de la robustesse d'estimateurs à noyau (cf. [6] p. 46–49 et [20] p. 90–98). On peut ainsi montrer que lorsque n tend vers l'infini, $\frac{1}{n}\sum_{i=1}^n K^y_{\Delta_n}(y_{t+1}-\tilde{y}^i_{t+1})$ tend p.s. vers $\frac{1}{n}\sum_{i=1}^n K^y_{\Delta_n}(y_{t+1}-\breve{y}^i_{t+1})$ où les $(\breve{y}^i_{t+1})$ sont les particules qui résultent de l'application du système (1) à des *particules virtuelles* $(\breve{x}^i_t,\breve{\theta}^i_t)$ distribuées selon la véritable densité $p_t(x,\theta|y_{1:t})$ inconnue. Alors, d'après le théorème de Prakasa Rao [17] sur la convergence *p.s.* de l'estimateur d'une densité $p\in L_2$ associé à un noyau positif de Parzen–Rosenblatt en tout point de continuité de p, on a : $\frac{1}{n}\sum_{i=1}^n K^y_{\Delta_n}(y_{t+1}-\breve{y}^i_{t+1})$ et donc $\frac{1}{n}\sum_{i=1}^n K^y_{\Delta_n}(y_{t+1}-\tilde{y}^i_{t+1})$, tendent vers $p(y_{t+1}|y_{1:t})$ *p.s.* □

2.2. *Estimation convergente de B_{12}*

Soit $B^n_{12}=\frac{p^n_1(Y)}{p^n_2(Y)}$, où $p^n_1(Y)$ et $p^n_2(Y)$ sont les estimateurs respectifs des vraisemblances marginales des modèles M_1 et M_2, définis tels que dans la proposition précédente.

Théorème 2.2. *Sous les hypothèses de la Proposition* 2.1*, B^n_{12} tend vers B_{12} p.s. lorsque n tend vers l'infini.*

Démonstration. Conséquence immédiate de la Proposition 2.1. □

3. Comparaison avec les estimateurs MCMC d'un facteur de Bayes

Les estimateurs les plus performants à ce jour sont ceux développés pour l'estimation du rapport des constantes de normalisation de deux densités de probabilité, problème qui généralise comme on l'a vu celui de l'estimation d'un facteur de Bayes. Il s'agit notamment des méthodes de l'Importance Sampling (IS), du Bridge sampling (BS) et du Ratio Importance Sampling (RIS) (cf. Chen et Shao [3]).

Soit $\pi_i(\theta)$, $i=1,2$, deux densités (par rapport à une mesure commune), de même dimension (pour l'instant), et connue chacune à une constante de normalisation près, c_i : $\pi_i(\theta)=\frac{q_i(\theta)}{c_i}$, $\theta\in\Theta_i\subset\mathbb{R}^p$, et $\Theta_1\cap\Theta_2\neq\emptyset$. On veut estimer $r=c_1/c_2$.

La méthode BS, la plus pratiquée, est basée sur l'identité du Bridge Sampling :

$$r=\frac{c_1}{c_2}=\frac{\mathrm{E}_{\pi_2}[q_1(\theta)\alpha(\theta)]}{\mathrm{E}_{\pi_1}[q_2(\theta)\alpha(\theta)]} \tag{3}$$

où $\alpha(\theta)$ est une fonction arbitraire définie sur $\Theta_1\cap\Theta_2$, telle que $0<|\int_{\Theta_1\cap\Theta_2}\alpha(\theta)\pi_1(\theta)\pi_2(\theta)\,\mathrm{d}\theta|<\infty$ (immédiat).

La méthode BS estime r de façon convergente par

$$\hat{r}_{\mathrm{BS}}=\frac{\frac{1}{n_2}\sum_{j=1}^{n_2}q_1(\theta_{2j})\alpha(\theta_{2j})}{\frac{1}{n_1}\sum_{j=1}^{n_1}q_2(\theta_{1j})\alpha(\theta_{1j})}$$

où θ_{ij}, $j=1,\dots,n_i$, $i=1,2$, sont des réalisations de π_1 et π_2 respectivement, obtenues le plus souvent par un algorithme MCMC de type Metropolis-Hastings [13]. Un choix optimal pour la fonction $\alpha(\theta)$ a été proposé par Meng et Wong [16]. Le cas de deux densités π_1 et π_2 de dimensions différentes se traite de façon analogue après extension paramétrique adéquate de la densité de plus faible dimension (Chen et Shao [3]).

La transposition de la méthode BS à l'estimation d'un facteur de Bayes B_{12} conduit facilement à : $\pi_i(\theta)=:\frac{p_i(\theta,Y)}{p_i(Y)}=p_i(\theta|Y)$, $i=1,2$ et l'estimateur BS de B_{12} s'écrit

$$B^{\mathrm{BS}}_{12}=\frac{\frac{1}{n_2}\sum_{j=1}^{n_2}p_1(\theta_{2j},Y)\alpha(\theta_{2j})}{\frac{1}{n_1}\sum_{j=1}^{n_1}p_2(\theta_{1j},Y)\alpha(\theta_{1j})}.$$

Dans le cas de deux modèles à espace d'état, des réalisations θ_{ij} de $p_i(\theta|Y)$, $j=1,\dots,n_i$, $i=1,2$, peuvent être obtenues facilement comme particules d'un filtre à convolution pour n suffisamment grand. Mais $p_i(\theta,Y)=$

$p_i(Y|\theta)p_i(\theta) = p_i(\theta|Y)p_i(Y),\ i = 1, 2$, est inconnue, comme le sont la vraisemblance $p_i(Y|\theta)$ et bien sûr $p_i(Y)$. Cependant le choix suivant de $\alpha(\theta)$, $\alpha(\theta) = [p_1(\theta|Y)p_1(Y) + p_2(\theta|Y)p_2(Y)]^{-1}$ conduit facilement à :

$$B_{12}^{\mathrm{BS}} = \frac{\frac{1}{n_2}\sum_{i=1}^{n_2} p_1(\theta_{2i}|Y)[p_1(\theta_{2i}|Y) + \frac{1}{B_{12}}p_2(\theta_{2i}|Y)]^{-1}}{\frac{1}{n_1}\sum_{i=1}^{n_1} p_2(\theta_{1i}|Y)[B_{12}p_1(\theta_{1i}|Y) + p_2(\theta_{1i}|Y)]^{-1}}.$$

A partir d'estimations par filtrage à convolution des valeurs des fonctions de densité inconnues $p_i(\theta|Y),\ i = 1, 2$, en les valeurs de particules $\tilde{\theta}_{ij},\ j = 1, \ldots, n_i,\ i = 1, 2$, B_{12}^{BS} peut s'évaluer par itérations successives partant d'une estimation initiale B_{12}^0 de B_{12} inconnu. L'étude de la convergence de cet estimateur n'est pas aisée. Il est de plus d'une complexité et d'une variabilité supérieures à celles de l'estimateur non paramétrique convergent B_{12}^n proposé au §3. En effet, en plus de la dépendance vis à vis des T ensembles de n particules utilisées de $t = 1$ à $t = T$ pour les estimations des densités $p_i(\theta|Y)$, $i = 1, 2$ (dont dépend également l'estimateur B_{12}^n), l'estimateur B_{12}^{BS} dépend de trois autres facteurs :

- les nombres n_1 et n_2 de tirages $\theta_{ij},\ j = 1, \ldots, n_i,\ i = 1, 2$ selon $p_i^n(\theta|Y), i = 1, 2$, estimations par filtrage des vraies densités $p_i(\theta|Y),\ i = 1, 2$, inconnues. Ces tirages sont utilisés pour l' approximation par la Loi des Grands Nombres de l'identité du Bridge Sampling (3).
- le choix de la valeur initiale B_{12}^0 à introduire dans le dénominateur de B_{12}^{BS} en place de B_{12} inconnu, pour initier le processus itératif.
- le nombre d'itérations requises pour atteindre un point fixe stable, dont la validité comme approximation de la valeur du facteur de Bayes reste conditionnée par les sources de variabilité précédentes.

Par ailleurs, les comparaisons effectuées sur différents modèles à espace d'état ont confirmé la supériorité de l'estimateur non paramétrique B_{12}^n (stabilité des estimations et temps de calcul).

Les autres méthodes d'estimations de facteur de Bayes par simulations MCMC, comme la méthode IS (cas particulier de la méthode BS pour $\alpha(\theta) = 1/p_2(\theta, Y)$), la méthode RIS, et de manière générale les méthodes pour lesquelles la connaissance des vraisemblances $p_i(Y|\theta)$ et donc celles des densités $p_i(\theta, Y)$, sont nécessaires et non substituables (comme on a vu qu'elles le sont pour la méthode BS), ne peuvent pas être appliquées aux modèles non linéaires à espace d'état de type (1).

4. Conclusion

Dans le cas de modèles non linéaires à espace d'état, la non disponibilité des fonctions de vraisemblance rend l'utilisation des méthodes classiques d'estimation de facteurs de Bayes par simulations MCMC, inadaptées. Un estimateur non paramétrique convergent des vraisemblances marginales de ces modèles, est par contre facilement disponible à partir du filtre à convolution de particules avec rééchantillonnage. Un estimateur convergent du facteur de Bayes s'en déduit immédiatement. Les essais numériques effectués sur des modèles à espace d'état de différentes complexités, confirment l'intérêt pratique de cet estimateur. De plus, l'intérêt de cette démarche dans le cas de modèles probabilistes généraux, à fonctions de vraisemblance accessibles, est également à considérer.

Références

[1] F. Bartolucci, L. Scaccia, A. Mira, Efficient Bayes factor estimation from the reversible jump output, Biometrika 93 (1) (2006) 41–52.
[2] D. Bosq, J.P. Lecoutre, Théorie de l'estimation fonctionnelle, Economica, Paris, 1987.
[3] M.H. Chen, Q.M. Shao, Estimating ratios of normalizing constants for densities with different dimensions, Statist. Sinica 7 (1997) 607–630.
[4] A.P. Dawid, Statistical theory: The prequential approach, J. Roy. Statist. Soc. Ser. A 147 (1984) 278–292.
[5] P. Del Moral, Feynman–Kac Formulae. Genealogical and Interacting Particle Systems with Applications, Springer-Verlag, New York, 2004.
[6] L. Devroye, A Course in Density Estimation, Birkhäuser, Boston, 1987.
[7] A. Doucet, N. de Freitas, N. Gordon, Sequential Monte Carlo Methods in Practice, Statistics for Engineering and Information Science, Springer, New York, 2001.
[8] A.E. Gelfand, A.F.M. Smith, Sampling-based approach to calculating marginal densities, J. Amer. Statist. Assoc. 85 (1990) 398–409.
[9] A.E. Gelman, X.L. Meng, Computing normalizing constants: from importance sampling to bridge sampling to path sampling, Statist. Sci. 13 (1998) 163–185.
[10] T. Gneiting, A.E. Raftery, Strictly proper scoring rules, prediction, and estimation, J. Amer. Statist. Assoc. 102 (2007) 359–378.

[11] P.J. Green, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, Biometrika 82 (1995) 711–732.
[12] C. Han, B.P. Carlin, MCMC methods for computing Bayes factors: a comparison review, J. Amer. Statist. Assoc. 96 (2001) 1122–1132.
[13] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1970) 97–109.
[14] H. Jeffreys, Theory of Probability, Oxford University Press, Oxford, 1961.
[15] R.E. Kass, A.E. Raftery, Bayes factors, J. Amer. Statist. Assoc. 90 (1995) 773–795.
[16] X.L. Meng, W.H. Wong, Simulating ratios of normalizing constants via a simple identity: a theoretical exploration, Statist. Sinica 6 (1996) 831–860.
[17] B.L.S. Prakasa Rao, Nonparametric Functional Estimation, Academic Press, Orlando, 1983.
[18] V. Rossi, Filtrage non linéaire par noyaux de convolution : application à un procédé de dépollution biologique, Thèse en science, Ecole Nationale Supérieure Agronomique de Montpellier, 2004.
[19] V. Rossi, J.P. Vila, Approche non paramétrique du filtrage de système non linéaire à temps discret et à paramètres inconnus, C. R. Acad. Sci. Paris, Ser. I 340 (2005) 759–764.
[20] V. Rossi, J.P. Vila, Nonlinear filtering in discrete time: a particle convolution approach, Pub. Inst. Stat. Univ. Paris L (3) (2006) 71–102.
[21] B.W. Silverman, Density Estimation, Chapman and Hall, London, 1986.
[22] H. Tanizaki, Nonlinear Filters, Springer-Verlag, New York, 1993.
[23] L. Tierney, Markov chains for exploring posterior distributions (with discussion), Ann. Statist. 22 (1994) 1701–1762.