

Analyse bayésienne de modèles markoviens d'évolution de ressources naturelles

Fabien Campillo* — Rivo Rakotozafy** — Vivien Rossi***

* INRIA/IRISA
Campus de Beaulieu
35042 Rennes
FRANCE
Fabien.Campillo@inria.fr

** Faculté des Sciences
Université de Fianarantsoa
301 Fianarantsoa
MADAGASCAR
rrakotoz@mail.univ-fianar.mg

*** INRIA/IRISA
Campus de Beaulieu
35042 Rennes
FRANCE
Vivien.Rossi@inria.fr

RÉSUMÉ. On applique deux méthodes de Monte Carlo à des modèles à espace d'état avec paramètres inconnus. La première est un algorithme de Monte Carlo par chaîne de Markov. La seconde est le filtrage particulaire. Nous comparons ces méthodes sur un modèle d'évolution de la biomasse dans une pêcherie.

ABSTRACT. One applies two Monte Carlo methods to state space models with unknown parameters. The first one is a Monte Carlo Markov Chain algorithm. The second one is the particle filtering. We compare these methods applied to a biomass evolution model for fisheries.

MOTS-CLÉS : Inférence bayésienne, Monte Carlo par chaînes de Markov, filtrage particulaire

KEYWORDS : Bayesian inference, Markov chain Monte Carlo, particle filtering

L'inférence bayésienne des ressources naturelles fait depuis quelques années appel aux modèles non linéaires à espace d'état. Associés aux méthodes de Monte Carlo par chaînes de Markov (MCMC) [6], ces modèles ont rencontré un certain succès car dans ce domaine les séries d'observations sont courtes avec une fréquence d'observation longue. Il est alors donc possible de faire appel à des méthodes non séquentielles, des méthodes "batch", comme les MCMC. Ces méthodes sont itératives et une des difficultés de leurs mises en œuvre réside dans le critère d'arrêt.

Ce domaine d'application est clairement bayésien. En effet, on dispose de peu de données et celles-ci sont très bruitées ("poor-data"). Il est donc nécessaire de faire appel aux compétences des spécialistes du domaine afin, d'une part, de proposer des systèmes dynamiques modélisant l'évolution des phénomènes considérés et, d'autre part, de déterminer les lois a priori correspondant aux paramètres inconnus du modèle.

Il est intéressant de faire appel à des méthodes de filtrage particulaire (PF) [5]. Dans leurs versions de base, ces méthodes ne sont pas itératives et sont séquentielles, elles traitent les observations au fur et à mesure de leur arrivée. Notons que cet aspect séquentiel n'a pas d'intérêt dans les applications concernées ici.

Bien que fondées sur le même principe de simulation de Monte Carlo, les méthodes MCMC et PF sont très différentes de par leur nature (théorème ergodique *vs* loi des grands nombres) ainsi que par leur mise en œuvre (itérative *vs* non itérative).

Nous présentons les modèles non linéaires à espace d'état qui sont en fait des modèles de Markov cachés à espace d'état général. Après avoir décrit les méthodes MCMC et PF, nous développons le problème étudié.

1. Modèles de Markov cachés

Considérons un modèle de Markov caché $(x_{0:T}, y_{1:T})$ à valeurs dans $\mathbb{R}^n \times \mathbb{R}^d$ et dépendant d'un paramètre θ inconnu. Le processus d'état non observé x_t est une chaîne de Markov homogène :

$$\text{de loi initiale} \quad \mu^\vartheta(dx) \stackrel{\text{déf}}{=} \mathbb{P}(x_0 \in dx | \theta = \vartheta),$$

$$\text{de noyau de transition} \quad Q^\vartheta(x, dx') \stackrel{\text{déf}}{=} \mathbb{P}(x_t \in dx' | \theta = \vartheta, x_{t-1} = x).$$

Le processus d'observation y_t est une suite de variables indépendantes conditionnellement à l'état x_t et au paramètre θ , et la loi conditionnelle de y_t ne dépend que de x_t et θ , i.e.

$$\mathbb{P}(y_{1:T} \in dy_{1:T} | \theta = \vartheta, x_{0:T} = x_{0:T}) = \prod_{t=1}^T \mathbb{P}(y_t \in dy_t | \theta = \vartheta, x_t = x_t).$$

On suppose de plus que la loi conditionnelle de y_t sachant x_t est absolument continue par rapport à la mesure de Lebesgue et on pose :

$$\mathbb{P}_\theta(y_t \in dy | \theta = \vartheta, x_t = x) = \psi_t^\vartheta(x, y) dy.$$

Les éléments du modèle, à savoir μ^ϑ , Q^ϑ et ψ_t^ϑ , dépendent d'un paramètre inconnu θ à valeurs dans \mathbb{R}^p . Ce paramètre est supposé indépendant de tous les autres termes et on se donne une loi a priori

$$\nu(d\vartheta) \stackrel{\text{déf}}{=} \mathbb{P}(\theta \in d\vartheta).$$

2. Méthodes de Monte Carlo par chaînes de Markov

Le but est de proposer, à partir de μ^ϑ , Q^ϑ , $\psi_{1:T}^\vartheta$ et ν , un algorithme de construction d'une chaîne de Markov ergodique $(x_{0:T}^{(k)}, \theta^{(k)})_{k \in \mathbb{N}}$ dont la loi invariante est la loi a posteriori

$$\pi_T(dx_{0:T}, d\vartheta | y_{1:T}) = \mathbb{P}(x_{0:T} \in dx_{0:T}, \theta \in d\vartheta | y_{1:T} = y_{1:T}). \quad (1)$$

Ainsi le théorème ergodique implique que la mesure empirique associée aux différents états explorés par la chaîne converge vers la loi cible π_T .

Nous présentons l'échantillonneur de Metropolis-Hastings hybride. Cette méthode fait appel aux lois marginales conditionnelles :

$$\pi_{t|-t}(dx_t | x_{-t}, \vartheta, y_{1:T}) \stackrel{\text{déf}}{=} \mathbb{P}(x_t \in dx_t | x_{-t} = x_{-t}, \theta = \vartheta, y_{1:T} = y_{1:T}).$$

Partant d'une configuration initiale $(x_{0:T}^{(0)}, \theta^{(0)})$, l'itération $k \rightarrow k+1$ consiste à choisir un indice t au hasard et de mettre à jour $x_t^{(k)}$ en deux étapes : On propose une nouvelle configuration selon une loi de proposition

$$\tilde{x}_t \sim \pi_{t|-t}^{\text{prop}}(dx_t | x_{-t}^{(k)}, \theta^{(k)}, y_{1:T}).$$

On pose $x_t^{(k+1)} = \tilde{x}_t$ avec probabilité

$$\alpha \stackrel{\text{déf}}{=} \frac{f_t(\tilde{x}_t)}{f_t(x_t^{(k)})} \wedge 1 \quad \text{où} \quad f_t(x_t) \stackrel{\text{déf}}{=} \frac{\pi_{t|-t}(dx_t | x_{-t}^{(k)}, \theta^{(k)}, y_{1:T})}{\pi_{t|-t}^{\text{prop}}(dx_t | x_{-t}^{(k)}, \theta^{(k)}, y_{1:T})},$$

et $x_t^{(k+1)} = x_t^{(k)}$ avec probabilité $1 - \alpha$. Il existe des lois marginales conditionnelles de proposition $\pi_{t|-t}^{\text{prop}}$ adaptées au cas étudié ici [1]. La mise à jour du paramètre θ se fait de façon similaire.

3. Méthodes de Monte Carlo séquentielles

Le contexte est ici légèrement différent, le but est de calculer séquentiellement une approximation particulière de

$$\rho_t(dx, d\vartheta | y_{1:t}) = \mathbb{P}(x_t \in dx, \theta \in d\vartheta | y_{1:t} = y_{1:t}). \quad (2)$$

Il suffit pour cela d'introduire un processus θ_t dont la dynamique est

$$\theta_t = \theta_{t-1}, \quad \theta_0 \sim \nu. \quad (3)$$

Le couple $(x_t, \theta_t)_{t=0:T}$ est markovien. On peut écrire le filtre bayésien optimal et séquentiel qui donne de façon récursive l'expression exacte de

$$\rho_t(dx, d\vartheta|y_{1:t}) = \mathbb{P}(x_t \in dx, \theta_t \in d\vartheta|y_{1:t} = y_{1:t}).$$

Cette expression n'est pas directement accessible, il est nécessaire de faire appel à une technique d'approximation. La technique rencontrant le plus de succès à l'heure actuelle est le filtrage particulaire [5].

En pratique, la partie "statique" (3) de l'équation d'état correspondant au paramètre est traitée à l'aide d'une approximation particulaire à noyau [3]. L'autre partie de l'équation d'état peut être traitée par la même technique ou par un simple filtre "bootstrap".

Partant d'une population initiale $(x_0^i, \theta_0^i)_{i=1:N}$ de N particules échantillonnées selon $\mu^\vartheta(dx) \times \nu(d\vartheta)$, l'itération $t - 1 \rightarrow t$ se fait en deux étapes :

mutation : on pose $x_{t-}^i \sim Q^{\theta_{t-}^i}(x_{t-1}^i, dx')$ et $\theta_{t-}^i \leftarrow \theta_{t-1}^i$ pour $i = 1 : N$.

redistribution : on pose $\omega_t^i \propto \psi^{\theta_{t-}^i}(x_{t-}^i, y_t)$ avec $\sum_{i=1}^N \omega_t^i = 1$ et

$$\begin{bmatrix} x_t^i \\ \theta_t^i \end{bmatrix} \sim \sum_{i=1}^N \omega_t^i \mathcal{N} \left(\begin{bmatrix} x_{t-}^i \\ \theta_{t-}^i \end{bmatrix}, H_t^N \right) \quad \text{indépendamment pour } i = 1 : N.$$

La matrice H_t^N est notamment fonction de la covariance empirique des particules prédites à la mutation [3].

4. Exemple : un modèle de pêche

On présente ici un exemple déjà proposé dans [1]. Les techniques d'estimation des stocks de pêche visent à évaluer l'impact de différents scénarios d'exploitation et de capture sur l'évolution de l'abondance d'un stock de poissons. Cette abondance s'exprime en poids total de poissons vulnérables à la pêche (biomasse). L'accroissement de cette biomasse intègre l'augmentation du poids moyen de chaque individu ainsi que le recrutement, c'est à dire l'arrivée chaque année de nouvelles générations de poissons. À l'inverse, cette biomasse décroît du fait de la capture (mortalité par pêche) ou d'autres causes (mortalité naturelle : prédation, maladies, etc.).

Pour chaque année $t = 1 \dots T$ d'une période de T années, B_t désigne la biomasse en début d'année, C_t la biomasse capturée en cours d'année et I_t un indice d'abondance du stock. Cet indice, mesuré chaque année, est la capture par unité d'effort (CPUE). Les

séries $C_{1:T}$ et $I_{1:T}$ sont données. Le but est de déterminer, au sein d'une classe de modèles, lequel s'ajuste le mieux aux observations.

À la suite des travaux de Meyer et Millar [9], on adopte une formulation à espace d'états non linéaire d'un modèle aux différences avec retard (delay difference model) de la dynamique de la biomasse et de l'observation de l'indice d'abondance de cette biomasse. Ce modèle a été introduit par Deriso [4] et généralisé par Schnute [10] (voir Hilborn et Walters [7] pour une présentation générale). Dans la représentation à espace d'états, les indices $I_{1:T}$ seront considérées comme les observations, $B_{1:T}$ comme les états (non observés) et $C_{1:T}$ comme les entrées.

Les paramètres du modèle sont ajustés aux données par la procédure de Metropolis-Hastings hybride décrite Section 2.

À l'année t la biomasse B_t est :

$$B_t \stackrel{\text{def}}{=} \sum_{a \geq k} w_a N_{a,t} \quad (4)$$

où $N_{a,t}$ est la taille de la population d'âge a à l'année t , w_a est le poids moyen d'un individu à l'âge a , k est l'âge de recrutement (âge auquel on estime que les poissons viennent s'ajouter à la biomasse). Le premier terme $R_t = w_k N_{k,t}$ dans la somme (4) est le recrutement de l'année t . On suppose les individus de la biomasse uniformément vulnérables à la mortalité naturelle et à la capture. L'évolution annuelle de chaque cohorte (population d'individus de la même classe d'âge) est donc décrite par :

$$N_{a,t} = s_{t-1} N_{a-1,t-1}, \quad (5)$$

où s_{t-1} est le taux de survie à la mortalité naturelle et à la pêche. L'évolution de w_a est donnée selon la relation classique de croissance du poids selon l'âge : $w_a = w_\infty (1 - e^{-\kappa(a-a_0)})$ pour $a \geq a_0$ où a_0 est l'âge auquel $w_{a_0} = 0$ et $\kappa > 0$. Cela conduit à :

$$w_a = (1 + \rho) w_{a-1} - \rho w_{a-2} \quad (6)$$

où $\rho = e^{-\kappa} \in (0, 1)$ est le taux de croissance. Les expressions (5) et (6) dans (4) conduisent à la récurrence :

$$B_t = (1 + \rho) s_{t-1} B_{t-1} - \rho s_{t-1} s_{t-2} B_{t-2} + R_t - \rho s_{t-1} w R_{t-1} \quad (7)$$

avec $w = w_{k-1}/w_k \in (0, 1)$.

On suppose que les causes de mortalité naturelle et de mortalité par capture agissent de façon indépendante sur l'ensemble des poissons, ainsi $s_t = s_t^M s_t^F$ où le taux de survie à la mortalité naturelle s_t^M est supposé constant ($s^M = e^{-M}$ avec $M > 0$), et où le taux de survie à la capture est $s_t^F = \frac{B_t - C_t}{B_t}$. Le recrutement est ensuite supposé constant, i.e. $R = R_t$. L'équation (7) devient donc :

$$B_t = (1 + \rho) e^{-M} \frac{B_{t-1} - C_{t-1}}{B_{t-1}} B_{t-1} - \rho e^{-2M} \frac{B_{t-1} - C_{t-1}}{B_{t-1}} \frac{B_{t-2} - C_{t-2}}{B_{t-2}} B_{t-2} \\ + R \left(1 - \rho e^{-M} w \frac{B_{t-1} - C_{t-1}}{B_{t-1}} \right).$$

Avant la première année $t = 1$, les captures sont supposées nulles $C_{t < 1} = 0$. La biomasse est supposée être à son équilibre jusqu'à l'année 1, i.e. $B_{t \leq 1} = K$ où K est un paramètre inconnu (la biomasse vierge). La biomasse à l'année $t = 2$ est donc :

$$B_2 = (1 + \rho - \rho e^{-M}) e^{-M} (B_1 - C_1) + R (1 - \rho e^{-M} w \frac{B_1 - C_1}{B_1}).$$

L'équation d'observation non bruitée est de la forme $I_t = q B_t$ pour $t = 1 \cdots T$, où I_t est un indice de biomasse relative et q est un coefficient de "capturabilité". Les processus de bruit d'état et d'observation sont supposés multiplicatifs et lognormaux, ce qui respecte mieux la nature du problème :

$$B_t = F_t(R, K, B_{t-1}, B_{t-2}) \times e^{\sigma_w W_t}, \quad (8)$$

$$I_t = q B_t \times e^{\sigma_v V_t}, \quad (9)$$

pour $t = 1 \cdots T$, $W_{1:T}$ et $V_{1:T}$ sont des bruits blancs gaussiens $N(0, 1)$. L'expression exacte de ces fonctions et ce modèle sont détaillés dans [1].

Pour revenir au modèle initial on pose $\theta \stackrel{\text{def}}{=} (K, R, q, \sigma_w^2, \sigma_v^2)$ et $x_t \stackrel{\text{def}}{=} \begin{pmatrix} B_t \\ B_{t-1} \end{pmatrix}$.

5. Conclusions et perspectives

Nous appliquons les algorithmes précédents à des données simulées comparables à celles utilisées dans Meyer et Millar [9]. Ces derniers font eux mêmes référence aux données de Pella–Tomlinson sur le thon albacore 1934–1967. Dans ces premiers essais, les variances σ_w^2 et σ_v^2 sont supposées connues.

Un résultat obtenu par la méthode de Metropolis–Hastings hybride est présenté dans la Figure 1. Il faut noter que l'algorithme converge très lentement : la chaîne tend très lentement vers son régime stationnaire. Dans la Figure 2, nous comparons le filtre particulaire et la méthode MCMC. Les résultats sont comparables, on note toutefois que le filtre particulaire présente un peu plus de variance. Ceci est cohérent avec le fait que le filtre particulaire donne une approximation de $\text{loi}(x_t, \theta | y_{1:t})$ alors que la méthode MCMC donne une approximation de $\text{loi}(x_t, \theta | y_{1:T})$.

En pratique, le critère d'arrêt de la méthode MCMC est difficile à mettre en place et la méthode MCMC nécessite beaucoup plus de temps calcul que la méthode PF. Afin d'accélérer la vitesse de convergence, il nous semble intéressant de proposer des méthodes MCMC en parallèle et en interaction [2]. De même, il est possible d'améliorer la partie mutation de la méthode PF et de développer un algorithme de lissage. Ces extensions sont en cours d'étude.

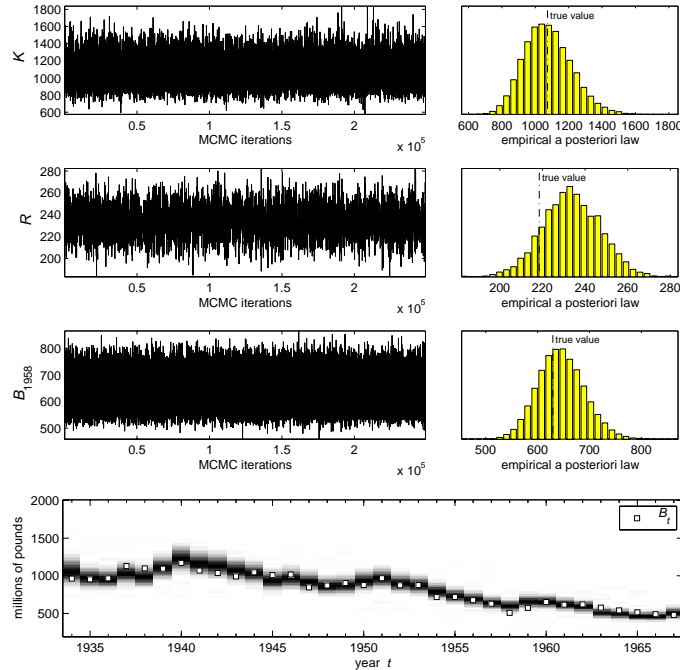


Figure 1. Méthode MCMC : • Les figures du haut à gauche présente les trajectoires des chaînes $k \mapsto (K^{(k)}, R^{(k)}, B_{1958}^{(k)})$, elles permettent de constater que l’algorithme MCMC a atteint son régime d’équilibre. • Les figures du haut à droite présentent les lois a posteriori empiriques correspondant respectivement à K , R , B_{1958} . • La figure du bas présente l’évolution de la biomasse $B_{1:T}$ ainsi, que pour chaque année t , la représentation en niveaux de gris de la loi a posteriori empirique associée.

6. Bibliographie

- [1] Fabien Campillo and Rivo Rakotozafy. MCMC for nonlinear/non-Gaussian state-space models, Application to fishery stock assessment. In *CAR’04*, Hammamet, Tunisia, 2004.
- [2] Fabien Campillo and Vivien Rossi. Parallel interacting MCMC’s for a class of hidden Markov models. En préparation.
- [3] Fabien Campillo and Vivien Rossi. Convolution particle filters for parameter estimation in general state-space models. Research Report 5939, INRIA, 06 2006.
- [4] Richard B. Deriso. Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 37 :268–282, 1980.

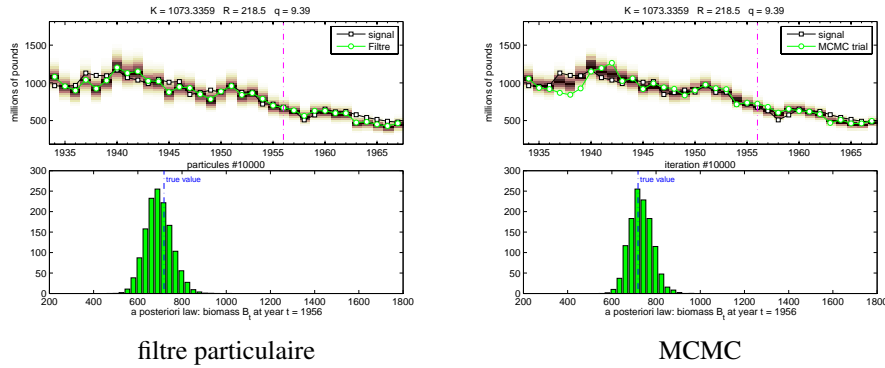


Figure 2. Comparaison filtrage particulaire (gauche) et MCMC (droite). En haut l'évolution de la biomasse $B_{1:T}$ ainsi, que pour chaque année t , la représentation en niveaux de gris de la loi a posteriori empirique associée. En bas, la densité empirique à l'année $t = 1956$.

- [5] Arnaud Doucet, Nando de Freitas, and Neil J. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.
- [6] Walter R. Gilks, Sylvia Richardson, and David J. Spiegelhalter, editors. *Markov Chain Monte Carlo in practice*. Chapman & Hall, London, 1995.
- [7] R. Hilborn and C. Walters. *Quantitative Fisheries Stock Assessment : Choice, Dynamics and Uncertainty*. Chapman and Hall, 1992.
- [8] D.K. Kimura, J.W. Balsiger, and D.H. Ito. Kalman filtering the delay difference equation : practical approaches and simulations. *Fishery Bulletin*, 94(4) :678–691, 1996.
- [9] Renate Meyer and Russell B. Millar. Bayesian stock assessment using a state-space implementation of the delay difference model. *Canadian Journal of Fisheries and Aquatic Sciences*, 56 :37–52, 1999.
- [10] Jon T. Schnute. A general theory for the analysis of catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences*, 42 :414–429, 1985.
- [11] H. Tanizaki. Non linear and non-Gaussian state–space modeling with Monte Carlo techniques : A survey and comparative study. In C.R. Rao and D.N. Shanbhag, editors, *Handbook of Statistics 21 : Stochastic Processes : Modeling and Simulation*, pages 871–929. North-Holland, 2003.