

REVUE DE STATISTIQUE APPLIQUÉE

BELKHEIR ESSEBBAR

Sur un modèle de série temporelle à conjoncture dans le tourisme

Revue de statistique appliquée, tome 52, n° 4 (2004), p. 55-70

http://www.numdam.org/item?id=RSA_2004__52_4_55_0

© Société française de statistique, 2004, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

SUR UN MODÈLE DE SÉRIE TEMPORELLE À CONJONCTURE DANS LE TOURISME

Belkheir ESSEBBAR

*Département de Mathématiques et Informatique
Faculté des Sciences de Rabat, Maroc
essebbar@fsr.ac.ma*

RÉSUMÉ

Dans ce travail, on propose un modèle de série temporelle construit à partir d'un processus de conjoncture dans le secteur du tourisme. Une analyse de ce modèle est faite avec une application aux données réelles du tourisme au Maroc. En particulier, une tentative de prévision à moyen et long terme, pour les arrivées des touristes internationaux aux postes frontières est donnée.

Mots-clés : *Série temporelle, Processus de conjoncture, Chaîne de Markov, Estimation, EMV, Moindres carrés, Problème inverse, MPM, Restauration, Prévision à long terme, Données touristiques.*

ABSTRACT

This work deals with a model of time series with conjuncture in Tourism. An analysis of the model is done as well as an application to tourism data. A tentative long term forecastings of the international demand Tourism in Morocco is given.

Keywords : *Time series, Conjunction process, Markov Chain, Estimation, MLE, Minimum squares, Inverse problem, MPM, Smoothing, Long term Forecasting, Tourism data.*

1. Introduction

La plupart des modèles utilisés pour les données touristiques sont basés sur les modèles ARMA (cf. Box et Jenkins (1976)) ou des modèles qui englobent diverses variables économiques (cf. Witt et Witt (1995)). Tous ces modèles traitent de la prévision à court terme (pas plus d'un an).

Les prévisions moyen et long terme, dans le secteur du tourisme et les secteurs connexes qui lui sont liés restent très aléatoire et incertaines. Cela est dû en partie à la difficulté de prévoir des événements de crise qui inversent les tendances de la conjoncture. Le dernier exemple en date est celui des événements du 11 Septembre 2001 qui frappent durement le secteur du tourisme.

En réalité, il y a plusieurs processus qui interagissent d'une façon dépendante ou indépendante entre eux : un processus stochastique (X_t) lié à une trajectoire (série

temporelle) et décrivant un phénomène dans le tourisme, tel le flux des touristes; un processus de conjoncture qui traduit la tendance dans le secteur. Ce dernier dépend de plusieurs sous-processus qui traduisent les conjonctures : locale, régionale, mondiale, etc. Ceci explique, en partie, la difficulté de la modélisation de certains phénomènes dans le tourisme, comme les arrivées aux postes frontières, les nuitées dans les hôtels classés, etc.

Un exemple d'un tel processus de conjoncture, (Y_t) , est celui qui décrit les états de conjoncture favorable ($Y_t = 1$) et conjoncture défavorable ($Y_t = 0$) à l'instant t . C'est un processus binaire à deux états 0 et 1. Ce modèle est connu sous le nom de Switching model ou modèle à double régime; on peut consulter Hamilton (1990, 1994) pour des applications dans la finance.

Ce travail sera organisé comme suit. Dans la section 2, on donne le modèle dont il sera question dans ce travail. Ainsi, on développe un modèle pour le processus de conjoncture (Y_t) dont on donne les caractéristiques essentielles. On construira par la suite un modèle pour le processus (X_t) . Dans la section 3, on aborde le problème statistique d'estimation des paramètres impliqués dans le modèle. On distinguera deux cas selon que le processus de conjoncture est observable ou non. Dans ce dernier cas on donne un algorithme récursif pour la restauration du processus de conjoncture (Y_t) . La section 4 sera consacrée à l'étude numérique par simulation et à l'application du modèle aux données réelles des Arrivées des touristes au Maroc.

2. Les Modèles

2.1. Modèle pour le processus (Y_t)

Pour le processus de conjoncture (Y_t) , considérons le modèle suivant. Soit (Y_t) une chaîne de Markov à deux états 0 et 1, homogène, de probabilités de transition $p_{ij} = P(Y_{t+1} = j | Y_t = i)$, Y_0 étant le point initial de la chaîne que l'on suppose donné. On posera par la suite pour les n premières observations de la chaîne

$$N_{ij} = \sum_{t=1}^n 1_{\{Y_{t-1}=i, Y_t=j\}}; \quad i, j \in \{0, 1\},$$

qui sont les statistiques qui comptent le nombre de transitions de l'état i à l'état j . Pour $Y = (Y_0, Y_1, \dots, Y_n)$ les n premières observations de la chaîne (Y_t) , la vraisemblance s'écrit

$$L(y) = \prod_{i,j} p_{ij}^{N_{ij}}; \quad y \in \{0, 1\}^n. \quad (1)$$

On déduit aisément l'estimateur du maximum de vraisemblance

$$\hat{p}_{01} = \frac{N_{01}}{N_{01} + N_{00}}; \quad \hat{p}_{10} = \frac{N_{10}}{N_{10} + N_{11}}.$$

La description des trajectoires d'un processus stochastique de conjoncture à deux états peut s'effectuer à l'aide des durées de séjour, D^i , dans l'état i , $i \in \{0, 1\}$. La loi de D^1 (resp. D^0) est une loi géométrique de paramètre p_{10} (resp. p_{01}).

On définit aussi la notion de persistance de conjoncture au k -ème instant de l'état i dans une séquence de i comme étant la probabilité que le séjour dans l'état i dure strictement plus de k instants, sachant qu'il a déjà duré k instants. Pour $i \in \{0, 1\}$, elle est notée par

$$q_k^i = P(Y_{t+k+1} = i | Y_t = 1 - i, Y_{t+1} = i, \dots, Y_{t+k} = i); \quad k \geq 1, t \geq 1.$$

Les persurances définies ci-dessus sont des caractéristiques importantes d'un modèle de série temporelle binaire et leur comportement (monotonie ou non, convergence,...) conditionne beaucoup son intérêt pour les applications. La persistance pour le modèle (Y_t) considéré est constante et il vaut $q_k^i = p_{ii}$.

Pour des applications autres que économique, on peut citer par exemple la climatologie (Essebbar et Le Breton (1985,1987)).

2.2. Modèle pour le processus (X_t)

Étant donné le processus binaire de conjoncture (Y_t), on construit le processus (X_t) comme suit. Considérons X_0 fixé et ($\epsilon_t; t \geq 1$) une suite de bruits blancs (variables aléatoires *i.i.d.* de loi $N(0, \sigma^2)$). On pose

$$X_t = (aY_t + b(1 - Y_t))X_{t-1} + \epsilon_t; \quad t \geq 1, \tag{2}$$

où a et b sont des réels tels que $a > 1$ et $0 < b < 1$. On voit que ce modèle est conditionné par (Y_t) et que le changement de tendances est expliqué par les valeurs a (cas de conjoncture favorable) et b (cas de conjoncture défavorable).

Il faut remarquer que ce processus est un processus autorégressif d'ordre 1 (AR(1)) avec «switching», qu'on notera SAR(1) (cf. Hamilton (1990, 1994)). Le SAR(1) diffère du modèle classique AR(1) car le coefficient de X_{t-1} est aléatoire.

Par construction, le SAR(1) modélise la tendance de la série temporelle en question, mais pas sa composante saisonnière. Ce modèle suppose alors que les données sont traitées au préalable en enlevant la partie saisonnière.

3. Estimation

Dans cette partie, on s'intéresse à la partie statistique d'estimation des paramètres intervenant dans le modèle SAR(1) donné par (2). On examinera les deux cas selon que le processus de conjoncture (Y_t) est observable ou pas.

3.1. Cas de conjoncture observable

Dans cette partie, On supposera que l'on dispose des observations $X = (X_0, \dots, X_n)$ et $Y = (Y_0, \dots, Y_n)$ et on se propose de faire l'estimation séparée des paramètres a, b, σ^2, p_{01} et p_{10} en deux temps.

Estimation de a et b : On utilise la méthode des moindres carrés. Soit le modèle SAR(1) donné par (2), *i.e.*

$$X_t = (aY_t + b(1 - Y_t))X_{t-1} + \epsilon_t; \quad t \geq 1.$$

Comme les ϵ_t s'écrivent

$$\epsilon_t = X_t - (aY_t + b(1 - Y_t))X_{t-1}; \quad t = 1, \dots, n.$$

Alors, on peut écrire la somme des carrés des ϵ_t :

$$S(a, b) = \sum_{t=1}^n \epsilon_t^2 = \sum_{t=1}^n \{X_t - (aY_t + b(1 - Y_t))X_{t-1}\}^2.$$

L'estimation de a et b par la méthode des moindres carrés consiste à trouver la solution au problème de minimisation suivant

$$(\hat{a}, \hat{b}) = \arg \min_{(a, b)} S(a, b).$$

On a le lemme suivant.

LEMME 1. – *L'estimateur par la méthode des moindres carrés est donné par*

$$\hat{a} = \frac{\sum_{t=1}^n Y_t X_{t-1} X_t}{\sum_{t=1}^n Y_t X_{t-1}^2}; \quad \hat{b} = \frac{\sum_{t=1}^n (1 - Y_t) X_{t-1} X_t}{\sum_{t=1}^n (1 - Y_t) X_{t-1}^2}.$$

Preuve. – Trouvons le point qui annule les dérivées de S . On a

$$\begin{aligned} \frac{\partial}{\partial a} S(a, b) &= -2 \sum_{t=1}^n \{X_t - (aY_t + b(1 - Y_t))X_{t-1}\} Y_t X_{t-1}, \\ \frac{\partial}{\partial b} S(a, b) &= -2 \sum_{t=1}^n \{X_t - (aY_t + b(1 - Y_t))X_{t-1}\} (1 - Y_t) X_{t-1}. \end{aligned}$$

Compte tenu de ce que $Y_t(1 - Y_t) = 0$, $Y_t^2 = Y_t$, $(1 - Y_t)^2 = 1 - Y_t$ puis que $Y_t = 0$ ou 1, on en déduit la solution exprimée par le lemme 1. On vérifie bien que la matrice des dérivées secondes est bien de déterminant positif et que $(\partial^2/\partial a^2)S$ et $(\partial^2/\partial b^2)S$ sont positifs. \square

Estimation de σ^2 , p_{01} et p_{10} : On suppose que les paramètres a et b sont fixes et on estime σ^2 , p_{01} et p_{10} par la méthode du maximum de vraisemblance (EMV). La vraisemblance associée aux observations X et Y est donnée par

$$L(x, y) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^n (x_t - (ay_t + b(1 - y_t))x_{t-1})^2\right\} L(y), \quad (3)$$

où $L(y)$ est donnée par (1), *i.e.*

$$L(y) = \prod_{i,j} p_{ij}^{N_{ij}}; \quad y \in \{0, 1\}^n,$$

avec $p_{00} = 1 - p_{01}$ et $p_{11} = 1 - p_{10}$.

L'estimation par la méthode du maximum de vraisemblance consiste à trouver le point $(\sigma^2, p_{01}, p_{10})$ qui maximise la vraisemblance $L(x, y)$ donnée par (3) ci-dessus. On a le lemme suivant.

LEMME 2. – *Étant donnés a et b alors l'EMV pour $(\sigma^2, p_{01}, p_{10})$ est donné par*

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n \{X_t - (aY_t + b(1 - Y_t))X_{t-1}\}^2,$$

$$\hat{p}_{01} = \frac{N_{01}}{N_{01} + N_{00}}; \quad \hat{p}_{10} = \frac{N_{10}}{N_{10} + N_{11}}.$$

Preuve. – Maximiser (3) revient à trouver le point qui annule les dérivées partielles suivantes

$$\frac{\partial}{\partial \sigma^2} \text{Log}L(x, y) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{t=1}^n \{x_t - (ay_t + b(1 - y_t))x_{t-1}\}^2,$$

$$\frac{\partial}{\partial p_{01}} \text{Log}L(x, y) = \frac{N_{01}}{p_{01}} - \frac{N_{00}}{1 - p_{01}},$$

$$\frac{\partial}{\partial p_{10}} \text{Log}L(x, y) = \frac{N_{10}}{p_{10}} - \frac{N_{11}}{1 - p_{10}}.$$

La solution donnée par le lemme 2 annule le système ci-dessus. On vérifie aisément que la matrice des dérivées secondes est de déterminant positif et que $(\partial^2/\partial(\sigma^2)^2)L$ est négatif. On en déduit alors l'EMV.□

On notera que tous les estimateurs proposés par les 2 lemmes ci-dessus sont fortement consistants.

3.2. Cas de conjoncture non observable

Dans cette partie, on supposera que le processus de conjoncture (Y_t) est non observable (partiellement ou complètement) et que le processus (X_t) est observable. Cette situation répond, généralement, à la réalité économique telle que la récession et l'expansion d'une économie (*cf.* Hamilton (1989)). Il faut donc estimer en plus du paramètre $\theta = (a, b, \sigma^2, p_{01}, p_{10})$, le processus de conjoncture (Y_t) . Le problème devient alors un problème inverse donné par le modèle SAR(1) à conjoncture non observable, *i.e.*

$$X_t = (aY_t + b(1 - Y_t))X_{t-1} + \epsilon_t; \quad t \geq 1,$$

où (Y_t) est un processus d'entrée qu'il faut estimer à partir du processus de sortie (X_t) .

De tels problèmes inverses sont connus en traitement du signal. On peut citer à titre d'exemples le problème de détection en théorie de communication (cf. Essebbar *et al.* (2001)) et celui de la restauration en imagerie (cf. par exemple Marroquin *et al.* (1987)).

Le cadre bayésien est plus adapté à ce type de problème. Le MAP (maximum *a posteriori*) et le MPM (maximum des probabilités marginales) sont les solutions bayésiennes les plus utilisées. Elles sont associées respectivement aux fonctions de perte $L(Y, \hat{Y}) = \mathbf{1}_{\{Y \neq \hat{Y}\}}$ et $L(Y, \hat{Y}) = \sum_i \mathbf{1}_{\{Y_i \neq \hat{Y}_i\}}$.

Le MPM est plus performant que le MAP (cf. Marroquin *et al.* (1987)) qui est donné par $\hat{Y}_i = 1$ si $P(Y_i = 1|X) > \mu$, où μ est non loin de $1/2$. Askar *et al.* (1981) en Théorie de la communication et Lee *et al.* (cf. Hamilton (1994)) en Économétrie ont donné un algorithme récursif qui permet le calcul des probabilités $P(Y_i = 1|X)$, $i = 1, \dots, n$. Dans l'appendice, on construit un algorithme [A2] adapté à notre problème inverse.

La restauration de (Y_t) suppose la connaissance de θ qui est un paramètre à estimer. On propose alors d'utiliser l'algorithme d'apprentissage suivant : $Y^{(0)}$ étant un processus d'initialisation, généralement on prend le processus de conjoncture empirique (voir étape 1 ci-dessous), à l'étape k , on a :

Algorithme [A1] :

- (i) Estimer $\theta^{(k)}$ à partir de X et $Y^{(k-1)}$ (section 3.1),
- (ii) Restauration de $Y^{(k)}$ par le MPM (Algorithme [A2]).

L'arrêt de cet algorithme se fait dès qu'il y a stabilisation de la solution. Pour les problèmes théoriques liés à la convergence de ce type d'algorithme d'apprentissage, on peut consulter par exemple Youness (1988). On montre dans nos simulations que cet algorithme converge rapidement (maximum cinq itérations pour des données simulées ou réelles) dans une fraction de seconde.

4. Étude numérique

On étudie la validation du modèle théorique SAR(1) proposé dans la section 2 et celle des estimateurs développés dans la partie statistique de l'estimation (section 3). Ainsi, dans un premier temps on travaille sur des données simulées afin de cerner les qualités des estimateurs. Dans un deuxième temps, on confronte notre modèle aux données réelles liées à certains phénomènes dans le tourisme tels que les arrivées des touristes internationaux aux postes frontières (Arrivées) et les nuitées dans les hôtels classés (Nuitées)

(i) Données simulées : On simule des trajectoires de processus de conjoncture (Y_t) étant donnés les paramètres de transition p_{01} et p_{10} et la valeur Y_0 d'initialisation. Juste après, on construit le processus (X_t) en se donnant X_0 le point de départ et une suite de bruits blancs (ϵ_t) de loi $N(0, \sigma^2)$.

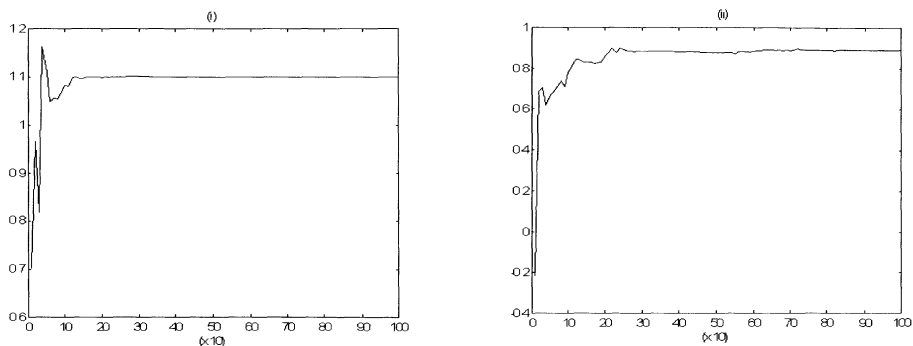


FIGURE 1
Vitesse de convergence des estimateurs \hat{a} (i) et \hat{b} (ii)

Par la suite on estime les paramètres du modèle SAR(1) selon que le processus de conjoncture (Y_t) est supposé observable ou non. On compare, alors les valeurs des paramètres d’entrée à celles de sortie. On cerne aussi le problème de la qualité de restauration du processus de conjoncture par le MPM et cela grâce à l’algorithme [A2] donné dans l’appendice et l’algorithme d’apprentissage [A1] ci-dessus. Ainsi, on calcule le nombre de points mal restaurés par le MPM avec $\mu = 1/2$.

TABLEAU 1
Paramètres d’entrée et estimés

	a	b	σ^2	p_{01}	p_{10}
Entrés	1.100	0.900	1,000	0.60	0.40
Estimés	1.099	0.910	1,010	0.595	0.392

Le Tableau 1 donne un exemple d’estimation avec les paramètres d’entrée et leurs estimations. La convergence de l’algorithme [A1] est très rapide (2 itérations en incluant l’algorithme [A2]) et la restauration du processus de conjoncture est de l’ordre de 100 %. Les figures 1 visualisent la vitesse de convergence des estimateurs \hat{a} et \hat{b} vers les vraies valeurs a et b .

(ii) Données réelles : Considérons la trajectoire du processus lié aux Arrivées mensuelles des touristes internationaux aux postes frontières de 1984 à 2001 (Tableau 6). Les données sont rangées comme suit : les lignes sont les années et les colonnes sont les mois. À partir de ces données qui sont brutes, on procède par les étapes suivantes :

Étape 1 : On construit le processus de conjoncture empirique ($Y_{t,j}^*$). À cause de la nature saisonnière des données, on travaille par mois (colonnes du tableau). On pose $Y_{t,j}^* = 1$ si au mois j et à l’année t il y a eu accroissement par rapport à l’année $t - 1$; sinon $Y_{t,j}^* = 0$. Ce processus sera utilisé, dans un premier temps, dans le modèle SAR(1) avec conjoncture observable. Ensuite, on l’utilisera comme processus d’initialisation dans le SAR(1) avec conjoncture non observable.

Étape 2 : On procède à l'estimation des paramètres du modèle comme suit. A cause des disparités saisonnières, on estime a_j , b_j et σ_j^2 à partir des données colonnes en utilisant les estimateurs donnés par les lemmes ci-dessus. Les paramètres de transition p_{01} et p_{10} sont estimés par la trajectoire qui comporte les mois des 17 années (toutes les lignes du tableau).

Étape 3 : Une tentative de reconstruction (prévision), $(\hat{X}_{t,j})$, du processus $(X_{t,j})$ est faite à partir des estimations des paramètres (cf. étape 2), du processus de conjoncture selon qu'il est observable $(Y_{t,j}^*)$ ou estimé (restauré) $(\hat{Y}_{t,j})$ (cf. étape 1) et à partir de la première ligne du tableau (initialisation du processus $(X_{t,j})$). On pose

$$\hat{X}_{t,j} = (a_j Y_{t,j} + b_j (1 - Y_{t,j})) \hat{X}_{t-1,j}, \quad t = 1, \dots, 17, \quad j = 1, \dots, 12, \quad (4)$$

avec $\hat{X}_{1,j} = X_{1,j}$, $j = 1, \dots, 12$. On utilisera la mesure suivante

$$\frac{\sqrt{\sum_{t,j} (\hat{X}_{t,j} - X_{t,j})^2}}{\sum_{t,j} X_{t,j}},$$

qui donne l'écart du modèle par rapport aux données réelles.

TABLEAU 2
Paramètres estimés : cas de SAR(1) avec conjoncture observable

j	1	2	3	4	5	6	7	8	9	10	11	12
a_j	1.16	1.15	1.20	1.19	1.12	1.14	1.10	1.15	1.09	1.16	1.27	1.21
b_j	0.80	0.73	0.84	0.85	0.88	0.86	0.83	0.87	0.75	0.85	0.83	0.84
$\sigma_j^2 \times 10^8$	3.16	3.63	8.98	3.66	4.53	4.22	8.44	17.78	7.94	8.80	8.03	7.84

TABLEAU 3
Paramètres estimés : cas de SAR(1) avec conjoncture non observable

j	1	2	3	4	5	6	7	8	9	10	11	12
a_j	1.14	1.15	1.13	1.17	1.06	1.06	1.08	1.12	1.06	1.16	1.23	1.18
b_j	0.79	0.71	0.82	0.86	0.55	0.57	0.78	0.86	0.47	0.86	0.83	0.48
$\sigma_j^2 \times 10^8$	4.14	3.49	11.60	4.62	2.40	4.31	9.48	19.80	4.44	9.22	9.15	8.72

SAR(1) avec conjoncture observable : Les résultats qui suivent sont obtenus par la modélisation des données réelles selon le modèle SAR(1) avec conjoncture observable. Les Tableaux 2 et 4 donnent les valeurs estimées des paramètres de

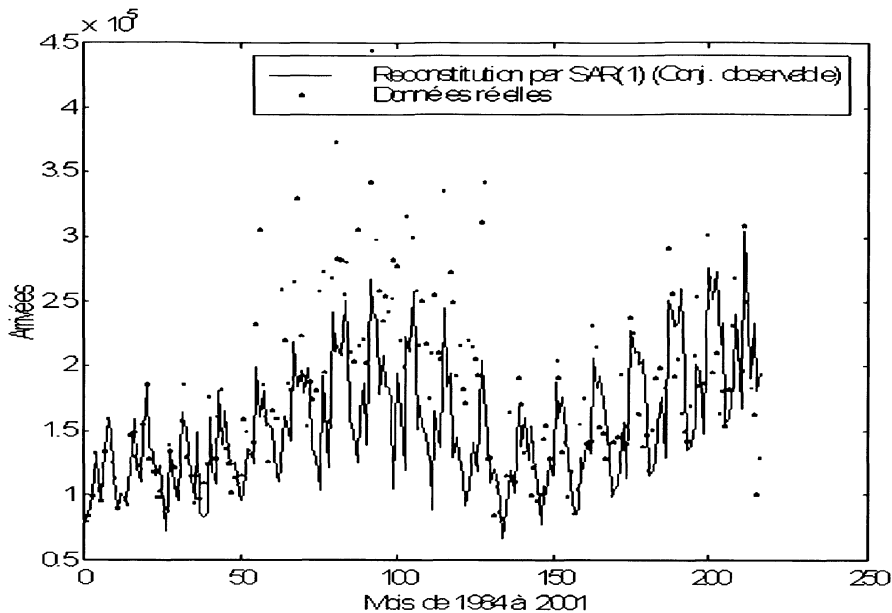


FIGURE 2

*SAR(1) avec conjoncture observable :
reconstitution mensuelle des arrivées de 1984 à 2001 (erreur = 2.2 %)*

TABLEAU 4
Probabilités de transition

p_{ij}	p_{01}	p_{10}
Conjoncture observable	0.33	0.17
Conjoncture non observable	0.29	0.09

(X_t) ($= (X_{t,j})$) et (Y_t^*) ($= (Y_{t,j}^*)$). Le Tableau 5 donne le processus de conjoncture (Y_t^*) . La figure 2 donne les trajectoires de (X_t) et du processus (\hat{X}_t) reconstruit par (4) de l'étape 3 ci-dessus. On remarque que l'écart entre le modèle et les données est de l'ordre de 2.2 % sur les 17 années.

La persistance de la conjoncture est de l'ordre $q^1 = p_{11} = 83 \%$ et $q^0 = p_{00} = 67 \%$. Les durées moyennes des séquences D^i dans l'état i sont $E(D^1) = 1/p_{10} = 5.88$ mois et $E(D^0) = 1/p_{01} = 3.03$ mois.

SAR(1) avec conjoncture non observable : La modélisation est faite ici selon le modèle SAR(1) avec conjoncture non observable (voir section 3.2 ci-dessus). Dans les Tableaux 3 et 4, on a les valeurs estimées des paramètres de (X_t) . Le Tableau 5 donne le processus de conjoncture restauré (\hat{Y}_t) par l'algorithme [A1]. La convergence

de ce dernier est très rapide (maximum cinq itérations en incluant l'algorithme [A2]). Le temps de calcul sur machine est de l'ordre de fractions de seconde. Dans la figure 3, on donne les trajectoires de (X_t) et du processus (\hat{X}_t) reconstruit à partir de (4). On remarque que l'écart s'est beaucoup amélioré, il est de l'ordre de 1.6 % sur les 17 années.

La persistance de la conjoncture est très haute dans les deux cas : $q^1 = 91\%$ et $q^0 = 71\%$. La durée moyenne d'une conjoncture favorable est de l'ordre $E(D^1) = 11.11$ mois. De même, la durée moyenne d'une conjoncture défavorable est $E(D^0) = 3.45$ mois. Ces durées moyennes vont nous permettre de construire des prévisions à long terme.

Les figures 4 et 5 donnent les prévisions mensuelles et annuelles sur les 30 années à venir. On remarque que la barre de 10 millions est atteinte en l'an 2026.

TABLEAU 5
Processus (Y_t^) et (\hat{Y}_t) des Arrivées*

Année/Mois	\hat{Y}_t	Y_t^*
	j f m a m j j a s o n d	j f m a m j j a s o n d
1985	1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1
1986	1 1 1 1 1 1 1 1 1 1 1 1 1	1 0 0 0 0 0 1 0 1 0 0 1
1987	0 1 1 1 1 1 1 1 1 1 1 1 1	0 1 0 1 1 1 1 0 1 1 1 0
1988	1 1 1 0 1 1 1 1 1 1 1 1 1	1 1 1 0 1 1 1 1 1 1 1 1 1
1989	1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1
1990	1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 0 1 1 0 1 1 1 1 1 1 1
1991	1 1 0 1 1 1 1 1 1 0 0 0	1 1 1 0 1 1 1 1 1 0 0 0
1992	1 1 0 1 1 1 1 0 1 0 0 0	1 1 0 1 0 0 0 0 1 0 0 0
1993	0 0 0 0 1 1 1 0 1 0 0 0	0 0 0 0 0 1 1 0 0 0 0 0
1994	0 0 0 0 1 1 0 0 0 0 0 0	0 0 1 0 0 0 0 0 0 0 0 0
1995	0 0 0 0 0 0 0 0 1 0 0 0	0 0 0 0 0 0 0 0 1 1 1 1
1996	0 1 1 0 1 1 1 0 1 1 1 1 1	1 1 1 0 1 1 1 1 1 1 0 0
1997	1 1 1 1 1 1 1 1 1 1 1 1 1	0 1 1 1 1 1 1 1 1 1 1 1 1
1998	1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 0 1 1 0 1 1 1 1 1 1 1
1999	1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1 1 1 1
2000	1 1 1 1 1 1 1 1 1 1 0 1	1 1 1 1 1 1 1 1 1 1 1 1 1
2001	1 1 1 1 1 1 1 1 1 0 0 0	1 1 1 1 1 1 1 0 0 0 0 0

TABLEAU 6 :
Arrivées des touristes de 1984 à 2001

AN	JANV.	FEVR.	MARS	AVRIL	MAI	JUIN	JUIL.	AOUT	SEPT.	OCT.	NOV.	DEC.	TOTAL
84	80834	84901	100847	133484	108871	97324	134463	160645	125685	114215	91635	102137	1335041
85	98403	93174	147277	150078	130771	123843	155116	186164	128185	119771	99317	104464	1536563
86	99444	90541	138716	134860	121192	106309	157725	185622	129699	115299	94916	115316	1489639
87	98475	109332	125069	176900	128522	128674	160918	182806	137020	124822	103124	115090	1590752
88	110141	115668	159148	149261	131654	141110	232750	306087	185748	161607	126717	165756	1985647
89	160231	159865	259614	219808	186936	182627	266337	330313	224013	193286	153968	188546	2525544
90	174290	181380	258242	273998	195619	156887	269814	373780	283480	282775	256780	281084	2988129
91	211679	203892	305311	216510	221220	202525	343414	445553	299003	259374	235222	253538	3197241
92	242911	252600	281730	277422	219627	200199	316450	434980	299579	259345	217002	250217	3252062
93	217521	175788	210601	254884	210341	206431	336028	401239	272508	249924	193360	217075	2945700
94	182918	171510	220457	215795	205461	193142	311052	342796	128592	129913	84678	107036	2293350
95	87289	81340	115447	164547	112511	110616	190475	171350	132939	135150	100357	122098	1524119
96	97434	101660	143636	153997	128667	116991	204986	191287	133655	142009	99001	119408	1632731
97	86630	105038	155994	175525	140460	142827	231448	215380	152765	148249	128334	140798	1823448
98	102348	141202	145281	194414	154458	140287	237140	226348	164194	163142	137848	146969	1953631
99	119821	150887	190958	198606	172380	183893	291261	256347	191642	205510	164347	149162	2274814
00	142758	170031	208298	253713	185336	186831	302754	261750	195781	210294	164239	181109	2462894
01	154726	182112	232093	269916	205538	196949	309401	250513	184059	162207	101925	129454	2378893

(*) Source : ministère du tourisme

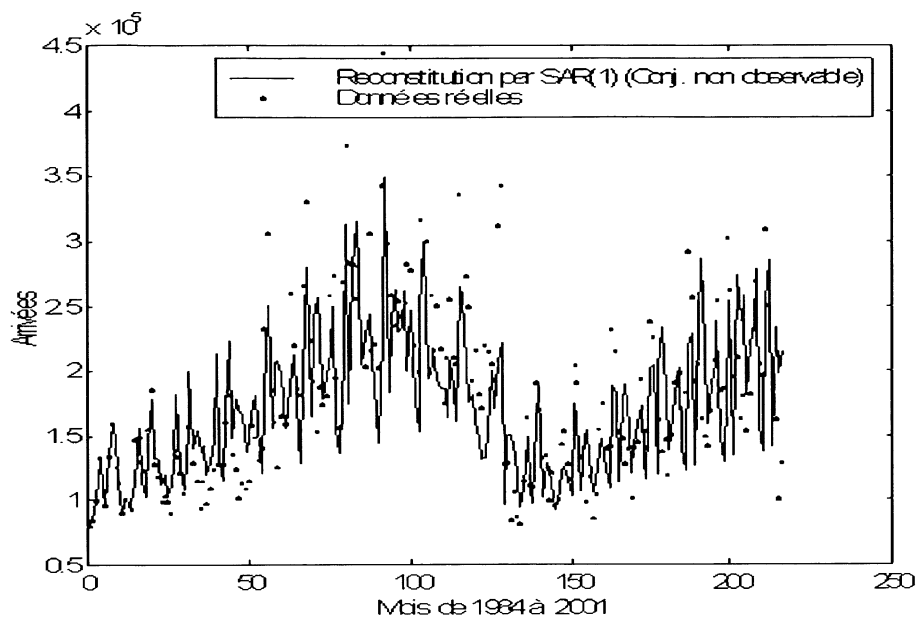


FIGURE 3

*SAR(1) avec conjoncture non observable :
reconstitution mensuelle des arrivées de 1984 à 2001 (erreur = 1.6 %)*

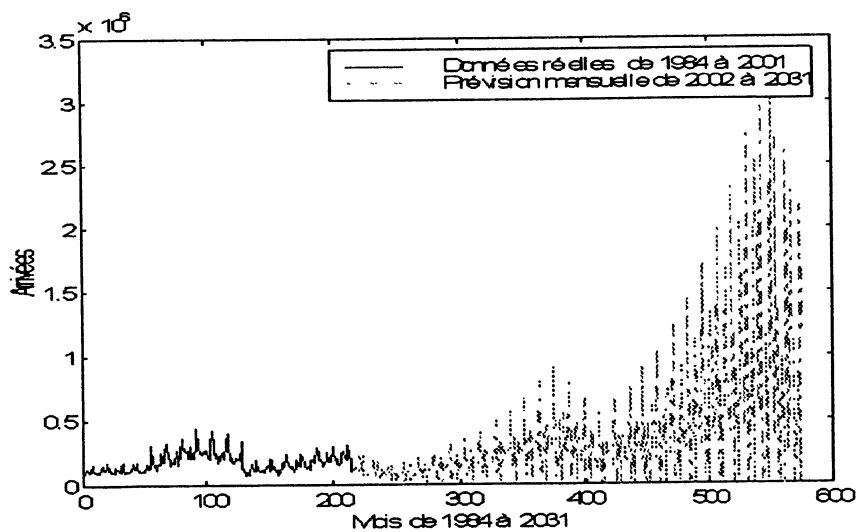


FIGURE 4

Prévision mensuelle avec SAR(1) (conjoncture non observable) de 2002 à 2031

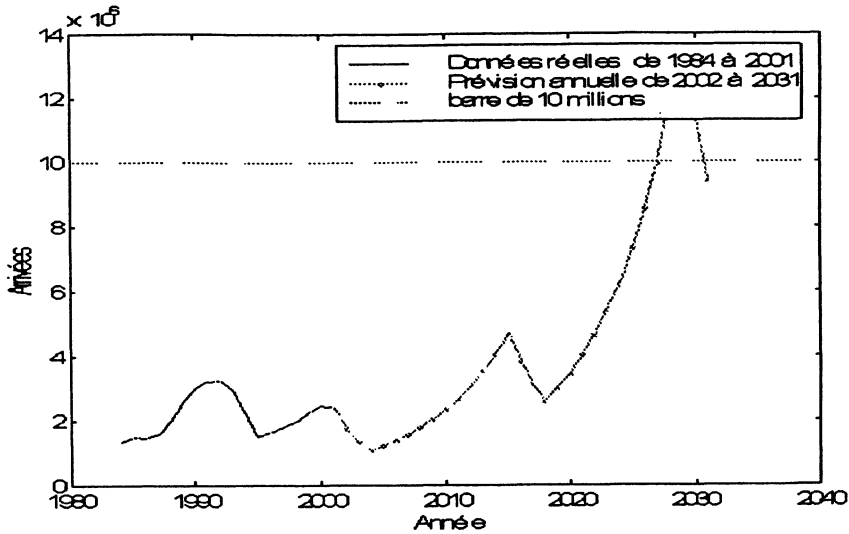


FIGURE 5
Prévision annuelle avec SAR(1) (conjoncture non observable) de 2002 à 2031

5. Conclusion

Le modèle qu'on vient d'étudier mathématiquement et numériquement semble avoir un intérêt certain dans la modélisation de certains phénomènes dans le tourisme, tels que le processus des arrivées internationales aux postes frontières.

L'introduction des techniques de télédétection en traitement de signal pour restaurer le processus de conjoncture s'est avérée concluante. Ainsi, on a pu améliorer la reconstitution du processus des arrivées mensuelles des touristes sur les 18 années (écart de 1.6 %). De même, on a pu améliorer les persistances des conjonctures et par suite leurs durées moyennes.

Il faut noter que les prévisions qu'on a données sont à prendre avec prudence et surtout la barre de 10 millions qui est prévue par le modèle vers l'an 2026.

Comme on l'a vu, le modèle SAR(1), par sa simplicité, se prête aisément à l'analyse mathématique et à son application à la modélisation de toutes les séries temporelles qui ont un double régime : régime haut et régime bas.

D'autres améliorations sont possibles en modélisant le processus de conjoncture par une chaîne de Markov d'ordre 2 et plus. On peut aussi prospecter la modélisation par un processus de renouvellement alterné à temps discret.

Appendice

Dans cette partie, on propose un algorithme (cf. Askar et Derin (1981)) qui permet le calcul récursif des probabilités $P(Y_i = 1|X)$, $i = 1, \dots, n$. Dans la suite, on utilisera les notations suivantes : $x_{(k)} = (x_1, \dots, x_k)$, $x_{(n-k)} = (x_{k+1}, \dots, x_n)$, $k = 1, \dots, n$; $B_{i|j} = P(Y_i = 1|X_{(j)} = x_{(j)})$, $i, j = 1, \dots, n$.

Algorithme [A2] : Pour $k = 1, \dots, n$, on a

$$\begin{aligned} B_{k|n} &= P(Y_k = 1|X_{(n)} = x_{(n)}) \\ &= \sum_{j=0}^1 P(Y_k = 1, Y_{k+1} = j|X_{(k)} = x_{(k)}, X_{(n-k)} = x_{(n-k)}) \\ &= P(Y_k = 1|Y_{k+1} = 1, X_{(k)} = x_{(k)})P(Y_{k+1} = 1|X_{(n)} = x_{(n)}) \\ &\quad + P(Y_k = 1|Y_{k+1} = 0, X_{(k)} = x_{(k)})P(Y_{k+1} = 0|X_{(n)} = x_{(n)}). \end{aligned}$$

On a

$$\begin{aligned} P(Y_k = 1|Y_{k+1} = 1, X_{(k)} = x_{(k)}) &= \frac{P(Y_{k+1} = 1|Y_k = 1)P(Y_k = 1|X_{(k)} = x_{(k)})}{P(Y_{k+1} = 1|X_{(k)} = x_{(k)})} \\ &= (1 - p_{10}) \frac{B_{k|k}}{B_{k+1|k}}, \end{aligned}$$

et

$$P(Y_k = 1|Y_{k+1} = 0, X_{(k)} = x_{(k)}) = p_{10} \frac{B_{k|k}}{1 - B_{k+1|k}}.$$

alors

$$B_{k|n} = B_{k|k} \left\{ (1 - p_{10}) \frac{B_{k+1|n}}{B_{k+1|k}} + p_{10} \frac{1 - B_{k+1|n}}{1 - B_{k+1|k}} \right\}, \quad k = n - 1, \dots, 1.$$

Le calcul de $B_{k|n}$ suit celui de $B_{k|k} = P(Y_k = 1|X_{(n)} = x_{(n)})$ et celui de $B_{k+1|k} = P(Y_{k+1} = 1|X_{(k)} = x_{(k)})$. On a

$$B_{1|0} = P(Y_1 = 1) = \frac{p_{01}}{p_{01} + p_{10}},$$

et

$$\begin{aligned} B_{k+1|k+1} &= P(Y_{k+1} = 1|X_{(k+1)} = x_{(k+1)}) \\ &= P(Y_{k+1} = 1|X_{(k)} = x_{(k)}, X_{k+1} = x_{k+1}) \\ &= \frac{f(x_{k+1}|Y_{k+1} = 1, X_{(k)} = x_{(k)})P(Y_{k+1} = 1|X_{(k)} = x_{(k)})}{P(X_{k+1} = x_{k+1}|X_{(k)} = x_{(k)})}, \end{aligned}$$

où $f(x_{k+1}|Y_{k+1} = 1, X_{(k)} = x_{(k)})$ est la densité de $(X_{k+1}|Y_{k+1} = 1, X_{(k)} = x_{(k)})$ qui est donnée par

$$f(x_{k+1}|Y_{k+1} = 1, X_{(k)} = x_{(k)}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x_{k+1} - ax_k)^2\right\}.$$

De même, on a

$$f(x_{k+1}|Y_{k+1} = 0, X_{(k)} = x_{(k)}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x_{k+1} - bx_k)^2\right\}.$$

Evaluons $P(X_{k+1} = x_{k+1}|X_{(k)} = x_{(k)})$, on a

$$\begin{aligned} P(X_{k+1} = x_{k+1}|X_{(k)} = x_{(k)}) &= P(X_{k+1} = x_{k+1}, Y_{k+1} = 1|X_{(k)} = x_{(k)}) \\ &+ P(X_{k+1} = x_{k+1}, Y_{k+1} = 0|X_{(k)} = x_{(k)}), \end{aligned}$$

avec pour $j \in \{0, 1\}$

$$\begin{aligned} P(X_{k+1} = x_{k+1}, Y_{k+1} = j|X_{(k)} = x_{(k)}) &= P(X_{k+1} = x_{k+1}|Y_{k+1} = j, X_{(k)} = x_{(k)}) \\ &\times P(Y_{k+1} = j|X_{(k)} = x_{(k)}). \end{aligned}$$

D'où, on obtient pour $k = 0$ à $n - 1$,

$$\begin{aligned} B_{k+1|k+1} \\ = \frac{B_{k+1|k} \exp\{-(x_{k+1} - ax_k)^2/2\sigma^2\}}{B_{k+1|k} \exp\{-(x_{k+1} - ax_k)^2/2\sigma^2\} + (1 - B_{k+1|k}) \exp\{-(x_{k+1} - bx_k)^2/2\sigma^2\}}. \end{aligned}$$

Les probabilités $B_{k+1|k}$ sont données par

$$\begin{aligned} B_{k+1|k} &= P(Y_{k+1} = 1|X_{(k)} = x_{(k)}) \\ &= \sum_{j=0}^1 P(Y_{k+1} = 1, Y_k = j|X_{(k)} = x_{(k)}) \\ &= P(Y_{k+1} = 1|Y_k = 1)P(Y_k = 1|X_{(k)} = x_{(k)}) \\ &+ P(Y_{k+1} = 1|Y_k = 0)P(Y_k = 0|X_{(k)} = x_{(k)}), \end{aligned}$$

d'où

$$B_{k+1|k} = (1 - p_{10})B_{k|k} + p_{01}(1 - B_{k|k}), \quad k = 0, \dots, n - 1.$$

Remarque : $B_{k+1|k}$ est la prévision à un pas de (Y_t) basée sur l'observation $X_{(k)}$ qui vérifie aussi

$$(1 - B_{k+1|k}, B_{k+1|k})' = Q \cdot (1 - B_{k|k}, B_{k|k})',$$

avec Q étant la matrice de transition de (Y_t) donnée par

$$Q = \begin{pmatrix} 1 - p_{01} & p_{10} \\ p_{01} & 1 - p_{10} \end{pmatrix}.$$

Remerciements : On tient à remercier les rapporteurs de leurs conseils qui ont permis l'amélioration de ce travail. On remercie aussi les fonds de recherches PROTARS P3T1/06 pour leur soutien financier et le service de DPCP du ministère du tourisme pour l'accès aux données.

Références

- ASKAR M., DERIN H. (1981), A recursive algorithm for the Bayes solution of the smoothing problem. *IEEE Trans. Automat. Contr.*, vol. AC26, 558-561.
- BOX G.E.P., JENKINS G.M. (1976), *Time series analysis : Forecasting and control (second edition)*. Holden Day : San Fransisco.
- ESSEBBAR B., LE BRETON A. (1985), A tentative Model of vector time series with binary components with application in climatology. *Model Choice*, Éditions de l'Université St-Louis, pp. 127-154, Bruxelles : Belgique.
- ESSEBBAR B., LE BRETON A. (1987), Sur un modèle de série chronologique vectorielle à composantes binaires en vue d'applications en climatologie. *Computer Methods and Water Resources*, Vol. 3, Comput. Mech., pp. 199-217, Southampton.
- ESSEBBAR B., HSAINA M. A. (2001), *Détection bayésienne d'un signal de canal via la perte de Baddeley*. GRETSI'01, Toulouse 10-13 Septembre 2001, France.
- HAMILTON J. D. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57, 357-84.
- HAMILTON J. D. (1990), Analysis of time series subject to changes in regime. *J. Econometrics*, 45, n° 1-2, pp. 39-70.
- HAMILTON J. D. (1994). *Time series analysis*. Princeton University Press : Princeton.
- MARROQUIN J., MITTER S., POGGIO T. (1987), *Probabilistic solutions of ill-posed problems in computational vision*, JASA, 82, pp. 76-89.
- ROBERT C. (1996), Méthodes de Monte Carlo par chaînes de Markov. Economica : Paris.
- WITT S.F., WITT C.A. (1995), Forecasting tourism demand : A review of empirical research. *International Journal of Forecasting*. 11, 447-475.
- YOUNESS L. (1988), Estimation and annealing for gibbsian fields. *Ann. Inst. Henri Poincaré*, Vol. 24, 269-294.