

GREQAM

**Groupement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III**

**Document de Travail
n° 2007-04**

LE CHANGEMENT STRUCTUREL DANS UN ENVIRONNEMENT MÉMOIRE LONGUE

**Mohamed BOUTAHAR
Et
Mustapha BELKHOUJA**

Mars 2007

DT-GREQAM

LE CHANGEMENT STRUCTUREL DANS UN ENVIRONNEMENT MÉMOIRE LONGUE

Mohamed BOUATAHAR*,

Mustapha BELKHOUSA†

GREQAM, Université de la Méditerranée 2 Rue de la charité, 13002 Marseille, France.

Résumé

Dans ce papier, nous considérons le problème d'estimation des dates de rupture et présentons un algorithme efficace pour obtenir les minimums globaux des sommes des carrés des résidus. Cet algorithme est basé sur le principe de la programmation dynamique et nécessite au plus des opérations de moindres carrés d'ordre $O(n^2)$ pour tout nombre de ruptures. Ensuite nous abordons le problème d'estimation du nombre de ruptures par les critères d'information, le test de Bai et Perron (1998) et la méthode de Lavielle (2004). Finalement, une analyse par la méthode de Monte Carlo donnera une idée sur le comportement des estimateurs et des tests en échantillons finis.

Mots clés : Changement structurel pur, dates de ruptures, estimation des moindres carrés pénalisés, mémoire longue, programmation dynamique, régimes multiples, sélection de modèle, tests d'hypothèses.

Abstract

In this paper, we consider the problem of estimation of the break dates and present an efficient algorithm in order to obtain global minimizers of the sum of squared residuals. This algorithm is based on the principle of dynamic programming and requires at most least-squares operations of order $O(n^2)$ for any number of breaks. We also study the estimation of the number of breaks by using the information criteria, the test of Bai and Perron (1998), and the method of Lavielle (2004). Finally, we perform a Monte Carlo study to analyse the behavior of estimators and tests in finite sample size.

Key words : Break dates, dynamic programming, hypothesis testing, long memory, model selection, multiple regimes, penalized least-squares estimate, pure structural change.

*E-mail : boutahar@lumimath.univ-mrs.fr

†E-mail : belkhouja_mustapha@yahoo.fr

1. INTRODUCTION

La majorité des efforts pionniers a été dédiée à la détection et la localisation d'un seul point de rupture au niveau de la moyenne d'une suite de variables aléatoires indépendantes et identiquement distribuées (voir parmi d'autres ; Hinkley (1970), Sen Srivastava (1975), Hawkins (1977), Bhattacharya (1987)). Plusieurs contributions récentes ont mis l'accent sur l'existence de plusieurs points de ruptures dans un processus aléatoire stationnaire ou non stationnaire. Bai et Perron (1998) ont proposé l'estimation par la méthode des moindres carrés, en présence d'une faible dépendance du terme d'erreur. D'un point de vue pratique, l'estimation par la méthode des moindres carrés possède un avantage considérable sur celle du maximum de vraisemblance pour la simple raison que la première n'exige pas la spécification de la distribution du terme d'erreur.

Le problème d'estimation du nombre de ruptures a été traité initialement par Yao (1988), qui a prouvé la consistance du critère de Schwarz quand le terme d'erreur est une suite i.i.d selon une loi de Gauss. Par la suite, Bai et Perron (1998) ont mis au point plusieurs tests. Parmi ceux-ci, le test de type sup Wald qui teste l'hypothèse nulle d'absence de ruptures contre l'hypothèse alternative d'un nombre arbitraire de ruptures. Le deuxième teste l'hypothèse de l ruptures contre $l + 1$.

Lavielle et Moulines (1996) adoptent une approche globale où tous les points de ruptures sont simultanément détectés en minimisant la fonction contraste pénalisée $Q_T(T, y) + \beta pen(K)$ dont le rôle est de localiser les points de ruptures aussi précis que possible. Le terme de pénalité $pen(K)$ ne dépend que de la dimension $K = m + 1$, qui est le nombre de régimes existants dans le modèle. Le paramètre de pénalisation β sert à faire l'ajustement entre la minimisation de $Q_{T,K}(T)$ (obtenue avec un K élevé) et la minimisation de $pen(K)$ (obtenue avec un K faible). Ce genre de fonction contraste peut être utile dans la pratique mais il reste le problème de choix d'une bonne fonction de pénalité et d'un bon coefficient β d'où l'intérêt de la procédure adaptative mise au point par Lavielle (2004).

Ce document est structuré ainsi ; la section 2 présente le modèle de changement structurel pur sur lequel nous allons travailler. La section 3 introduit l'algorithme basé sur le principe de la programmation dynamique qui permet d'estimer efficacement le modèle pour tout nombre de ruptures. La section 4 présente les méthodes d'estimation du nombre de ruptures par les critères d'informations, le test de Bai et Perron (1998) ainsi que la méthode de Lavielle (2004). La section 5 présente les expériences de simulations, les analyses et les résultats de chaque méthode d'estimation du nombre et des dates de ruptures. Quelques remarques et conclusions sont contenues dans la section 6.

2. LE MODÈLE

Nous considérons le modèle de changement structurel avec m ruptures ($m+1$ régimes) :

$$y_t = \delta_j + u_t, t = T_{j-1} + 1, \dots, T_j \quad (2.1)$$

pour $j = 1, \dots, m + 1$, $T_0 = 0$ et $T_{m+1} = n$, y_t est la variable observée, $\delta = (\delta_1, \dots, \delta_{m+1})'$ est le vecteur des coefficients avec $\delta_i \neq \delta_{i+1}$ ($1 \leq i \leq m$), et u_t est le terme d'erreur. Les dates de ruptures (T_1, \dots, T_m) sont considérées comme inconnues pour $i = 1, \dots, m$, et nous avons $\lambda_i = T_i/n$ les fractions de ruptures avec $0 < \lambda_1 < \dots < \lambda_m < 1$. Nous sommes dans le cas d'un modèle de changement structurel pur où tous les coefficients varient dans le temps. Bai et Perron(1998) imposent quelques restrictions sur les valeurs possibles des dates de ruptures, en effet, ils définissent un ensemble pour une certaine valeur arbitraire, petite et positive notée η tel que :

$\Lambda_\eta = \{(\lambda_1, \dots, \lambda_m); |\lambda_{i+1} - \lambda_i| \geq \eta, \lambda_1 \geq \eta, \lambda_m \leq 1 - \eta\}$ dans le but de contraindre les dates de ruptures d'être asymptotiquement différentes et limitées par les valeurs extrêmes de l'échantillon. La méthode d'estimation considérée est celle basée sur le principe des moindres carrés proposé par Bai et Perron (1998). Cette méthode est décrite comme suit; pour chaque partition (T_1, \dots, T_m) , notée $\{T_j\}$, les estimations sont obtenues d'abord en minimisant la somme des carrés des résidus $\sum_{i=1}^{m+1} \sum_{j=T_{i-1}+1}^{T_i} (y_j - \delta_i)^2$.

Soient $\hat{\delta}(\{T_j\})$ les estimations résultantes, en les substituant dans la fonction objectif nous aurons une somme notée Q_T :

$$\begin{aligned} Q_T(T_1, \dots, T_m) &= \min_{(\delta_1, \dots, \delta_{m+1})} \sum_{i=1}^{m+1} \sum_{j=T_{i-1}+1}^{T_i} (y_j - \delta_i)^2 \\ &= \sum_{i=1}^{m+1} \sum_{j=T_{i-1}+1}^{T_i} (y_j - \bar{y}(T_{i-1}, T_i))^2 \end{aligned} \quad (2.2)$$

$$\text{où } \bar{y}(T_{i-1}, T_i) = \frac{1}{T_i - T_{i-1}} \sum_{j=T_{i-1}+1}^{T_i} y_j .$$

Les estimations des dates de ruptures sont telles que :

$(\hat{T}_1, \dots, \hat{T}_m) = \arg \min_{(T_1, \dots, T_m)} Q_T(T_1, \dots, T_m)$, avec $T_i - T_{i-1} \geq [\eta n]$, ainsi les estimateurs des dates de ruptures sont les minimums globaux de la fonction objectif. Finalement les estimateurs des paramètres de régression sont associés aux dates de ruptures estimées c'est à dire $\hat{\delta} = \hat{\delta}(\{\hat{T}_i\})$. Pour le calcul, nous utilisons l'algorithme efficace développé par Bai et Perron (1996) et basé sur le principe de la programmation dynamique qui permet l'obtention des minimums globaux en utilisant un nombre de sommes des carrés des résidus d'ordre $O(n^2)$ pour tout $m \geq 2$.

3. LA METHODE DE CALCUL DES MINIMUMS GLOBAUX

Bai et Perron (1998), considèrent un algorithme efficace qui se base sur le principe de la programmation dynamique dans le but de calculer les estimations des points de ruptures comme des minimums globaux de la somme des carrés des résidus. L'algorithme utilise des opérations de moindres carrés d'ordre $O(n^2)$ au plus, pour toute valeur m , tout à fait différent d'une grille standard dont la procédure nécessite des opérations de moindres carrés d'ordre $O(n^m)$. L'idée de base de cette méthode, utilisée pour des cas particuliers, n'est pas nouvelle et a été déjà considérée par Guthery (1974), Bellman et Roth (1969) et Fisher (1958). Cette méthode est peu utilisée, au moins au niveau de la littérature économétrique, malgré son grand intérêt.

3.1. La matrice triangulaire des sommes des carrés des résidus

L'approche est réalisée avec un échantillon de taille n , le nombre total de régimes possibles est au plus égale à $n(n+1)/2$, il est donc d'ordre de $O(n^2)$.

	d a t e											s o r t i e													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	a	a	a	a	*	*	*	*	*	*	*	*	*	*	*	b	b	b	b	b	b	b	b	b	b
2		a	a	a	a	c	c	c	c	c	c	c	c	c	c	b	b	b	b	b	b	b	b	b	b
3			a	a	a	a	c	c	c	c	c	c	c	c	c	b	b	b	b	b	b	b	b	b	b
4				a	a	a	a	c	c	c	c	c	c	c	c	b	b	b	b	b	b	b	b	b	b
5					a	a	a	a	c	c	c	c	c	c	c	b	b	b	b	b	b	b	b	b	b
6						a	a	a	a	*	*	*	*	*	*	*	*	*	*	*	*	b	b	b	b
7							a	a	a	a	*	*	*	*	*	*	*	*	*	*	*	b	b	b	b
d	8							a	a	a	a	*	*	*	*	*	*	*	*	*	*	b	b	b	b
a	9								a	a	a	a	*	*	*	*	*	*	*	*	*	b	b	b	b
t	10									a	a	a	a	*	*	*	*	*	*	*	*	b	b	b	b
e	11										a	a	a	a	*	*	*	*	*	*	*	b	b	b	b
	12											a	a	a	a	*	*	*	*	*	*	b	b	b	b
e	13												a	a	a	a	*	*	*	*	*	b	b	b	b
n	14													a	a	a	a	*	*	*	*	b	b	b	b
t	15														a	a	a	a	*	*	*	*	b	b	b
r	16															a	a	a	a	*	*	b	b	b	b
é	17																a	a	a	a	b	b	b	b	
e	18																	a	a	a	a	b	b	b	
	19																		a	a	a	a	b	b	
	20																			a	a	a	a	b	
	21																				a	a	a	a	
	22																					a	a	a	
	23																						a	a	
	24																							a	
	25																								a

NOTES :

Les numéros verticaux représentent la date d'entrée dans le régime et les numéros horizontaux représentent la date de sortie du régime.

a : indique que cette date ne peut pas être considérée comme une date de rupture car le régime correspondant contient un nombre d'observations inférieur à 5.

b : indique que cette date ne peut pas être considérée comme une date de rupture car il n'y aura plus de place pour 3 régimes de taille 5.

c : indique que cette date ne peut pas être considérée comme une date de rupture car il n'y aura pas de place pour le régime de taille 5 précédent.

* : indique que ce régime est admissible.

FIGURE 1. un exemple de la matrice triangulaire des sommes des carrés des résidus avec

$$n = 25, m = 2 \text{ et } h = \lceil \eta n \rceil = 5.$$

Les axes horizontaux représentent la date d'entrée dans un régime quant aux axes verticaux ils représentent la date de sortie. Chaque entrée représente une somme des carrés des résidus correspondant au régime concerné. La somme globale des carrés des résidus pour toute partition $(\hat{T}_1, \dots, \hat{T}_m)$, et pour toute valeur m doit être impérativement une des combinaisons linéaires de ces $n(n+1)/2$ sommes des carrés des résidus. L'algorithme de ce programme dynamique peut être considéré comme une procédure efficace pour faire la comparaison entre toutes les combinaisons possibles et atteindre à la fin le minimum de la somme globale des carrés des résidus.

Dans la pratique, moins que $n(n+1)/2$ régimes sont admissibles. D'abord une certaine taille minimale de chaque régime est imposée, dans notre cas c'est $h = \lceil \eta n \rceil$. Cela implique une réduction de $(h-1)n - (h-2)(h-1)/2$ dans le nombre de régimes considérés (voir Figure 1). Il y a d'autres réductions qui peuvent être envisagées comme par exemple imposer à chaque régime de ne pas dépasser une certaine taille afin de permettre l'existence de m régimes supplémentaires, ce qui implique une réduction de $h^2m(m+1)/2$. Finalement la date d'entrée du premier régime ne peut être que 1 et cela réduit encore le nombre de régimes considérés de $n(h-1) - mh(h-1) - (h-1)^2 - h(h-1)/2$. Nous discutons après comment cette matrice triangulaire des sommes des carrés des résidus, peut être construite et utilisée dans le but d'obtenir les minimums globaux dans le cas d'un modèle de changement structurel pur.

3.2. L'algorithme de la programmation dynamique

Une fois que les sommes des carrés des résidus, des régimes concernés, sont calculées et stockées, l'approche de programmation dynamique sert à préciser la partition qui atteint la minimisation globale de toutes les sommes des carrés des résidus. Cette méthode procède essentiellement selon un examen séquentiel d'une partition contenant un seul point de rupture. Soit $SSR(\{n_{m,T}\})$ la somme des carrés des résidus associée à la partition optimale contenant m ruptures sur les n observations. La partition

optimale résoud le problème récursif suivant :

$$SSR(\{n_m, T\}) = \min_{mh \leq j \leq n-h} [SSR(\{T_{m-1, j}\}) + SSR(\{j+1, n\})] \quad (3.1)$$

plus explicitement nous pouvons écrire cette relation ainsi :

$$SSR(\{n_m, T\}) = \min_{mh \leq j_1 \leq n-h} [SSR(j_1+1, n) + \min_{(m-1)h \leq j_2 \leq j_1-h} [SSR(j_2+1, j_1) + \min_{(m-2)h \leq j_3 \leq j_2-h} [SSR(j_3+1, j_2) + \dots \min_{h \leq j_m \leq j_{m-1}-h} [SSR(1, j_m) + SSR(j_m+1, j_{m-1})] \dots]]]$$

Nous commençons la procédure par évaluer la partition optimale avec un seul point de rupture pour tous les sous-échantillons qui ont une date d'entrée 1 et une date de sortie allant de h jusqu'à $n - mh$. Ainsi la première étape est de stocker un ensemble de $n - (m+1)h + 1$ partitions optimales avec un seul point de rupture accompagnées de leurs sommes des carrés des résidus. Chacune des partitions optimales correspond à un sous-échantillon dont la date de sortie est comprise entre $2h$ et $n - (m-1)h$.

La seconde étape consiste à chercher une partition optimale avec deux points de ruptures. En effet, nous cherchons la combinaison minimale en faisant la somme des sommes des carrés des résidus du premier ensemble et celles du deuxième ensemble des régimes possibles pour chaque date de sortie. Le résultat est un ensemble de $n - (m+1)h + 1$ partitions optimales à deux points de ruptures. Telles partitions correspondent aux sous-échantillons dont les dates de sortie sont comprises entre $3h$ et $n - (m-2)h$. La méthode est exécutée séquentiellement jusqu'à l'obtention d'un ensemble de $n - (m+1)h + 1$ partitions optimales avec m points de ruptures. L'étape finale consiste à trouver parmi ces partitions optimales avec m points de ruptures celle qui nous mène à la valeur minimale des sommes des carrés des résidus.

Il est important de noter que, dans la pratique, cette méthode est très rapide. En effet, le seul élément coûteux dans le calcul est la construction de la matrice triangulaire des sommes des carrés des résidus de tous les régimes possibles. Le temps de recherche de la partition optimale est négligeable par rapport à celui consacré à tout le calcul, c'est à dire qu'il n'y a pas une grande différence au niveau du coût entre obtenir les minimums globaux avec dix et deux points de ruptures.

4. ESTIMATION DU NOMBRE DE RUPTURES

4.1. La sélection du modèle par les critères d'information

Pour déterminer la dimension d'un modèle, nous pouvons utiliser les critères d'information. L'idée principale est que l'ajout d'un point de rupture réduit forcément la somme des carrés des résidus et par conséquent nous devons pénaliser cette addition. La pénalité doit obliger l'estimateur du nombre des dates de ruptures à converger rapidement vers la vraie valeur afin d'assurer les propriétés asymptotiques des autres estimateurs qui en dépendent. En effet, une lourde pénalité réduit énormément la probabilité d'une surestimation mais pas excessivement car nous risquons de sousestimer le modèle. Le terme pénalisant diffère d'un critère à un autre et comme nous le verrons dans les simulations cela a un effet sur les résultats.

Yao (1988) suggère l'utilisation du critère d'informations Bayésien défini comme suit :

$$BIC(m) = \ln \left(Q_T \left(\hat{T}_1, \dots, \hat{T}_m \right) / n \right) + p^* \ln(n) / n \quad (4.1)$$

où $p^* = (m+1)q + m + p$ est le nombre de paramètres inconnus. Il montre que l'estimateur du nombre de ruptures \hat{m} est un estimateur consistant (pour une suite de variables aléatoires normales avec un changement au niveau de la moyenne) de la vraie valeur m , pourvu que $m \leq M$, avec M supposée connue.

Un autre critère proposé par Yao et Au (1989) est le suivant :

$$YIC(m) = \ln \left(Q_T \left(\hat{T}_1, \dots, \hat{T}_m \right) / (n - m) \right) + mC_n/n \quad (4.2)$$

où $\{C_n\}$ est une suite qui satisfait la propriété suivante : $C_n n^{-2/r} \rightarrow \infty$ et $C_n/n \rightarrow 0$ quand $n \rightarrow \infty$ pour tout entier positif r sachant que le moment d'ordre $2r$ du terme d'erreur est fini pour tout $r \geq 3$. Dans la partie simulation, nous utilisons la suite $C_n = 0.368n^{0.7}$ suggérée par Liu et al (1997).

Liu et al (1997) ont proposé le critère de Schwarz modifié :

$$MIC(m) = \ln \left(Q_T \left(\hat{T}_1, \dots, \hat{T}_m \right) / (n - p^*) \right) + p^* c_0 [\ln(n)]^{2+\delta_0} / n \quad (4.3)$$

tel que $\delta_0 = 0.1$ et $c_0 = 0.299$ en se basant sur la performance de l'estimateur du nombre de changements obtenu par ce critère après plusieurs expériences de simulation. L'estimateur \hat{m} du nombre de ruptures est obtenu en minimisant ces critères d'informations, étant donné la valeur maximale M et sachant que dans notre cas $p = 0$ et $q = 1$. Il faut noter que ces critères d'informations ne tiennent pas compte de la présence de corrélation entre les termes d'erreur.

4.2. La méthode de Bai et Perron (1998)

4.2.1. Le test de l ruptures contre $l+1$

Cette section présente un test d'hypothèse nulle de l ruptures contre l'hypothèse alternative de l'existence d'une rupture supplémentaire. Le test se base sur la différence entre la somme des carrés des résidus obtenue avec l ruptures et celle obtenue avec $l + 1$ ruptures. La distribution limite de la statistique de ce test est malheureusement difficile à obtenir, par conséquent Bai et Perron (1998) adoptent une stratégie différente. Pour le modèle avec l ruptures, les points de ruptures estimés et notés par $\widehat{T}_1, \dots, \widehat{T}_l$, sont obtenus par une minimisation globale de la somme des carrés des résidus. La stratégie de Bai et Perron (1998) consiste à tester pour chacun des $(l + 1)$ régimes (obtenus en utilisant la partition estimée $\widehat{T}_1, \dots, \widehat{T}_l$) la présence d'une rupture supplémentaire.

L'idée est d'effectuer $(l + 1)$ tests de l'hypothèse nulle d'absence de changement structurel contre l'hypothèse alternative de présence d'un changement. Ce test est appliqué à chaque régime i contenant les observations entre la date $\widehat{T}_{i-1} + 1$ et la date \widehat{T}_i ($i = 1, \dots, l + 1$) tout en respectant la convention $\widehat{T}_0 = 0$ et $\widehat{T}_{l+1} = n$. Nous rejetons l'hypothèse nulle, c'est à dire nous optons pour un modèle avec $(l + 1)$ ruptures si la valeur minimale de la somme des carrés des résidus, après avoir inclus une rupture supplémentaire, régime par régime, est suffisamment inférieure à celle du modèle avec l ruptures. Les dates de ruptures choisies sont par conséquent celles associées à cette valeur minimale. Pour plus de détails le test est défini comme suit :

$$F_T(l + 1/l) = \left\{ Q_T(\widehat{T}_1, \dots, \widehat{T}_l) - \min_{1 \leq i \leq l+1} \inf_{\lambda \in \Lambda_{i,\eta}} Q_T(\widehat{T}_1, \dots, \widehat{T}_{i-1}, \lambda, \widehat{T}_i, \dots, \widehat{T}_l) \right\} / \widehat{\sigma}^2 \quad (4.4)$$

où $\Lambda_{i,\eta} = \left\{ \lambda; \widehat{T}_{i-1} + (\widehat{T}_i - \widehat{T}_{i-1})\eta \leq \lambda \leq \widehat{T}_i - (\widehat{T}_i - \widehat{T}_{i-1})\eta \right\}$ et $\widehat{\sigma}^2$ est un estimateur consistant de σ^2 sous l'hypothèse nulle. Nous pouvons obtenir la valeur critique de ce test pour les différentes valeurs de l à partir de la distribution de la fonction $G_{q,\eta}(x)$ (voir Bai et Perron (1998), page 60). Après plusieurs simulations effectuées par Bai et Perron, les valeurs critiques, pour $\eta = 0.05$ et $q = 1$, sont données dans la table 1.

	l									
$1 - \alpha$	0	1	2	3	4	5	6	7	8	9
0.90	8.02	9.56	10.45	11.07	11.65	12.07	12.47	12.70	13.07	13.34
0.95	9.63	11.14	12.16	12.83	13.45	14.05	14.29	14.50	14.69	14.88
0.975	11.17	12.88	14.05	14.50	15.03	15.37	15.56	15.73	16.02	16.39
0.99	13.58	15.03	15.62	16.39	16.60	16.90	17.04	17.27	17.32	17.61

TABLE 1. Les valeurs critiques asymptotiques du test $F_T(l + 1/l)$

4.2.2. La méthode d'estimation séquentielle

La stratégie de cette méthode est de commencer par l'estimation d'un modèle avec un petit nombre de ruptures que nous pensons être nécessaire ou bien d'un modèle avec zéro ruptures, ensuite nous ajoutons un point de rupture au régime associé à un rejet obtenu par le test $F_T(l + 1/l)$. Cette procédure est répétée tout en augmentant séquentiellement le nombre de ruptures l jusqu'à ce que le test échoue à rejeter l'hypothèse nulle c'est à dire l'existence d'un point de rupture additionnel. La distribution limite de ce test reste la même quand nous utilisons les dates de ruptures estimées par la minimisation globale. Le nombre final de ruptures est donc égal au nombre de rejets obtenus par le test $F_T(l + 1/l)$ plus le nombre de ruptures utilisé initialement. L'avantage de cette méthode sur celle des critères d'informations est que celle-ci peut permettre la présence de corrélation entre les termes d'erreur.

4.3. La méthode de Lavielle (2004)

4.3.1. La fonction de pénalité pour le problème de changement structurel

Soit $K = m + 1$ le nombre de régimes, quand ce nombre est inconnu nous pouvons l'estimer en minimisant une version pénalisée de

$$\begin{aligned}
 Q_{T,K}(T_1, \dots, T_{K-1}) &= \min_{(\delta_1, \dots, \delta_K)} \sum_{i=1}^K \sum_{j=T_{i-1}+1}^{T_i} \left(y_j - \hat{\delta}_i \right)^2 \\
 &= \sum_{i=1}^K \sum_{j=T_{i-1}+1}^{T_i} \left(y_j - \bar{y}(T_{i-1}, T_i) \right)^2
 \end{aligned} \tag{4.5}$$

où $\bar{y}(T_{i-1}, T_i) = \frac{1}{T_i - T_{i-1}} \sum_{j=T_{i-1}+1}^{T_i} y_j$.

Soit $pen(T_1, \dots, T_{K-1})$ une fonction qui augmente avec le nombre de régimes K et soit $\hat{T} = (\hat{T}_1, \dots, \hat{T}_{K-1})$ la suite des dates de ruptures qui minimisent :

$$H(T) = Q_{T,K}(T) + \beta_n pen(K), \quad (4.6)$$

β_n est une fonction de n qui tend vers 0, lorsque n tend vers l'infini et nous supposons que $K(\hat{T})$ converge en probabilité vers le vrai nombre de régimes $K(T)$. D'autres auteurs suggèrent des fonctions différentes de pénalité $pen(K)$ selon les modèles qu'ils choisissent. Par définition, $pen(K)$ augmente lorsque le nombre de régimes K augmente. Pour des raisons pratiques et suivant les critères d'informations AIC et Schwarz, Lavielle et Moulines (1996) suggèrent la fonction de pénalité la plus simple $pen(K) = K(T)$.

4.3.2. Un choix adaptatif du paramètre de pénalisation

Etant donné la fonction contraste $Q_{T,K}(T)$ et la fonction de pénalité $pen(K)$, le problème se réduit au choix du paramètre β . Soit K_{MAX} le nombre maximum de régimes possible et pour tout $1 \leq K \leq K_{MAX}$, soit :

$\zeta_K = \{T = (T_0, \dots, T_K), \text{ tel que } T_0 = 0 < T_1 < T_2 < \dots < T_{K-1} < T_K = n\}$ l'ensemble de tous les modèles de dimension K . Par définition le meilleur modèle de dimension K est celui qui minimise la fonction contraste $Q_{T,K}$, ainsi on aura $\hat{T}_K = \arg \min_{T \in \zeta_K} Q_{T,K}(T)$. Pour le calcul nous utilisons l'algorithme efficace, développé par Bai et Perron (1998), basé sur le principe de la programmation dynamique et qui permet l'obtention des minimums globaux en utilisant un nombre de sommes des carrés des résidus d'ordre $O(n^2)$ pour tout $m \geq 2$. Donc,

$$\begin{aligned} \hat{T}(\beta) &= \arg \min_T \{Q_{T,K}(T) + \beta_n pen(K)\} \\ &= \hat{T}_{\hat{K}(\beta)} \end{aligned} \quad (4.7)$$

où $\hat{K}(\beta) = \arg \min_{K \geq 1} \{Q_{\hat{T},K} + \beta pen(K)\}$

La manière dont la solution $\hat{K}(\beta)$ varie avec le paramètre de pénalisation β est présentée dans la proposition suivante :

Proposition : il existe une suite $K_1 = 1 < K_2 < \dots$, et une suite $\beta_0 = \infty > \beta_1 > \dots$, avec $\beta_i = \frac{Q_{T,K_i} - Q_{T,K_{i+1}}}{pen_{K_{i+1}} - pen_{K_i}}$, $i \geq 1$ tel que $\hat{K}(\beta) = K_i, \forall \beta \in (\beta_i, \beta_{i-1})$. La suite estimée \hat{T} ne devrait

pas beaucoup dépendre du coefficient pénalisant β c'est à dire un léger changement de ce dernier n'implique pas une différence radicale au niveau de la solution \hat{T} . Cette stabilité par rapport au choix du paramètre β sera assurée si et seulement si nous retenons l'intervalle le plus large $([\beta_i, \beta_{i-1}], i \geq 1)$.

Pour résumer, Lavielle propose la procédure suivante :

1. Pour $K = 1, 2, \dots, K_{\max}$, nous calculons \hat{T}_K , $Q_{T,K} = Q_T(\hat{T}_K)$ et $pen_K = pen(\hat{T}_K)$.
2. Nous calculons les suites (K_i) , (β_i) et les longueurs (l_i) des intervalles $([\beta_i, \beta_{i-1}])$.
3. Nous retenons la plus grande valeur de K_i tel que $l_i \gg l_j$ pour $j > i$.

Le fait de choisir l'intervalle le plus large, sousestime souvent le nombre de changements. En effet, cet intervalle correspond à un très petit nombre de ruptures et nous n'aurons que le changement le plus radical avec cette pénalité, ce qui explique pourquoi nous devrions mieux chercher la dimension la plus élevée K_i tel que $l_i \gg l_j$ pour $j > i$ c'est à dire l_j est significativement inférieure à l_i pour $j > i$ et non choisir la dimension K_i associée à la longueur l_i la plus élevée. Cette méthode n'est pas tout à fait pratique pour nos expériences de simulations parce que nous serons obligés d'étudier les réalisations cas par cas, ce qui demande énormément de temps. Pour résoudre ce problème, Lavielle (2004) a mis au point une procédure automatique pour estimer le nombre de régimes K .

4.3.3. La procédure automatique pour estimer le nombre de régimes K

Lavielle (2004) a automatisé l'étape 3 de la procédure, que nous avons évoqué, pour des raisons pratiques, et il propose l'algorithme suivant pour estimer la dimension K :

1. Pour tout $1 \leq K \leq K_{MAX}$, soit $\tilde{Q}_{T,K} = \frac{Q_{T,K_{MAX}} - Q_{T,K}}{Q_{T,K_{MAX}} - Q_{T,1}} (K_{MAX} - 1) + 1$ avec $\tilde{Q}_{T,1} = K_{MAX}$ et $\tilde{Q}_{T,K_{MAX}} = 1$. Cette suite décroît avec une pente de moyenne égale à -1 .

2. Pour tout $2 \leq K \leq K_{MAX} - 1$, soit $D_K = \tilde{Q}_{T,K-1} - 2\tilde{Q}_{T,K} + \tilde{Q}_{T,K+1}$ et $D_1 = \infty$, l'estimation du nombre de régimes est donné par : $\hat{K} = \max \{1 \leq K \leq K_{MAX} - 1 \text{ tel que } D_K > S\}$.

\hat{K} est défini comme la plus grande valeur de K tel que la dérivée seconde de $Q_{T,K}$ est supérieure à un certain seuil S . Si aucune dérivée seconde n'est supérieure à S , nous concluons qu'il n'y a pas de changement structurel et que $\hat{K} = 1$. Malheureusement, nous ne pouvons pas obtenir la distribution de probabilité de la statistique D_K et S ne peut pas être calculé comme un quantile de cette distribution. Néanmoins, après plusieurs et diverses expériences numériques, Lavielle propose la valeur 0.75 pour S . Les valeurs inférieures à 0.75 souvent surestiment le nombre de régimes, cependant les valeurs supérieures sousestiment ce nombre.

5. LES EXPERIENCES DE SIMULATIONS

Nous considérons le modèle (1) avec $m = 2$, $\lambda_1 = 0.25$, $\lambda_2 = 0.5$, $\delta_1 = 2$, $\delta_2 = 0$ et $\delta_3 = 1$. Le terme d'erreur u_t est un bruit gaussien fractionnaire tel que $(1 - L)^d u_t = \varepsilon_t$ où $\varepsilon_t \rightsquigarrow N(0, 1)$ et $|d| < 0.5$, ainsi u_t est un processus stationnaire au second ordre avec une moyenne nulle et dont la fonction de densité spectrale vérifie :

$$f(\omega) \approx C\omega^{-2d} \text{ lorsque } \omega \rightarrow 0^+ \quad (5.1)$$

où d est le paramètre de mémoire longue et C est une constante.

Une réalisation du processus Y avec $T=1000$, $d=0.30$ et les points de changement $\lambda_1 = 0.25$ et $\lambda_2 = 0.5$.

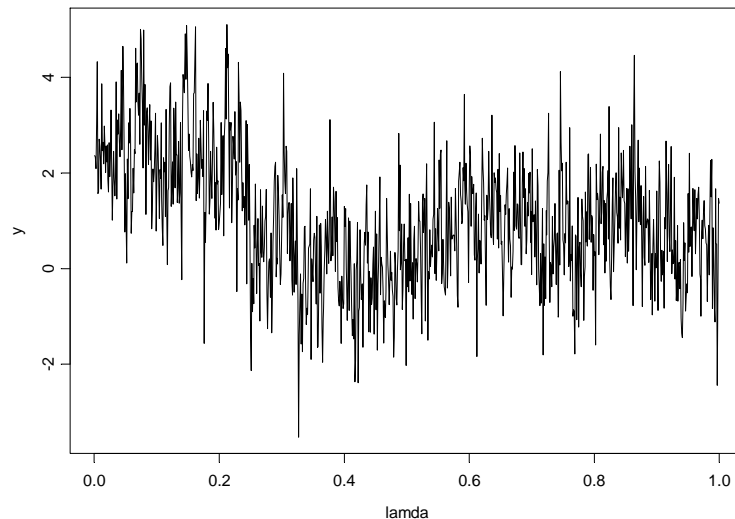


FIGURE 2. Une réalisation de Y , $T=1000$, $d=0.30$ et les points de rupture $\lambda_1 = 0.25$ et $\lambda_2 = 0.5$.

Le nombre minimal d'observations dans chaque régime est égale à $[\eta n]$ avec $\eta = 0.05$. Le nombre maximal de ruptures permis est $M = 7$. On considère trois tailles d'échantillon $n = \{500, 1000, 2000\}$ et cinq valeurs pour le paramètre d'intégration $d = \{0.10; 0.20; 0.30; 0.40; 0.49\}$. Tous les résultats de simulations reportés sont basés sur 100 réalisations et se trouvent dans l'annexe.

Expérience 1 : Analyse des critères d'information

Nous estimons un modèle avec changement dans la moyenne par les critères d'informations BIC, MIC et YIC. Les résultats sont résumés dans les tables 2, 3 et 4, avec différentes valeurs de d et de n . Dans le cas où $n = 500$, les résultats sont assez similaires pour les critères d'informations MIC et YIC, mais lorsque n augmente MIC agit comme BIC c'est à dire il choisissent le nombre maximum

de changements possibles quand d augmente et nous nous approchons de la non-stationarité. Pour $d = 0.49$, nous remarquons que la fréquence du choix de $\hat{m} = 2$ a tendance à baisser quand n augmente pour MIC et YIC, par contre pour BIC elle est toujours nulle. Donc, d'après les trois tables, nous remarquons que pour tous les critères d'information, la fréquence du choix de $\hat{m} = 2$ est très raisonnable quand $d = 0.10$ mais elle baisse quand d augmente, d'un autre côté la fréquence du choix $\hat{m} = M$ augmente quand n augmente et elle est toujours la plus importante, sauf pour le cas de YIC où l'estimateur du nombre de ruptures a gardé une distribution sur l'ensemble $\{0, 1, \dots, M\}$ assez stable, sans favoriser M , quand n augmente. Ainsi, une lourde pénalité réduit la tendance de choisir le nombre maximum de ruptures possibles. Ces conclusions sont en quelque sorte attendues puisque comme nous l'avons déjà dit, les critères d'informations ne tiennent pas compte de la présence de corrélation entre les termes d'erreur.

Expérience 2 : Méthode $F_T(l + 1/l)$ de Bai et Perron (1998)

Avec cette méthode, la corrélation et l'hétéroscédasticité des résidus sont permises. Les résultats sont présentés dans les tables 5, 6 et 7 avec différents choix de d et n et quatre valeurs pour le niveau de significativité α . De point de vue variation de α , nous remarquons que ça n'a pas beaucoup d'influence sur les résultats, donc nous allons nous concentrer sur la variation de n et de d uniquement. Lorsque d augmente, nous remarquons que la fréquence du choix de $\hat{m} = 2$ ne cesse de baisser mais elle reste tout de même la plus importante dans la distribution de l'estimateur du nombre de ruptures jusqu'à $d = 0.40$, donc c'est assez bon comme résultat pour un paramètre d'intégration élevé. A partir de cette valeur c'est à dire pour $d = 0.49$, la plus grande fréquence correspond à $\hat{m} = 1$ et c'est tout à fait compréhensible puisque nous frôlons la non-stationarité. Quant à la variation de la taille de l'échantillon, nous remarquons que lorsque n augmente, la fréquence du choix $\hat{m} = 2$ augmente à son tour jusqu'à $d = 0.30$, par la suite cette fréquence commence plus ou moins à baisser. D'après ce qui précède, nous pouvons conclure que ces résultats sont nettement meilleurs et plus raisonnables que ceux obtenus par les critères d'information cependant ils restent insatisfaisants dans le cas de forte dépendance.

Expérience 3 : Analyse de la procédure automatique de Lavielle(2004)

Les résultats sont donnés dans la table 8 avec différentes valeurs de d et n . Lorsque d augmente, la fréquence du choix $\hat{m} = 2$ ne cesse de baisser mais elle reste la plus importante dans la distribution de l'estimateur du nombre de ruptures pour toute valeur de d , y compris $d = 0.49$ ce qui manifeste l'avantage de cette méthode sur celle de Bai et Perron. Lorsque n augmente, nous voyons bien que la fréquence du choix de $\hat{m} = 2$ augmente aussi, et ça se produit pratiquement pour toute valeur de d

d'où l'amélioration apportée par cette méthode aux résultats.

Expérience 4 : Estimation des dates de ruptures

Une fois que nous avons estimé le nombre de ruptures par les méthodes présentées précédemment, nous estimons les dates de ruptures grâce à l'algorithme de la programmation dynamique, après avoir calculé les minimums globaux à partir de la matrice triangulaire pour $m = 2$ et $\eta = 0.05$. Les distributions empiriques des estimateurs des points de ruptures $\hat{\lambda}_1$ et $\hat{\lambda}_2$, sont données par les histogrammes des figures 3, 4, 5, 6, et 7 pour des valeurs différentes de n et d . Nous remarquons d'après les graphiques que la fréquence de détection de la première date de rupture est plus élevée que celle de la seconde, pour toute valeur d , ajoutons à ça que sa distribution est plus concentrée autour de la vraie valeur λ ce qui veut dire que cette première date est estimée avec plus de précision puisque elle possède une amplitude de changement plus large. Un autre élément, aussi important, est que les distributions des dates de ruptures ont tendance à être de plus en plus dispersées lorsque d augmente et varie entre 0.10 et 0.49 pour n fixe. Nous remarquons aussi que les distributions des dates de ruptures se concentrent de plus en plus autour des vraies valeurs $\lambda_1 = 0.25$ et $\lambda_2 = 0.5$ quand $n \rightarrow \infty$. Donc nous pouvons conclure à la fin que la présence de la mémoire longue dans le processus affecte la précision des estimations des dates de ruptures.

6. CONCLUSION

Ce document a étudié le problème de sélection du nombre de ruptures dans un modèle dont le terme d'erreur est un processus $I(d)$. Pour les critères d'information, les résultats se détériorent au fur et à mesure que nous augmentons le coefficient d'intégration d , et nous remarquons une tendance à choisir le nombre maximum de changements possibles puisque les critères d'information ne tiennent pas compte de la présence de corrélation entre les termes d'erreur. Pour remédier à cette défaillance, nous avons expérimenté d'autres méthodes qui fonctionnent sous des conditions plus générales. En utilisant le test $F_T(l + 1/l)$ de Bai et Perron, la fréquence du choix de la vraie valeur m est la plus élevée, sauf pour $d = 0.49$ par contre avec la procédure automatique mise au point par Lavielle, la fréquence du choix de la vraie valeur m est la plus élevée pour toute valeur d , ainsi cette méthode est efficace même si nous sommes très proche de la non-stationarité. Afin d'estimer les dates de ruptures, nous avons eu recours à un algorithme efficace pour obtenir les minimums globaux des sommes des carrés des résidus et nous avons pu conclure que la présence de la mémoire longue dans le processus affecte la précision des estimations de ces dates. En conclusion, il est intéressant d'étendre notre travail sur des modèles plus généraux tel que les modèles avec changement structurel partiel.

RÉFÉRENCES

- [1] Bai, J. (1994a). Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis*, 15, 453-472.
- [2] Bai, J. and Perron, P.(1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66, 47-78.
- [3] Bai, J. and Perron, P.(2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18, 1-22.
- [4] Berran, J. (1994). A Goodness-of-fit test for time series with long range dependence. *Journal of The Royal Statistical Society, Ser. B*, 54, 749-760.
- [5] Berran, J. (1994). *Statistics for long memory processes*. London : Chapman and Hall.
- [6] Boutahar, M. and Jouini, J.(2004). Structural change in the long-memory environment. *Document de travail n°2004 – 42*.
- [7] Boutahar, M. and Jouini, J.(1996). Spuriousness of informations creteria in selecting the number of changes. *Document de travail n°03A17*.
- [8] Diebold, F. X., and Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics* , 105, 131-159.
- [9] Engle, R. F. And Smith, A. D. (1999). Stochastic permanent breaks. *Review of Economics and Statistics*, 81, 553-574.
- [10] Fox , R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics*, 14, 517-532.
- [11] Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis* , 4, 15-39.
- [12] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. *Probability Theory and Related Fields*, 86, 87-104.
- [13] Granger , C. W. J. and Hyung, N. (1999). Occasional structural breaks and long memory. *Discussion paper 99-14, University of California, San Diego*.
- [14] Grenander, U. (1954). On estimation of regression coefficients in the case of autocorrelated disturbance. *Annals of Mathematical Statistics*, 25, 252-272.
- [15] Hassler, U., and Wolters, J. (1995). Long Memory in Inflation Rates : International Evidence. *Journal of Business and Economic Statistics*, 13, 37-45.

- [16] Hwang, S. (1996). Structural breaks in AR(1) process and persistence. *Manuscript, Faculty of Finance, Cass Business School*.
- [17] Lavielle, M. and Moulines, E. (1996). Least squares estimations of an unknown number of shifts in a time series. *Journal of Time Series Analysis*, 21, No. 1.
- [18] Lavielle, M. (1999). Detection of multiple changes in a sequence of dependent variables, stoch. *Proc. Appl.* 83, 79-102.
- [19] Lavielle, M. (2004). Using penalized contrasts for the change-point problem. *Elsevier*.
- [20] Perron, P. (1997). Estimation de modèles avec changements structurels multiples. *Actualité Economique*, 73, 457-505.
- [21] Schwarz, G. (1987). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- [22] Yajima, Y. (1988). On estimation of regression model with long-memory stationary errors. *Annals of Statistics*, 16, 791-807.
- [23] Yajima, Y. (1988). Asymptotic properties of the LSE in a regression model with long-memory stationary errors. *Annals of Statistics*, 19, 158-177.
- [24] Yao, Y.C. (1988). Estimating the number of change-points via schwarz's criterion. *Statistics and Probability Letters*, 6, 181-189.
- [25] Yao, Y.C, and Au, S. T. (1989). Least squares estimations of a step function. *Sankhya, Ser. A*, 51, 370-381.

ANNEXE : RÉSULTATS DE SIMULATIONS

TABLE 2 : Pourcentage du nombre de ruptures \hat{m} avec différentes valeurs de d , $n=500$.

d \hat{m}	BIC					MIC					YIC				
	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.01	0.00	0.00	0.10	0.02	0.11	0.10	0.03	0.01	0.02	0.11	0.08	0.03
2	0.81	0.33	0.06	0.01	0.00	0.99	0.92	0.69	0.31	0.14	0.99	0.89	0.66	0.29	0.12
3	0.10	0.31	0.17	0.03	0.01	0.00	0.05	0.18	0.35	0.25	0.00	0.08	0.19	0.37	0.19
4	0.08	0.25	0.24	0.10	0.03	0.00	0.01	0.01	0.17	0.23	0.00	0.01	0.03	0.17	0.26
5	0.01	0.06	0.18	0.14	0.04	0.00	0.00	0.01	0.06	0.19	0.00	0.00	0.01	0.06	0.21
6	0.00	0.04	0.19	0.31	0.17	0.00	0.00	0.00	0.01	0.12	0.00	0.00	0.00	0.03	0.13
7	0.00	0.01	0.15	0.41	0.75	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.06

TABLE 3 : Pourcentage du nombre de ruptures \hat{m} avec différentes valeurs de d , $n=1000$.

d \hat{m}	BIC					MIC					YIC				
	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.00	0.01	0.06	0.04	0.00
2	0.69	0.21	0.01	0.00	0.00	0.99	0.88	0.48	0.11	0.05	1.00	0.96	0.60	0.25	0.09
3	0.25	0.25	0.03	0.00	0.00	0.01	0.12	0.35	0.26	0.03	0.00	0.03	0.31	0.35	0.12
4	0.05	0.27	0.14	0.00	0.01	0.00	0.00	0.10	0.28	0.18	0.00	0.00	0.03	0.25	0.36
5	0.01	0.16	0.23	0.06	0.00	0.00	0.00	0.03	0.19	0.25	0.00	0.00	0.00	0.09	0.20
6	0.00	0.06	0.20	0.12	0.02	0.00	0.00	0.01	0.13	0.26	0.00	0.00	0.00	0.02	0.11
7	0.00	0.05	0.39	0.82	0.97	0.00	0.00	0.00	0.02	0.23	0.00	0.00	0.00	0.00	0.12

TABLE 4 : Pourcentage du nombre de ruptures \hat{m} avec différentes valeurs de d , $n=2000$.

d \hat{m}	BIC					MIC					YIC				
	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.02
2	0.67	0.09	0.00	0.00	0.00	1.00	0.89	0.40	0.04	0.01	1.00	0.99	0.83	0.22	0.08
3	0.23	0.21	0.01	0.00	0.00	0.00	0.11	0.29	0.07	0.06	0.00	0.01	0.14	0.35	0.22
4	0.08	0.22	0.03	0.00	0.00	0.00	0.00	0.20	0.14	0.04	0.00	0.00	0.02	0.27	0.21
5	0.02	0.18	0.04	0.01	0.00	0.00	0.00	0.10	0.28	0.15	0.00	0.00	0.00	0.08	0.18
6	0.00	0.18	0.23	0.07	0.00	0.00	0.00	0.01	0.30	0.25	0.00	0.00	0.00	0.02	0.14
7	0.00	0.12	0.69	0.92	1.00	0.00	0.00	0.00	0.17	0.49	0.00	0.00	0.00	0.01	0.15

TABLE 5 : Pourcentage du nombre de ruptures \hat{m} avec différentes valeurs de α et de d , $n=500$.

$1 - \alpha$	0.90					0.95				
d \hat{m}	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.04	0.01	0.00
2	0.93	0.61	0.40	0.26	0.11	0.98	0.72	0.48	0.31	0.19
3	0.07	0.33	0.38	0.33	0.30	0.02	0.24	0.33	0.38	0.34
4	0.00	0.05	0.17	0.26	0.32	0.00	0.03	0.15	0.23	0.26
5	0.00	0.01	0.01	0.13	0.15	0.00	0.01	0.00	0.06	0.12
6	0.00	0.00	0.00	0.01	0.08	0.00	0.00	0.00	0.01	0.08
7	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.01
$1 - \alpha$	0.975					0.99				
d \hat{m}	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.01	0.04	0.22	0.00	0.00	0.01	0.04	0.21
1	0.04	0.12	0.22	0.28	0.40	0.04	0.14	0.23	0.29	0.39
2	0.96	0.88	0.66	0.47	0.26	0.96	0.86	0.65	0.48	0.22
3	0.00	0.00	0.11	0.19	0.12	0.00	0.00	0.11	0.18	0.14
4	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.01	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 6 : Pourcentage du nombre de ruptures \hat{m} avec différentes valeurs de α et de d , $n=1000$.

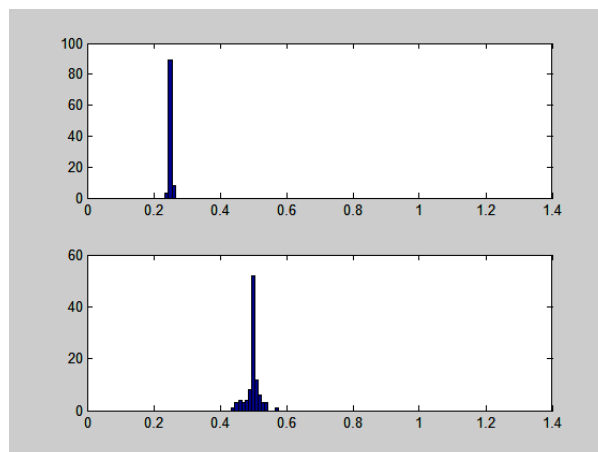
$1 - \alpha$	0.90					0.95				
d \hat{m}	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.79	0.61	0.41	0.01	0.07	0.87	0.55	0.21	0.04	0.08
3	0.21	0.33	0.37	0.19	0.08	0.13	0.35	0.32	0.23	0.12
4	0.00	0.05	0.19	0.27	0.24	0.00	0.09	0.24	0.31	0.28
5	0.00	0.01	0.03	0.20	0.29	0.00	0.01	0.20	0.23	0.31
6	0.00	0.00	0.00	0.15	0.15	0.00	0.00	0.03	0.15	0.15
7	0.00	0.00	0.00	0.18	0.17	0.00	0.00	0.00	0.04	0.06
$1 - \alpha$	0.975					0.99				
d \hat{m}	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.01	0.00	0.08	0.00	0.00	0.01	0.00	0.08
1	0.00	0.02	0.08	0.25	0.51	0.00	0.02	0.08	0.26	0.52
2	1.00	0.96	0.71	0.41	0.21	1.00	0.97	0.74	0.42	0.21
3	0.00	0.02	0.17	0.25	0.13	0.00	0.01	0.16	0.25	0.12
4	0.00	0.00	0.03	0.08	0.06	0.00	0.00	0.01	0.07	0.06
5	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 7 : Pourcentage du nombre de ruptures \hat{m} avec différentes valeurs de α et de d , $n=2000$.

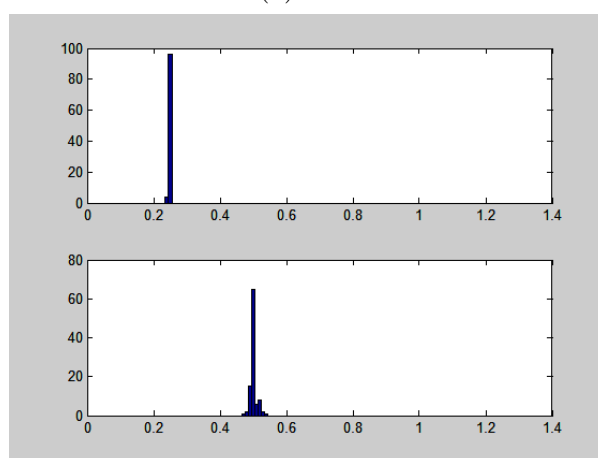
$1 - \alpha$	0.975					0.99				
d \hat{m}	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.01	0.03	0.16	0.00	0.00	0.01	0.03	0.16
1	0.00	0.00	0.03	0.17	0.36	0.00	0.00	0.04	0.17	0.36
2	1.00	1.00	0.77	0.42	0.20	1.00	1.00	0.77	0.44	0.20
3	0.00	0.00	0.14	0.22	0.17	0.00	0.00	0.14	0.22	0.18
4	0.00	0.00	0.05	0.14	0.10	0.00	0.00	0.04	0.13	0.09
5	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.01
6	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$1 - \alpha$	0.90					0.95				
d \hat{m}	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.78	0.20	0.07	0.05	0.05	0.84	0.26	0.10	0.06	0.06
3	0.20	0.37	0.16	0.07	0.10	0.16	0.44	0.19	0.08	0.11
4	0.02	0.33	0.36	0.10	0.11	0.00	0.27	0.35	0.15	0.11
5	0.00	0.10	0.23	0.78	0.13	0.00	0.03	0.28	0.17	0.16
6	0.00	0.00	0.03	0.00	0.23	0.00	0.00	0.03	0.30	0.23
7	0.00	0.00	0.15	0.00	0.38	0.00	0.00	0.05	0.24	0.33

TABLE 8 : Pourcentage du nombre de ruptures \hat{m} avec différentes valeurs de n et de d.

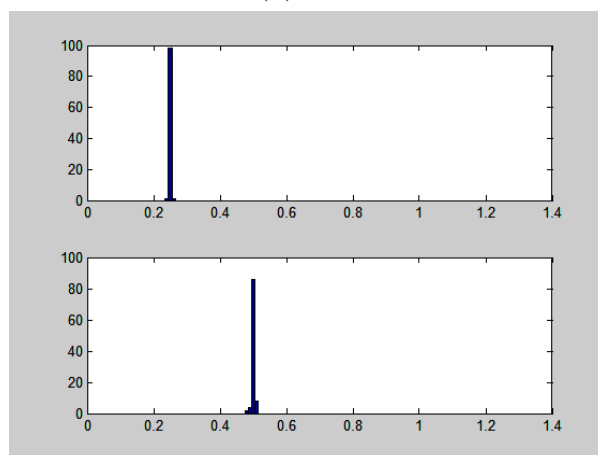
n	500					1000					2000				
d \hat{m}	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49	0.10	0.20	0.30	0.40	0.49
0	0.00	0.00	0.00	0.02	0.03	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.02	0.01
1	0.01	0.14	0.16	0.32	0.24	0.00	0.04	0.19	0.26	0.32	0.00	0.01	0.14	0.24	0.27
2	0.98	0.85	0.69	0.45	0.50	1.00	0.96	0.74	0.51	0.44	1.00	0.99	0.84	0.56	0.41
3	0.00	0.01	0.12	0.20	0.17	0.00	0.00	0.07	0.14	0.15	0.00	0.00	0.02	0.16	0.24
4	0.01	0.00	0.02	0.01	0.05	0.00	0.00	0.00	0.04	0.07	0.00	0.00	0.00	0.01	0.05
5	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.01	0.02
6	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00



(a)

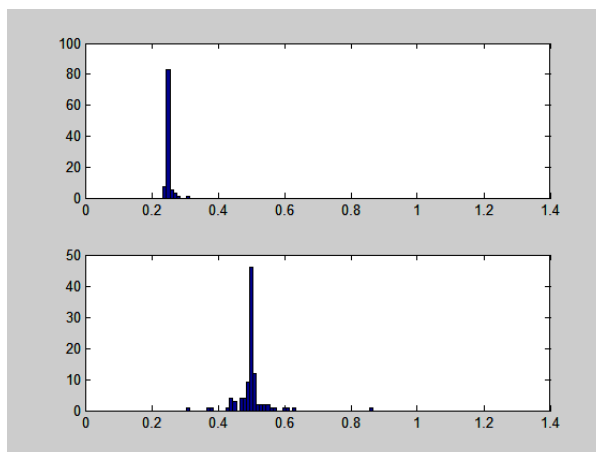


(b)

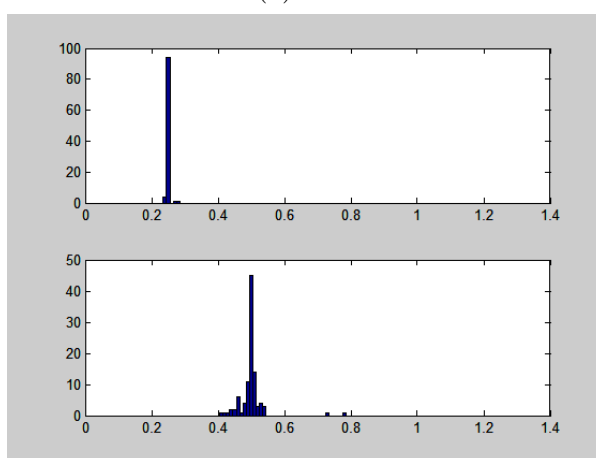


(c)

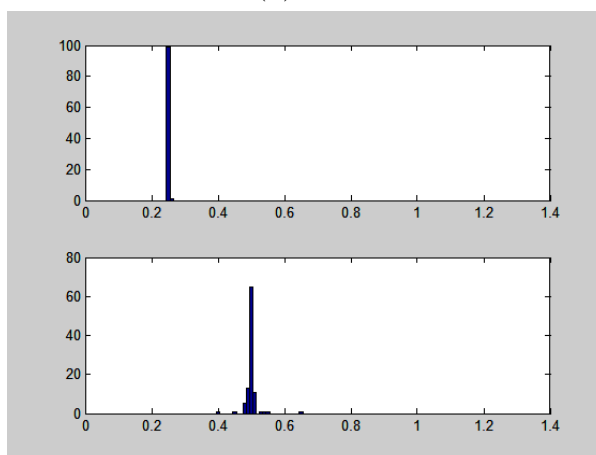
FIGURE 3. La distribution empirique des points de ruptures estimés, $d=0.10$,
(a) : $n=500$, (b) : $n=1000$ et (c) : $n=2000$



(a)

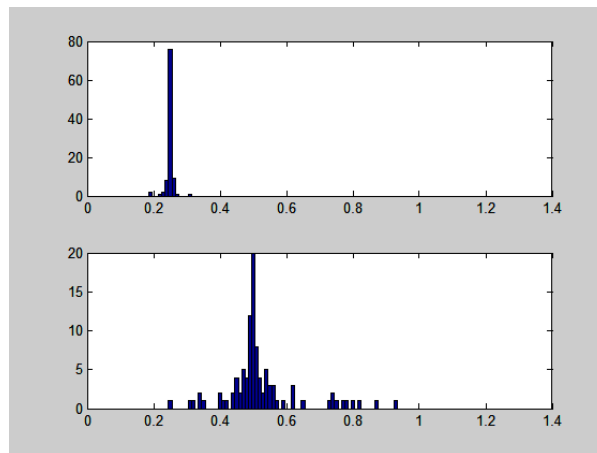


(b)

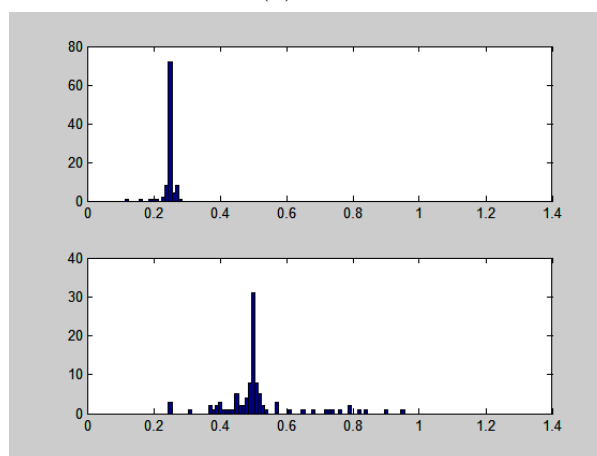


(c)

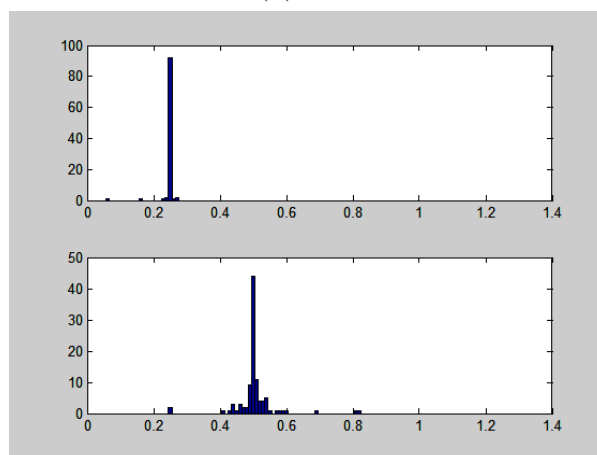
FIGURE 4. La distribution empirique des points de ruptures estimés, $d=0.20$,
(a) : $n=500$, (b) : $n=1000$ et (c) : $n=2000$



(a)

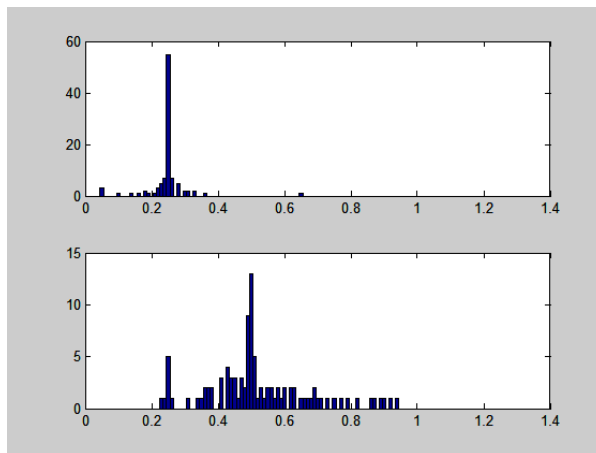


(b)

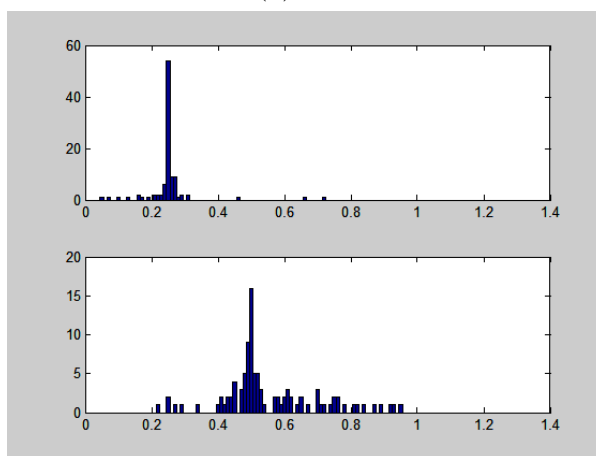


(c)

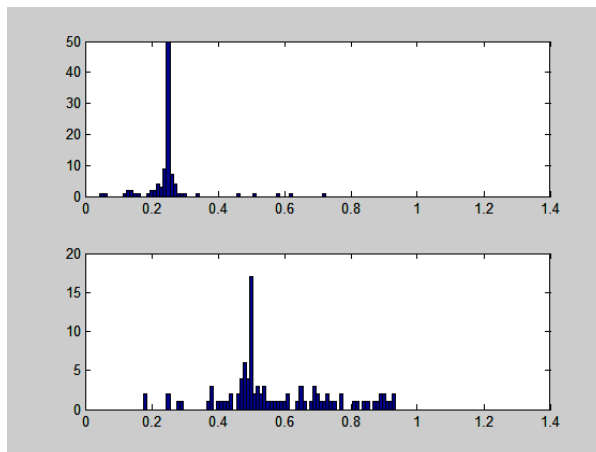
FIGURE 5. La distribution empirique des points de ruptures estimés, $d=0.30$,
 (a) : $n=500$, (b) : $n=1000$ et (c) : $n=2000$



(a)

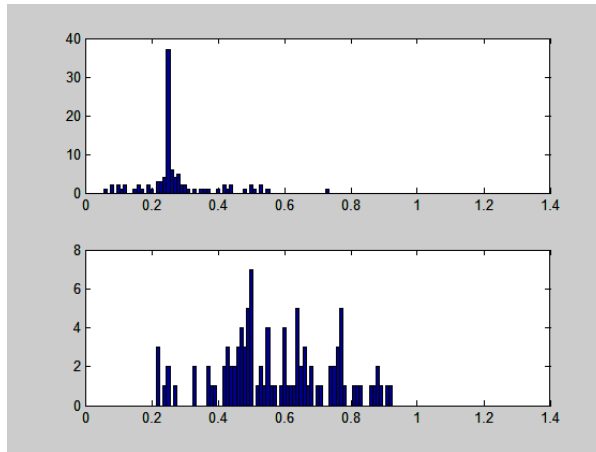


(b)

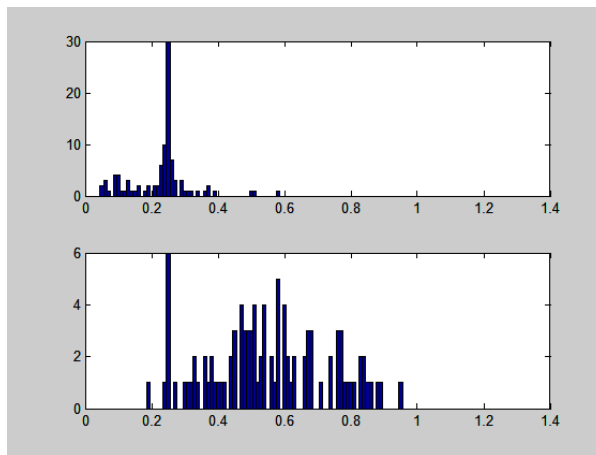


(c)

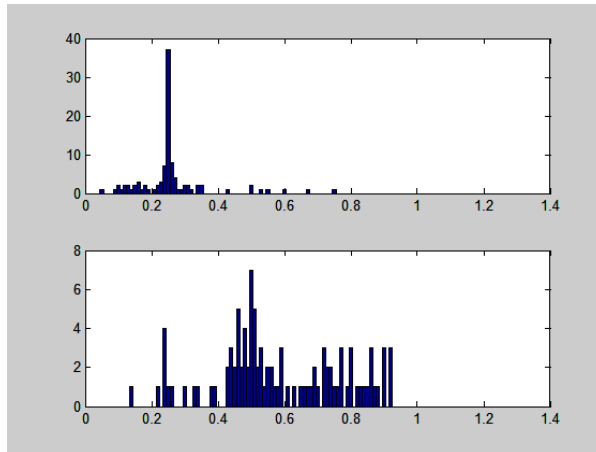
FIGURE 6. La distribution empirique des points de ruptures estimés, $d=0.40$,
(a) : $n=500$, (b) : $n=1000$ et (c) : $n=2000$



(a)



(b)



(c)

FIGURE 7. La distribution empirique des points de ruptures estimés, $d=0.49$,
 (a) : $n=500$, (b) : $n=1000$ et (c) : $n=2000$