An analysis of the influence of some prior specifications in the identification of change points via product partition model

R. H. Loschi [a,1] and F. R. B. Cruz [b,2]

[a] *Departamento de Estatística, Universidade Federal de Minas Gerais, 31270-901 - Belo Horizonte - MG, Brazil*

[b] *Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst, MA 01003, USA*

Abstract

In this paper, we consider the product partition model for the estimation of normal means and variances of a sequence of observations that experiences changes in these parameters at unknown times. The estimates of the parameters by using product partition model are the weighted average of the estimates based in blocks (groups) of observations by the posterior relevance of these blocks which depends on the prior cohesions. We implement the Barry and Hartigan's method to this change point problem and propose an easy-to-implement modification to their method. We use the Yao's prior cohesions and investigate the influence of different prior distributions to the parameter that indexes these cohesions in the product estimates. A comparison between the estimates obtained by using both these methods and those provided by using the Yao's method is done considering different settings for its application. We apply the three methods presented in this paper to stock market data. The results seem to indicate that the method proposed is competitive and also that the prior specifications influence in the product estimates.

Key words: Change points, product partition model, relevance, Student-t distribution, Yao's cohesions

[1] Corresponding author, e-mail loschi@est.ufmg.br.

[2] On sabbatical leave from the *Departamento de Estatística, Universidade Federal de Minas Gerais, 31270-901 - Belo Horizonte - MG, Brazil.*

1 Introduction

The product partition model (PPM) proposed by Hartigan (1990) is a good way to model uncertain about a sequence of random quantities, if the prior opinion about it discloses the existence of blocks of observations produced by some judgment of similarities among these observations, as well as independence among the different blocks. In particular, the PPM is an useful tool to analyze change points problems.

The PPM establishes that the random partition produced by the change points has a prior product distribution, and assumes that, given the partition, the parameters in different blocks have independent prior distributions (Barry and Hartigan, 1992). Consequently, the posterior estimates of these parameters (product estimates) are the weighted average of the estimates in each block by the posterior probability that the block appears in the partition, which is called the posterior relevance of the block (see details in Section 2). In general, large number of computations are involved on the product estimates.

The product estimates of the mean of normal random variables with common variance is considered in detail by Barry and Hartigan (1993) using a Gibbs sampling scheme. Barry and Hartigan (1993) compare their estimates with those obtained by Chernoff and Zacks (1964), Yao (1984) and by considering the Schwartz criterion (Yao, 1988). Barry and Hartigan (1993) conclude that their method provide more accurate estimates if outliers are observed and that Yao's method usually demand more computational time. Recently, Crowley (1997) provides a new implementation of the Gibbs sampling in order to solve the problem of estimating normal means by using the general PPM. Loschi et al. (1999) extend some results from Crowley (1997) and Barry and Hartigan (1993) by using the PPM to identify multiple change points in the mean and variance of normal data. Loschi et al. (1999) assume the prior cohesions proposed by Yao (1984) and consider Yao's algorithm to compute the product estimates.

This paper discusses the same change point problem considered by Loschi et al. (1999). We also assume the Yao's prior cohesions which depend on the probability that a change occurs at any time. We implement the method established by Barry and Hartigan (1993) and propose an easy-to-implement modification of this method. Similar estimates are obtained as well as the computational time involved in their calculation. We also compare our propose and the Barry and Hartigan's method with Yao's method. Yao's method usually demands less computational time and provides similar results for the product estimates in the application we consider. This contradicts Barry and Hartigan (1993) statements. The ultimate goal is to compare all methods presented here in different contexts in order to perceive the influence in the product estimates of the prior specifications to the parameter involved in the Yao's cohesions.

This paper is organized as follows. Section 2 briefly reviews the PPM introduced by Barry and Hartigan (1992). Section 3 presents inferential solutions to identify change points for random variables which are normally distributed, given the means and variances, according to Loschi et al. (1999). In Section 4 we introduce the computational procedure to calculate the posterior relevances based in a Gibbs sampling approach and review the Yao's algorithm and the method proposed by Barry and Hartigan (1993). In Section 5, we apply the methods to the two most important Brazilian indexes, "*Índice Geral da Bolsa de Valores de São Paulo*" (IBOVESPA) and "*Índice da Bolsa de Valores de Minas Gerais, Espírito Santo e Brasília*" (IBOVMESB), comparing their performance under different prior specifications.

2 Product partition models

In this section, we present a brief revision of the product partition model (PPM), introduced by Barry and Hartigan (1992) to identify multiple change points in a sequence of variables observed at consecutive points in time. A more general definition to PPM can be found in Hartigan (1990).

Let $X_1, \ldots, X_n$ be a observed time series and consider the index set $I = \{1, \ldots, n\}$. Consider a random partition $\rho = \{i_0, i_1, \cdots, i_b\}$ of the set $I \cup \{0\}$, $0 = i_0 < i_1 < \cdots < i_b = n$, and a random variable B to represent the number of blocks in ρ. Consider that each partition divides the sequence $X_1, \ldots, X_n$ into $B = b$ contiguous subsequences, which will be denoted here by $\mathbf{X}_{[i_{r-1}i_j]} = (X_{i_{r-1}+1}, \ldots, X_{i_r})'$, $r = 1, \ldots, b$. Let $c_{[ij]}$ be the prior cohesion associated with the block $[ij] = \{i+1, \ldots, j\}$, $i, j \in I \cup \{0\}$, $j > i$, which represents the degree of similarity among the observations in $\mathbf{X}_{[ij]}$

Hence, we say that the random quantity $(X_1, \ldots, X_n; \rho)$ follows a PPM, denoted by $(X_1, \ldots, X_n; \rho) \sim PPM$, if:

i) the prior distribution of ρ is the following product distribution:

$$P(\rho = \{i_0, \ldots, i_b\}) = \frac{\Pi_{j=1}^{b} c_{[i_{j-1}i_j]}}{\sum_{\mathcal{C}} \Pi_{j=1}^{b} c_{[i_{j-1}i_j]}}, \tag{1}$$

where $\mathcal{C}$ is the set of all possible partitions of the set I into b contiguous blocks with endpoints $i_1, \ldots, i_b$, satisfying the condition $0 = i_0 < i_1 < \ldots < i_b = n$, for all $b \in I$;

ii) conditionally on $\rho = \{i_0, \ldots, i_b\}$, the sequence $X_1, \ldots, X_n$ has the joint density given by:

$$f(X_1,\ldots,X_n|\rho=\{i_0,\ldots,i_b\}) = \Pi_{j=1}^{b} f_{[i_{j-1}i_j]}(\mathbf{X}_{[i_{j-1}i_j]}), \tag{2}$$

where $f_{[ij]}(\mathbf{X}_{[ij]})$ is the density of the random vector, called data factor, $\mathbf{X}_{[ij]} = (X_{i+1},\ldots,X_j)'$.

Consequently, if $(X_1,\ldots,X_n;\rho) \sim$ PPM, the number of blocks B in ρ has a prior distribution given by:

$$P(B=b) \propto \sum_{\mathcal{C}_1} \Pi_{j=1}^{b} c_{[i_{j-1}i_j]}, \; b \in I, \tag{3}$$

where $\mathcal{C}_1$ is the set of all partitions of I in b contiguous blocks with endpoint $i_1,\ldots,i_b$, satisfying the condition $0 = i_0 < i_1 < \ldots < i_b = n$.

As shown in Barry and Hartigan (1992), the posterior distributions of ρ and B have the same form of the prior distribution, where the posterior cohesion for the block $[ij]$ is given by

$$c^{*}_{[ij]} = c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]}). \tag{4}$$

In the parametric approach to PPM, a sequence of unknown parameters $\theta_1,\ldots,\theta_n$, such that, conditionally in $\theta_1,\ldots,\theta_n$, the sequence of random variables $X_1,\ldots,X_n$ has conditional marginal densities $f_1(X_1|\theta_1),\ldots,f_n(X_n|\theta_n)$, respectively, is considered. The prior distribution of $\theta_1,\ldots,\theta_n$ is constructed as follows. Given a partition $\rho = \{i_0,\ldots,i_b\}$, $b \in I$, we have that $\theta_i = \theta_{[i_{r-1}i_r]}$ for every $i_{r-1} < i \leq i_r$, $r = 1,\ldots,b$, and that $\theta_{[i_0 i_1]},\ldots,\theta_{[i_{b-1}i_b]}$ are independent, with $\theta_{[ij]}$ having (block) prior density $\pi_{[ij]}(\theta)$, $\theta \in \Theta_{[ij]}$, where $\Theta_{[ij]}$ is the parameter space corresponding to the common parameter, say, $\theta_{[ij]} = \theta_{i+1} = \ldots = \theta_j$, which indexes the conditional density of $\mathbf{X}_{[ij]} = (X_{i+1},\ldots,X_j)'$. In this case, we consider that two observations X_i and X_j, $i \neq j$, are in the same block, if they are identically distributed. Thus, in this approach to PPM, the predictive distribution $f_{[ij]}(X_{[ij]})$, which appeared in (2), can be obtained as follows:

$$f_{[ij]}(\mathbf{X}_{[ij]}) = \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta, \tag{5}$$

The goal is to obtain the marginal posterior distributions of the parameters ρ, B, and θ_k, $k = 1,\ldots,n$. The posterior distributions of ρ and B are obtained as described before and considering the joint density given in (5). Barry and

Hartigan (1992) have shown that the posterior distributions of θ_k is given by:

$$\pi(\theta_k|X_1,\ldots,X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]}\pi_{[ij]}(\theta_k|\mathbf{X}_{[ij]}),\ k=1,\ldots,n, \tag{6}$$

and the posterior expectation (or product estimate) of θ_k is given by:

$$E(\theta_k|X_1,\ldots,X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]}E(\theta_k|\mathbf{X}_{[ij]}),\ k=1,\ldots,n, \tag{7}$$

where $r^*_{[ij]}$ denotes the posterior relevance for the block $[ij]$, that is:

$$r^*_{[ij]} = P([ij]\in\rho|X_1,\ldots,X_n),$$

which, in the situation introduced by Barry and Hartigan (1993) and briefly described in this section, become:

$$r^*_{[ij]} = \frac{\lambda_{[0i]}c^*_{[ij]}\lambda_{[jn]}}{\lambda_{[0n]}}, \tag{8}$$

with $\lambda_{[ij]} = \sum\Pi_{k=1}^{b}c^*_{[i_{k-1}i_k]}$, where the summation is over all partitions of $\{i+1,\ldots,j\}$ in b blocks with endpoints $i_0,i_1,\ldots,i_b$ satisfying the condition $i=i_0<i_1<\ldots<i_b=j$.

Figure 1 shows a graphical representation of the relationships between the random objects in the PPM. The conditional dependence between the objects is represented by arrows liking them. The conditional independence of $X_1,\ldots,$ X_n, given $\theta_1,\ldots,$ θ_n, is represented by the absence of links between the X_i. The conditional independence of the parameters $\theta_{[ij]}$, given ρ, is represented by the absence of links between the groups of parameters $\theta_{i+1},\cdots,\theta_j$.

Notice that, since the PPM is formulated to allow changes in the parameters through the time, it is a kind of dynamical model (see Pole, West and Harrison, 1994). If we consider the general definition of the PPM (Hartigan, 1990; Crowley, 1997), we can also observe that the threshold model with delay parameter equal to zero and m regimes is the particular PPM with

$$P(\rho=\{S_1,\ldots,S_b\}) \propto \begin{cases} c(S_1).\ldots.c(S_b), & \text{if } b=m,\\ 0, & \text{otherwise,}\end{cases}$$

where $S_1,\ldots,S_b$ is any partition of the set $I=\{1,\ldots,n\}$. Notice that when the product prior distribution is non-zero, only the partitions that provide m (fixed) blocks are considered.

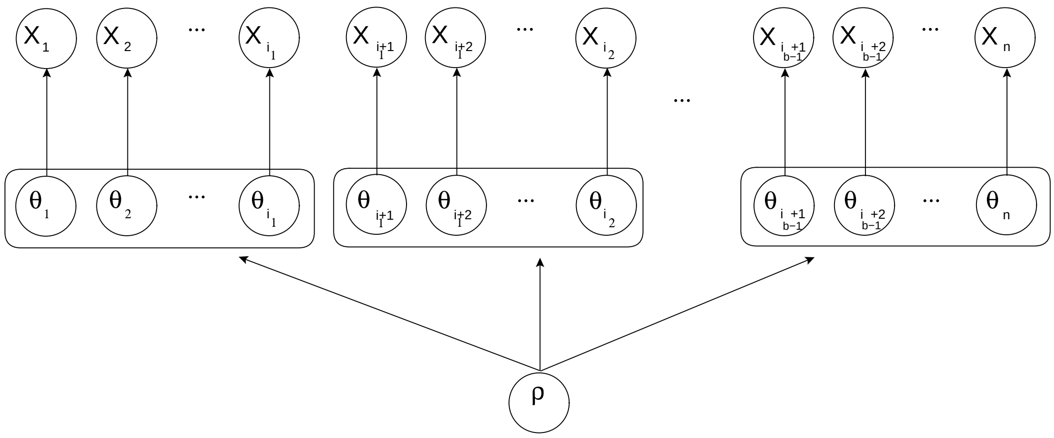


Fig. 1. Graphical Representation of the PPM

3 Posterior estimates for the normal means and variances

In order to specify the PPM for the normal case, Loschi et al. (1999) assume that there is a sequence of unknown parameters $\theta_1 = (\mu_1, \sigma_1^2), \ldots, \theta_n = (\mu_n, \sigma_n^2)$, such that $X_k|\mu_k, \sigma_k^2 \sim \mathcal{N}(\mu_k, \sigma_k^2)$, $k = 1, \ldots, n$, and that they are independent.

It is also assumed that each common parameter $\theta_{[ij]} = (\mu_{[ij]}, \sigma_{[ij]}^2)$, related to the block $[ij]$, has the conjugate normal-inverted-gamma prior distribution denoted by:

$$(\mu_{[ij]}, \sigma_{[ij]}^2) \sim \mathcal{NIG}(m_{[ij]}, v_{[ij]}; a_{[ij]}/2, d_{[ij]}/2),$$

that is,

$$\mu_{[ij]}|\sigma_{[ij]}^2 \sim \mathcal{N}(m_{[ij]}, v_{[ij]}\sigma_{[ij]}^2) \quad \text{and} \quad \sigma_{[ij]}^2 \sim \mathcal{IG}(a_{[ij]}/2, d_{[ij]}/2), \tag{9}$$

where $\mathcal{IG}(a, d)$ denotes the inverted-gamma distribution with parameters a and d, $m_{[ij]} \in \mathcal{R}$, and $a_{[ij]}$, $d_{[ij]}$ and $v_{[ij]}$ are positive values. Hence, the conditional distribution of $\theta_{[ij]} = (\mu_{[ij]}, \sigma_{[ij]}^2)$, given the observations in $\mathbf{X}_{[ij]}$, is the normal-inverted-gamma distribution given by:

$$(\mu_{[ij]}, \sigma_{[ij]}^2)|\mathbf{X}_{[ij]} \sim \mathcal{NIG}(m_{[ij]}^*, v_{[ij]}^*; a_{[ij]}^*/2, d_{[ij]}^*/2), \tag{10}$$

where

$$\left.\begin{aligned} m^*_{[ij]} &= \tfrac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \tfrac{m_{[ij]}}{(j-i)v_{[ij]}+1}, \\ v^*_{[ij]} &= \tfrac{v_{[ij]}}{(j-i)v_{[ij]}+1}, \\ d^*_{[ij]} &= d_{[ij]} + j - i, \\ a^*_{[ij]} &= a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]}), \end{aligned}\right\} \tag{11}$$

with

$$\bar{X}_{[ij]} = \frac{1}{j-i}\sum_{r=i+1}^{j} X_r,$$

$$q_{[ij]}(\mathbf{X}_{[ij]}) = \sum_{r=i+1}^{j}(X_r - \bar{X}_{[ij]})^2 + \frac{(j-i)(\bar{X}_{[ij]} - m_{[ij]})^2}{(j-i)v_{[ij]}+1}.$$

Consequently, it follows from (10) that, given $X_{[ij]}$, the conditional marginal densities of $\mu_{[ij]}$ and $\sigma^2_{[ij]}$ are, respectively:

$$\mu_{[ij]}|\mathbf{X}_{[ij]} \sim t(m^*_{[ij]}, v^*_{[ij]}, a^*_{[ij]}, d^*_{[ij]}) \text{ and } \sigma^2_{[ij]}|\mathbf{X}_{[ij]} \sim \mathcal{IG}(a^*_{[ij]}/2, d^*_{[ij]}/2), \tag{12}$$

for which it is observed that

$$E(\mu_{[ij]}|X_{[ij]}) = m^*_{[ij]} \quad (\text{if } d^*_{[ij]} > 1) \tag{13}$$

and

$$E(\sigma^2_{[ij]}|X_{[ij]}) = \frac{a^*_{[ij]}}{d^*_{[ij]} - 2} \quad (\text{if } d^*_{[ij]} > 2). \tag{14}$$

The interested reader may find details in O'Hagan (1994).

From (7), (13) and (14), it follows that the product estimates for the parameters μ_k and σ^2_k, $k = 1, \ldots, n$, are given by:

$$\hat{\mu}_k = E(\mu_k|X_1, \ldots, X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]} m^*_{[ij]} \quad (\text{if } d^*_{[ij]} > 1) \tag{15}$$

and

$$\hat{\sigma}_k^2 = E(\sigma_k^2|X_1,\ldots,X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]}\frac{a^*_{[ij]}}{d^*_{[ij]}-2} \quad (\text{ if} d^*_{[ij]} > 2), \tag{16}$$

respectively, where $m^*_{[ij]}$, $a^*_{[ij]}$ and $d^*_{[ij]}$ are defined as in (11).

Let $\mathbf{1}_n$ be the $n \times 1$ vector of ones and $\mathbf{I}_n$ the $n \times n$ identity matrix. The posterior relevances $r^*_{[ij]}$ can be obtained from (8) and (4) where the random vector $\mathbf{X}_{[ij]}$ follows a $(j-i)$-dimensional Student-t distribution denoted by $\mathbf{X}_{[ij]} \sim t_{j-i}(\mathbf{m}_{[ij]}, \mathbf{V}_{[ij]}; a_{[ij]}, d_{[ij]})$ with density function given by

$$\begin{aligned} f(\mathbf{X}_{[ij]}) = c(d_{[ij]}, j-i) a_{[ij]}^{d_{[ij]}/2} |\mathbf{V}_{[ij]}|^{-1/2} \times \\ \{a_{[ij]} + (\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})' \mathbf{V}_{[ij]}^{-1} (\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})\}^{-(d_{[ij]}+j-i)/2}, \end{aligned} \tag{17}$$

where $c(d,k) = \Gamma[\frac{d+k}{2}]\{\Gamma[\frac{d}{2}]\ \pi^{\frac{k}{2}}\}^{-1}$ and $\mathbf{m}_{[ij]} = m_{[ij]}\mathbf{1}_{j-i}$ and $\mathbf{V}_{[ij]} = \mathbf{I}_{j-i} + v_{[ij]}\mathbf{1}_{j-i}\mathbf{1}'_{j-i}$.

The algorithm of this normal case is depicted in Figure 2.

4 Computational procedures

Notice from (15) and (16) that high computational efforts are demanded to calculate the product estimates. To simplify these calculations some procedures were proposed in the literature. In this section, we review the computational approach developed by Yao (1984) and introduce a Gibbs sampling scheme to compute the posterior relevances (and, consequently, the product estimates). We also describe the Barry and Hartigan (1993) method to calculate the product estimates of the means and variances by using PPM.

We will assume the prior cohesions suggested by Yao (1984) and defined below, since for the PPM shown in this paper, the cohesions can be interpreted as transition probabilities in the Markov chain defined by the endpoints of the blocks in ρ.

Let p, $0 \leq p \leq 1$, be the probability that a change occurs at any instant in the sequence. Therefore, the prior cohesion for block $[ij]$ is given by:

$$c_{[ij]} = \begin{cases} p(1-p)^{j-i-1}, & \text{if } j < n, \\ (1-p)^{j-i-1}, & \text{if } j = n, \end{cases} \tag{18}$$

algorithm

  read $X_1, \ldots, X_n$

  for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**

    $f_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta$

    $c^*_{[ij]} \leftarrow c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]})$

  end for

  compute $\begin{cases} \lambda_{[0j]}, \forall j = 0, \ldots, n; \\ \lambda_{[in]}, \forall i = 1, \ldots, n-1; \end{cases}$

  for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**

    $r^*_{[ij]} \leftarrow \frac{\lambda_{[0i]} c^*_{[ij]} \lambda_{[jn]}}{\lambda_{[0n]}}$

    $\bar{X}_{[ij]} \leftarrow \frac{1}{j-i} \sum_{r=i+1}^{j} X_r$

    $m^*_{[ij]} \leftarrow \frac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \frac{m_{[ij]}}{(j-i)v_{[ij]}+1}$

    $v^*_{[ij]} \leftarrow \frac{v_{[ij]}}{(j-i)v_{[ij]}+1}$

    $d^*_{[ij]} \leftarrow d_{[ij]} + j - i$

    $q_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \sum_{r=i+1}^{j} (X_r - \bar{X}_{[ij]})^2 + \frac{(j-i)(\bar{X}_{[ij]} - m_{[ij]})^2}{(j-i)v_{[ij]}+1}$

    $a^*_{[ij]} \leftarrow a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})$

  end for

  for $k = 1$ **to** n **do**

    $E(\mu_k|X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} m^*_{[ij]}$

    $E(\sigma_k^2|X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \frac{a^*_{[ij]}}{d^*_{[ij]}-2}$

  end for

  write $E(\mu_1), E(\sigma_1^2), \ldots, E(\mu_n), E(\sigma_n^2)$

end algorithm

Fig. 2. Main Algorithm (μ and σ $\mathcal{N}$ormal Case)

for all $i,\ j \in I$, $i < j$, which corresponds to the probability that a new change takes place after $j - i$ instants, given that a change has taken place at instant i.

Consequently, for the normal case presented in Section 3, the posterior cohesion of the block $[ij]$ become:

$$c^*_{[ij]} = \begin{cases} \frac{p(1-p)^{j-i-1} c(d_{[ij]}, j-i) a_{[ij]}^{d_{[ij]}/2}}{(1+(j-i)v_{[ij]})^{1/2} \{a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})\}^{(d_{[ij]}+j-i)/2}}, & \text{if } j < n \\ \frac{(1-p)^{j-i-1} c(d_{[ij]}, j-i) a_{[ij]}^{d_{[ij]}/2}}{(1+(j-i)v_{[ij]})^{1/2} \{a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})\}^{(d_{[ij]}+j-i)/2}}, & \text{if } j = n, \end{cases} \tag{19}$$

4.1 Yao's algorithm

Let $\lambda_{[ij]}$ be the summation presented in (8) where $c^*_{[ij]}$ is the posterior cohesions given in (19). Hence, the exact posterior relevances given in (8), can be obtained by using the following recursive algorithm:

$$\left.\begin{aligned}
\lambda_{[00]} &= 1, \\
\lambda_{[01]} &= c^*_{[01]}, \\
\lambda_{[0j]} &= c^*_{[0j]} + \textstyle\sum_{t=1}^{j-1} \lambda_{[0t]} c^*_{[tj]}, \ \forall j = 2, \ldots, n, \\
\lambda_{[(n-1)n]} &= c^*_{[(n-1)n]}, \\
\lambda_{[in]} &= c^*_{[in]} + \textstyle\sum_{t=i+1}^{n-1} \lambda_{[tn]} c^*_{[it]}, \ \forall i = 1, \ldots, n-2, \\
\lambda_{[nn]} &= 1.
\end{aligned}\right\} \tag{20}$$

Notice that in spite of the simplifications introduced by Yao (1984), high computational efforts are still demanded in PPM. In Section 4.2, we propose a procedure to simplify the computation of the posterior relevances and, consequently, the product estimates using the transformation suggested by Barry and Hartigan (1993). We also describe the Barry and Hartigan's method to obtain the product estimates of normal means and variances.

4.2 Methods based in Gibbs sampling schemes

In order to implement the Barry and Hartigan's method and the proposed method, the following Gibbs sampling scheme is considered.

4.2.1 Gibbs sampling scheme to PPM

Let assume the auxiliary random quantity U_i suggested by Barry and Hartigan (1993) which reflects whether or not a change point occurs at the time i, that is:

$$U_i = \begin{cases} 1, \text{ if } \theta_i = \theta_{i+1}, \\ 0, \text{ if } \theta_i \neq \theta_{i+1}, \end{cases}$$

$i = 1, \ldots, n-1$. Notice that the random partition ρ is perfectly identified by considering vectors $\mathbf{U} = (U_1, \ldots, U_{n-1})$ of these random quantities.

Each vector $(U_1^s, \ldots, U_{n-1}^s)$, $s \geq 1$, is generated by using the Gibbs sampling as follows. Starting with an initial values $(U_1^0, \ldots, U_{n-1}^0)$ of $\mathbf{U}$, at step s, the r-th element U_r^s is generated from the conditional distribution:

$$U_r | U_1^s, \ldots, U_{r-1}^s, U_{r+1}^{s-1}, \ldots, U_{n-1}^{s-1}; X_1, \ldots, X_n,$$

$r = 1, \ldots, n-1$. In order to generate the samples above, it is sufficient to consider the following ratio:

$$R_r = \frac{P(U_r = 1 | A_s; X_1, \ldots, X_n)}{P(U_r = 0 | A_r^s; X_1, \ldots, X_n)},$$

$r = 1, \ldots, n-1$, where $A_r^s = \{U_1^s = u_1, \ldots, U_{r-1}^s = u_{r-1}, U_{r+1}^{s-1} = u_{r+1}, \ldots, U_{n-1}^{s-1} = u_{n-1}\}$. Consequently, the criterion of choosing the values U_i^s, $i = 1, \ldots, n-1$ becomes:

$$U_r^s = \begin{cases} 1, & \text{if } R_r \geq \frac{1-u}{u} \\ 0, & \text{otherwise,} \end{cases}$$

where $r = 1, \ldots, n-1$ and $u \sim \mathcal{U}(0,1)$.

Consider the prior cohesions given in (18) and assume that p has the prior distribution $\pi(p)$. Let $\mathcal{C}$ be the set of all partitions of the set I into b contiguous blocks with endpoint $i_0, \ldots, i_b$ satisfying the condition $0 = i_0 < i_1 < \ldots < i_b = n$, $b \in I$ and consider $\mathcal{C}_1 \subset \mathcal{C}$ the subset of all partitions that contain the block $[ij] = \{i+1, \ldots, j\}$. Thus, each value U_r^s, $s \geq 1$, $r = 1, \ldots, n-1$, can be generated by using

$$R_r = \frac{f_{[xy]}(X_{[xy]}) \int_0^1 p^{b-2}(1-p)^{n-b+1} d\pi(p)}{f_{[xr]}(X_{[xr]}) f_{[ry]}(X_{[ry]}) \int_0^1 p^{b-1}(1-p)^{n-b} d\pi(p)}, \tag{21}$$

where:

$$x = \begin{cases} \max\{i, \ \textbf{s.t.:}\ 0 < i < r, U_i^s = 0\}, \text{ if } U_i^s = 0, \text{ for} & \\ \qquad\qquad\qquad\qquad \text{some } i \in \{1, \ldots, r-1\}, & \\ 0, \qquad\qquad\qquad\qquad \text{otherwise,} & \end{cases} \tag{22}$$

and

$$y = \begin{cases} \min\{i, \ \textbf{s.t.:} \ r < i < n, U_i^{s-1} = 0\}, & \text{if } U_i^{s-1} = 0, \ \text{for} \\ & \text{some } i \in \{r+1, \ldots, n-1\}, \\ n, & \text{otherwise.} \end{cases} \tag{23}$$

According to Loschi et al. (2001a), if p has the beta prior distribution with $\alpha > 1$ and $\beta > 1$ parameters, denoted by $p \sim \mathcal{B}(\alpha, \beta)$, the value R_r given in (21) become:

$$R_r = \frac{f_{[xy]}(X_{[xy]})\Gamma(n+\beta-b+1)\Gamma(b+\alpha-2)}{f_{[xr]}(X_{[xr]})f_{[ry]}(X_{[ry]})\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}, \ b = 1, \ldots, n, \tag{24}$$

where x and y is obtained as in (22) and (23), respectively.

If no prior distribution is considered to p, the value R_r used in the generation of U_r^s can be obtained by using the following ratio:

$$R_r = \frac{c^*_{[xy]}}{c^*_{[xr]}c^*_{[ry]}}, \tag{25}$$

where x and y are given in (22) and (23) (see details in Loschi et al., 1999). Notice that in this case, as well as in the Yao's method, we are essentially considering a degenerate prior distribution to p, which put all its mass probability to a single value of p.

If the normal case described in Section 3 is considered, the joint density $f_{[ij]}(X_{[ij]})$ is the Student-t distribution given in (17) and the posterior cohesions $c^*_{[ij]}$ are given in (19).

4.2.2 Barry and Hartigan's method

Barry and Hartigan (1993) obtain the product estimates of μ_k and σ_k^2, $k = 1, \ldots, n$, as follows. For each partition $(U_1^s, \ldots, U_{n-1}^s)$, $s \geq 1$, the estimates (per blocks) given in (13) and (14), that is,

$$\hat{\mu}_{ks} = m^*_{[ij]} \text{ and}$$

$$(\hat{\sigma}^2)_{ks} = \frac{a^*_{[ij]}}{d^*_{[ij]} - 2},$$

for $r = i+1, \ldots, j$, $i, j = 0, 1, \ldots, n$, $i < j$, are computed. The product estimates of μ_k and σ_k^2, $k = 1, \ldots, n$, are approximated, respectively, by

$$\hat{\mu}_k = \frac{\sum_{s=1}^{T} \hat{\mu}_{ks}}{T} \text{ and}$$

$$\hat{\sigma}_k^2 = \frac{\sum_{s=1}^{T} (\hat{\sigma}^2)_{ks}}{T},$$

where T is the net size of the generated sample.

The Barry and Hartigan's algorithm is presented in Figure 3.

4.2.3 Proposed method

We obtain the product estimates as follows. Generate a sample of size T of the vector $(U_1, \ldots, U_{n-1})$. The estimate of the posterior relevance of the block $[ij]$, $i, j = 1, \ldots, n, i < j$, is computed as follows:

$$\hat{r}^*_{[ij]} = \frac{M}{T}, \tag{26}$$

where M is the number of vectors in the generate sample where is observed $U_i = 0$, $U_{i+1} = \ldots, U_{j-1} = 1$ and $U_j = 0$. The product estimates of μ_k and σ_k^2, $k = 1, \ldots, n$, is obtained substituting (26) in (15) and (16), respectively.

Figure 4 shows the algorithm described above, in pseudo-language.

4.3 Remarks on the algorithms

There is no essential difference between our method and Barry and Hartigan's method. However, the proposed algorithm is immediately linked with the theoretical statements because, like Yao's algorithm, we propose a method to calculate the posterior relevances. From Figure 3, we also notice that our method is easy to implement and its efficiency is similar to Barry and Hartigan's method (see the comparison in Section 5).

5 Implementation and analysis

The aim of this section is to compare the different algorithms described in Section 4. The method we propose and the Barry and Hartigan's method are implemented considering three different prior specifications for the parameter

algorithm
    read $X_1, \ldots, X_n$
    for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
        $f_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta$
        $c^*_{[ij]} \leftarrow c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]})$
    end for
    for $k = 1$ **to** SAMPLES **do**
        generate U^k
    end for
    for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
        $\bar{X}_{[ij]} \leftarrow \frac{1}{j-i} \sum_{r=i+1}^{j} X_r$
        $m^*_{[ij]} \leftarrow \frac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \frac{m_{[ij]}}{(j-i)v_{[ij]}+1}$
        $v^*_{[ij]} \leftarrow \frac{v_{[ij]}}{(j-i)v_{[ij]}+1}$
        $d^*_{[ij]} \leftarrow d_{[ij]} + j - i$
        $q_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \sum_{r=i+1}^{j}(X_r - \bar{X}_{[ij]})^2 + \frac{(j-i)(\bar{X}_{[ij]}-m_{[ij]})^2}{(j-i)v_{[ij]}+1}$
        $a^*_{[ij]} \leftarrow a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})$
    end for
    for $k = 1$ **to** n **do**
        $\mu\text{aux} \leftarrow 0$
        $\sigma^2_{\text{aux}} \leftarrow 0$
        for all i, j, l **such that**
            $U^l_i = 0, U^l_{i+1} = \cdots = U^l_k = \cdots = U^l_j = 1, U^k_{j+1} = 0$
                $\mu\text{aux} \leftarrow \mu\text{aux} + m^*_{[ij]}$
                $\sigma^2_{\text{aux}} \leftarrow \sigma^2_{\text{aux}} + \frac{a^*_{[ij]}}{(d^*_{[ij]}-2)}$
        end for
        $E(\mu_k|X_1, \ldots, X_n) \leftarrow \mu\text{aux}/\text{SAMPLES}$
        $E(\sigma^2_k|X_1, \ldots, X_n) \leftarrow \sigma^2_{\text{aux}}/\text{SAMPLES}$
    end for
    write $E(\mu_1), E(\sigma^2_1), \ldots, E(\mu_n), E(\sigma^2_n)$
end algorithm

Fig. 3. Barry and Hartigan Method (μ and σ $\mathcal{N}$ormal Case)

p. Firstly, we assume that p is a fixed value arbitrarily chosen. After that, we consider that p has a beta prior distribution. Finally, we assume a non-informative prior distribution for p.

As it is well known, Yao's method does not consider prior distributions for p. That is, to use Yao's method, we ought to be sure about the value of p. Then, to fairly compare the methods, we consider different values of p. Beta distributions whose modal values are close to the selected p are also assumed.

Our applications focus on the identification of multiple change points in the

$$
\begin{array}{l}
\textbf{algorithm} \\
\quad \textbf{read } X_1, \ldots, X_n \\
\quad \textbf{for all } i, j \in \{0, \ldots, n\} \textbf{ such that } i < j \textbf{ do} \\
\qquad f_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta \\
\qquad c^*_{[ij]} \leftarrow c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]}) \\
\quad \textbf{end for} \\
\quad \textbf{for } k = 1 \textbf{ to } \text{SAMPLES} \textbf{ do} \\
\qquad \textit{generate } U^k \\
\quad \textbf{end for} \\
\quad \textbf{for all } i, j \in \{0, \ldots, n\} \textbf{ such that } i < j \textbf{ do} \\
\qquad r^*_{[ij]} \leftarrow \text{proportion of samples such that} \\
\qquad \qquad \qquad U^k_i = 0, U^k_{i+1} = \cdots = U^k_{j-1} = 1, U^k_j = 0 \\
\quad \textbf{end for} \\
\quad \textbf{for all } i, j \in \{0, \ldots, n\} \textbf{ such that } i < j \textbf{ do} \\
\qquad \bar{X}_{[ij]} \leftarrow \frac{1}{j-i} \sum_{r=i+1}^{j} X_r \\
\qquad m^*_{[ij]} \leftarrow \frac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \frac{m_{[ij]}}{(j-i)v_{[ij]}+1} \\
\qquad v^*_{[ij]} \leftarrow \frac{v_{[ij]}}{(j-i)v_{[ij]}+1} \\
\qquad d^*_{[ij]} \leftarrow d_{[ij]} + j - i \\
\qquad q_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \sum_{r=i+1}^{j} (X_r - \bar{X}_{[ij]})^2 + \frac{(j-i)(\bar{X}_{[ij]} - m_{[ij]})^2}{(j-i)v_{[ij]}+1} \\
\qquad a^*_{[ij]} \leftarrow a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]}) \\
\quad \textbf{end for} \\
\quad \textbf{for } k = 1 \textbf{ to } n \textbf{ do} \\
\qquad E(\mu_k|X_1, \ldots, X_n) \leftarrow \sum_{0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} m^*_{[ij]} \\
\qquad E(\sigma^2_k|X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \frac{a^*_{[ij]}}{d^*_{[ij]}-2} \\
\quad \textbf{end for} \\
\quad \textbf{write } E(\mu_1), E(\sigma^2_1), \ldots, E(\mu_n), E(\sigma^2_n) \\
\textbf{end algorithm}
\end{array}
$$

Fig. 4. Proposed Algorithm (μ and σ $\mathcal{N}$ormal Case)

mean (expected return) and variance (volatility) in stock market return series. We consider the two most important Brazilian indexes, "*Índice Geral da Bolsa de Valores de São Paulo*" (IBOVESPA) and "*Índice da Bolsa de Valores de Minas Gerais, Espírito Santo e Brasília*" (IBOVMESB), from January, 1991 to August, 1999.

As usual in finance, a return series is defined by using the transformation $R_t = (P_t - P_{t-1})/P_{t-1}$, where P_t is the closing price in the month t. IBOVESPA and IBOVMESB return series are plotted all together in Figure 5.

From Figure 5, it is noticeable that IBOVESPA and IBOVMESB series present a similar behavior, suggesting the existence of some changes in the mean and variance of the returns in both series. Despite the similarities, IBOVMESB

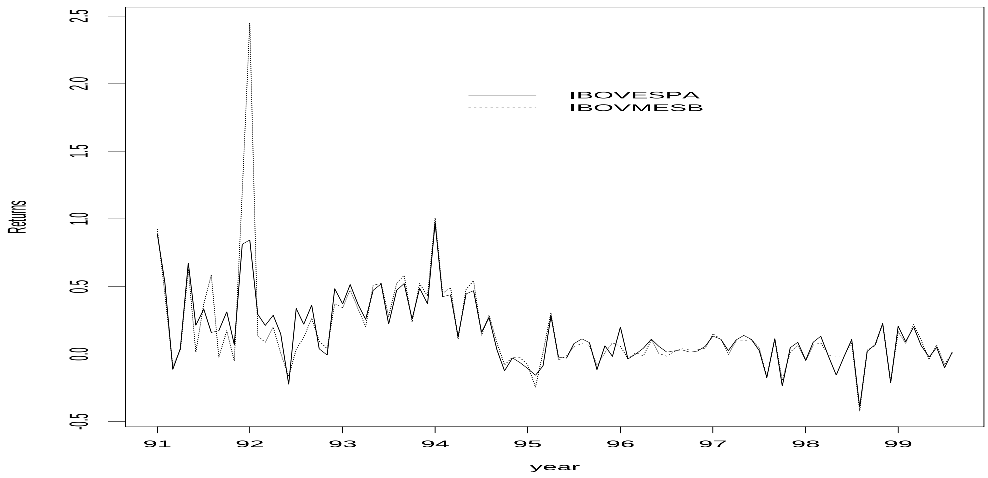


Fig. 5. IBOVESPA and IBOVMESB Return Series

series presents a considerably different return in January, 1992. We are also interested in identifying which method would work better in such a situation.

We suppose that returns are conditionally independent and distributed according to the normal distribution $\mathcal{N}(\mu_{[ij]}, \sigma^2_{[ij]})$, and adopt the natural conjugate prior distribution for the parameters $\mu_{[ij]}$ and $\sigma^2_{[ij]}$ which, in this case, is a normal-inverted-gamma distribution.

In accordance to the specifications of Loschi et al. (2001b), the following normal-inverted-gamma prior distribution is adopted to describe the uncertainty on the parameter $(\mu_{[ij]}, \sigma^2_{[ij]})$ for both indexes:

$$\mu_{[ij]}|\sigma^2_{[ij]} \sim \mathcal{N}(0, \sigma^2_{[ij]}), \text{ and } \sigma^2_{[ij]} \sim \mathcal{IG}\left(\frac{0.01}{2}, \frac{4}{2}\right).$$

A small number of change points is expected for both indexes which implies that p should assume lower values with higher probability. Despite this, to observe the influence of the prior specifications of p, we will consider $p = 0.1, 0.5, 0.9$ and also the following beta prior distributions of p: $p \sim \mathcal{B}(5, 50)$, $p \sim \mathcal{B}(50, 50)$ and $p \sim \mathcal{B}(50, 5)$.

In the Gibbs sampling scheme, we generate 10,000 samples of $\mathbf{U}$ vector with dimension 103, starting from a vector of zeros, and discharged the initial 4,000 iterations. A lag of 10 is selected in order to avoid correlation among vectors, which means that we worked with a net sample size of 600. Other combination of sample sizes, lags and burn -in periods were considered (not shown) and the results indicated that the convergence is always fast and the autocorrelation is low. See details about Gibbs sampling practical implementations in Gamerman (1997), Robert and Casella (1999) and Gilks et al. (1996).

5.1 *Results analysis*

The algorithms presented in Section 4 were coded in C++ and they are available from the authors upon request. All tests were performed in a PC-like computer, 166 MHz, 32 MB RAM, running Windows 98, and using the freely available C++ compiler DJGPP (http://www.delorie.com/djgpp).

5.1.1 *IBOVESPA case*

In this section, we present the posterior estimates of the mean return and the volatility of IBOVESPA series obtained by using all methods described in Section 4.

Figures 6, 7 and 8 show the product estimates obtained by Yao's method, Barry and Hartigan's method, as well as the method here proposed. Once Yao's method considers a fixed value for p rather than priors, p was set to the modal value of the beta prior distribution considered. The computational time was around 25 seconds for the last two methods considered. It is noticeable that the three methods produce similar posterior estimates in all cases.

We also tested fixed p's for all three methods (results not shown). The estimates obtained by Yao's, Barry and Hartigan's, and the proposed method are all equal. We also observe that, in this particular case in which no priors were considered for p, the results were similar to the results obtained assuming a beta prior distribution concentrated in values close to the modal value p.

Since the use of non-informative prior distributions is common in Bayesian statistics, it is important to discuss the performance of Barry and Hartigan's method and the proposed method under such an assumption. Thus, assuming that p has an uniform prior distribution in the $(0, 1)$ interval, denoted by $p \sim \mathcal{U}(0, 1)$, the product estimates depicted in Figure 9 were obtained. For comparison purposes, Figure 9 also displays the estimates obtained by Yao's method, assuming a $p = 0.9$.

As in the other scenarios considered before, we notice that the product estimates obtained by using Barry and Hartigan's method and our method are coincident. However, in this case, the posterior estimates indicate the existence of a high number of change points, as we observe in those cases in which the prior specifications for p consider more mass to high values. It is also important to notice that the estimates tend to be equal for all three methods, if a non-informative prior distribution is assumed and we consider values of p close to one for Yao's.

Figure 10 shows the product estimates for mean and variance obtained by the

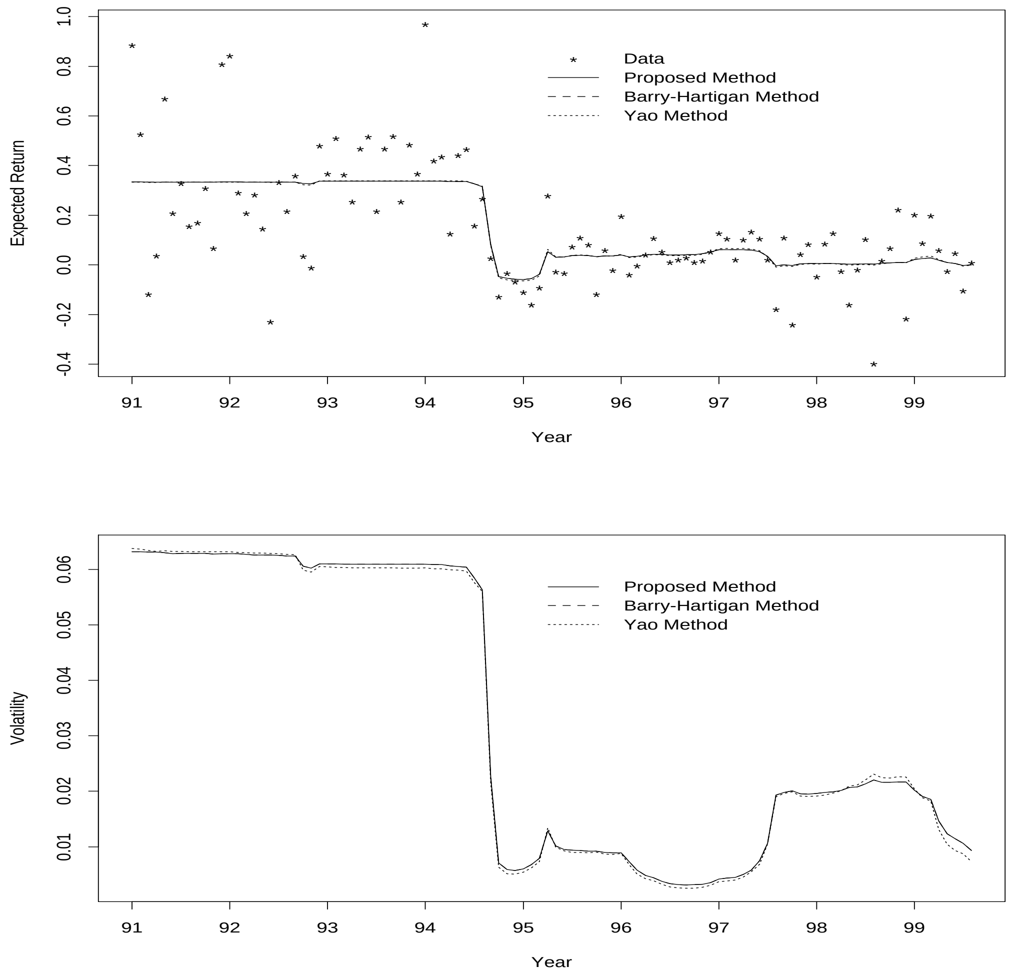


Fig. 6. Product Estimates to the Expected Returns and Volatilities of IBOVESPA - $p \sim \mathcal{B}(5, 50)$

proposed method. In both cases, it is assumed that $p = 0.1$, $p \sim \mathcal{B}(5, 50)$, and $p \sim \mathcal{U}(0, 1)$. We notice that the adoption of this beta prior distribution produces the same product estimates then those obtained by considering a fixed value to p. Other beta prior distributions which concentrate most of the mass in lower values of p were also considered, producing very similar estimates (see an example in Figure 11). Similar behavior (not shown) was observed for the other scenes.

5.1.2 IBOVMESB case

The IBOVESPA and IBOVMESB series present similar behavior but in IBOVMESB series we observe a considerably different return in January, 1992 (see Figure 5). Barry and Hartigan (1993) conclude that, considering a com-

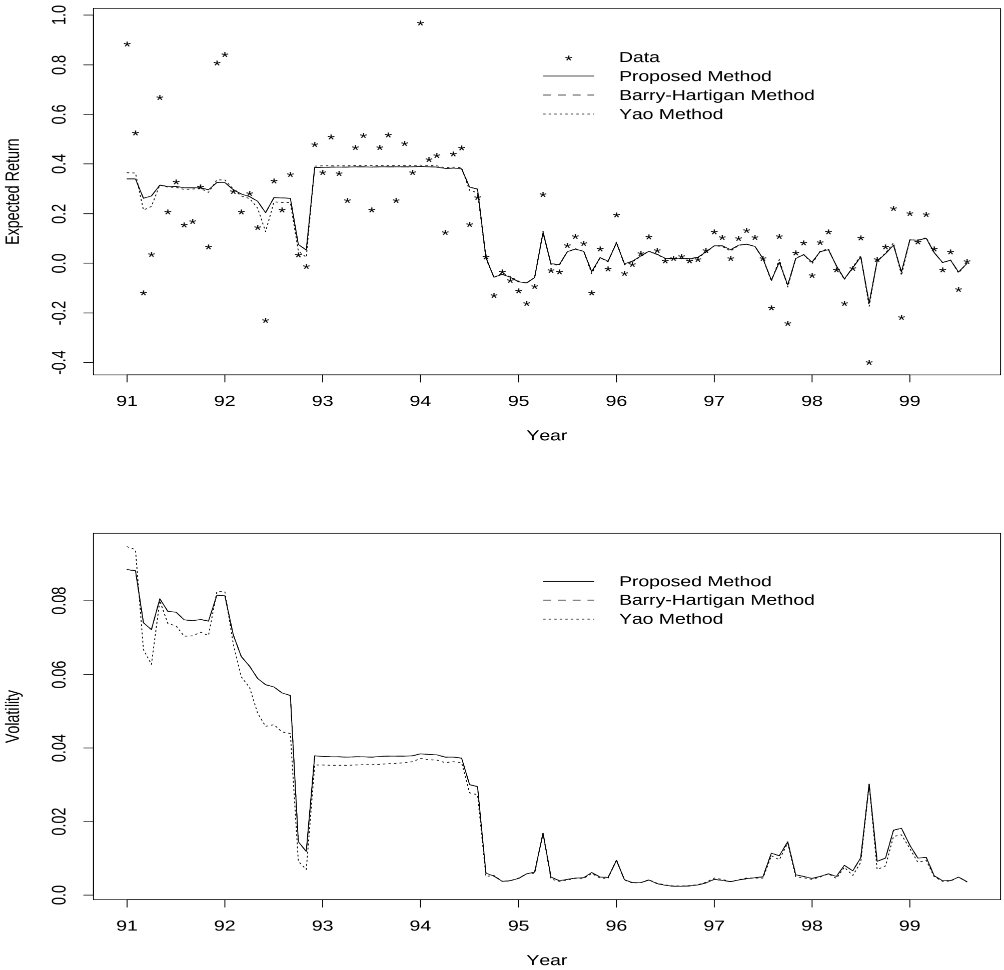


Fig. 7. Product Estimates to the Expected Returns and Volatilities of IBOVESPA - $p \sim \mathcal{B}(50, 50)$

pletely Bayesian approach, in which hiperprior distributions are assumed, their method works better than Yao's method, in the identification of atypical observations.

Figures 12 and 13 present the product estimates considering $p \sim \mathcal{B}(5, 50)$ and $p \sim \mathcal{U}(0, 1)$, respectively, for our method and Barry and Hartigan's method. For Yao's method, we consider $p = 0.1$ in all circumstances. Notice that, if the beta prior distribution is assumed, our method and Barry and Hartigan's method works better then Yao's method in estimating the mean and variance of the atypical observation occurred in January, 1992. Yao's method also provide higher estimates to the variance during 1991. From January, 1992 on, the same product estimates were obtained for all methods. The same analysis can be done if no prior distribution is assumed. Similarly to IBOVESPA series, if an uniform prior distribution is considered (see Figure 13) or if the prior spec-

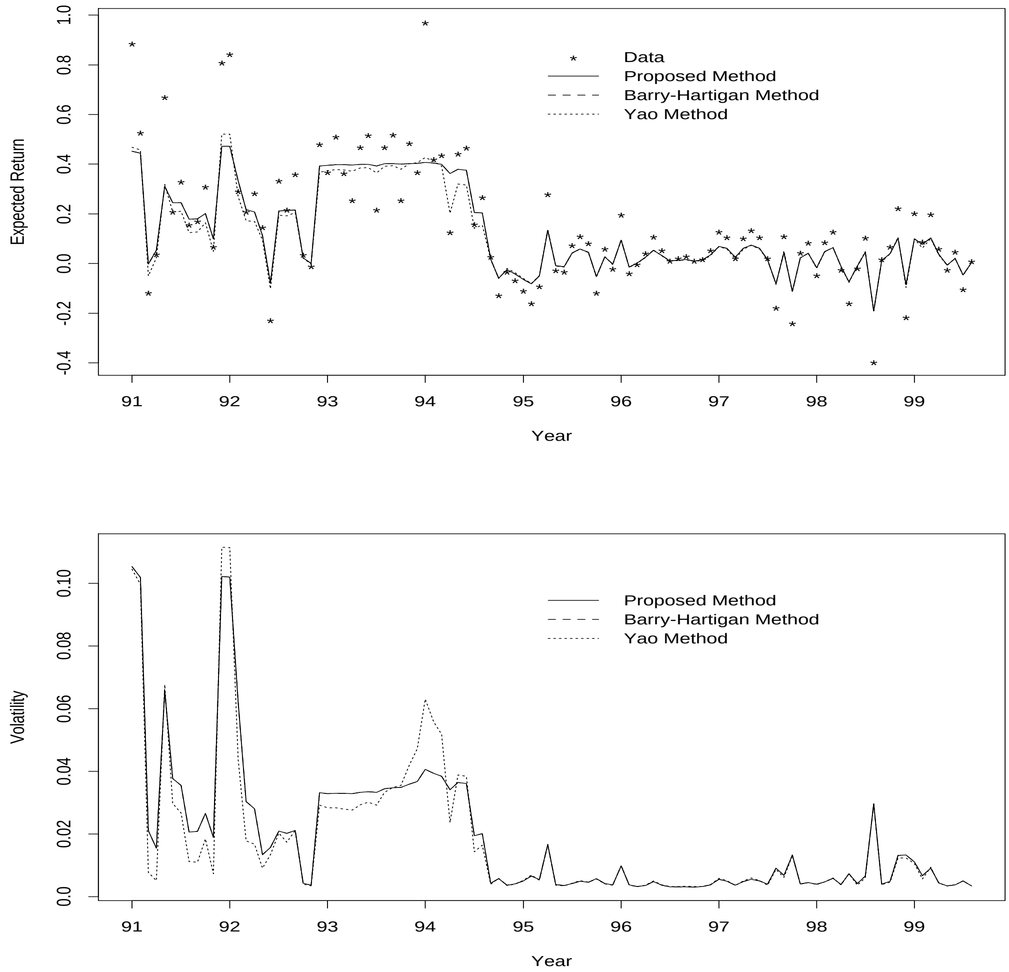


Fig. 8. Product Estimates to the Expected Returns and Volatilities of IBOVESPA - $p \sim \mathcal{B}(50, 5)$

ifications that consider high mass to high values of p are considered, almost all points are identified as a change point and the atypical observation is not identified.

5.2 Data analysis

According to Loschi et al. (2001b), a small number of change points is expected for the Brazilian stock markets. Then, to analyze the data, we consider prior specifications for p which considers most mass in small values.

We notice from Figures 6 and 12, for example, that change points observed in IBOVESPA and IBOVMESB series typically occur at the same time and that the changes are in the same direction. However, some differences in the be-

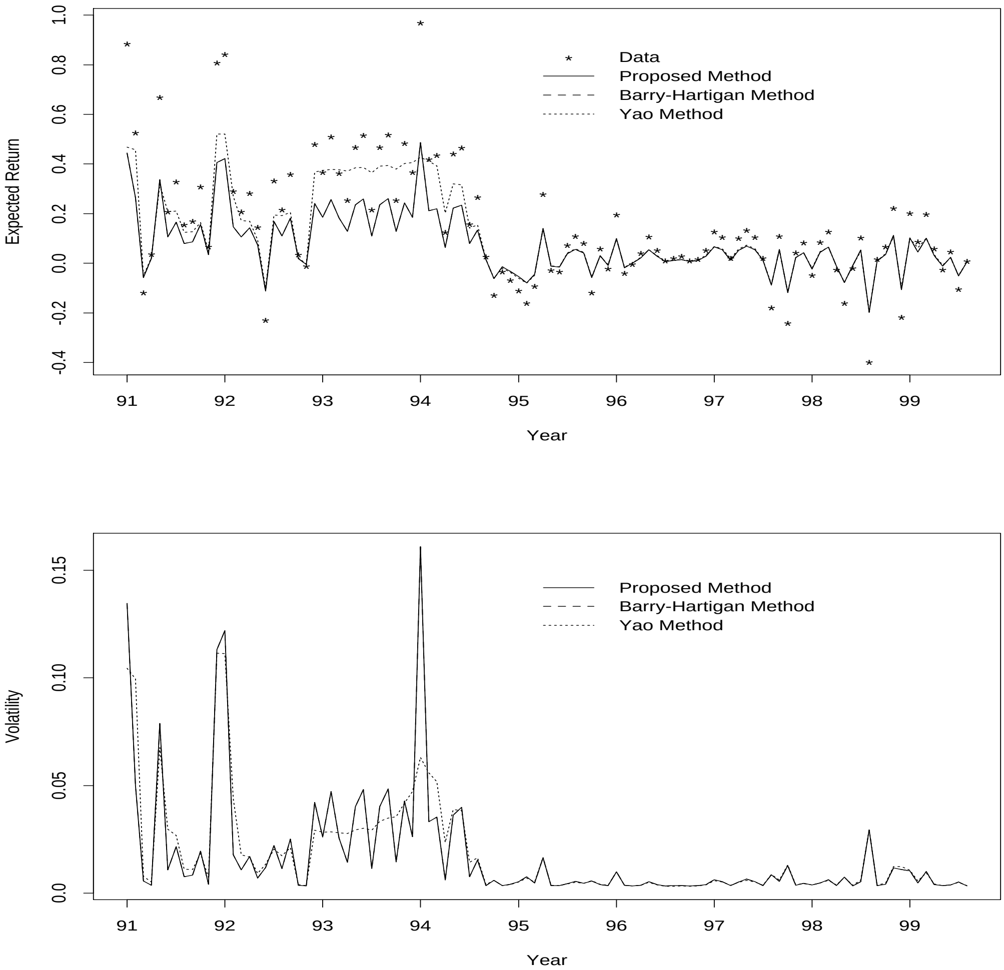


Fig. 9. Product Estimates to the Expected Returns and Volatilities of IBOVESPA - $p \sim \mathcal{U}(0,1)$ and $p = 0.9$

havior of these series are observed. The two changes observed in IBOVMESB series, in August and October, 1991, do not occur in IBOVESPA series. These change points could be related to the sale of USIMINAS, a very important state steel company, located in Minas Gerais state. In October, 1991, USIMINAS was sold for a private group. The beginning of the crisis in the Fernando Collor's government in March, 1992, which culminate with his impeachment, in December of the same year, could be the events that produced the change points in IBOVMESB series, around these two months. Unlike the initial expectations, these important historical facts do not seem to produce changes in the behavior of IBOVESPA series.

In July, 1999, Russia's crisis could have produced the change in the IBOVMESB series. However, we do not observe changes in the IBOVESPA series within that period. This different behavior could be explained by the

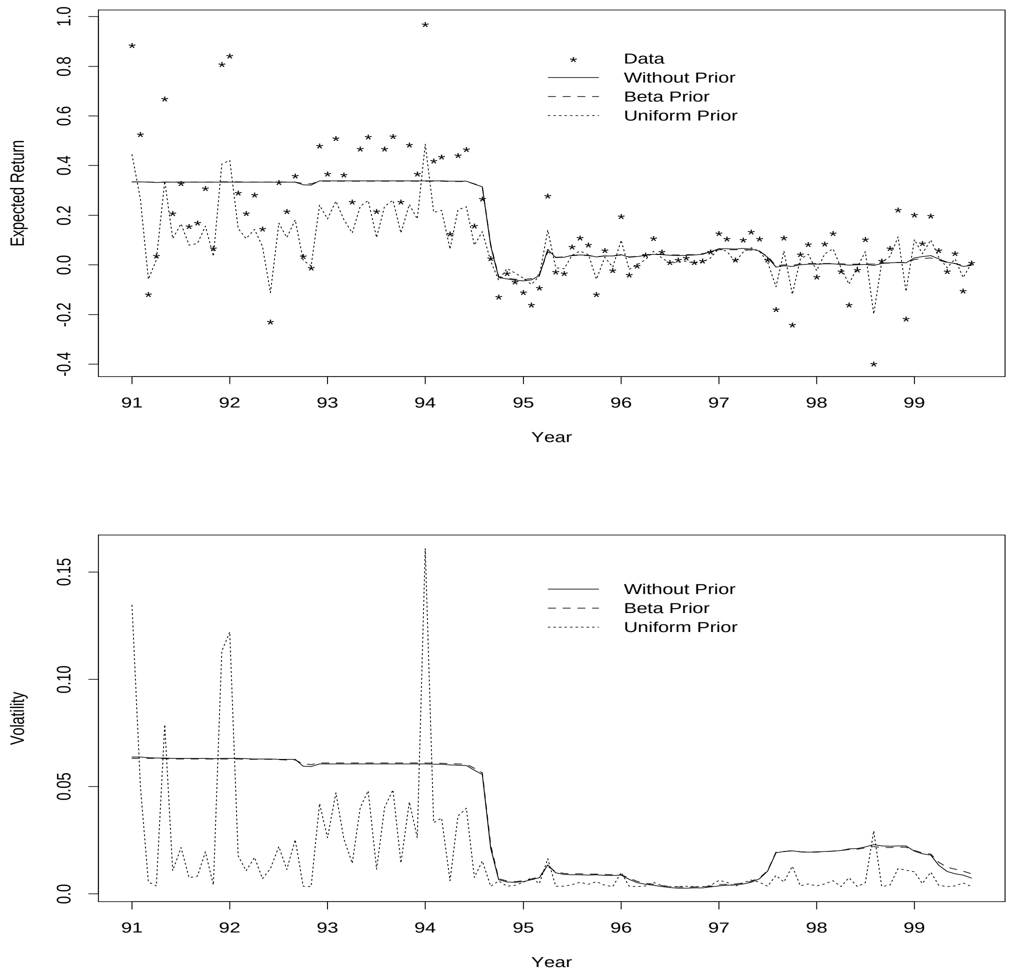


Fig. 10. Product Estimates to the Expected Returns and Volatilities of IBOVESPA - Proposed Method

policy adopted by Brazilian government during Asia's crisis, in August, 1997, and because IBOVESPA is the main indicator of Brazilian economy, incorporating the benefits of the government policies more immediately.

A new currency, the Real, was introduced in July, 1994. The Real period has presented lower expected returns and volatilities than the previous period. Mexico, and Asia's crises might be responsible for the market warm-up observed, in January, 1995 and August, 1997, respectively. We notice that the periods when higher volatility was observed during the Real period have been smaller than in the preceding period. Some political actions of the Minas Gerais State Governor, in January, 1999, could be associated with the decrease of the expected returns and volatilities of both indexes, from this period on.

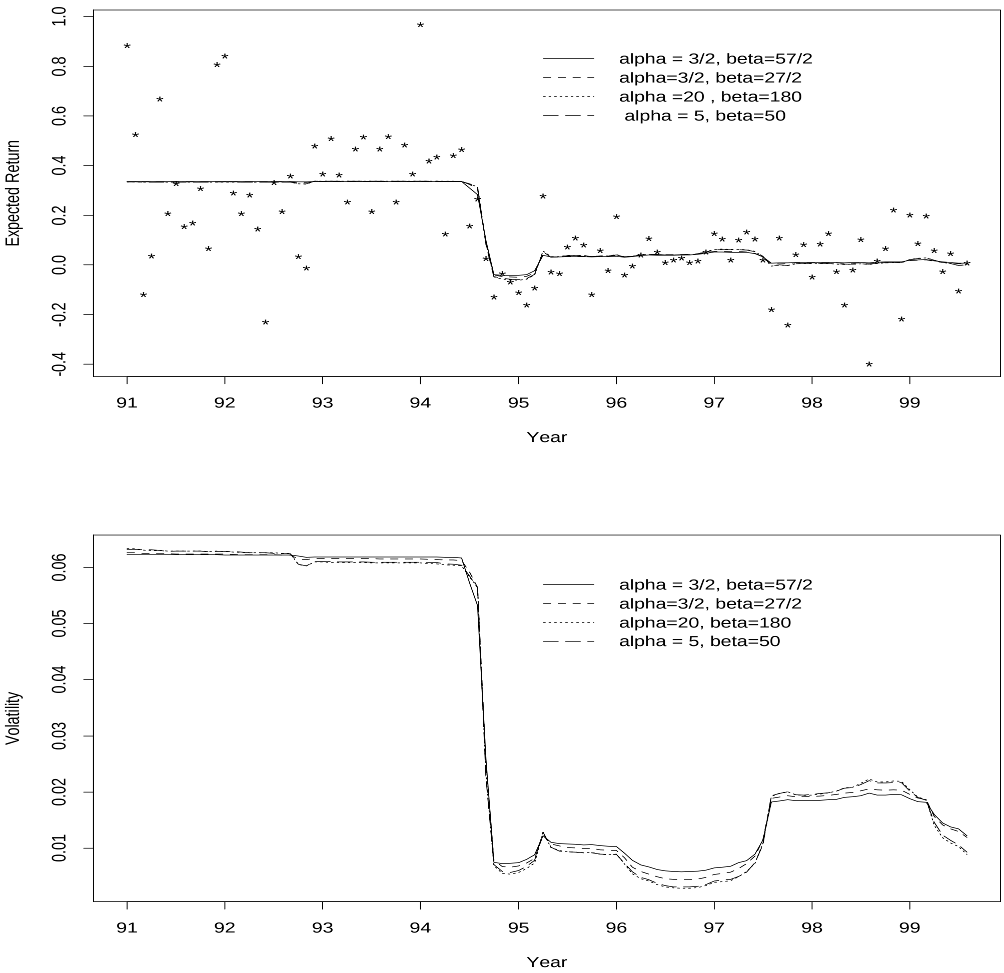
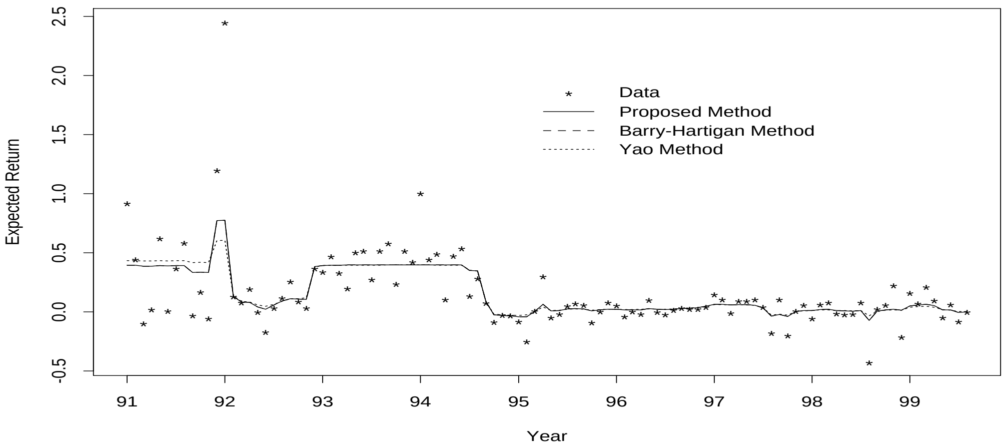


Fig. 11. Product Estimates Considering Different Beta Prior Distributions - IBOVESPA (Proposed Method)

6 Final Comments

We proposed an easy-to-implement method to compute posterior relevances and, consequently, to calculate the product estimates, considering different prior specifications to the parameter that indexes Yao's cohesions. We applied this method to identify change points in normal means and variances. We also implemented Barry and Hartigan's method to the same change point problem. Both methods were compared with Yao's method within different settings. The effect of different prior specifications in the product estimates were studied.

We conclude that our method and Barry and Hartigan's method produce the same posterior estimates and take the same computational time. Contradicting the statements of Barry and Hartigan (1993), we obtained that, in general,



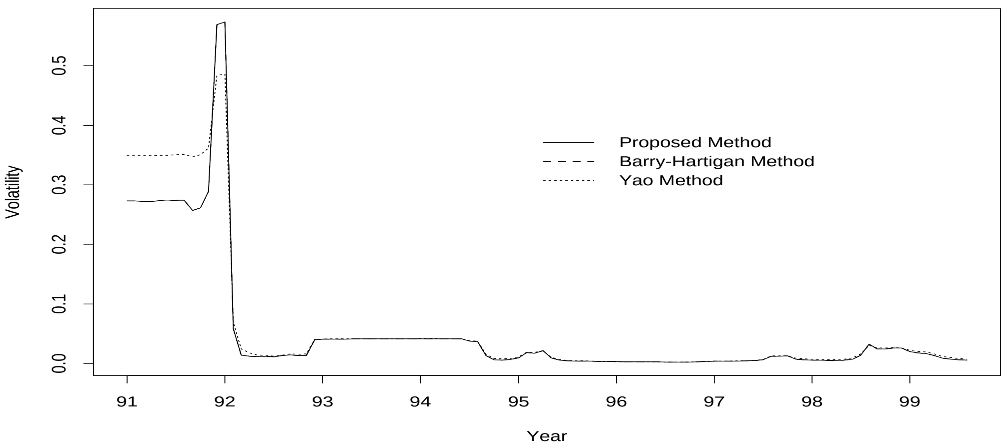


Fig. 12. Product Estimates to the Expected Returns and Volatilities of IBOVMESB - $p \sim \mathcal{B}(5, 50)$ and $p = 0.1$

Yao's method produces comparable estimates taking less computational time for every scenario we consider. However, our conclusion about the efficiency of Yao's method in the identification of atypical observations agree with the statements of Barry and Hartigan (1993). Our method and Barry and Hartigan's method are more sensitive to the presence of outliers.

Other advantage of our method and Barry and Hartigan's method is the possibility of eliciting prior distributions to p, that can give more flexibility in the use of the PPM. For example, we observe that the product estimates tend to be close if beta distributions having the most of their mass concentrated in small values are considered.

The use of non-informative prior distributions to p as well as the use of prior specifications to p that consider high mass to high values of p identify almost

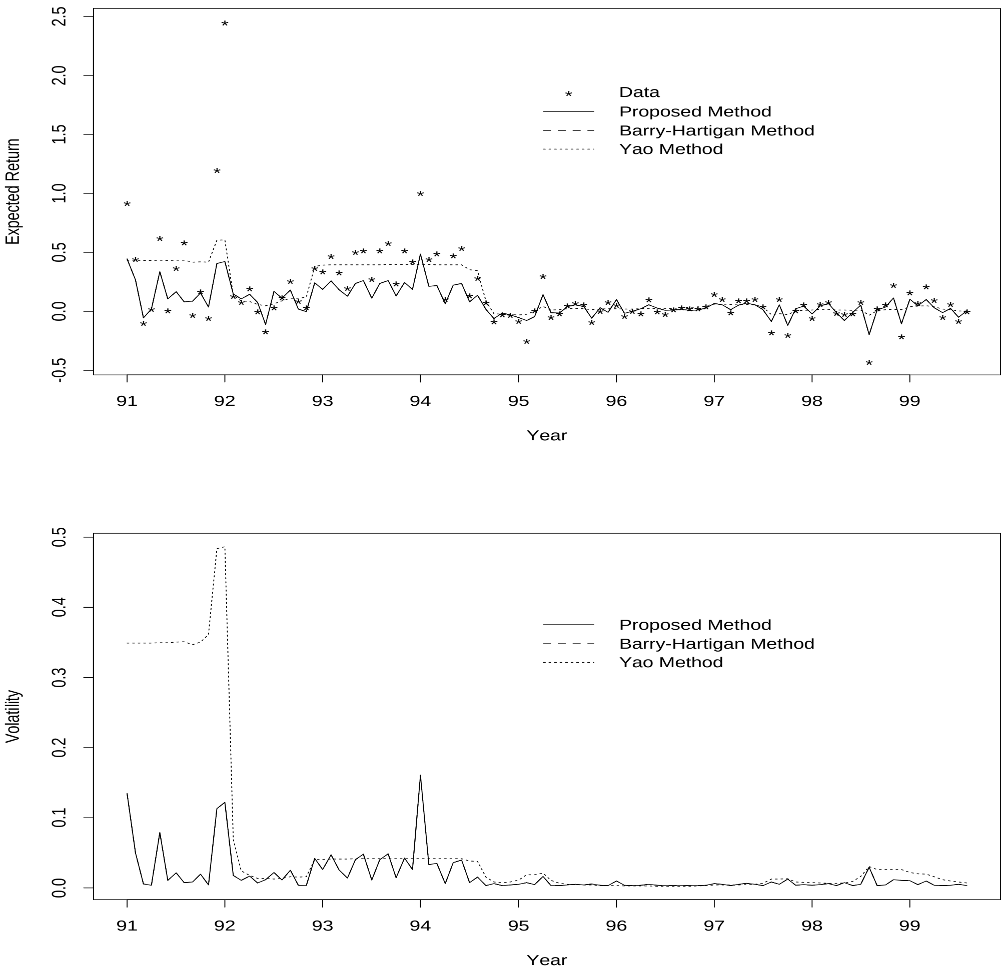


Fig. 13. Product Estimates to the Expected Returns and Volatilities of IBOVMESB - $p \sim \mathcal{U}(0, 1)$ and $p = 0.1$

every point as a change point, which for the Brazilian stock market is not an appropriated choice. We also notice that by using these prior specifications the methods do not identify the atypical observations.

Finally, we observe that the PPM works well in the identification of change points in the Brazilian stock market if an appropriated prior specification is done.

Acknowledgments

The authors hereby acknowledge: CAPES, CNPq (grants 301809/96-8 and 201046/94-6), FAPEMIG (grants CEX 855/98 and CEX 795/00), and PRPq-UFMG (grants 4081-UFMG/RTR/FUNDO/PRPq/99 and 3893-UFMG/RTR/FUNDO/PRPq/98), for a partial allowance to this project; Dani Gamerman and the referees, for their valuable comments and suggestions to improve this paper.

References

Barry, D. and J. A. Hartigan, Product partition models for change point problems, *The Annals of Statistics*, **20**(1) (1992) 260–279.

Barry, D. and J. A. Hartigan, A Bayesian Analysis for change point problems, *Journal of the American Statistical Association*, **88**(421)(1993) 309–319.

Chernoff, H. and S. Zacks, Estimating the current mean of normal distribution which is subject to changes in time, *Annals of Mathematical Statistics*, **35** (1964) 999–1018.

Crowley, E. M., Product partition models for normal means, *Journal of the American Statistical Association*, **92**(437) (1997) 192–198.

Gamerman, D. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference* (Chapman and Hall, London, UK, 1997).

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (Chapman & Hall, New York, USA, 1996).

Hartigan. J. A., Partition models, *Communication in Statistics - Theory and Method*, **19**(8) (1990) 2745–2756.

Loschi, R. H., F. R. B. Cruz, P. L. Iglesias, and R. B. Arellano-Valle, A Gibbs sampling scheme to product partition model: An application to change point problems, *Technical Report RTP 04/01* (Departamento de Estatística, UFMG, Belo Horizonte, Brazil, 2001a).

Loschi, R. H., P.L. Iglesias, and R. B. Arellano-Valle, Bayesian detection of change points in the Chilean stock market, in: *Proceedings of the Annual Meeting of American Statistical Association. Section on Bayesian Statistical Science*, (Baltimore, MD, USA, 1999) 160–165.

Loschi, R. H., G. G. Moreira, and P. L. Iglesias, A comparative analysis of BOVESPA and BOVMESB indexes using the Bayesian approach, *unpublished manuscript*, (2001b).

O'Hagan, A., *Kendall's Advanced Theory of Statistics 2A*, chapter Bayesian Inference, (John Wiley & Sons, New York, NY, 1994).

Pole, A., M. West, and J. Harrison,*Applied Bayesian Forecasting and Time series Analysis*, (Chapman and Hall, New York, USA, 1994).

Robert, C. P. and G. Casella, *Monte Carlo Statistical Methods*, (Springer, New York, USA, 1999).

Yao. Y., Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches, *The Annal of Statistics*, **12**(4) (1984) 1434–1447.

Yao. Y., Estimating the number of change-points by Schwarz's criteria, *Statistics and Probability Letters*, **6** (1988) 181–187.