Advances in Complex Systems, Vol. 5, No. 4 (2002) 371–387

APPLYING THE PRODUCT PARTITION MODEL TO THE IDENTIFICATION OF MULTIPLE CHANGE POINTS

R. H. LOSCHI*

Departamento de Estatística, Universidade Federal de Minas Gerais, 31270-901 - Belo Horizonte - MG, Brazil
E-mail: loschi@est.ufmg.br

F. R. B. CRUZ†

Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst MA 01003, USA
E-mail: fcruz@ecs.umass.edu

Received 3 October 2001
Revised 15 June 2002
Accepted 23 July 2002

The multiple change point identification problem may be encountered in many subject areas, including disease mapping, medical diagnosis, industrial control, and finance. One appealing way of tackling the problem is through the product partition model (PPM), a Bayesian approach. Nowadays, practical applications of Bayesian methods have attracted attention perhaps because of the generalized use of powerful and inexpensive personal computers. A Gibbs sampling scheme, simple and easy to implement, is used to obtain the estimates. We apply the algorithm to the analysis of two important stock market data in Brazil. The results show that the method is efficient and effective in analyzing change point problems.

Keywords: Beta prior distribution, relevance, Student-t distribution, Yao's cohesions.

1. Introduction

Many subject areas share interest in change-point identification problems, including disease mapping, medical diagnosis, industrial control, and finance. For instance, Fig. 1 shows two important stock market indexes in Brazil, expressed in terms of monthly returns, from January, 1991, to August, 1999, namely IBOVESPA (*Índice Geral da Bolsa de São Paulo*) and IBOVMESB (*Índice da Bolsa de Valores de Minas Gerais, Espírito Santo e Brasília*). For such series, it is important to know, e.g. for the evaluation of financial risks [21], whether or not multiple change points

*Corresponding author.
†On sabbatical leave from the *Departamento de Estatística*, *Universidade Federal de Minas Gerais*, 31270-901 - Belo Horizonte - MG, Brazil.

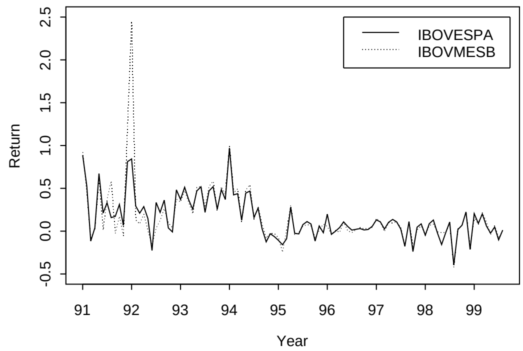


Fig. 1. Return series (IBOVESPA and IBOVMESB).

occurred in the expected means (returns) and in the expected variances (volatilities) of the returns.

Some possible tools already considered to tackle this important problem include either Bayesian [13,22,26] and non-Bayesian approaches [12,14,27]. In particular, this paper is concerned with a Bayesian approach to the multiple change-point identification problem, more specifically about methods based on the well-known Product Partition Model (PPM).

The PPM was introduced by Hartigan [11], as a generalization of several models [13,22,26]. One of the advantages of using the PPM is that the number of change points in the series is a random variable in opposition to models that consider it fixed, as in threshold models [6]. Later, Barry and Hartigan [2,3] and Crowley [7] considered the PPM for the identification of multiple change points in the normal means. By using previous results [15], Loschi et al. [18] extended the PPM to the identification of multiple change points in the means and variances of normal data. Loschi et al. [18] obtained the product estimates by means of a recursive algorithm from Yao [28] and proposed a Gibbs sampling scheme to obtain the posterior distributions for the number of change points and for the instants when the changes occurred. Loschi et al. [17] also extended the PPM to include a prior specification for the parameter p that describes the probability of having a change point at each instant of the time and they proposed a Gibbs sampling scheme to obtain the posterior relevances involved in the product estimate computation. As another contribution to the subject, Loschi and Cruz [16] asserted the adequacy of some prior distributions for the PPM.

The aim of this paper is to apply the PPM to the identification of multiple change points in the means, μ, and variances, σ^2, of normal data, assumed a prior distribution to parameters μ, σ^2, and p. More specifically, the PPM will be applied to the analysis of two important indexes of the Brazilian stock market, Fig. 1.

This paper is organized as follows. Section 2 provides a brief note on Bayesian statistics, reviews the parametric approach to the PPM, presents inferential solutions to identify multiple change points for random variables normally distributed, given the means and variances, and details a Gibbs sampling scheme to implement the PPM. In Section 3, we discuss the results obtained for the Brazilian stock market indexes. Section 4 concludes the paper with final remarks and future topics for investigation.

2. Statistical Model

This section starts reviewing briefly the Bayesian approach to inference. Then, the statistical models are presented.

2.1. *Bayesian Statistics*

From a Bayesian point of view, a random quantity is an unknown quantity that can vary and assume different values from trial to trial (e.g. a random variable) or simply it can be a fixed quantity about which there is little (or there is not) available information (e.g. a parameter). The uncertainty about random quantities is summarized by probability distributions as usual but in the Bayesian approach uncertainty is also viewed as a subjective (or personal) evaluation.

Besides the information obtained from the data, also considered by the classical school of statistics, the Bayesian approach considers other sources of information to solve problems of inference. Let θ be a parameter of interest that assumes values in the parametric space Θ. Denote by $\mathcal{H}$ the previous information and knowledge about Θ. Based on $\mathcal{H}$, the uncertainty about θ is summarized by the prior distribution $p(\theta|\mathcal{H})$. The prior distribution is the inference about θ before observing any data. If Θ is a finite set, this inference represents the chance of occurrence of each value θ in Θ.

In general, $\mathcal{H}$ does not contains all relevant information about the parameter. In this case, the prior distribution is not a good inference for θ. If the information content of $\mathcal{H}$ is not enough, an experimental design must be performed to obtain further information about the parameter. Suppose that a random variable X related to θ can be observed. Before observing X, the uncertainty about this quantity is summarized by the likelihood function $p(X|\theta, \mathcal{H})$, supposed known a fixed value of θ. Notice that the likelihood function gives the probability of occurrence of each particular sample x of X, in the case that θ is the true value for the parameter. After performing the experiment, the prior opinion about θ must be updated using the new information x. The usual tool to update the prior distribution is the Bayes theorem. Thus, the posterior distribution for θ is given by

$$p(\theta|x, \mathcal{H}) = \frac{p(x|\theta, \mathcal{H})p(\theta|\mathcal{H})}{p(x|\mathcal{H})}, \tag{2.1}$$

where the prior predictive distribution $p(x|\mathcal{H})$ is calculated using the following expression:

$$p(x|\mathcal{H}) = \int_{\Theta} p(X|\theta, \mathcal{H})p(\theta|\mathcal{H})d\theta. \tag{2.2}$$

The posterior distribution computed in Eq. (2.1) describes the uncertainty about θ after observing the data, that is, $p(\theta|X, \mathcal{H})$ is the posterior inference about θ. Thus, point (Bayes) estimates for θ can be obtained computing summaries from the posterior distribution. In this case, the resulting estimates depend on the loss function we choose. For example, if we consider the square error loss function, the Bayes estimator for θ is the mean of the posterior distribution of θ. From the posterior distribution we can also construct interval estimates — the credibility interval — or perform hypothesis tests. More about Bayesian statistics can be found in the literature [8,4,23,24]

2.2. *Product Partition Model*

The PPM is a Bayesian model to treat change-point problems which consider special forms for the likelihood function as well as for the prior distribution for the parameter ρ that denotes the instants when the change points have taken place.

In the parametric approach to the PPM [3], we consider that the sequence of random variables $X_1, \ldots, X_n$, conditionally in $\theta_1, \ldots, \theta_n$, has conditional marginal densities $f_1(X_1|\theta_1), \ldots, f_n(X_n|\theta_n)$. We also assume that, given the partition $\rho = \{i_0, \ldots, i_b\}$ of the set $I \cup \{0\}$, $I = \{1, \ldots, n\}$, $b \in I$, such that $0 = i_0 < i_1 < \cdots < i_b = n$, we have that $\theta_i = \theta_{[i_{r-1} i_r]}$ for every $i_{r-1} < i \leq i_r$, for $r = 1, \ldots, b$, and that $\theta_{[i_0 i_1]}, \ldots, \theta_{[i_{b-1} i_b]}$ are independent, with $\theta_{[ij]}$ having (block) prior density $\pi_{[ij]}(\theta)$, where $\theta \in \Theta_{[ij]}$ and $\Theta_{[ij]}$ is the parameter space corresponding to the common parameter, say, $\theta_{[ij]} = \theta_{i+1} = \ldots = \theta_j$, which indexes the conditional density of $\mathbf{X}_{[ij]} = (X_{i+1}, \cdots, X_j)'$. Denote by $c_{[ij]}$, $i, j \in I \cup \{0\}, i < j$, the prior cohesions associated with the block $[ij]$. The prior cohesions are a personal choice and, in the analysis of time series, they can be interpreted as the transition probability of having a change in the instant j given that a change has occurred in the instant i.

In the parametric 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, $(X_1, \ldots, X_n, \rho)$ follows the PPM if:

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

$$P(\rho = \{i_0, \ldots, i_b\}) = \frac{\Pi_{j=1}^{b} c_{[i_{j-1} i_j]}}{\sum_{\mathcal{C}} \Pi_{j=1}^{b} c_{[i_{j-1} i_j]}}, \tag{2.3}$$

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

$$f(X_1, \ldots, X_n|\rho) = \Pi_{j=1}^{b} f_{[i_{j-1}i_j]}(\mathbf{X}_{[i_{j-1}i_j]}), \tag{2.4}$$

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

$$f_{[ij]}(\mathbf{X}_{[ij]}) = \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta. \tag{2.5}$$

Assuming the PPM and a square error loss function, Barry and Hartigan [3] state that the posterior expectation (or product estimate) of θ_k is given by

$$E(\theta_k|X_1, \ldots, X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r_{[ij]}^{*} E(\theta_k|\mathbf{X}_{[ij]}), \tag{2.6}$$

for $k = 1, \ldots, n$, in which the posterior relevance for the block $[ij]$ is given by

$$r_{[ij]}^{*} = \frac{\lambda_{[0i]}c_{[ij]}^{*}\lambda_{[jn]}}{\lambda_{[0n]}}, \tag{2.7}$$

where $c_{[ij]}^{*} = c_{[ij]}f_{[ij]}(\mathbf{X}_{[ij]})$ is the posterior cohesion for the block $[ij]$ and $\lambda_{[ij]} = \sum \Pi_{k=1}^{b} c_{[i_{k-1}i_k]}^{*}$, and 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 $i = i_0 < i_1 < \cdots < i_b = j$. The posterior relevance $r_{[ij]}^{*}$ is the posterior probability that the block $[ij]$ is in the partition ρ, which is computed by means of Eq. (2.1).

The other parameter considered is the number of blocks B (or the number of change points, $B-1$) in ρ. If the PPM is assumed, the posterior distribution of B is given by

$$P(B = b|X_1, \ldots, X_n) \propto \sum_{\mathcal{C}} \Pi_{j=1}^{b} c_{[i_{j-1}i_j]}^{*}. \tag{2.8}$$

From Eq. (2.1), we obtain that the posterior distribution of ρ has the same form of the prior distribution given in (2.3), considering the posterior cohesions $c_{[ij]}^{*}$.

2.3. *Normal Case*

For the normal case [18], it is assumed that $\theta_1 = (\mu_1, \sigma_1^2), \ldots, \theta_n = (\mu_n, \sigma_n^2)$, 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 the common parameter $\theta_{[ij]} = (\mu_{[ij]}, \sigma_{[ij]}^2)$, related to the block $[ij]$, has the following conjugate normal-inverted-gamma prior distribution:

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

that is,

$$\begin{cases} \mu_{[ij]}|\sigma_{[ij]}^2 \sim \mathcal{N}(m_{[ij]}, v_{[ij]}\sigma_{[ij]}^2), \\ \sigma_{[ij]}^2 \sim \mathcal{IG}(a_{[ij]}/2, d_{[ij]}/2), \end{cases} \tag{2.9}$$

where $m_{[ij]} \in \mathbb{R}$, $a_{[ij]}$, $d_{[ij]}$, and $v_{[ij]}$ are positive values, and $\mathcal{IG}(a,d)$ denotes the inverted-gamma distribution with parameters a and d, plotted in Fig. 3, with the following density function (for details, see the book of O'Hagan [24]):

$$f(\sigma^2) = \frac{(a/2)^{d/2}}{\Gamma(d/2)} (\sigma^2)^{-\frac{d+2}{2}} \exp\left\{-\frac{a}{2\sigma^2}\right\}, \quad \sigma^2 > 0.$$

By means of Eq. (2.2), we prove that 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

$$\begin{aligned} f(\mathbf{X}_{[ij]}) &= K(d_{[ij]}, j-i) \times a_{[ij]}^{d_{[ij]}/2} \times |\mathbf{V}_{[ij]}|^{-1/2} \times \\ &\left\{a_{[ij]} + \left[(\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})' \mathbf{V}_{[ij]}^{-1} (\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})\right]\right\}^{(d_{[ij]}+j-i)/2}, \end{aligned} \tag{2.10}$$

where $K(d,k) = \Gamma[\frac{(d+k)}{2}]/\{\Gamma[\frac{d}{2}]\pi^{k/2}\}$, $\mathbf{m}_{[ij]} = m_{[ij]}\mathbf{1}_{j-i}$, $\mathbf{V}_{[ij]} = \mathbf{I}_{j-i} + v_{[ij]}\mathbf{1}_{j-i}\mathbf{1}'_{j-i}$, $\mathbf{1}_n$ is the $n \times 1$ vector of ones, and $\mathbf{I}_n$ denotes the $n \times n$ identity matrix (see more about Student-t distribution in the paper of Arellano-Valle and Bolfarine [1]).

From Eq. (2.1), the conditional marginal densities of $\mu_{[ij]}$ and $\sigma^2_{[ij]}$, given $X_{[ij]}$, are respectively:

$$\begin{cases} \mu_{[ij]} | \mathbf{X}_{[ij]} \sim t(m^*_{[ij]}, v^*_{[ij]}, a^*_{[ij]}, d^*_{[ij]}), \\ \sigma^2_{[ij]} | \mathbf{X}_{[ij]} \sim \mathcal{IG}(a^*_{[ij]}/2, d^*_{[ij]}/2), \end{cases} \tag{2.11}$$

where

$$\begin{cases} m^*_{[ij]} = \frac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \frac{m_{[ij]}}{(j-i)v_{[ij]}+1}, \\ v^*_{[ij]} = \frac{v_{[ij]}}{(j-i)v_{[ij]}+1}, \\ d^*_{[ij]} = d_{[ij]} + j - i, \\ a^*_{[ij]} = a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]}), \end{cases} \tag{2.12}$$

with

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

and

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

Consequently, it is observed that

$$\begin{cases} \hat{\mu}_{[ij]} = E(\mu_{[ij]}|X_{[ij]}) = m^*_{[ij]}, \quad \text{if } d^*_{[ij]} > 1, \\ \hat{\sigma}^2_{[ij]} = E(\sigma^2_{[ij]}|X_{[ij]}) = \frac{a^*_{[ij]}}{d^*_{[ij]}-2}, \text{ if } d^*_{[ij]} > 2. \end{cases} \tag{2.13}$$

From Eq. (2.6) and (2.13), it follows that the product estimates for the parameters μ_k and σ_k^2, $k = 1, \ldots, n$, are given by

$$\begin{cases} \hat{\mu}_k = E(\mu_k|X_1, \ldots, X_n) = \displaystyle\sum_{i=0}^{k-1}\sum_{j=k}^{n} r_{[ij]}^* m_{[ij]}^*, & \text{if } d_{[ij]}^* > 1, \\ \hat{\sigma}_k^2 = E(\sigma_k^2|X_1, \ldots, X_n) = \displaystyle\sum_{i=0}^{k-1}\sum_{j=k}^{n} r_{[ij]}^* \frac{a_{[ij]}^*}{d_{[ij]}^* - 2}, & \text{if } d_{[ij]}^* > 2. \end{cases} \tag{2.14}$$

The posterior relevances $r_{[ij]}^*$ can be obtained from Eq. (2.7) and (2.10).

For the PPM considered in this paper, the prior cohesions proposed by Yao [28] are assumed, that is, for block $[ij]$ we consider

$$c_{[ij]} = \begin{cases} p(1-p)^{j-i-1}, & \text{if } j < n, \\ (1-p)^{j-i-1}, & \text{if } j = n, \end{cases} \tag{2.15}$$

for all $i, j \in I$, $i < j$, where p, $0 < p < 1$, is the probability that a change occurs at any instant in the sequence. Notice that, as a truncated geometric distribution is memoryless, this choice imply that we are assuming that the past change points are noninformative about the future change points.

Consequently, for the normal case, the posterior cohesion for the block $[ij]$ becomes

$$c_{[ij]}^* = \begin{cases} \dfrac{p(1-p)^{j-i-1} K(d_{[ij]}, j-i) a_{[ij]}^{d_{[ij]}/2}}{(1+(j-i)v_{[ij]})^{1/2}\{a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})\}^{(d_{[ij]}+j-i)/2}}, & \text{if } j < n, \\[2ex] \dfrac{(1-p)^{j-i-1} K(d_{[ij]}, j-i) a_{[ij]}^{d_{[ij]}/2}}{(1+(j-i)v_{[ij]})^{1/2}\{a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})\}^{(d_{[ij]}+j-i)/2}}, & \text{if } j = n. \end{cases} \tag{2.16}$$

2.4. *A Gibbs Sampling Scheme Applied to the PPM*

An extraordinary array of problems in Bayesian inference has been solved by Markov chain Monte Carlo (MCMC) methods since when the seminal paper by Gelfand and Smith [10] illustrated how easily a variety of intractable problems could be approximately solved. This ease of use led to an explosion of research since complex Bayesian models without analytical solution could be tractable by MCMC methods. Recent results and overviews of the research in this area includes the papers by MacEachern and Peruggia [19], Robert [25], and Besag et al. [5], to cite a few. In particular, our purpose is to use Gibbs sampling as a generation scheme of posterior distributions. We will now describe briefly the method to estimate the posterior distributions of B and ρ, and the posterior relevance of each block $[ij]$, as proposed by Loschi et al. [17].

Let assume the auxiliary random quantity U_i which reflects whether or not a change point occurred at the time i, that is

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

where $i = 1, \ldots, n-1$.

The vector $\mathbf{U}^k = (U_1^k, \ldots, U_{n-1}^k)$ at the kth step is generated by using the Gibbs sampling as follows.

Considering a beta prior distribution for p, that is, $p \sim \mathcal{B}(\alpha, \beta)$, $\alpha, \beta > 0$, it is sufficient to consider the ratio given by the following expression, in order to generate the vectors $\mathbf{U}^k$'s:

$$R_r = \frac{f_{[xy]}(X_{[xy]})\Gamma(n+\beta-b+1)\Gamma(b+\alpha-2)}{f_{[xr]}(X_{[xr]})f_{[ry]}(X_{[ry]})\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}, \tag{2.17}$$

where:

$$x = \begin{cases} \boxed{\begin{array}{ll} \max i & \\ \textbf{s.t.:} & 0 < i < r, \\ & U_i^k = 0, \end{array}} & \begin{array}{l} \text{if there is } U_i^k = 0, \text{ for} \\ \text{some } i \in \{1, \ldots, r-1\}, \end{array} \\ 0, & \text{otherwise,} \end{cases}$$

and

$$y = \begin{cases} \boxed{\begin{array}{ll} \min i & \\ \textbf{s.t.:} & r < i < n, \\ & U_i^{k-1} = 0, \end{array}} & \begin{array}{l} \text{if there is a } U_i^{k-1} = 0, \text{ for} \\ \text{some } i \in \{r+1, \ldots, n-1\}, \end{array} \\ 0, & \text{otherwise,} \end{cases}$$

since the rth element at the kth step U_r^k is generated from the conditional distribution:

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

where $r = 1, \ldots, n-1$, starting from an initial vector $\mathbf{U}^0 = (U_1^0, \ldots, U_{n-1}^0)$.

Notice that, in the normal case, $f_{[ij]}(X_{[ij]})$ is the Student-t distribution given in Eq. (2.10). Consequently, the criterion for choosing the values U_r^k becomes

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

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

We remark that the posterior relevance of the block $[ij]$, for $i < j$, used in Eq. (2.6) to estimate θ_k, can be obtained by computing the proportion of samples that presents $U_i^k = 0$, $U_{i+1}^k = \ldots = U_{j-1}^k = 1$, and $U_j^k = 0$. The random quantity ρ is perfectly identified by considering a vector of these random quantities. Consequently, we can estimate the posterior probability for each particular partition in b contiguous blocks, $\rho = \{i_0, i_1, \ldots, i_b\}$. We also remark that it is possible to use this

algorithm
    read *all prior specifications*
    read $X_1, \ldots, X_n$
    for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
        $c^*_{[ij]} \leftarrow$ Eq. (2.16)
    end for
    for $k = 1$ **to** SAMPLES **do**
        generate $\mathbf{U}^k$
    end for
    for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
        $r^*_{[ij]} \leftarrow$ proportion of samples such that
$$U^k_i = 0, U^k_{i+1} = \cdots = U^k_{j-1} = 1, U^k_j = 0$$
    end for
    for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
$$\bar{X}_{[ij]} \leftarrow \tfrac{1}{j-i} \sum_{r=i+1}^{j} X_r$$
$$m^*_{[ij]} \leftarrow \tfrac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \tfrac{m_{[ij]}}{(j-i)v_{[ij]}+1}$$
$$v^*_{[ij]} \leftarrow \tfrac{v_{[ij]}}{(j-i)v_{[ij]}+1}$$
$$d^*_{[ij]} \leftarrow d_{[ij]} + j - i$$
$$q_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \sum_{r=i+1}^{j} (X_r - \bar{X}_{[ij]})^2 + \tfrac{(j-i)(\bar{X}_{[ij]} - m_{[ij]})^2}{(j-i)v_{[ij]}+1}$$
$$a^*_{[ij]} \leftarrow a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})$$
    end for
    for $k = 1$ **to** n **do**
$$E(\mu_k | X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} m^*_{[ij]}$$
$$E(\sigma^2_k | X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \tfrac{a^*_{[ij]}}{d^*_{[ij]} - 2}$$
    end for
    write $E(\mu_1), E(\sigma^2_1), \ldots, E(\mu_n), E(\sigma^2_n)$
    write B^k
end algorithm

Fig. 2. PPM Gibbs sampling algorithm.

procedure to estimate the posterior distribution of B (or the posterior distribution of the number of change points, $B-1$) by noticing that

$$B^k = 1 + \sum_{i=1}^{n-1}(1 - U_i^k),$$

which completes the algorithm, shown in pseudo-code in Fig. 2.

3. Application to the Brazilian Indexes

In this section, we will focus on the identification of multiple change points in the expected means (returns) and variances (volatilities) of two important Brazilian indexes, IBOVESPA and IBOVMESB, available from the authors upon request. As usual in finance, these return series were defined by the transformation $R_k = (P_k - P_{k-1})/P_{k-1}$, where P_k is the closing price at the kth month. From Fig. 1, it seems that both series are similar and some changes along the years are suggested for the means and variances of the returns. We will see now how the PPM can be of help in capturing the time series structure.

3.1. *Remarks on Modeling Means and Variances*

The observations are assumed to be conditionally independent and distributed according to the normal distribution $\mathcal{N}(\mu_{[ij]}, \sigma^2_{[ij]})$, and we 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. These assumptions are not too restrictive, since the normal distribution is appropriated for many practical applications and it was used to successfully model stock market data in the past [13]. Additionally, the normal-inverted-gamma distribution is rich enough to describe the uncertainty about the parameters under many practical circumstances. As a consequence of these assumptions, the returns are distributed according to a Student-t distribution, which is appropriate to describe the stock behavior of emerging markets since the Student-t distribution discloses a structure of correlation amongst the returns and has heavier tails than the normal distribution.

For the present case, it seems reasonable to consider that

$$\begin{cases} v_{[ij]} = 1.0, \\ a_{[ij]} = 1.0, \\ d_{[ij]} = 4.0, \end{cases} \tag{3.1}$$

and $m_{[ij]}$ being considered the general average of the respective series, for all $i, j \in I$.

In other words, from Eq. (2.9), we are considering

$$\begin{cases} \mu_{[ij]} | \sigma^2_{[ij]} \sim \mathcal{N}(0.1569, \sigma^2_{[ij]}), \\ \sigma^2_{[ij]} \sim \mathcal{IG}(1/2, 4/2), \end{cases} \tag{3.2}$$

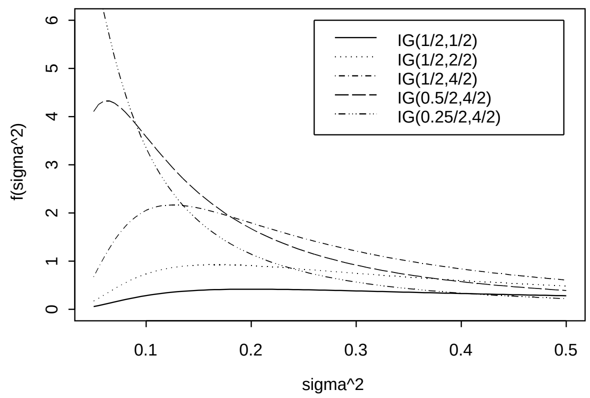


Fig. 3. Probability function $\mathcal{IG}(a/2, d/2)$.

for the IBOVESPA and

$$\begin{cases} \mu_{[ij]} | \sigma^2_{[ij]} \sim \mathcal{N}(0.1668, \sigma^2_{[ij]}), \\ \sigma^2_{[ij]} \sim \mathcal{IG}(1/2, 4/2), \end{cases} \tag{3.3}$$

for the IBOVMESB.

About this subjective choice, we found reasonable to consider for those series that the returns are not null on average as supposed by Hawkins [12]. If someone has a different point of view, the parameter $m_{[ij]}$ should be conveniently adjusted. Concerning the volatility, the inverted-gamma distribution considered, $\mathcal{IG}(1/2, 4/2)$, plotted in Fig. 3, concentrates its mass in a low value and it is also as flat as our uncertainty about this parameter. These specifications can be supported by the fact that the Brazilian market is an emerging market, more susceptible to the political scenario than developed markets [21]. Other similar settings for $a_{[ij]}$ and $d_{[ij]}$ where considered but the results (not shown) do not differ very much.

3.2. *Remarks on Modeling Change Points*

And how about the change points? Is it really appropriated to model the prior cohesions the way we did in Eq. (2.15)? It seems that the answer is yes, since it is reasonable to assume that changes in the behavior of Brazilian stock return series are a consequence of the receipt of non previously anticipated information (see Loschi [15], for a in-depth view upon unpredictability), so that past changes points are non-informative concerning future change points (see Mandelbrot [20]).

Thus, one last decision that has to be made concerns the probability p of having a change. If we assumed a high value for p, we would be previously considering that there are small blocks of returns. Consequently, we would be assuming the existence of a high number of change points. However, we will assume $p \sim \mathcal{B}(2, 32)$, plotted

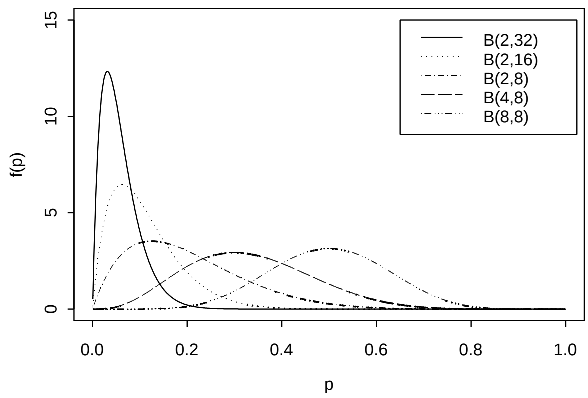


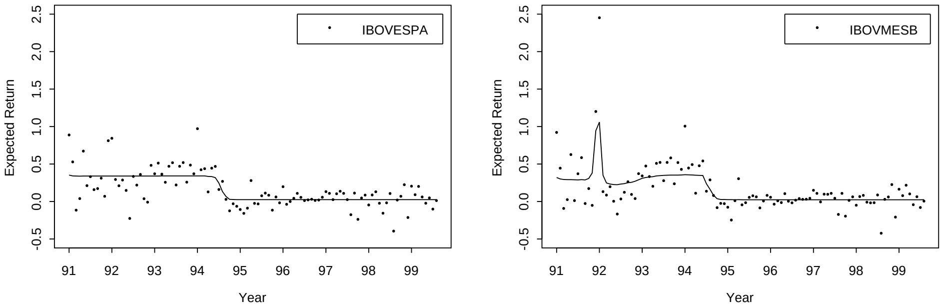
Fig. 4. Probability function $\mathcal{B}(\alpha, \beta)$.

in Fig. 4, because this distribution concentrates most of its mass in small values of p and it is reasonable to expect a small number of changes in both IBOVESPA and IBOVMESB. This assumption is also in agreement with the remarks by Loschi and Cruz [16].

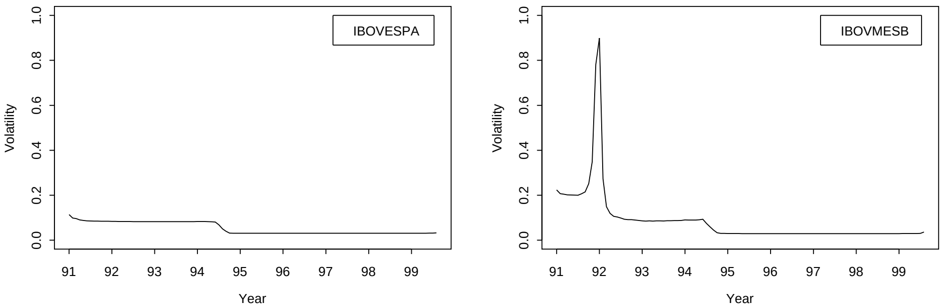
3.3. *Computational Results*

Because of its computationally intensive nature, the algorithm presented in Fig. 2 was coded in C++ (available from the authors). All tests, performed in a PC, Pentium processor 166 MHz, 32 MB RAM, took around 25 seconds of CPU time. In order to estimate the posterior relevances $r^*_{[ij]}$ and the posterior distribution of B (or the number of change points, $B - 1$), we generate 10,000 samples of 0-1 values with dimension 103, starting from a sequence of zeros. We discarded the initial 4,000 iterations and we selected a lag of ten to avoid correlation, what means that we worked with a net sample size of 600. Discussions about the number of iterations to be discharged and the lag to be taken can be easily found in the literature (see, for instance, the book of Gamerman [9]).

Figure 5 contrasts posterior estimates for the expected returns and expected volatilities for IBOVESPA and IBOVMESB at each instant of the time. We notice that the change points occurred typically at the same time and in the same direction for both series. However, against our initial expectations, considerable differences in the behavior of these series were observed. One of the changes observed in IBOVMESB, around 92, did not occur in IBOVESPA. These change points could be related to the privatization of the USIMINAS, a important state steel company located in Minas Gerais state, and to the impeachment of the Brazilian president. These important historical facts however do not seem to produce any change in the behavior of IBOVESPA. This different behavior could be explained by the fact that IBOVESPA is the main indicator of the Brazilian economy, incorporating im-



a) Expected returns.



b) Expected volatilities.

Fig. 5. Posterior estimates.

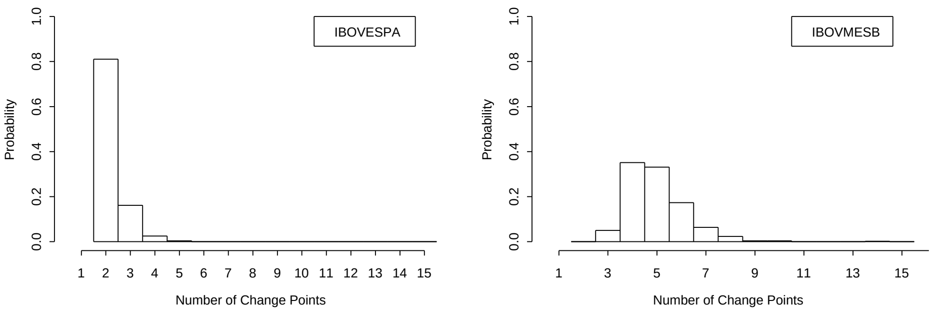


Fig. 6. Posterior distribution of B.

mediately the benefits of the Brazilian government policies. The latest Brazilian currency, the Real, was introduced in July, 1994, and this fact is well captured by the PPM. The Real period has presented lower expected returns and volatilities than the previous period. Both indexes seem to have similar behavior in the Real period.

Figure 6 shows the posterior distributions of B, the number of blocks that occurred for each index, and Table 1 presents the respective descriptive statistics. We

Table 1. Descriptive Statistics for the Posterior Distribution of B

	Mean	Median	Mode	StDev	Q1	Q3	Min	Max
IBOVESPA	2.22	2	2	0.49	2	2	2	5
IBOVMESB	4.96	5	4	1.21	4	6	3	14

Table 2. Some Partitions and their Posterior Probabilities

	partition	posterior probability
IBOVESPA	$\rho = \{0$, Jul. 94, Aug. 99$\}$	0.32000
	$\rho = \{0$, Jun. 94, Aug. 99$\}$	0.16833
	$\rho = \{0$, Aug. 94, Aug. 99$\}$	0.15167
	$\rho = \{0$, Sep. 94, Aug. 99$\}$	0.10667
	$\rho = \{0$, May 94, Aug. 99$\}$	0.04333
IBOVMESB	$\rho = \{0$, Nov. 91, Jan. 92, Aug. 94, Aug. 99$\}$	0.05500
	$\rho = \{0$, Nov. 91, Jan. 92, Jun. 94, Aug. 99$\}$	0.05500
	$\rho = \{0$, Nov. 91, Jan. 92, Jul. 94, Aug. 99$\}$	0.02833
	$\rho = \{0$, Nov. 91, Jan. 92, Sep. 94, Aug. 99$\}$	0.02333
	$\rho = \{0$, Nov. 91, Feb. 92, Sep. 94, Aug. 99$\}$	0.01667

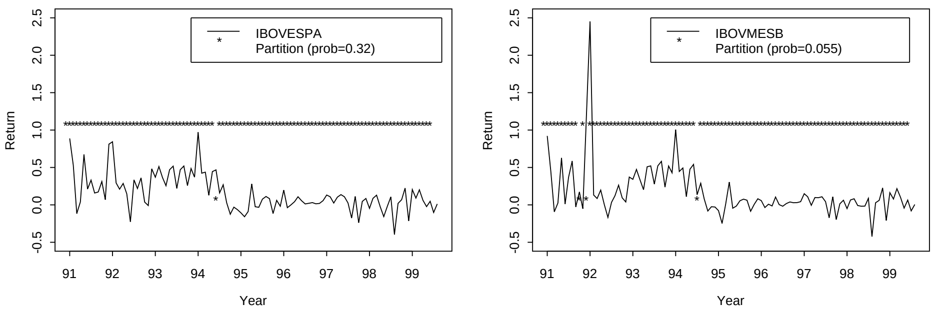


Fig. 7. The most probable partition.

notice that most of their masses are concentrated in small values, as expected. For IBOVESPA, however, the posterior distribution is more concentrated (its standard deviation is 0.49, which is lower then 1.21 — the standard deviation for IBOVMESB) and concentrates its mass on smaller values than in the IBOVMESB case (which can be perceived by the smaller values for the mean, mode, and median). Thus, we can conclude that the IBOVESPA series comes from a more stable market.

Finally, Table 2 show the five most probable partitions for IBOVESPA and IBOVMESB and their respective probabilities. We should notice that it is more probable that IBOVESPA suffered only one change (in July, 1994) and that IBOVMESB suffered three changes (in November, 1991, January, 1992, and August, 1994). The most probable partitions for each index are shown in Fig. 7. The symbol '$*$' plotted in zero means that a change occurred at that point. Here, we can also observe that more changes occurred in IBOVMESB than in IBOVESPA.

4. Final Remarks and Future Directions

We described the PPM, particularly to analyze time series, and stressed its importance for the identification of multiple change points. We detailed a Gibbs sampling scheme that avoided some of the computational difficulties of the PPM. The algorithm was coded, tested, and proved to be an efficient and useful tool for the analysis of financial data, since, in the examples considered, the results satisfactorily explained the behavior of the series.

We concluded that the IBOVESPA and IBOVMESB series had a very similar behavior and could probably be suffering influences from the same local events. We noticed that both indexes presented clusters in the expected returns and volatilities, and presented a small number of change points. These same conclusions were also driven from the Chilean stock market [18], disclosing the similarities among the Brazilian and Chilean markets. São Paulo and Minas Gerais states are two of the most important economies in Brazil, thus having high political influence. Hence, as Minas Gerais is the strongest economy involved in IBOVMESB, the similarities observed in the behavior of the IBOVESPA and IBOVMESB series are justified.

Some open questions remains. Would it be possible to find even simpler implementations for the product partition model? How big would the treatable series be? How well does the methodology fit for other subject areas? These and other similar questions are interesting and relevant topics for future research in this area.

Acknowledgments

R. H. Loschi acknowledges CAPES, FAPEMIG, PRPq-UFMG (grants 3893-UFMG/RTR/PRPq/FUNDO/98 and 4801-UFMG/ RTR/ FUNDO/ PRPq/ RECEMDOUTORES/00) and Fundación Andes, Chile, for a partial allowance for her research. The research of F. R. B. Cruz was partially funded by the CNPq (*Conselho Nacional de Desenvolvimento Científico e Tecnológico*) of the Ministry for Science and Technology of Brazil, grants 301809/96-8 and 201046/94-6, the FAPEMIG (*Fundação de Amparo à Pesquisa do Estado de Minas Gerais*), grants CEX-289/98 and CEX-855/98, and PRPq-UFMG, grant 4081-UFMG/RTR/FUNDO/PRPq/99.

References

[1] R. B. Arellano-Valle and H. Bolfarine. On some characterizations of the t-distribution, *Statistics & Probability Letters* **25**, 79–85 (1995).
[2] D. Barry and J. A. Hartigan. Product partition models for change point problems, *The Annals of Statistics* **20** (1), 260–279 (1992).
[3] D. Barry and J. A. Hartigan. A Bayesian analysis for change point problem, *Journal of the American Statistical Association* **88** (421), 309–319 (1993).
[4] J. M. Bernardo and A. F. M. Smith. *Bayesian Theory*, (John Wiley & Sons, Chichester, England, 1994), 1st edition.

[5] J. Besag, D. H. P. Green, and K. Mengersen. Bayesian computation and stochastic systems, *Statistical Science* **10** 3–66 (1995).

[6] C. W. S. Chen and J. C. Lee. Bayesian inference of threshold autoregressive models, *Journal of Time Series Analysis* **16** (5), 483–492 (1995).

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

[8] M. H. DeGroot. *Probability and Statistics* (Addison-Wesley, California, USA, 1989), 2nd edition.

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

[10] A. E. Gelfand and A. F. M. Smith. Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**, 398–409 (1990).

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

[12] D. M. Hawkins. Fitting multiple change-point models to data, *Computational Statistics & Data Analysis* **37** (3), 323–341 (2001).

[13] D. A. Hsu. A Bayesian robust detection of shift in the risk structure of stock market returns, *Journal of the American Statistical Association* **77** (2), 29–39 (1982).

[14] V. K. Jandhyala, S. B. Fotopoulos, and D. M. Hawkins. Detection and estimation of abrupt changes in the variability of a process, *Computational Statistics & Data Analysis* **40** (1), 1–19 (2002).

[15] R. H. Loschi. *Imprevistos e suas Conseqüências (Unpredictabilities and Consequences)*. PhD thesis, Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil, 1998 (in Portuguese).

[16] R. H. Loschi and F. R. B. Cruz. An analysis of the influence of some prior specifications in the identification of change points via product partition model, *Computational Statistics & Data Analysis* **39** (4), 477–501 (2002).

[17] R. H. Loschi, 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, *Computers & Operations Research* **30** (3), 463–482 (2003).

[18] R. H. Loschi, P. L. Iglesias, and R. B. Arellano-Valle. Bayesian detection of change points in the Chilean stock market, In *Proc. of the Section on Bayesian Statistical Science*, pages 160–165, Baltimore, MD, USA, 1999. Annual Meeting of American Statistical Association.

[19] S. N. MacEachern and M. Peruggia. Subsampling the Gibbs sampler: Variance reduction, *Statistics & Probability Letters* **47**, 91–98 (2000).

[20] B. Mandelbrot. The variation of certain speculative prices, *Journal of Business* **36**, 394–419 (1963).

[21] B. V. Mendes. Computing robust risk measures in emerging markets using extreme value theory, *Emerging Markets Quarterly* **4** (2), 25–42 (2000).

[22] U. Menzefricke. A Bayesian analysis of a change in the precision of a sequence of independent normal random variables at an unknown time point, *Applied Statistics* **30** (2), 141–146 (1981).

[23] H. S. Migon and D. Gamerman. *Statistical Inference: An Integrated Approach* (Arnold, London, Great Britain, 1999).

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

[25] C. Robert. Convergence control methods for Markov chain Monte Carlo algorithms, *Statistical Science* **10**, 231–253 (1995).

[26] A. F. M. Smith. A Bayesian approach to inference about a change-point in a sequence of random variables, *Biometrika*, **62** (2), 407–416 (1975).

[27] D. Stauffer. Percolation models of financial market dynamics, *Advances in Complex Systems* **4** (1), 19–27 (2001).

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