A Gibbs Sampling Scheme to the Product Partition Model: An Application to Change-Point Problems

R. H. Loschi [a,1], F. R. B. Cruz [b,2], P. Iglesias [c], R. B. Arellano-Valle [c]

[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. E-mail: fcruz@ecs.umass.edu*

[c] *Facultad de Matematicas, Pontificia Universidad Catolica de Chile, Casilla 306, Santiago 22, Chile. E-mail: {pliz,reivalle}@mat.puc.cl*

Abstract

This paper extends previous results for the classical product partition model (PPM) applied to the identification of multiple change points in the means and variances of time series. Prior distributions for these two parameters and for the probability p that a change takes place at a particular period of time are considered and a new scheme based on Gibbs sampling to estimate the posterior relevances of the model is proposed. The resulting algorithm is applied to the analysis of two Brazilian stock market data. The computational experiments seem to indicate that the algorithm runs fast in common PC-like machines and it may be a useful tool for analyzing change-point problems.

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.*

Scope and Purpose

The problem of change-point identification is encountered in many subject areas, including disease mapping, medical diagnosis, industrial control, and finance. A Bayesian way to tackle the problem is through the well-known product partition model (PPM) introduced by Hartigan in the early 90's. Nowadays, the PPM is still attracting researchers' attention because of its flexibility and the spreading use of the powerful personal computers that make it possible to deal with its inherent computational complexity. In this paper, the PPM is tailored to the identification of change points both in the means and variances of time series, assuming that, given these parameters, the data are normally distributed. We extend some previous works by considering a non-degenerate prior distribution to the probability p of having a change point at a particular period of time. An original Gibbs sampling scheme is also developed to compute the product estimates and, consequently, to attack the difficult resulting model which is applied to the identification of change points in the expected returns (means) and volatilities (variances) of two important stock market data in Brazil. The computational results seem to indicate that method is effective and efficient, making it possible useful inferences. In addition, the method is simple and easy to implement.

1 Introduction

The identification of change points is important in many data analysis problems, such as disease mapping, medical diagnosis and industrial control. This problem also arises in stock market analysis. Indeed, Figure 1 shows the IBOVESPA ("*Índice Geral da Bolsa de São Paulo*") and the IBOVMESB ("*Índice da Bolsa de Valores de Minas Gerais, Espírito Santo e Brasília*") series, two of the most important Brazilian indexes. Both are expressed in terms of the returns calculated on closing prices recorded monthly. Inferences on the instants when changes occurred in the expected returns (means) and in the volatilities (variances) of such series, for example, allows the identification of events that could have produced the changes, helping decision makers in the future under similar situations.

Bayesian approaches for the change-point problem have been presented by several authors. For example, Menzefricke [2] considers the problem of making inferences about a change point in the precision of normal data with unknown mean. A single change point in the functional form of the distribution is explored by Hsu [3], who considers the class of the exponential-power distributions Box and Tiao [4] for treating the problem. Hsu [3] and Menzefricke [2] apply their methodologies to stock market prices (see also Smith [5]). Stephens

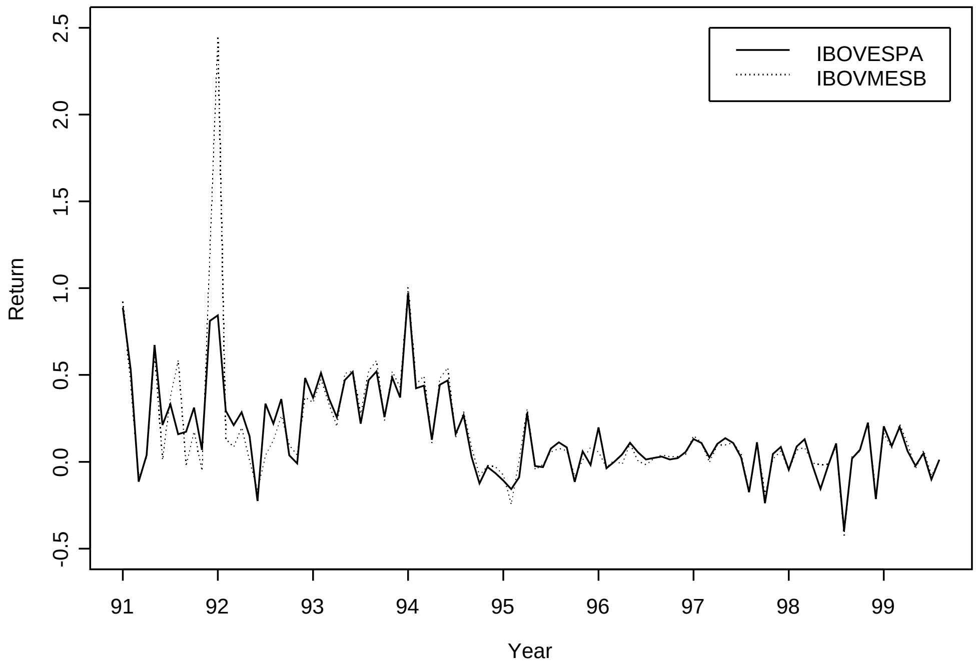


Fig. 1. IBOVESPA and IBOVMESB Return Series

[6] discusses the discrete multiple change-point problem and the continuous single-change point problems, which is illustrated by considering some kidney transplant data. Stephens [6] also focuses on the computational complexity involved in the change-point identification.

Later, Hartigan [1] proposes the Product Partition Model (PPM), which generalizes most of the situations described above. The PPM allows the identification of multiple change points in the parameters as well as in the distribution function itself. Besides, the PPM assumes that the number of change points and also the instants when the changes occur are random variables, which makes the PPM more interesting and flexible than those models that consider the number of changes as fixed (threshold models and the method based on maximum likelihood estimation considered by Hawkins [7], for example). The PPM is considered by Barry and Hartigan [8] to identify multiple change points in normal means with common variance. Recently, Crowley [9] provides a new implementation of the Gibbs sampling scheme and also consider an empirical-Bayes approach in order to solve the problem of estimating normal means by using the PPM. Quintana and Iglesias [10] present a decision-theoretical approach formulation to the PPM and connect the PPM to the the Dirichlet process. The PPM is also used by Loschi et al. [11] to identify multiple change points both in the means and variances of normal data. Loschi et al. [11] consider the recursive algorithm proposed by Yao [12] to obtain the posterior relevances (and, consequently, the product estimates of the normal means and variances) and implement a Gibbs sampling scheme to estimate the posterior distributions of the number of change points and the instants when changes occur. Loschi et al. [11] also consider the prior cohesion defined by Yao [12], which depends on the probability p that a change takes place at any time, assuming a degenerate prior distribution for that probability.

This paper extends Loschi et al.'s results by assuming rather a non-degenerate prior distribution to the parameter p involved in the Yao's cohesions and by providing an original procedure to evaluate the posterior relevances based on a Gibbs sampling scheme. The algorithm proposed is applied to identify multiple changes in the mean returns and in the volatilities of the IBOVESPA and IBOVMESB series to illustrate the use of the method, and to point out some events that could have produced the changes.

The paper is organized as follows. Section 2 briefly reviews the PPM and presents inferential solutions to identify change points in random variables normally distributed, given the means and variances. Section 2 also presents the exact posterior relevances, the posterior distributions of the random partition generated by the change points, and the posterior distribution of the number of change points in the partition. In Section 3, a Gibbs sampling scheme is proposed to compute the posterior relevances, the posterior distributions of the number of blocks in the partition generated by the change points, and the posterior distribution of this random partition. In Section 4, the algorithm is applied to the identification of change points in the mean returns and in the volatilities of the IBOVESPA and IBOVMESB. Section 5 closes the paper with final remarks and future topics for investigation.

2 Statistical Models

In this section, the PPM and some preliminary results are presented, as shown in Barry and Hartigan [13, 8]. Details are given on how the PPM may be tailored to identify multiple change points in the means and variances of normal data, as presented by Loschi et al. [11]. Then, original results are presented on how to compute the posterior relevances of the model and the exact posterior distributions of the blocks and the number of change points, considering the beta prior distribution for the parameter p (probability of having a change) involved in the Yao's cohesions.

2.1 The PPM

Let $X_1, \ldots, X_n$ be a observed time series. Consider a random partition ρ of the set $I = \{1, \ldots, n\} \cup \{0\}$ and a random variable B that represents the number of blocks in ρ. Consider that each partition $\rho = \{i_0, i_1, \ldots, i_b\}$, $0 = i_0 < i_1 < \cdots < i_b = n$, divides the sequence $X_1, \ldots, X_n$ into $B = b$, $b \in I$, contiguous subsequences, which will be denoted 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 to 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, it is said 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 end points $i_1, \cdots, i_b$, satisfying the condition $0 = i_0 < i_1 < \ldots < i_b = n$, $b \in I$;

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, $\mathbf{X}_{[ij]} = (X_{i+1}, \ldots, X_j)'$.

Notice that the PPM described above describes the uncertainty about the random object $(X_1, \ldots, X_n; \rho)$, if the prior opinion about this object discloses the existence of blocks of observations produced by some judgment of similarities (in some sense) among these observations, as well as independence among the different blocks.

Also note that 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 $I \cup \{0\}$ in b (fix) contiguous blocks.

As shown in Barry and Hartigan [13], 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 $c^*_{[ij]} = c_{[ij]} f_{[ij]}(\mathbf{X}_{[ij]})$. That is, the PPM induces some kind of conjugacy.

In the parametric approach to the 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. In this case, it is considered that two observations X_i and X_j, such that $i \neq j$, are in the same block, if it is believed that they are identically distributed. Thus, in this approach to the PPM, the predictive distribution $f_{[ij]}(X_{[ij]})$, which appeared in (2), can be obtained as follows:

$$f_{[ij]}(\mathbf{X}_{[ij]}) = \int\limits_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\ \pi_{[ij]}(\theta) d\theta, \tag{4}$$

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 $\mathbf{X}_{[ij]} = (X_{i+1}, \ldots, X_j)'$.

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 $\pi_{[ij]}(\theta)$, $\theta \in \Theta_{[ij]}$.

Hence, the goal is to obtain the marginal posterior distributions of the parameters ρ, B, and θ_k, $k = 1, \ldots, n$. Barry and Hartigan [13] have shown that the posterior distributions of θ_k is given by:

$$\pi(\theta_k|X_1, \ldots, X_n) = \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]}\ \pi_{[ij]}(\theta_k|\mathbf{X}_{[ij]}), \tag{5}$$

for $k = 1, \ldots, n$, and the posterior expectation of θ_k is given by:

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

for $k = 1, \ldots, n$, where $r^*_{[ij]}$ denotes the posterior relevance for the block $[ij]$, that is:

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

In pseudo-language, Figure 2 shows the PPM to solve the change-point identification problem.

algorithm
(1) **read** $X_1, \ldots, X_n$
for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
(2) $r^*_{[ij]} \leftarrow P([ij] \in \rho | X_1, \ldots, X_n)$
end for
for $k = 1$ **to** n **do**
(3) $E(\theta_k | X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \; E(\theta_k | \mathbf{X}_{[ij]})$
end for
(4) **write** $E(\theta_1), \ldots, E(\theta_n)$
end algorithm

Fig. 2. The PPM

2.2 The Normal PPM, the Mean and Variance Case

To specify the PPM for normal data, it is assumed 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 $X_k | \mu_k, \sigma_k^2 \sim N(\mu_k, \sigma_k^2)$, $k = 1, \ldots, n$, and that the parameters are independent. It is also assumed that each common parameter $\theta_{[ij]} = (\mu_{[ij]}, \sigma^2_{[ij]})$, related to the block $[ij]$, has the conjugate normal-inverted-gamma prior distribution denoted by:

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

that is,

$$\left.\begin{array}{l} \mu_{[ij]} | \sigma^2_{[ij]} \sim N(m_{[ij]}, v_{[ij]} \sigma^2_{[ij]}) \text{ and} \\ \sigma^2_{[ij]} \sim IG(a_{[ij]}/2, d_{[ij]}/2), \end{array}\right\} \tag{7}$$

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

$$(\mu_{[ij]}, \sigma^2_{[ij]}) | \mathbf{X}_{[ij]} \sim NIG(m^*_{[ij]}, v^*_{[ij]}; a^*_{[ij]}/2, d^*_{[ij]}/2), \tag{8}$$

where

$$\left.\begin{array}{l} m^*_{[ij]} = \frac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \frac{m_{[ij]}}{(j-i)v_{[ij]}+1}, \\ v^*_{[ij]} = \frac{v_{[ij]}}{(j-i)v_{[ij]}+1}, \\ d^*_{[ij]} = d_{[ij]} + j - i, \\ a^*_{[ij]} = a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]}), \end{array}\right\} \quad (9)$$

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 follows from Eq. (8) that, given $X_{[ij]}$, the conditional marginal densities of $\mu_{[ij]}$ and $\sigma^2_{[ij]}$ are, respectively:

$$\left.\begin{array}{l} \mu_{[ij]}|\mathbf{X}_{[ij]} \sim t(m^*_{[ij]}, v_{[ij]}, a^*_{[ij]}, d^*_{[ij]}) \text{ and} \\ \sigma^2_{[ij]}|\mathbf{X}_{[ij]} \sim IG(a^*_{[ij]}/2, d^*_{[ij]}/2), \end{array}\right\} \quad (10)$$

for which it is observed that (the interested reader may find details in O'Hagan [14]):

$$\left.\begin{array}{l} E(\mu_{[ij]}|X_{[ij]}) = m^*_{[ij]}, \ \text{ if } d^*_{[ij]} > 1, \ \text{and} \\ E(\sigma^2_{[ij]}|X_{[ij]}) = \frac{a^*_{[ij]}}{d^*_{[ij]}-2}, \text{ if } d^*_{[ij]} > 2. \end{array}\right\} \quad (11)$$

From Eq. (6) and (11), it follows that the posterior estimates for the parameters μ_k and σ^2_k are given by:

$$\left.\begin{array}{l} E(\mu_k|X_1,\ldots,X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]} m^*_{[ij]}, \ \text{ if } d^*_{[ij]} > 1, \ \text{and} \\ E(\sigma^2_k|X_1,\ldots,X_n) = \sum_{i=0}^{k-1}\sum_{j=k}^{n} r^*_{[ij]} \frac{a^*_{[ij]}}{d^*_{[ij]}-2}, \ \text{if} d^*_{[ij]} > 2. \end{array}\right\} \quad (12)$$

algorithm
  read $X_1, \ldots, X_n$
  for all $i, j \in \{0, \ldots, n\}$ **such that** $i < j$ **do**
    $r^*_{[ij]} \leftarrow P([ij] \in \rho | X_1, \ldots, X_n)$
    $\bar{X}_{[ij]} \leftarrow \frac{1}{j-i} \sum_{r=i+1}^{j} X_r$
    $m^*_{[ij]} \leftarrow \frac{(j-i)v_{[ij]}\bar{X}_{[ij]}}{(j-i)v_{[ij]}+1} + \frac{m_{[ij]}}{(j-i)v_{[ij]}+1}$
    $v^*_{[ij]} \leftarrow \frac{v_{[ij]}}{(j-i)v_{[ij]}+1}$
    $d^*_{[ij]} \leftarrow d_{[ij]} + j - i$
    $q_{[ij]}(\mathbf{X}_{[ij]}) \leftarrow \sum_{r=i+1}^{j} (X_r - \bar{X}_{[ij]})^2 + \frac{(j-i)(\bar{X}_{[ij]} - m_{[ij]})^2}{(j-i)v_{[ij]}+1}$
    $a^*_{[ij]} \leftarrow a_{[ij]} + q_{[ij]}(\mathbf{X}_{[ij]})$
  end for
  for $k = 1$ **to** n **do**
    $E(\mu_k | X_1, \ldots, X_n) \leftarrow \sum_{0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \; m^*_{[ij]}$
    $E(\sigma^2_k | X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \; \frac{a^*_{[ij]}}{d^*_{[ij]}-2}$
  end for
  write $E(\mu_1), E(\sigma^2_1), \ldots, E(\mu_n), E(\sigma^2_n)$
end algorithm

Fig. 3. The Normal PPM, Mean and Variance Case

respectively, with $k = 1, \ldots, n$, where $m^*_{[ij]}$, $a^*_{[ij]}$ and $d^*_{[ij]}$ are defined as in Eq. (9).

Figure 3 shows the normal PPM, for the mean and variance case.

2.3 The Exact Posterior Distributions of ρ and B and the Posterior Relevances $r^*_{[ij]}$

As one may already have noticed, the algorithm presented in Figure 3 gives no details at all on how to compute the posterior relevance $r^*_{[ij]}$. Assuming only the existence of contiguous blocks, the prior cohesions, as defined by Yao [12], can be interpreted as the transition probabilities in the Markov chain defined by the endpoints of the blocks in the partition ρ. Let $0 \leq p \leq 1$ be the probability that a change occurs at any instant in the sequence. Thus, the prior cohesion for block $[ij]$, $c_{[ij]}$, corresponds to the probability that a new change takes place after $j - i$ instants, given that a change has taken place at the instant i, that is:

$$c_{[ij]} = \begin{cases} p(1-p)^{j-i-1}, \text{ if } j < n, \\ (1-p)^{j-i-1}, \quad \text{if } j = n, \end{cases} \tag{13}$$

for all $i, j \in I$, such that $i < j$.

Consequently, from Eq. (1), one can obtained that the prior distribution of ρ takes the form:

$$P(\rho = \{i_0, i_1, \ldots, i_b\}) = p^{b-1}(1-p)^{n-b}, b \in I,$$

and from Eq. (2), it follows that the prior distribution of the random variable B is given by:

$$P(B = b) = C_{b-1}^{n-1} p^{b-1}(1-p)^{n-b}, \quad \forall\, b \in I,$$

where C_{b-1}^{n-1} is the number of distinct partitions of I into b contiguous blocks.

Assume that p has the beta prior distribution with $\alpha > 0$ and $\beta > 0$ parameters, denoted by $p \sim B(\alpha, \beta)$. Let $\mathcal{C}$ be the set of all partitions of the set I into b contiguous blocks with endpoints $i_0, \ldots, i_b$, satisfying the condition $0 = i_0, \ldots, i_b = n$, $b \in I$ and consider $\mathcal{C}_1 \subset \mathcal{C}$ the subset of all partitions that contain the block $[ij] = \{i+1, \ldots, j\}$. Thus, since $\alpha > 1$ and $\beta > 1$, the posterior distribution of the random partition ρ is given by:

$$P(\rho = \{i_0, i_1, \ldots, i_b\} | X_1, \ldots, X_n) = \frac{\left\{\Pi_{j=1}^{b} f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]})\right\}}{\sum_{\mathcal{C}} \left\{\Pi_{j=1}^{b} f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]})\right\}} \times \frac{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}. \tag{14}$$

The posterior probability of the event $B = b$, $b \in I$, is given by multiplying the posterior probability in Eq. (14) by C_{b-1}^{n-1}. Notice that the posterior distributions of ρ and B do not have the product distribution presented in Section 2.1 (as obtained by Loschi et al. [11]).

The exact posterior relevance $r_{[ij]}^*$ to the block $[ij]$, for $i < j$, can be calculated as follows:

$$r^*_{[ij]} = \frac{\sum_{\mathcal{C}_1} \Pi^k_{j=1,i_k=i} f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]}) f_{[ij]}(X_{[ij]})}{\sum_{\mathcal{C}} \Pi^b_{j=1} f_{[i-j-1i_j]}(X_{[i_{j-1}i_j]})} \times$$

$$\Pi^b_{j=k+2,i_{k+1}=j} f_{[i_{j-1}i_j]}(X_{[i_{j-1}i_j]}) \times$$

$$\frac{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}. \tag{15}$$

Denote by $\mathbf{1}_n$ the n-dimensional vector of ones and let $\mathbf{I}_n$ be the $n \times n$-dimensional identity matrix. If the PPM presented in Section 2.2 is assumed which consider conditionally normally distributed data, it follows that each block of observations $\mathbf{X}_{[ij]}$ has the $(j-i)$-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:

$$f_{[ij]}(\mathbf{X}_{[ij]}) = \frac{\Gamma[(d_{[ij]}+j-i)/2]}{\Gamma[d_{[ij]}/2]\pi^{k/2}} a_{[ij]}^{d_{[ij]}/2} |\mathbf{V}_{[ij]}|^{-1/2} \times$$

$$\{a_{[ij]} + (\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})' \mathbf{V}^{-1}_{[ij]} (\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})\}^{-(d_{[ij]}+j-i)/2}, \tag{16}$$

where $\mathbf{m}_{[ij]} = m_{[ij]}\mathbf{1}_{j-i}$ and $\mathbf{V}_{[ij]} = \mathbf{I}_{j-i} + v_{[ij]}\mathbf{1}_{j-i}\mathbf{1}'_{j-i}$ (see more about Student-t distribution in Arellano-Valle and Bolfarine [15]).

Notice that in spite of the advantages introduced by the PPM in the identification of multiple change points (the number of change points is not previously fixed), the exact calculation of the posterior distributions of ρ and B, as well as the posterior relevances $r^*_{[ij]}$, demands such a high computational effort that it is unlikely that the PPM would be of practical interest in the analysis of large data sets. In Section 3, Loschi et al.'s computational approach to find the posterior distributions of ρ and B is shown and adapted to the beta prior situation and a new Gibbs sampling scheme to overcome the difficulties of computing the posterior relevances is proposed.

3 Gibbs Sampling Scheme Applied to the PPM

Gibbs Sampling is a Monte Carlo Markov Chain scheme proposed by Geman and Geman [16] and adapted to Bayesian statistics by Gelfand and Smith [17]. In particular, Gibbs sampling provides a posterior distribution generation scheme.

In order to estimate the posterior distributions of ρ and B and also the posterior relevances $r^*_{[ij]}$, the transformation suggested by Barry and Hartigan [8] is used which assumes the auxiliary random quantity U_i that reflects whether or not a change point occurred at the time i, that is:

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

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

Notice that the random quantity ρ is immediately identified once the vector $\mathbf{U} = (U_1, U_2, \ldots, U_{n-1})$ is known. Consequently, the posterior probability of each particular partition $\rho = \{i_0, i_1, \ldots, i_b\}$, into b contiguous blocks, can be estimated from the number of $\mathbf{U}$'s for which this particular value of ρ is found. It is also possible to use the $\mathbf{U}$'s to estimate the posterior distribution of B (or the posterior distribution of the number of change points $B-1$) simply by noticing that (see details in Loschi et al. [11]):

$$B = 1 + \sum_{i=1}^{n-1} (1 - U_i). \tag{17}$$

The posterior relevances can be estimated by using the following procedure. Generate a sample of $\mathbf{U}$'s of size T. The estimate of the posterior relevance $r^*_{[ij]}$, for $i, j = 1, \ldots, n$, such that $i < j$, can be computed as follows:

$$\hat{r}^*_{[ij]} = \frac{M_{[ij]}}{T}, \tag{18}$$

where $M_{[ij]}$ is the number of $\mathbf{U}$'s for which the pattern $U_i = 0$, $U_{i+1} = \cdots = U_{j-1} = 1$ and $U_j = 0$ is observed.

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

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

for $r = 1, \ldots, n-1$. To generate the $\mathbf{U}^k$'s, it is sufficient to consider the ratios given by the following expressions Ross [18]:

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

for $r = 1, \ldots, n-1$, where $A_r^k = \{U_1^k = u_1, \ldots, U_{r-1}^k = u_{r-1}, U_{r+1}^{k-1} = u_{r+1}, \ldots, U_{n-1}^{k-1} = u_{n-1}\}$. Hence, considering a beta prior distribution for p, it results that:

$$R_r = \frac{f_{[xy]}(X_{[xy]})\Gamma(n+\beta-b+1)\Gamma(b+\alpha-2)}{f_{[xr]}(X_{[xr]})f_{[ry]}(X_{[ry]})\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}, \tag{19}$$

where:

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

and

$$x = \begin{cases} \min\ i \\ \textbf{s.t.:}\ r < i < n, \\ \qquad U_i^{k-1} = 0, \quad \text{if there is an } U_i^{k-1} = 0, \\ \qquad\qquad\qquad \text{for some } i \in \{r+1, \ldots, n-1\}, \\ n, \qquad\qquad\quad \text{otherwise.} \end{cases}$$

Notice that, in the normal case, $f_{[ij]}(X_{[ij]})$ is the Student-t distribution given in Eq. (16). Consequently, the criterion for choosing the values $(U_1^k, \ldots, U_{n-1}^k)$ becomes:

$$U_r^k = \begin{cases} 1, \text{ if } R_r \geq \frac{1-u}{u} \\ 0, \text{ otherwise,} \end{cases} \tag{20}$$

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

This completes the procedure proposed. The algorithm in pseudo-code is presented in Figure 4.

algorithm
  read $X_1, \ldots, X_n$
  for $k = 1$ **to** SAMPLES **do**
    generate 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
      $U_i^k = 0, U_{i+1}^k = \cdots = U_{j-1}^k = 1, U_j^k = 0$
  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**
    $E(\mu_k | X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} m^*_{[ij]}$
    $E(\sigma_k^2 | X_1, \ldots, X_n) \leftarrow \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \frac{a^*_{[ij]}}{d^*_{[ij]} - 2}$
  end for
  write $E(\mu_1), E(\sigma_1^2), \ldots, E(\mu_n), E(\sigma_n^2)$
end algorithm

Fig. 4. The Proposed Normal PPM for μ and σ^2

4 Application to Two Important Brazilian Indexes

In this section, the focus is on the identification of multiple change points in the means (expected or mean returns) and variances (volatilities) of the IBOVESPA and IBOVMESB series. Both time series, available from the authors, are expressed in terms of the returns calculated on closing prices, recorded monthly. 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 price in the month t. The IBOVESPA and IBOVMESB return series are plotted in Figure 1, from which it is noticeable that they present a similar behavior, suggesting the existence of some changes in the means and variances of the returns in both of them. The purpose is to verify whether or not, within the period considered, January, 1991 to August 1999, the two return series present change points in the expected returns and in the volatilities.

4.1 The Data Analysis

The same prior cohesions and distributions are considered to describe the initial uncertain for both series, although the IBOVMESB series seem to present lower variances, as one could see from Figure 1. These choice were done as reported by Loschi et al. [11], for the Chilean market. These specifications can be supported by the fact that the Brazilian market is also an emerging market and, like the Chilean market, more susceptible to the political scenario than developed markets (see Mendes [19]). As for the Chilean market, we also assume that changes in the behavior of the Brazilian stock return series are a consequence of the receipt of not previously anticipated information (more about unpredictability can be found in Loschi [20]), so that past change points are non-informative about future change points (see Mandelbrot [21]). Hence, the prior cohesions presented in Eq. (13), which imply that the sequence of change points establishes a discrete renewal process, with occurrence times geometric and identically distributed, are also an adequate choice for the Brazilian stock market.

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

In accordance to Loschi et al.'s specifications for the Chilean stock market [11], the following normal-inverted-gamma prior distribution is adopted to describe the uncertainty on the parameter $(\mu_{[ij]}, \sigma^2_{[ij]})$ for both indexes:

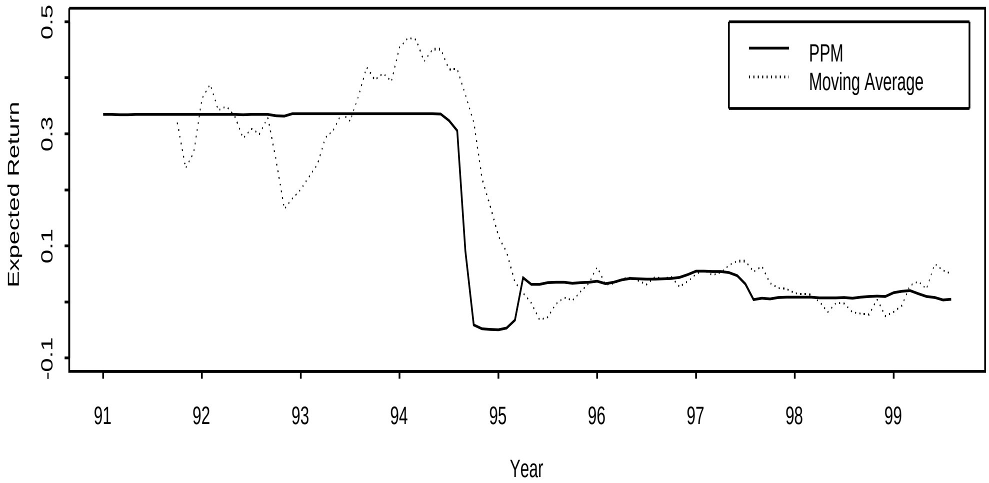
$$\left.\begin{array}{l} \mu_{[ij]}|\sigma^2_{[ij]} \sim N(0, \sigma^2_{[ij]}), \text{ and} \\ \sigma^2_{[ij]} \sim IG\left(\frac{0.01}{2}, \frac{4}{2}\right). \end{array}\right\}$$

Since a small number of changes is expected in both series, a beta distribution which concentrates most of its mass on small values needs to be considered as prior distribution of p. The following distribution is considered:

$$p \sim \mathcal{B}\left(\frac{3}{2}, \frac{57}{2}\right).$$

To estimate the posterior relevances $r^*_{[ij]}$ and the posterior distribution of B

(or the number of change points $B - 1$), 50,000 samples of 0-1 values with dimension 103, starting from a sequence of zeros were generated. The initial 5,000 iterations were discarded and a lag of ten was selected to avoid correlation. That means that a net sample size of 4,500 was used. Discussion on the number of iterations to be discarded, as well as the lag to be taken, can be easily found in the literature by the interested reader (see Gamerman [22], for example).



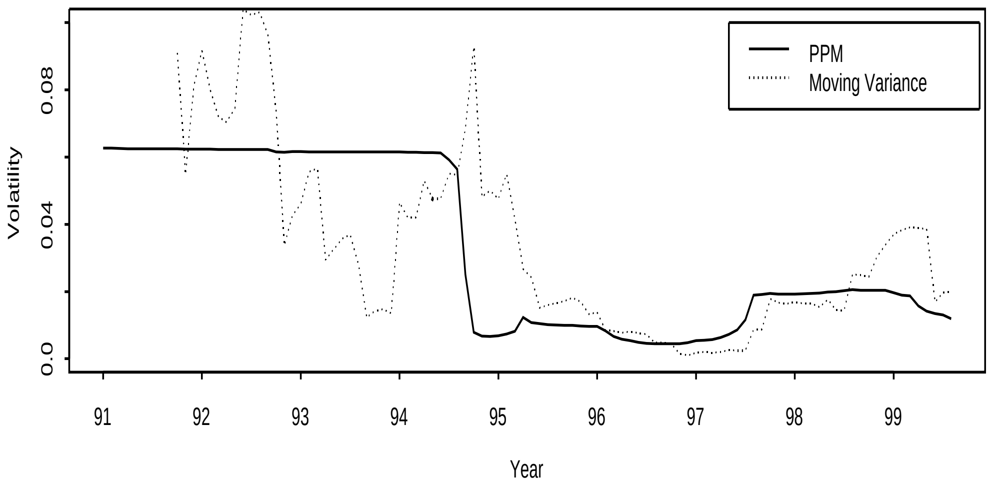


Fig. 5. IBOVESPA's Posterior Estimates

The algorithm in Figure 4 was coded in C^{++}, with the settings mentioned above, and it is available upon request. All tests were performed in PC-like

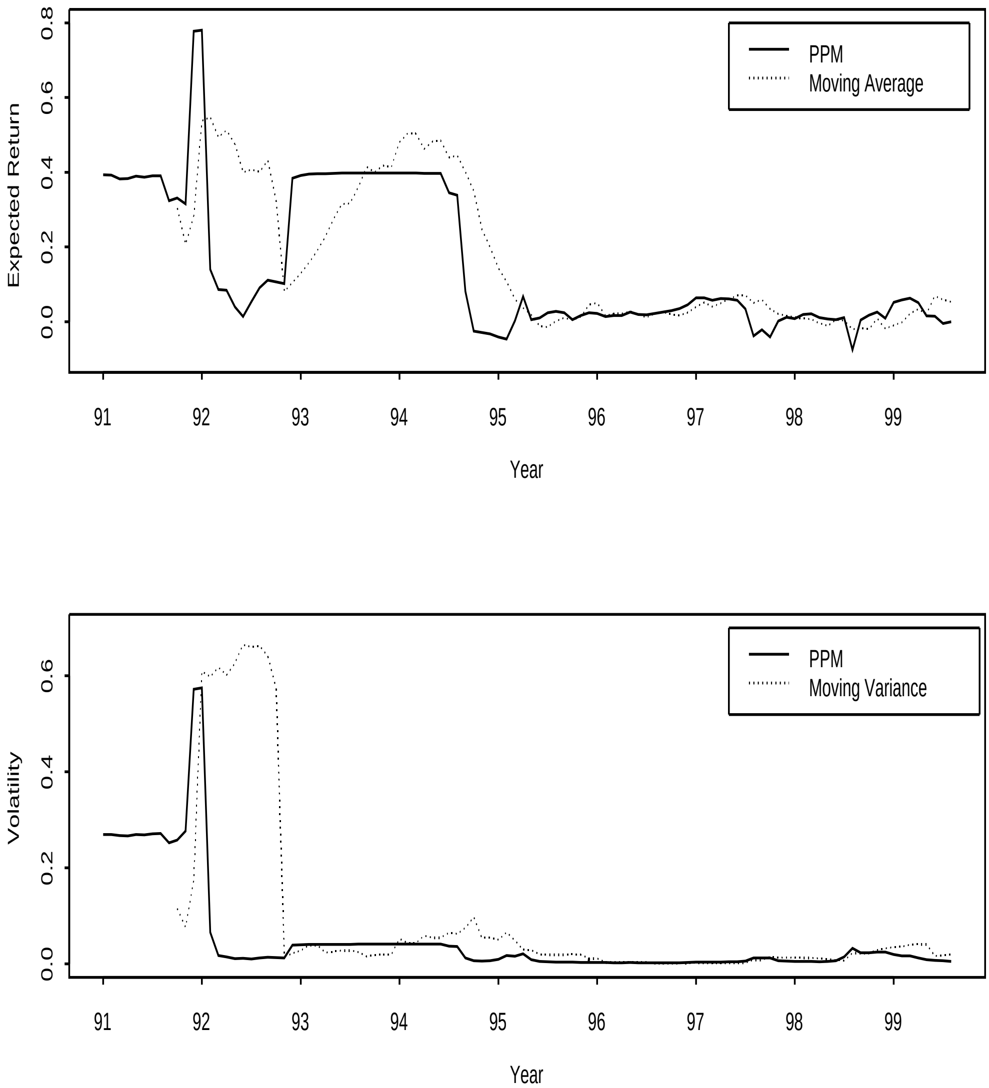


Fig. 6. IBOVMESB's Posterior Estimates

machine, 166 MHz, 32 MB RAM, and using the free C^{++} compiler DJGPP (url: http://www.delorie.com/djgpp). All tests took less than 10 minutes of CPU time.

Figures 5 and 6 present the posterior estimates (solid lines) of the monthly mean returns and volatilities for the IBOVESPA and IBOVMESB series, respectively. These estimates are contrasted with the order 10 arithmetic moving averages (dotted lines) for means and variances. It is noticed that the estimates obtained using the PPM are similar to the respective naïve estimates.

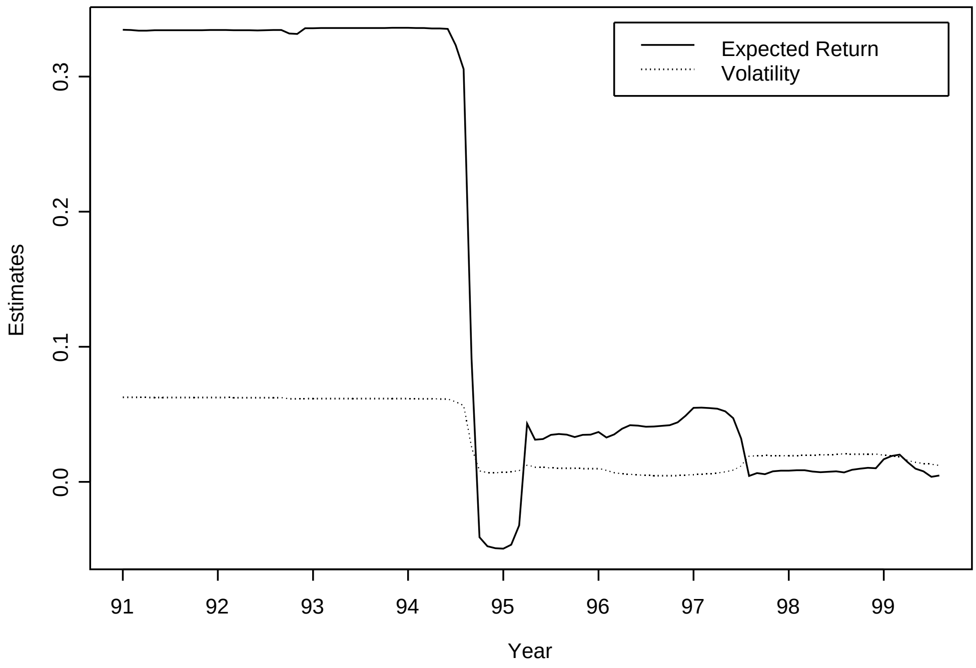


Fig. 7. IBOVESPA's Joint Behavior of Expected Returns and Volatilities

Figure 7 presents the posterior estimates of the IBOVESPA's expected returns (solid line) and volatilities (dotted line). A similar comparison is presented in Figure 8 for the IBOVMESB. Additionally, Figures 7 and 8 show that more changes occurred in the expected returns than in the volatilities in both series and that typically changes in the volatilities are followed by changes in the expected returns for both indexes, which can also be seen in the dispersion diagrams in Figures 9 and 10.

Figure 11 and 12 show respectively the expected return posterior estimates and the volatility posterior estimates for both series. It is noticed that typically, change points observed in the IBOVESPA and IBOVMESB series occur at the same time and that the changes are in the same direction. However, some differences in the behavior of these series are observed. The two changes observed in IBOVMESB series, in August, 1991 and in October, 1991, do not occur in IBOVESPA series. These change points could be related to the USIMINAS privatization, a important steel company located in Minas Gerais state. 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. Against the initial expectations, these important political events do not seem to produce changes in the behavior of IBOVESPA series.

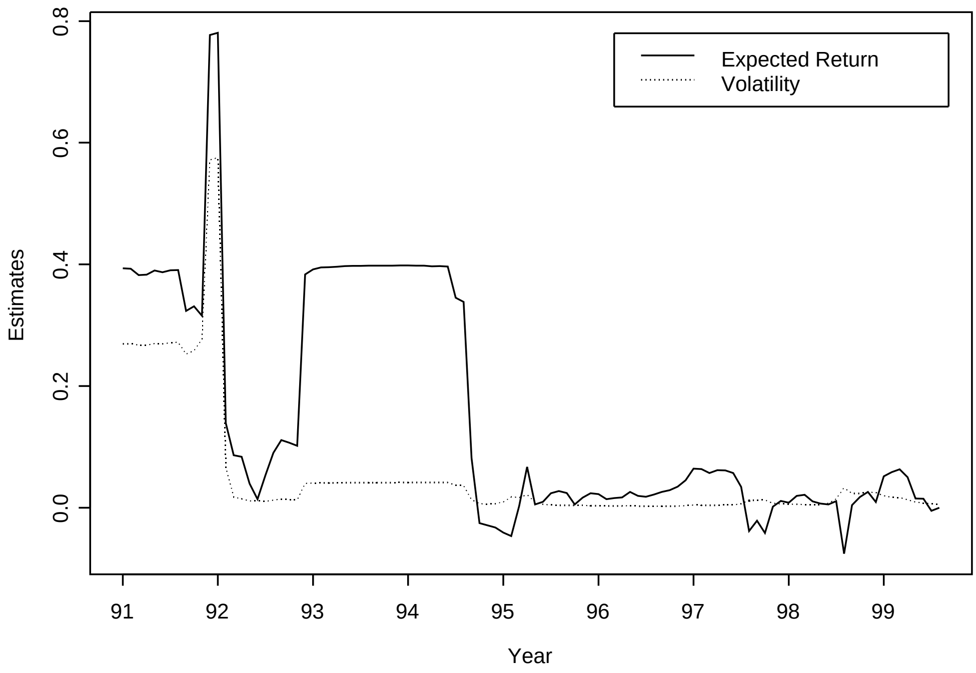


Fig. 8. IBOVMESB's Joint Behavior of Expected Returns and Volatilities

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.

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

Figure 13 shows the posterior distribution of the number of change points that occurs in each index. We notice that the posterior distributions of the number of change points for both indexes concentrate most of their mass on small values as expected. However, the posterior distribution of the number of change points for IBOVESPA series are more concentrated and typically concentrate their mass on smaller values than the IBOVMESB series, which

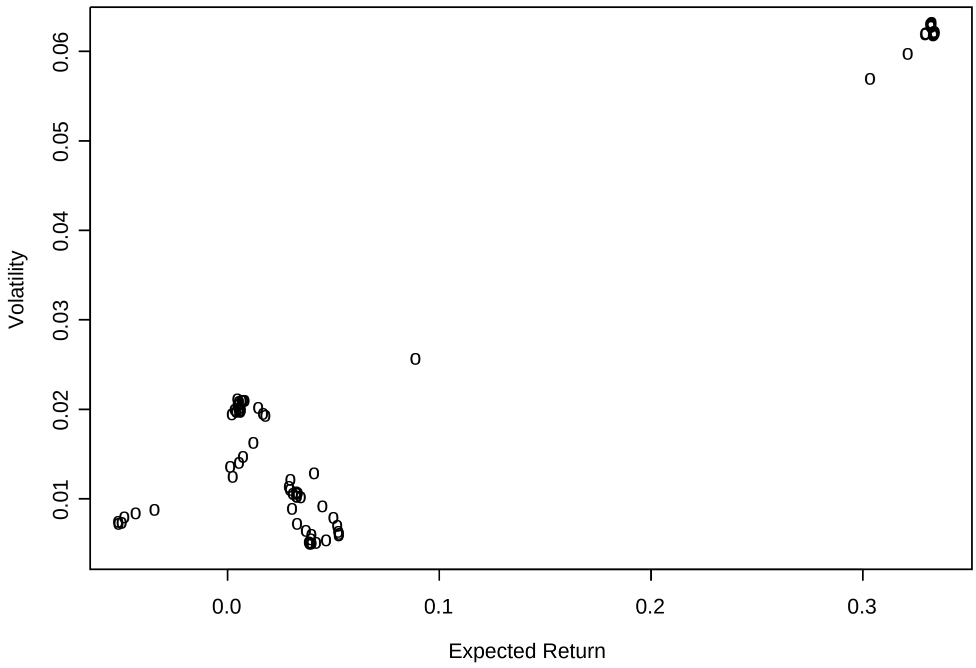


Fig. 9. IBOVESPA's Expected Returns × Volatilities Dispersion Diagram

means that the former series comes from a more stable market.

5 Final Remarks and Future Directions

The classical Product Partition Model (PPM) was described and its importance for change-point identification problems in time series analysis was stressed. The PPM was tailored to the analysis of multiple change points in the means and variances of normal data, assuming a prior specifications for these parameters and for the parameter p that is the probability of having a change in a period of time. A new scheme based on Gibbs sampling was proposed to implement the PPM that avoided its inherent computational hardness. The algorithm was coded in C^{++} and it was made available upon request. Two important Brazilian indexes were analyzed and the method seemed to explain satisfactorily their behavior, if a change-point analysis is required.

It was concluded that the IBOVESPA and IBOVMESB series have a very similar behavior and could probably suffer the influences of the same non-local events. It was noticed that both indexes have presented clusters in the expected returns and volatilities, as well as a small number of change points. These same conclusions were also driven for the Chilean stock market by Loschi et al. [11],

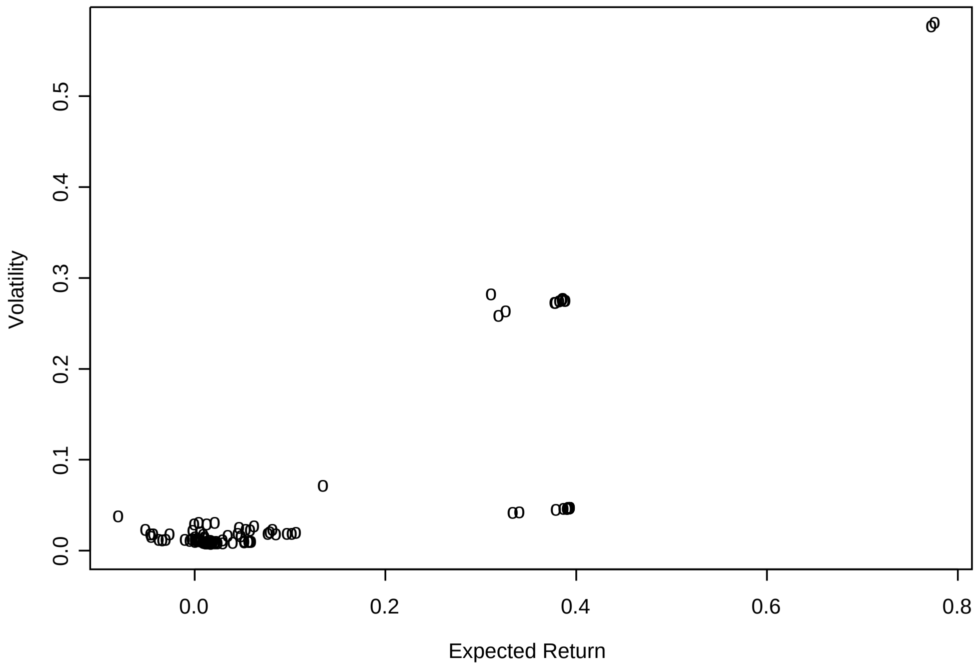


Fig. 10. IBOVMESB's Expected Returns × Volatilities Dispersion Diagram

disclosing the similarities that exist in the behavior of Brazilian and Chilean markets. São Paulo and Minas Gerais states are two of the most important economies in Brazil, thus having a high political influence. Hence, as Minas Gerais is the strongest economy involved in the IBOVMESB, the similarities observed in the behavior of IBOVESPA and IBOVMESB are justified.

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

Acknowledgments

The authors hereby acknowledge the Brazilian agencies CNPq (grants 301809/96-8 and 201046/94-6), FAPEMIG (grants CEX 855/98 and CEX 795/00), and PRQp-UFMG (grants 4081-UFMG/RTR/FUNDO/PRPq/99 and 3893-UFMG/RTR/FUNDO/PRPq/98), as well as the Chilean agencies FONDECYT (grants 8000004, 1971128, and 1990431) and Fundación Andes (grant C 13680-4), for the support to their research.

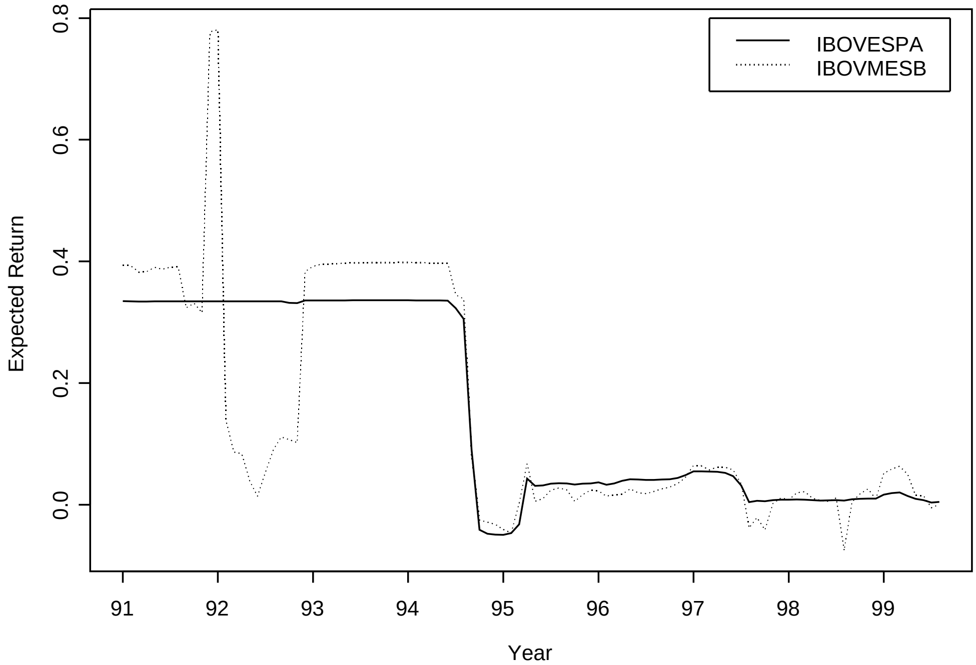


Fig. 11. Expected Return Posterior Estimates

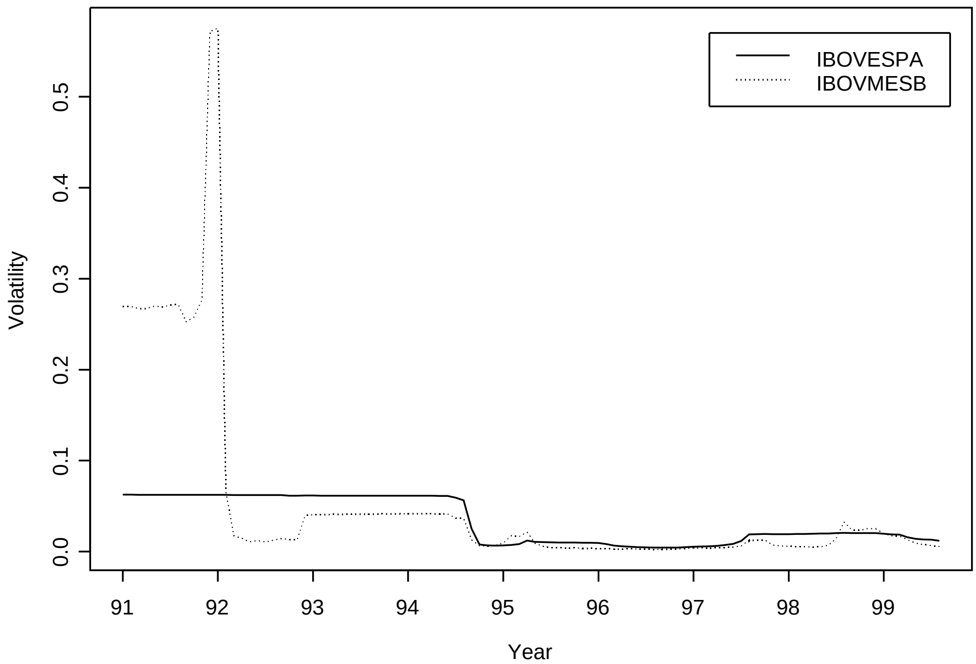


Fig. 12. Volatility Posterior Estimates

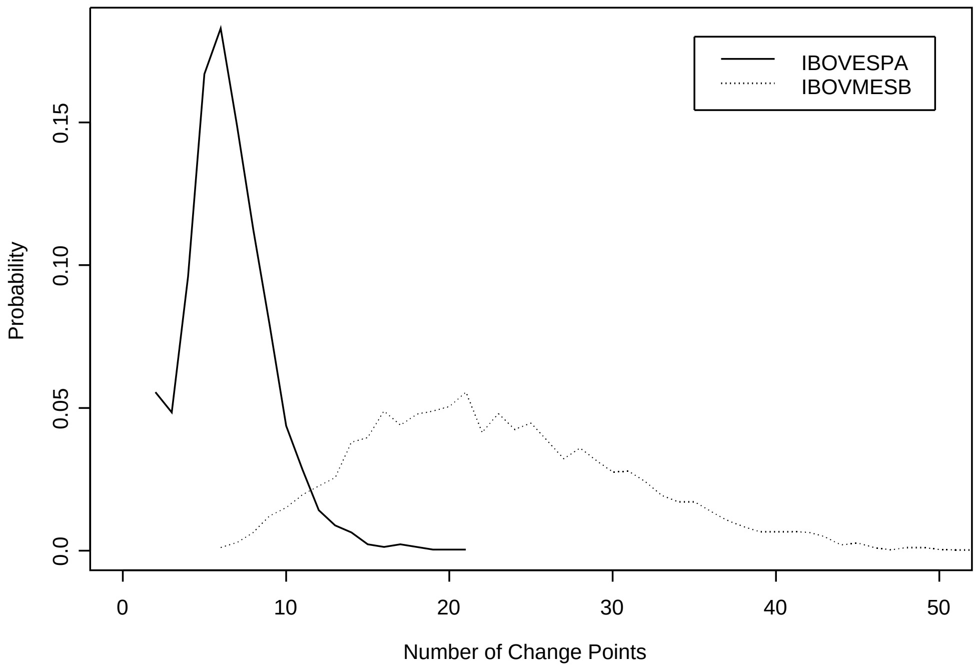


Fig. 13. Posterior Distribution of the Number of Change Points

References

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

[2] 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.

[3] 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.

[4] G. E. P Box and G.C. Tiao. *Bayesian Inference in Statistical Analysis.* Addison-Wesley, New York, USA, 1973.

[5] 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.

[6] D. A. Stephens. Bayesian retrospective multiple-changepoint identification. *Applied Statistics*, 43(1):159–178, 1994.

[7] D. M. Hawkins. Fitting multiple change-point models to data. *Computational Statistics & Data Analysis*, 1:323–341, 2001.

[8] D. Barry and J. A. Hartigan. A Bayesian analysis for change point problem. *Journal of the American Statistical Association*, 88(421):309–319, 1993.

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

[10] F. A. Quintana and P. L. Iglesias. Non-parametric Bayesian clustering and product partition models. Technical Report PUC/FM-06/2000, Departamento de Estadistica, Pontificia Universidad Catolica de Chile, 2000.

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

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

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

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

[15] R. B. Arellano-Valle and H. Bolfarine. On some characterizations of the t-distribution. *Statistics & Probability Letters*, 25:79–85, 1994.

[16] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[17] 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.

[18] S. Ross. *A First Course in Probability*. Prentice-Hall, New Jersey, USA, 5th edition, 1997.

[19] B. V. M. Mendes. Computing robust risk measures in emerging equity markets using extreme value theory. *Emerging Markets Quarterly*, pages 24–41, 2000.

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

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

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

Biographical Sketches

Rosangela Helena Loschi received her Doctor degree in Statistics from *Universidade de São Paulo*, São Paulo, Brazil. Her research areas include Bayesian statistics and she has published in *Computational Statistics & Data Analysis*, *Test* and *Journal of the Chilean Statistical Society*. Currently, she is an associate professor in the *Departamento de Estatítica* at *Universidade Federal de Minas Gerais*, Belo Horizonte, Brazil. E-mail: loschi@est.ufmg.br.

Frederico Rodrigues Borges da Cruz received his Doctor degree in Computer Science from *Universidade Federal de Minas Gerais*, Belo Horizonte, Brazil. His research areas include computational statistics and operations research. His papers have appeared in *Computer & Operations Research* and *European Journal of Operations Research*. Currently, he is an associate professor in the *Departamento de Estatítica* at *Universidade Federal de Minas Gerais*, Belo Horizonte, Brazil. E-mail: fcruz@est.ufmg.br.

Pilar Iglesias received her Doctor degree in Statistics from *Universidade de São Paulo*, São Paulo, Brazil. Her research areas include Bayesian statistics and decision theory. Her papers have appeared in *Bayesian Statistics 6*, *Test*, *Journal of Statistical Planning and Inference* among others. Currently, she is an associate professor in the *Facultad de Matematicas* at *Pontificia Universidad Catolica de Chile*, Santiago, Chile. E-mail: pliz@mat.puc.cl.

Reinaldo Arellano Valle received his Doctor degree in Statistics from *Universidade de São Paulo*, São Paulo, Brazil. His research areas include Bayesian statistics and asymptotic theory. His papers have appeared in *Statistics & Probability Letters*, *Test*, *Communication in Statistics* among others. Currently, he is an associate professor in the *Faculdad de Matematicas* at *Pontificia Universidad Catolica de Chile*, Santiago, Chile. E-mail: reivalle@mat.puc.cl.