Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

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

R. H. Loschi, F. R. B. Cruz, P. Iglesisas and R. B. Arellano-Valle

Relatório Técnico
RTP-05/2001

Relatório Técnico
Série Pesquisa

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

R. H. Loschi[a,*], F. R. B. Cruz[a], P. Iglesisas[b], R. B. Arellano-Valle[b]

[a]*Departamento de Estatística,*
Universidade Federal de Minas Gerais,
31270-901 - Belo Horizonte - MG, Brazil
E-mail: {loschi,fcruz}@est.ufmg.br

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

January, 2001

*To whom all correspondences should be addressed at: Caixa Postal 702, Departamento de Estatística - UFMG, 30123-970 - 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. One appealing way to tackle the problem is by means of the product partition model. The product partition model applied to the change point identification has attracted researchers' attention, particularly nowadays because the spread use of the powerful personal computers that make it possible to deal with its inherent computational complexity. To attack this difficult problem, a Gibbs sampling scheme is develop in this paper which is applied to the analysis of two very important stock market data in Brazil. The computational results show that method is effective and efficient, making possible useful inferences. In addition, the method is simple and easy to implement.

Abstract — This paper extends some previous results concerning the product partition model (PPM) by proposing a new scheme to estimate the posterior relevances of the model and by considering a prior distribution of the probability that a change takes place in any time. The algorithm was coded and applied to the identification of multiple change points in Brazilian stock market data. The algorithm performed well and proved to be a useful tool for analyzing change point problems.

Keywords: change points, product partition model, relevance, Student-*t* distribution, Yao's cohesions.

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. Inferences on the instants when changes in the volatility occurred, for example, allows the identification of events that could produce the changes, helping decision makers in the future, under similar situations.

This paper considers the Bayesian analysis of the multiple change point problem considering the product partition model (PPM) introduced by Hartigan [1]. Bayesian approaches for the change point problems has 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 [4] for treating the problem. Hsu [3] and Menzefricke [2] applied their methodologies to stock market prices (see also Smith [5]). Stephens [6] discusses the discrete multiple change point problem and the continuous single change point problems, which is illustrated considering some kidney transplant data. Stephens 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. The identification of multiple change points in the mean of normal random variables with common variance is considered in detail by Barry and Hartigan [7]. Recently, Crowley [8] provides a new implementation of the Gibbs sampling scheme in order to solve the problem of estimating normal means by using the PPM. The PPM is also used by Loschi *et al.* [9] to identify multiple change points in the mean and variance of normal data. The results obtained by Loschi *et al.* extend some results presented by Crowley [8] and Barry and Hartigan [7] and are applied to the Chilean stock market data. In the PPM model presented by Loschi *et al.* [9], it is assumed that the cohesion is that defined by Yao [10], which depends on the probability of a change takes place at any time.

The aim of this paper is to apply the PPM to the identification of multiple change points in both the mean μ and also the variance σ^2 of normal data, assuming a prior distribution to the parameter involved in the Yao's cohesions. We also provide a different procedure to evaluate the posterior relevances necessary in some posterior estimates by using the PPM.

This paper is organized as follows. Section 2 briefly reviews the PPM introduced by Hartigan [1] and presents inferential solutions to identify change points for random variables which are normally distributed, given the means and variances, according to Loschi *et al.* [9]. Section 2 also presents the exact posterior relevances and posterior distributions of the random partition generated by the change points and of the number of change points in the partition. We consider the cohesions proposed by Yao [10] and a prior distribution for the parameter involved in these cohesions. In Section 3, we describe the Gibbs sampling scheme proposed to compute (i) the posterior relevances, (ii) the posterior distributions of the number of blocks in the partition generated by the change points and (iii) the posterior distribution of this random partition. In Section 4, we apply the results to identify change points in the mean returns as well as in the volatilities of two important Brazilian economic indexes. Section 5 closes the paper with final remarks and future topics for investigation.

2 Statistical Models

In this section, we give a brief of the product partition model (PPM), introduced by Hartigan [1], and its implementation in identifying multiple change points in the mean and variance of normal data [9]. The exact posterior relevances are presented. The exact posterior distributions of the random partition generated by the change points and of the number of change points in the partition are also presented. We consider Yao's cohesions [10] and a beta prior distribution for the parameter involved in these cohesions.

We shall now present the definition of PPM and some preliminary results concerning this model, as given by Barry and Hartigan [11, 7].

2.1 The PPM

Let $X_1, \ldots, X_n$ be a observed time series. Consider a random partition $\rho = \{i_0, i_1, \ldots, i_b\}$ of the set $I = \{1, \ldots, n\} \cup \{0\}$, $0 = i_0 < i_1 < \ldots < i_b = n$, 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 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 $\mathbf{X}_{[ij]}$ [1].

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 end points $i_1, \ldots, i_b$, 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:

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

where $\mathcal{C}$ is defined as in (1) and K' is an appropriate constant.

As shown in Barry and Hartigan [11], 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 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, we consider that two observations X_i and X_j, $i \neq j$, are in the same block, if we believe that 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:

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

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 [11] have shown that the posterior distributions of θ_k is given by:

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

and the posterior expectation of θ_k is given by:

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

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

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

2.2 The Normal PPM

To specify the PPM for the normal case, Loschi *et al.* [9] 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 $X_k | \mu_k, \sigma_k^2 \sim N(\mu_k, \sigma_k^2)$, $k = 1, \ldots, n$, and that 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,

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

where $IG(a, d)$ 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 $\theta_{[ij]} = (\mu_{[ij]}, \sigma^2_{[ij]})$, 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

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

with

$$\begin{array}{rcl} \bar{X}_{[ij]} & = & \dfrac{1}{j-i} \sum_{r=i+1}^{j} X_r, \\ q_{[ij]}(\mathbf{X}_{[ij]}) & = & \sum_{r=i+1}^{j} (X_r - \bar{X}_{[ij]})^2 + \dfrac{(j-i)(\bar{X}_{[ij]} - m_{[ij]})^2}{(j-i)v_{[ij]} + 1}. \end{array}$$

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

$$\mu_{[ij]}|\mathbf{X}_{[ij]} \sim t(m^*_{[ij]}, v_{[ij]}, a^*_{[ij]}, d^*_{[ij]}) \text{ and } \sigma^2_{[ij]}|\mathbf{X}_{[ij]} \sim IG(a^*_{[ij]}/2, d^*_{[ij]}/2), \tag{10}$$

for which it is observed that

$$E(\mu_{[ij]}|X_{[ij]}) = m^*_{[ij]} \ \ (\text{if } d^*_{[ij]} > 1) \tag{11}$$

and

$$E(\sigma^2_{[ij]}|X_{[ij]}) = \frac{a^*_{[ij]}}{d^*_{[ij]} - 2} \ \ (\text{if } d^*_{[ij]} > 2). \tag{12}$$

The interested reader may find details in O'Hagan [12].

From (6), (11) and (12), it follows that the posterior estimates for the parameters μ_k and σ^2_k are given by:

$$E(\mu_k|X_1, \ldots, X_n) = \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} m^*_{[ij]} \ \ (\text{if } d^*_{[ij]} > 1) \tag{13}$$

and

$$E(\sigma^2_k|X_1, \ldots, X_n) = \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} \frac{a^*_{[ij]}}{d^*_{[ij]} - 2} \ \ (\text{ if} d^*_{[ij]} > 2), \tag{14}$$

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

Now, we shall present the exact posterior distributions of the random partition ρ and the number of blocks B in the partition. This cohesions considered are those defined by Yao [10]. The exact calculation of the posterior relevances are also shown.

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

As we are focusing on the PPM assuming only the existence of contiguous blocks, the prior cohesions can be interpreted as the transition probabilities in the Markov chain defined by the endpoints of the blocks in the partition ρ. We assume the prior cohesions suggested by Yao [10] as presented below.

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 $[ij]$ corresponds to the probability that a new change takes place after $j - i$ instants, given that a change has taken place at instant i, that is:

$$c_{[ij]} = \begin{cases} p(1-p)^{j-i-1}, & \text{if } j < n, \\ (1-p)^{j-i-1}, & \text{if } j = n, \end{cases} \tag{15}$$

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

Consequently, from (1) we obtain 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 (2), it follows that the prior distribution of the random variable B is given by:

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

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 and denoted by $p \sim \mathcal{B}(\alpha, \beta)$. 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 $0 = i_0, \ldots, i_b = n$, $b \in I$ 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:

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

$$\frac{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}. \tag{16}$$

The posterior probability of the event $B = b$, $b \in I$, is given by multiplying the posterior probability in (16) by C_{b-1}^{n-1}. Notice that the posterior distributions of ρ and B do not have a product distribution as presented in Section 2.1 and obtained by Loschi *et al.* [9], concerning the normal case.

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

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

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

$$\frac{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}{\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}. \tag{17}$$

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 we consider the PPM presented in Section 2.2 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:

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

$$\{a_{[ij]} + (\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})' \mathbf{V}^{-1}_{[ij]} (\mathbf{X}_{[ij]} - \mathbf{m}_{[ij]})\}^{-(d_{[ij]}+j-i)/2}, \tag{18}$$

where $\mathbf{m}_{[ij]} = m_{[ij]} \mathbf{1}_{j-i}$ and $\mathbf{V}_{[ij]} = \mathbf{I}_{j-i} + v_{[ij]} \mathbf{1}_{j-i} \mathbf{1}'_{j-i}$.

Notice that the exact calculation of the posterior distributions of ρ and B as well as the posterior relevances demands great computational efforts. In Section 3, we propose a computational approach to find these posterior distributions, which is based on the sample generated by using the Gibbs sampling approach [13].

3 A Gibbs Sampling Scheme Applied to PPM

Gibbs Sampling is a Monte Carlo Markov Chain (MCMC) scheme proposed by Geman and Geman [14] and adapted to Bayesian statistics by Gelfand and Smith [13]. In particular, Gibbs sampling provides a posterior distribution generation scheme.

To estimate the posterior distributions of ρ and B and also the posterior relevances of each block $[ij]$, we consider the transformation suggested by Barry and Hartigan [7]. Let assume the auxiliary random quantity U_i which reflects whether or not a change point occurred at the time i, that is:

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

$i = 1, \ldots, n-1$. Thus, 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, $\rho = \{i_0, i_1, \ldots, i_b\}$. It is also possible to use the above procedure to estimate the posterior distribution of B (or the posterior distribution of the number of change points $B-1$) noticing that:

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

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

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

$r = 1, \ldots, n-1$. To generate the vectors above, it is sufficient to consider the ratios given by the following expressions [15]:

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

$r = 1, \ldots, n-1$, 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, we have 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:

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

and

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

Notice that, in the normal case, $f_{[ij]}(X_{[ij]})$ is the Student-t distribution given in (18). Consequently, the criterion for choosing the values $(U_1^k, \ldots, U_{n-1}^k)$ becomes:

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

$r = 1, \ldots, n-1$, where $u \sim \mathcal{U}(0,1)$. This completes the procedure to estimate the posterior distributions of the random partition ρ and of the number of blocks B.

The posterior relevance of the block $[ij]$, $i < j$, used in (6) to estimate θ_k can be obtained by considering how many samples present $U_i = 0$, $U_{i+1} = \ldots, U_{j-1} = 1$ and $U_j = 0$.

In spite of using the transformation suggested by Barry and Hartigan [7], the procedure we propose here to estimate θ_k is entirely new and simpler, because it does not use the Gibbs sampling to generate samples from the θ_k distribution. Another general Markov sampling technique to obtain product estimates is also described by Crowley [8].

4 Application to the Two Most Important Brazilian Indexes

In this section, we focus on the identification of multiple change points in the mean (expected or mean return) and variance (volatility) of the two most important Brazilian indexes, *'Indice Geral da Bolsa de São Paulo* (IBOVESPA) and *'Indice da Bolsa de Valores de Minas Gerais, Espírito Santo e Brasília* (IBOVMESB). Both are expressed in terms of the returns calculated on closing prices, recorded monthly. We apply the methodology developed in the previous sections to analyze the behavior of these indexes from January, 1991 to August, 1999. These time series are available from the authors upon request or dire ctly through the anonymous ftp site at address **ftp://ftp.est.ufmg.br/pub/loschi/pub/gibbs**.

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. IBOVESPA and IBOVMESB return series are plotted together in Figure 1.

Figure 1 goes around here

From Figure 1, it is noticeable that 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. Our purpose is to show that within the period considered, the two return series present several change points in their volatility and expected return. That is, we show that IBOVESPA and IBOVMESB series possess volatility and expected return clusters. The posterior distributions of the number of change points which occur in IBOVESPA and IBOVMESB return series within this period are presented. A comparative analysis points out that the number of changes in both indexes has similar behavior and that changes in IBOVESPA and IBOVMESB series occur in general at the same time and in the same direction. Some political and economical events are identified and pointed out as possible causes of the change points we have found.

4.1 Data Analysis

The algorithm presented in the previous sections was coded in $C++$ and it is available from the authors upon request or directly through the anonymous ftp site at address **ftp://ftp.est.ufmg.br/pub/loschi/pub/gibbs**.

We consider the same prior cohesions and distributions to describe the initial uncertain for both IBOVESPA and IBOVMESB series, although the former seems to present lower variances, as one could see from Figure 1. These choice were done as reported by Loschi *et al.* [9], for the Chilean market. These specifications can be supported by the fact that the Brazilian market is also an emerging market and, like Chilean market, more susceptible to the political scenario than developed markets [16]. As for the Chilean market, we also assume that changes in the behavior of Brazilian stock return series are a consequence of the receipt of not previously anticipated information, so that past change points are non-informative concerning future change points (see Mandelbrot [17]). Hence, the prior cohesions presented in (15), 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.

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 Loschi *et al.*, [9], specifications for the Chilean stock market, the following normal-inverted-gamma prior distribution is adopted to describe the uncertainty on the parameter $(\mu_{[ij]}, \sigma^2_{[ij]})$ for both indexes:

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

Since a small number of changes is expected to both IBOVESPA and IBOVMESB series, a beta distribution which concentrates most of its mass in small values needs to be considered as prior distribution of p. We consider the following distribution:

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

To estimate the posterior relevances $r^*_{[ij]}$ and the posterior distribution of B (or the number of change points $B-1$), we generate 50,000 samples of 0-1 values with dimension 103, starting from a sequence of zeros. We discarded the initial 5,000 iterations and, to avoid correlation, we selected a lag of ten, that means that we worked with a net sample size of 4,500. Discussion about the number of iterations to be discharged, as well as the lag to be taken, can be easily found in the literature (*e.g.* Gamerman [18]). All tests were performed in PC, 166 MHz, 32 MB RAM, taking less than 10 minutes of CPU time.

Figures 2 and 3 present the posterior estimates (solid lines) of the monthly mean returns and volatilities for IBOVESPA and IBOVMESB series, respectively. These estimates are contrasted with the arithmetic moving averages (dotted lines) of order 10 for the means and variances. We can notice that the estimates obtained using the PPM are similar to the respective naïve estimates.

Figures 2 and 3 go around here

Figure 4 presents the posterior estimates of the expected returns (solid line) and volatilities (dotted line) of IBOVESPA series. A similar comparison is presented in Figure 5 for the estimates obtained for IBOVMESB series.

Figures 4 and 5 go around here

Figures 4 and 5 show that more changes occurred in the expected returns rather than in the volatility in both IBOVESPA series and IBOVMESB series. Typically, changes in the volatility are followed by changes in the expected return for both indexes. (See also the dispersion diagrams in Figures 6 and 7).

Figures 6 and 7 go around here

In Figure 8, the expected return posterior estimates for IBOVESPA series are contrasted with the correspondent estimates for IBOVMESB series. A similar analysis for the volatility posterior estimates is presented in Figure 9.

Figures 8 and 9 go around here

We notice that, typically, change points observed in 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 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. Unlikewise 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 policy adopted by Brazilian government during Asias'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.

Figure 10 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 IBOVMSB series, which means that the former series comes from a more stable market.

Figure 10 goes around here

5 Final Remarks and Future Directions

We have proposed a new scheme to implement the PPM that avoid its computational difficulties. We have described the PPM and stressed the importance of change point problems, particularly to analyze time series. The algorithm was coded and proved to be efficient and useful for the finance area. In the application considered, the results obtained seems to explain the behavior of BOVESPA and BOVMESB indexes satisfactorily, if a change point analysis is required.

We conclude that IBOVESPA and IBOVMESB series have a very similar behavior and could probably suffer the influences of the same non-local events. We notice that both indexes present expected return clusters and volatility clusters, and also a small number of change points. These same conclusions were also driven for the Chilean stock market [9], 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 great political influence. Hence, as Minas Gerais is the strongest economy involved in BOVMESB index, the similarities observed in the behavior of BOVESPA and BOVMESB indexes are justified.

Some open questions remains. Would it be possible to find even simpler implementations for the product partition model? How sensitive to the prior statement the results are? 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

The authors hereby acknowledge the Brazilian agencies CAPES, FAPEMIG, grant CEX-855/98, PRQp-UFMG, and CNPq, grant 301809/96-8, as well as the Chilean agency FONDECYT, grants 8000004, 1971128, 1990431, for a partial support to this project.

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 struture 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. Barry and J. A. Hartigan. Product partition models for change point problems. *Journal of the American Statistical Association*, 88(421):309–319, 1993.

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

[9] 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.

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

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

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

[13] 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.

[14] 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.

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

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

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

[18] 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 operations research. 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 *Location Science*, *Computer & Operations Research*, *European Journal of Operations Research*, and *Journal of Statistical Computation & Simulation.* 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. Currently, she is an associate professor in the *Faculdad 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. 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.

Figure Captions

Figure 1: IBOVESPA and IBOVMESB Return Series

Figure 2: Posterior Estimates to the Expected Returns and Volatilities - IBOVESPA

Figure 3: Posterior Estimates to the Expected Returns and Volatilities - IBOVMESB

Figure 4: Joint Behavior of Expected Returns and Volatilities - IBOVESPA

Figure 5: Joint Behavior of Expected Returns and Volatilities - IBOVMESB

Figure 6: Dispersion Diagrams - Expected Returns $\times$ Volatilities (IBOVESPA)

Figure 7: Dispersion Diagrams - Expected Returns $\times$ Volatilities (IBOVMESB)

Figure 8: Posterior estimates for the Expected Returns of IBOVESPA and IBOVMESB

Figure 9: Posterior estimates for the Volatilities of IBOVESPA and IBOVMESB

Figure 10: Posterior Distribution of the Number of Change Points in IBOVESPA and IBOVMESB Series

Figures

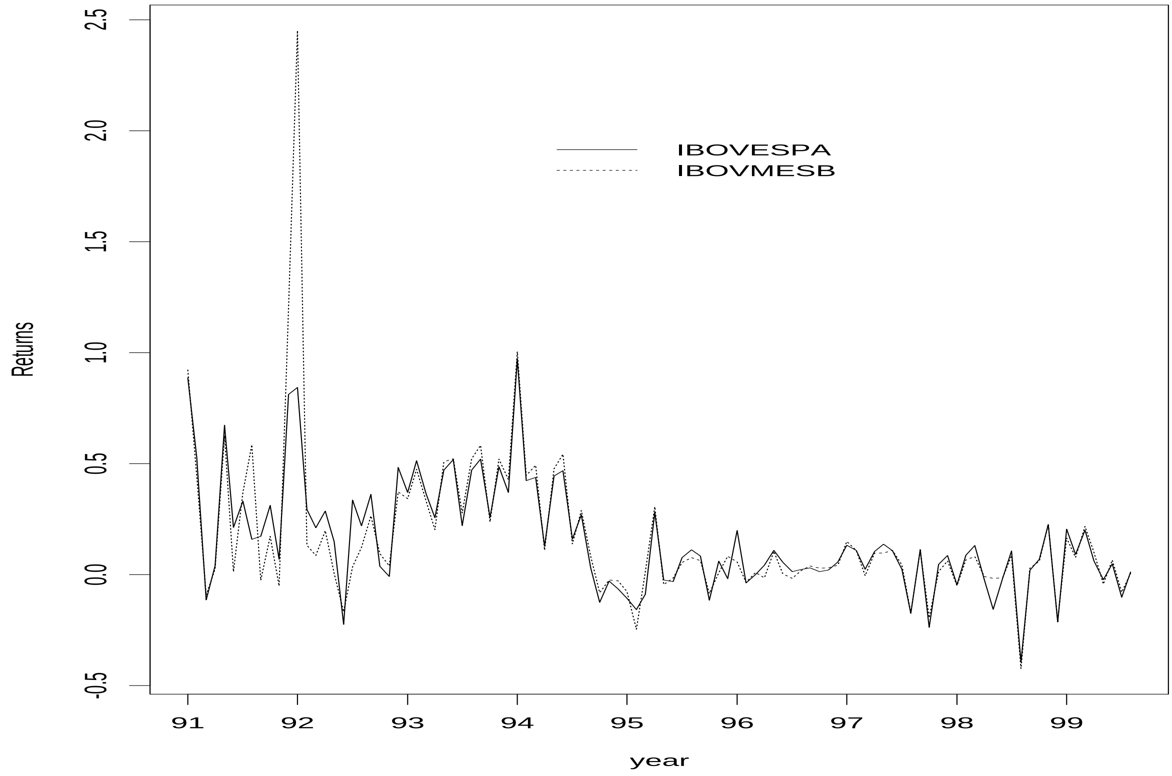

Figure 1:

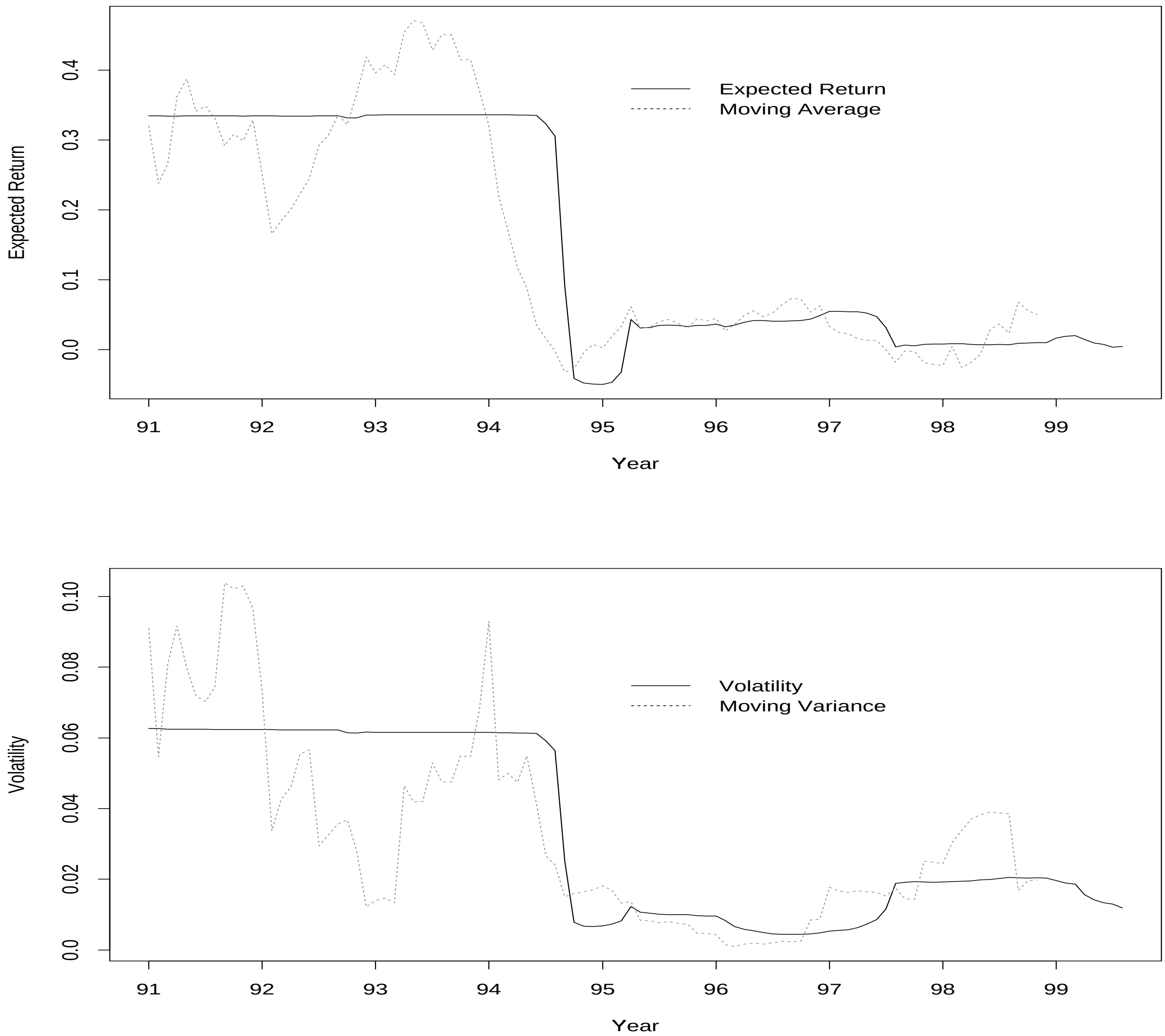
Expected Return
Moving Average
Expected Return
0.0
0.1
0.2
0.3
0.4
91
92
93
94
95
96
97
98
99
Year
Volatility
Moving Variance
Volatility
0.0
0.02
0.04
0.06
0.08
0.10
91
92
93
94
95
96
97
98
99
Year

Figure 2:

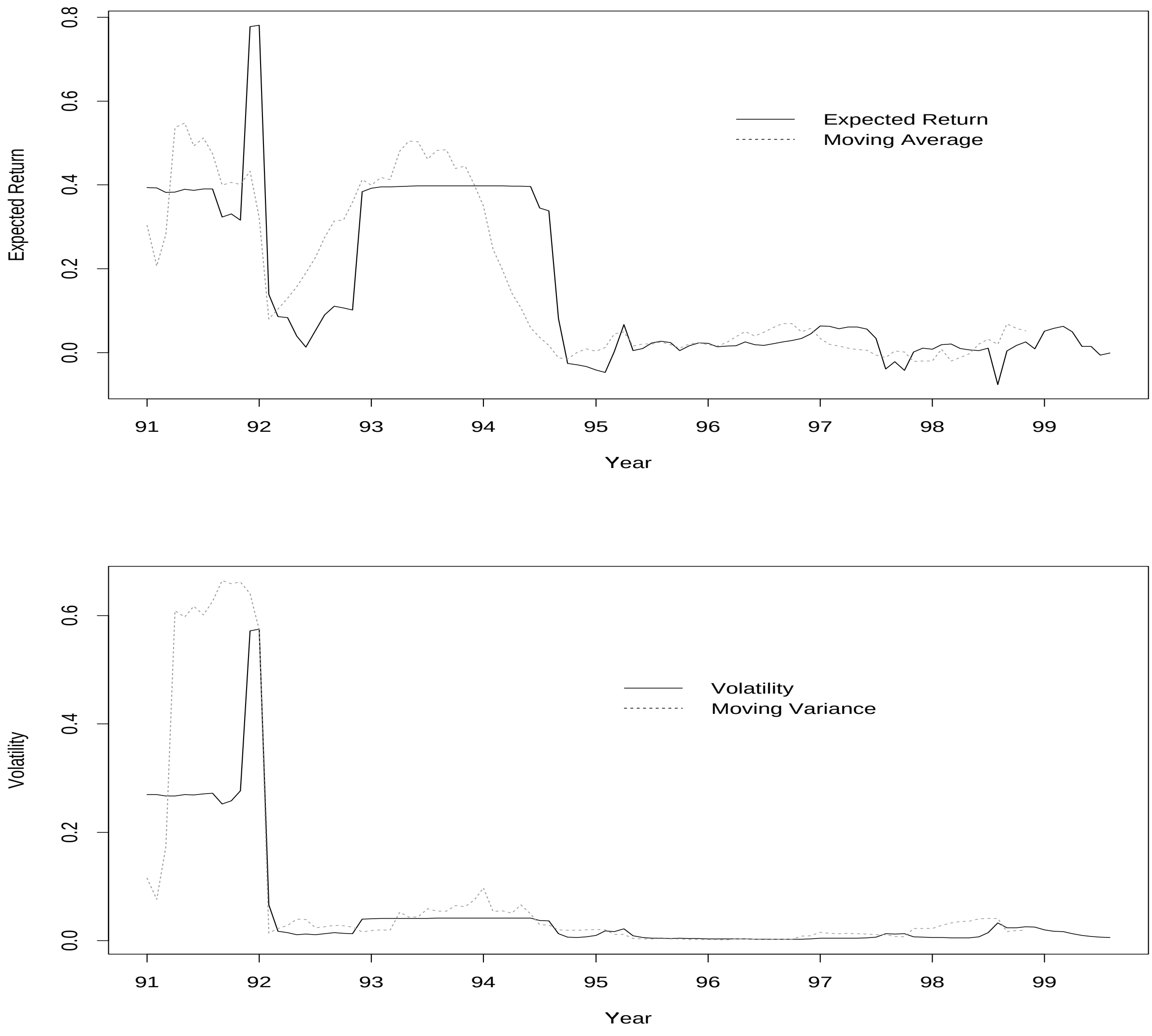
0.8
0.6
0.4
0.2
0.0
Expected Return
Expected Return
Moving Average
91
92
93
94
95
96
97
98
99
Year
0.6
0.4
0.2
0.0
Volatility
Volatility
Moving Variance
91
92
93
94
95
96
97
98
99
Year

Figure 3:

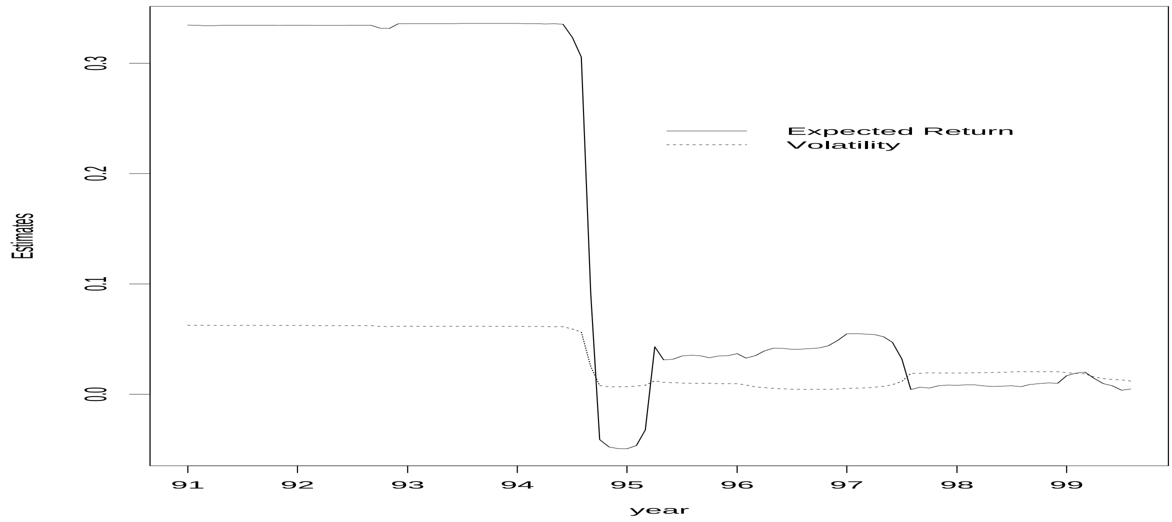

Figure 4:

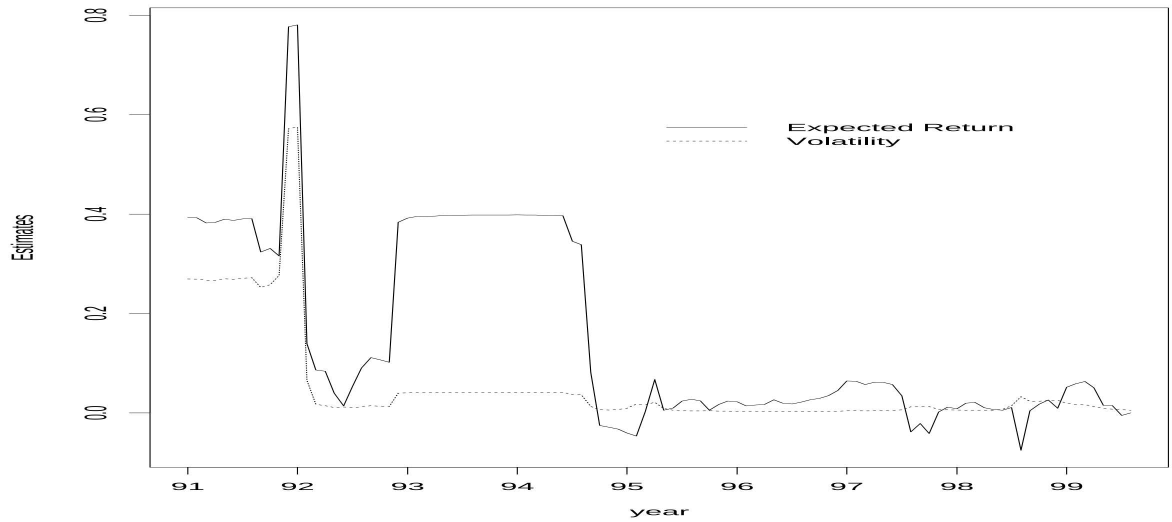

Figure 5:

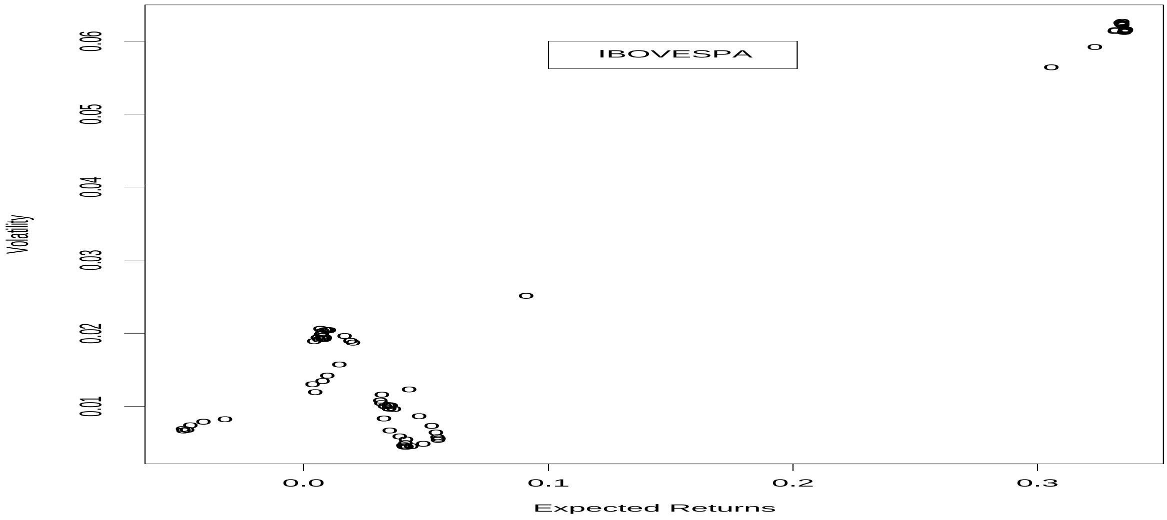

Figure 6:

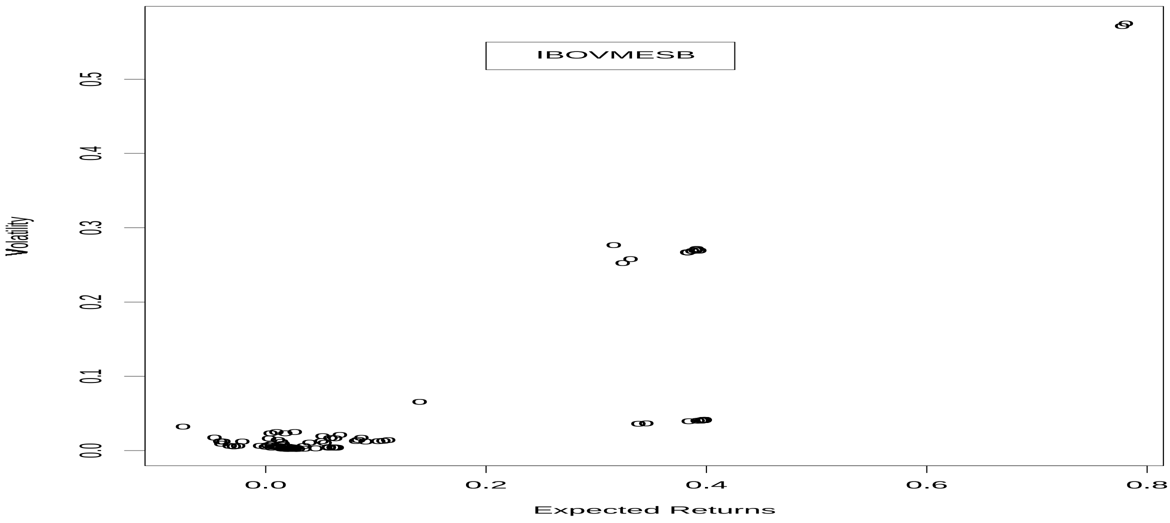

Figure 7:

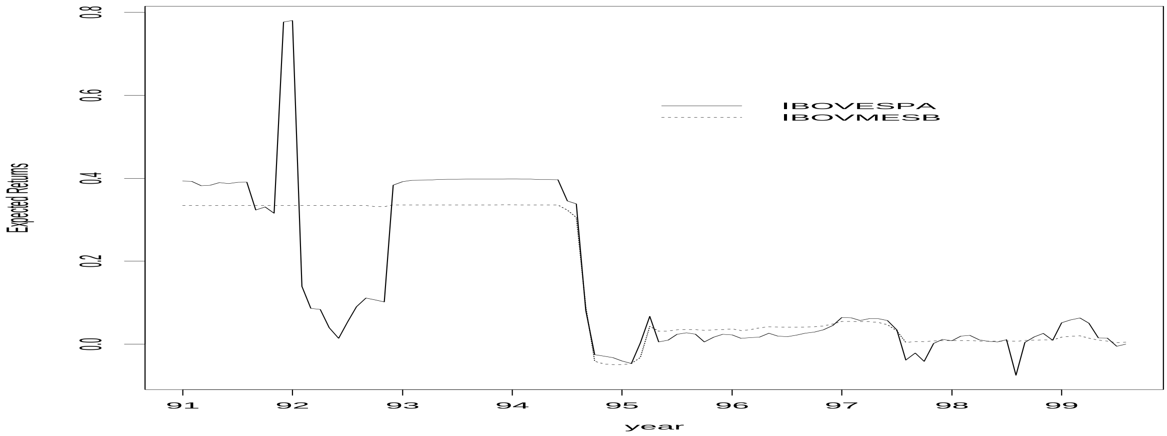

Figure 8:

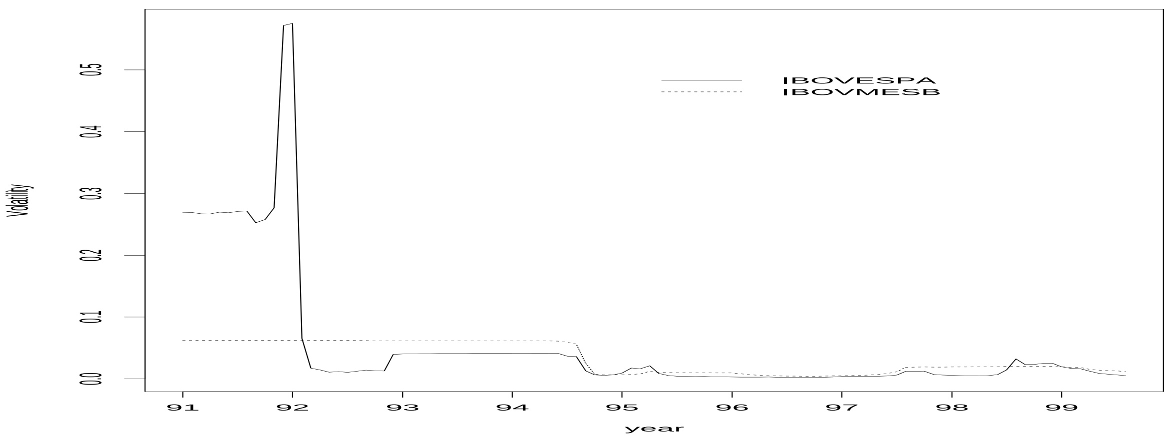

Figure 9:

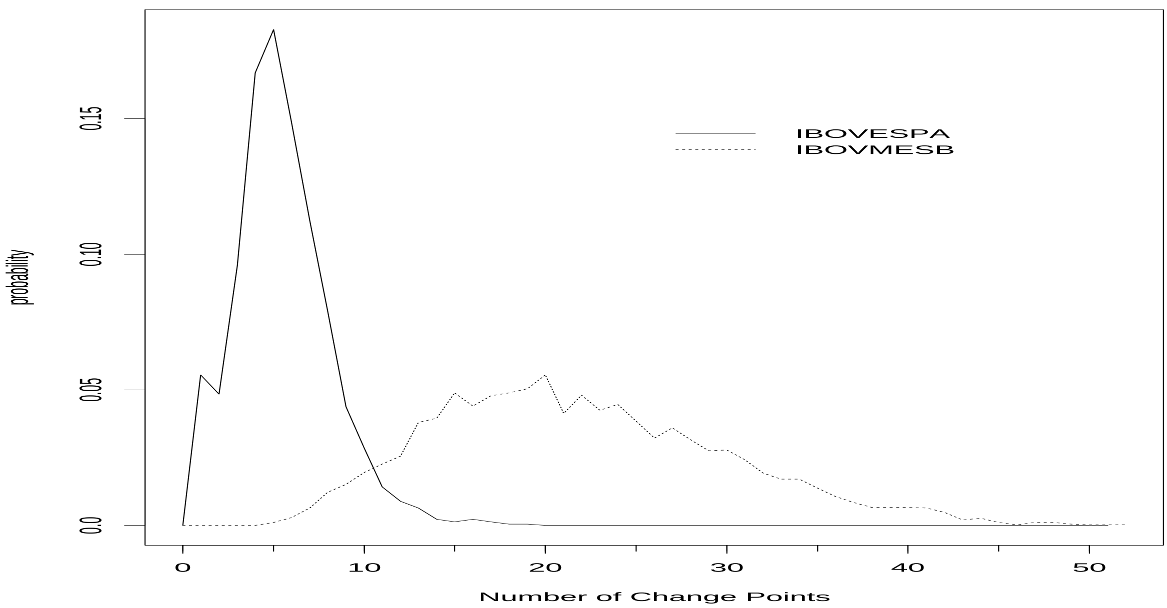
IBOVESPA
IBOVMESB
probability
0.0
0.05
0.10
0.15
0
10
20
30
40
50
Number of Change Points

Figure 10: