Extension to the Product Partition Model: Computing the Probability of a Change

R. H. Loschi and F. R. B. Cruz [1]

*Department of Statistics, Universidade Federal de Minas Gerais,
31270-901 - Belo Horizonte - MG, Brazil*

Abstract

The well-known product partition model (PPM) is considered for the identification of multiple change points in the means and variances of normal data sequences. In a natural fashion, the PPM may provide product estimates of these parameters at each instant of time, as well as the posterior distributions of the partitions and the number of change points. Prior distributions are assumed for the means, variances, and for the probability p that each individual time is a change point. The PPM is extended to generate the posterior distribution of p and the posterior probability that each instant of time is a change point. A Gibbs sampling scheme is used to compute all estimates of interest. The methodology is applied to an important time series from the Brazilian stock market. A sensitivity analysis is performed assuming different prior specifications of p.

Key words: Change point, Gibbs sampler, random partitions.

1 Introduction

Developed by Hartigan (1990), the product partition model (PPM) is a dynamic model useful to the analysis of change point problems. The PPM introduced more flexibility into the analysis of these problems since it considers the number of change points as a random variable, in opposition to several well-known methods to identify structural changes that assume the number of change points as known, e.g. threshold models and the method considered by Hawkins (2001). As shown by Barry and Hartigan (1992), by applying the

---

[1] Corresponding author. E-mail fcruz@est.ufmg.br. This research was developed while the author was a visiting professor at the Department of Mechanical & Industrial Engineering, University of Massachusetts, Amherst MA 01003, USA.

PPM one can easily obtain product estimates for the parameters of interest at each individual time, the posterior distribution of the random partition, and also the posterior distribution of the number of change points.

Barry and Hartigan (1993) and Crowley (1997) applied the PPM to the identification of multiple change points in normal means and used a Gibbs sampling approach to obtain the product estimates. However, it would be no trouble to extend the PPM to identify multiple changes in both means and variances of normal data and also to apply a Gibbs sampling scheme to compute the posterior distributions of the random partition and the posterior distributions of the number of change points. Details can be seen in the papers by Loschi and Cruz (2002) and by Loschi et al. (2003).

Concerning its timeliness, recently Quintana and Iglesias (2003) provided a theoretical decision approach to change point problems and linked the PPM to the Dirichlet process. In one of the most popular versions of the PPM, only contiguous blocks are allowed and in such cases the prior cohesions, usually Yao's (1984) cohesions, are a truncated geometric distribution with parameter p. Then, Loschi et al. (2003) were the first to assume a prior distribution for p, in a successful application of the PPM to change-point analysis. However, in spite of all flexibility that a prior distribution for p may provide, the product estimates may be considerably influenced by its specifications (see Loschi and Cruz, 2002).

In this paper, the aim is to extend the PPM as to obtain the posterior distribution of p and the posterior probabilities that each individual time is a change point. A Gibbs sampling scheme will be used to estimate the posterior distribution of p and to estimate the posterior probability that each instant is a change point. As it will be seen, these extensions not only enrich a data analysis provided by the PPM but also may be a useful tool to decision-makers since in general the posterior distribution of the random partition as originally defined by Barry and Hartigan (1992) put small mass in each partition. Finally, we will see a successful application of the PPM to the analysis of an important series of returns of the Brazilian stock market, the BOVESPA index (*Índice da Bolsa de Valores do Estado de São Paulo*). Also a sensitivity analysis will be shown considering different prior specifications of p.

The paper is organized as follows. In Section 2, the PPM and related results are presented. The PPM is applied to identify multiple change points in normal means and variances. A new procedure to obtain the posterior distribution of p and the posterior probability that each instant is a change point is proposed. In Section 3, we describe the computational method usually applied to obtain the product estimates of means and variances, as well as a Gibbs sampling scheme to compute all posterior distributions and probabilities aforementioned. In Section 4, the methodology is applied to the BOVESPA index

and a sensitivity analysis is provided. Finally, Section 5 closes the paper with concluding remarks.

2   Statistical Models

2.1   The Product Partition Model

Let $X_1, \ldots, X_n$ be a data sequence and consider the index set $I = \{1, \ldots, n\}$. Consider a random partition $\rho = \{i_0, i_1, \cdots, i_b\}$ of set I, $0 = i_0 < i_1 < \cdots < i_b = n$, and a random variable B which denotes the number of blocks in ρ. Consider that each partition divides the data sequence into b contiguous subsequences, which will be denoted here by $\mathbf{X}_{[i_{(r-1)}i_r]} = (X_{i_{(r-1)}+1}, \ldots, X_{i_r})^{\mathrm{T}}$, for $r = 1, \ldots, b$. Let $c_{[ij]}$ be the prior cohesion associated with block $[ij] = \{i + 1, \ldots, j\}$, for $i, j \in I \cup \{0\}$, and $j > i$, that represents the degree of similarity among the observations in $\mathbf{X}_{[ij]}$ and that can be interpreted as transition probabilities in the Markov chain defined by the change points.

In this paper Yao's (1984) cohesions are considered. Let p be the probability that a change occurs at any instant in the sequence. Therefore, the prior cohesion for block $[ij]$ is given by:

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

for all $i, j \in I$, $i < j$, which corresponds to the probability that a new change takes place after $j - i$ instants, given that a change has taken place at the instant i. These prior cohesions imply that the sequence of change points establishes a discrete renewal process with occurrence times identically distributed with geometric distribution. Such cohesions are appropriate when it is reasonable to assume that the past change points are noninformative about the future change points, which is of use for many practical applications.

Let $\theta_1, \ldots, \theta_n$ be a sequence of unknown parameters conditional on which the random variables $X_1, \ldots, X_n$ have marginal densities $f_1(X_1|\theta_1), \ldots, f_n(X_n|\theta_n)$, respectively. The prior distributions of $\theta_1, \ldots, \theta_n$ are built as follows. Given a partition ρ, one has that $\theta_i = \theta_{[i_{(r-1)}i_r]}$, for every $i_{(r-1)} < i \leq i_r$, and that $\theta_{[i_0 i_1]}, \ldots, \theta_{[i_{(b-1)}i_b]}$ are independent, with $\theta_{[ij]}$ having prior density $\pi_{[ij]}(\theta)$, $\theta \in \Theta_{[ij]}$, in which $\Theta_{[ij]}$ is the parameter space.

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

Hartigan, 1992):

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

$$P(\rho = \{i_0, i_1, \ldots, i_b\}) = \frac{\Pi^b_{j=1} c_{[i_{(j-1)}i_j]}}{\sum_{\mathcal{C}} \Pi^b_{j=1} c_{[i_{(j-1)}i_j]}} \qquad (2)$$

in which $\mathcal{C}$ is the set of all possible partitions of set I 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) conditional on $\rho = \{i_0, i_1, \ldots, i_b\}$, the sequence $X_1, \ldots, X_n$ has the joint density given by:

$$f(X_1, \ldots, X_n | \rho = \{i_0, i_1, \ldots, i_b\}) = \Pi^b_{j=1} f_{[i_{(j-1)}i_j]}(\mathbf{X}_{[i_{(j-1)}i_j]}), \qquad (3)$$

in which $f_{[ij]}(\mathbf{X}_{[ij]}) = \int_{\Theta_{[ij]}} f_{[ij]}(\mathbf{X}_{[ij]}|\theta)\pi_{[ij]}(\theta)d\theta$ is the density of the random vector, called data factor.

Consequently, if $(X_1, \ldots, X_n; \rho) \sim$ PPM and a square error loss function is assumed, the posterior expectation (or the product estimate) of parameter θ_k is given by (Barry and Hartigan, 1993):

$$E(\theta_k | X_1, \ldots, X_n) = \sum_{i=0}^{k-1} \sum_{j=k}^{n} r^*_{[ij]} E(\theta_k | \mathbf{X}_{[ij]}), \qquad (4)$$

for $k = 1, \ldots, n$, in which $r^*_{[ij]} = P([ij] \in \rho | X_1, \ldots, X_n)$ denotes the posterior relevance of block $[ij]$.

If Yao's (1984) cohesions are assumed, the conditional prior distributions for ρ and B are given, respectively, by:

$$P(\rho = \{i_0, i_1, \ldots, i_b\}|p) = p^{b-1}(1-p)^{n-b}, \quad b \in I, \qquad (5)$$

for every partition $\{i_0, i_1, \ldots, i_b\}$, satisfying $0 = i_0 < i_1 < \cdots < i_b = n$, and,

$$P(B = b|p) = C^{n-1}_{b-1} p^{b-1}(1-p)^{n-b}, \ b \in I, \qquad (6)$$

in which C^{n-1}_{b-1} denotes the number of distinct partitions of I into b contiguous blocks.

Additionally, if a prior distribution $\pi(p)$ is assumed for p, it follows that the posterior distribution of ρ and the posterior distribution of the number of blocks B assume, respectively, the forms (Loschi et al., 2003):

4

$$P(\rho = \{i_0, i_1, \ldots, i_b\}|X_1, \ldots, X_n) =$$

$$\Pi_{j=1}^n f(X_{i_{(j-1)}i_j}) \int_0^1 p^{b-1}(1-p)^{n-b}\pi(p)dp, \tag{7}$$

$$P(B = b|X_1, \ldots, X_n) =$$

$$C_{b-1}^{n-1}\Pi_{j=1}^n f(X_{i_{(j-1)}i_j}) \int_0^1 p^{b-1}(1-p)^{n-b}\pi(p)dp. \tag{8}$$

For further details on how the PPM can be tailored for normal means and variances, the reader is referred to the paper by Loschi et al. (2003).

Remark: *Notice that if p has beta distribution with parameters α and β, $p \sim \mathcal{B}(\alpha, \beta)$, the number of change points $B - 1$ has beta-binomial distribution with parameters $n - 1$, α and β. In this case, the prior mean and variance of B are given, respectively, by:*

$$E(B) = (n-1)\frac{\alpha}{\alpha+\beta} + 1;$$

$$Var(B) = (n-1)\frac{\alpha\beta(\alpha+\beta+n-1)}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Assuming in the prior evaluation that we believe that the number of change points in the sequence is small, we should choose the hiperparameters α and β such that the prior distribution of p is concentrated in small values. For example, we can take small values for α and large values for β as it will be considered in Section 4.

2.2 PPM Extensions

In general, the posterior distribution of ρ will not provide a good idea about when changes occurred since each value of ρ usually will receive low mass. More informative for a decision-maker is to obtain the posterior probability of each instant to be a change point. This posterior probability can be derived as follows.

Let $\mathcal{C}_k$ be the subset of $\mathcal{C}$ that contains all partitions that include the kth instant as a change point, that is, each partition in $\mathcal{C}_k$ assume the form $\{i_0, \ldots, i_{(l-1)}, i_l = k, i_{(l+1)}, \ldots, i_b\}$ for any $l \in I$. The event A_k denotes that the kth instant is a change point, for $k = 2, \ldots, n$. Thus:

$$P(A_k|X_1, \ldots, X_n) =$$
$$= \sum_{\mathcal{C}_k} P(\rho = \{i_0, \ldots, i_{(l-1)}, i_l = k-1, i_{(l+1)}, \ldots, i_b\}|X_1, \ldots, X_n)$$
$$\propto \sum_{\mathcal{C}_k} \Pi_{j=1}^{(l-1)} c^*_{[i_{(j-1)} i_j]} c^*_{[i_{(l-1)}(k-1)]} c^*_{[(k-1) i_{(l+1)}]} \Pi_{j=l+1}^b c^*_{[i_{(j-1)} i_j]}. \qquad (9)$$

Similarly, the posterior probability for two specific instants i, j, $i \neq j$, to be change points can be obtained by computing the probability of $A_i \cap A_j$, and so on.

Let assume that p has prior distribution $\pi(p)$. It follows that the posterior probability for A_k is given by:

$$P(A_k|X_1, \ldots, X_n) = \sum_{\mathcal{C}_k} \Pi_{j=1}^n f(X_{i_{(j-1)} i_j}) \int_0^1 p^{b-1} (1-p)^{n-b} \pi(p) dp, \qquad (10)$$

for all $k = 2, \ldots, n$.

3 Computational Methods

3.1 Introduction

Suppose that p has prior distribution $\pi(p)$ and that, given ρ, $\theta_k \in [ij]$, for $k = 1, \ldots, n$, and $i, j \in I$, $i < j$. Let $\mathbf{X}_{[0n]} = (X_1, \ldots, X_n)$ and $\theta = (\theta_1, \ldots, \theta_n)$ and denote by θ_{-k} the vector $(\theta_1, \ldots, \theta_{k-1}, \theta_{k+1}, \ldots, \theta_n)$. The full conditional distributions of p, ρ, and θ_k are given, respectively, by:

$$\pi(p|\rho, \theta, \mathbf{X}_{[0n]}) \propto p^{b-1}(1-p)^{n-b} \pi(p),$$
$$\pi(\rho|p, \theta, \mathbf{X}_{[0n]}) \propto \left(\Pi_{j=1}^b f_{[i_{(j-1)} i_j]}(X_{[i_{(j-1)} i_j]}) \right) p^{b-1}(1-p)^{n-b},$$
$$\pi(\theta_k|\rho, p, \theta_{-k}, \mathbf{X}_{[0n]}) \propto f_{[ij]}(\theta_k|X_{[ij]}),$$

for $k = 1, \ldots, n$. Notice that it is not easy to sample directly from the full conditional distribution of ρ. In the next section, a method to sampler from the previous distributions is described.

6

3.2 Gibbs Sampling Scheme to the PPM

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

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

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

Each partition $(U_1^s, \ldots, U_{n-1}^s)$, $s \geq 1$, is generated by using Gibbs sampling. Starting from an initial value $(U_1^0, \ldots, U_{n-1}^0)$, the rth element at step s, U_r^s, is generated from the conditional distribution:

$$U_r | U_1^s, \ldots, U_{r-1}^s, U_{r+1}^{s-1}, \ldots, U_{n-1}^{s-1}, p^{(s-1)}, \theta^{(s-1)}; \mathbf{X}_{[0n]},$$

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

In order to generate the samples of $\mathbf{U}$ above, it is enough to consider the following ratio:

$$R_r = \frac{P(U_r = 1 | V_r^s, p^{(s-1)}, \theta^{(s-1)}; \mathbf{X}_{[0n]})}{P(U_r = 0 | V_r^s, p^{(s-1)}, \theta^{(s-1)}; \mathbf{X}_{[0n]})}$$

for $r = 1, \ldots, n-1$, in which $V_r^s = \{U_1^s = u_1, \ldots, U_{r-1}^s = u_{r-1}, U_{r+1}^{s-1} = u_{r+1}, \ldots, U_{n-1}^{s-1} = u_{n-1}\}$.

For the present case, in which p has a beta prior distribution $p \sim \mathcal{B}(\alpha, \beta)$ with parameters $\alpha > 1$ and $\beta > 1$, the value R_r becomes:

$$R_r = \frac{f_{[xy]}(X_{[xy]})\Gamma(n+\beta-b+1)\Gamma(b+\alpha-2)}{f_{[xr]}(X_{[xr]})f_{[ry]}(X_{[ry]})\Gamma(b+\alpha-1)\Gamma(n+\beta-b)}, \tag{11}$$

for $b = 1, \ldots, n$, in which:

$$x = \begin{cases} \max\{i, \text{ s.t. } : 0 < i < r, U_i^s = 0\}, & \text{if } U_i^s = 0, \text{ for some} \\ & i \in \{1, \ldots, r-1\} \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

and

$$y = \begin{cases} \min\{i, \text{ s.t. } : r < i < n, U_i^{s-1} = 0\}, & \text{if } U_i^{s-1} = 0, \text{for some} \\ & i \in \{r+1, \ldots, n-1\} \quad (13) \\ n, & \text{otherwise.} \end{cases}$$

Consequently, the criterion of choosing the values U_i^s, $i = 1, \ldots, n-1$, becomes:

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

for $r = 1, \ldots, n-1$, in which $u \sim \mathcal{U}(0,1)$. Further details on how one could derive the product estimates and the posterior distribution of the number of blocks can be found in the paper by Loschi et al. (2003).

3.3 Gibbs Sampling Scheme to the PPM Extensions

Each sample of the posterior distribution of p is generated from the following beta distribution:

$$p^s | \rho, \theta, \mathbf{X}_{[0n]} \sim \mathcal{B}(\alpha + b^s - 1, n + \beta - b^s), \tag{14}$$

for $s \geq 1$, in which b^s is the number of blocks in the sth vector $\mathbf{U}$ which is obtained by noticing that the number of blocks in ρ is given by:

$$B = 1 + \sum_{i=1}^{n-1} (1 + U_i). \tag{15}$$

Similarly, the estimates of the posterior probability of the kth instant to be a change point is:

$$P(A_k) = \frac{N}{T}, \tag{16}$$

for $k = 2, \ldots, n$, in which N is the number of vectors for which it is observed that $U_{k-1} = 0$ and T is the total number of vectors generated.

8

4 Application

All the algorithms described in the previous sections were coded in C++ and are available from the authors upon request. In the experiments here presented, the algorithms were run by a PC, Pentium processor 166 MHz, 32 MB RAM. The CPU times were around 25 seconds. For the Gibbs sampling scheme, 4,600 samples were generated. From a graph (not shown) it was seen that after only 100 samples the simulation reached the steady state. Thus, the initial 100 samples were discharged for *burn-in*. Additionally, since no correlation was found at the sequence of simulated values, a lag of one was chosen. One can easily find in the literature a discussion about the number of iterations to be discharged and the lag to be taken. For example, we suggest the book by Gamerman (1997).
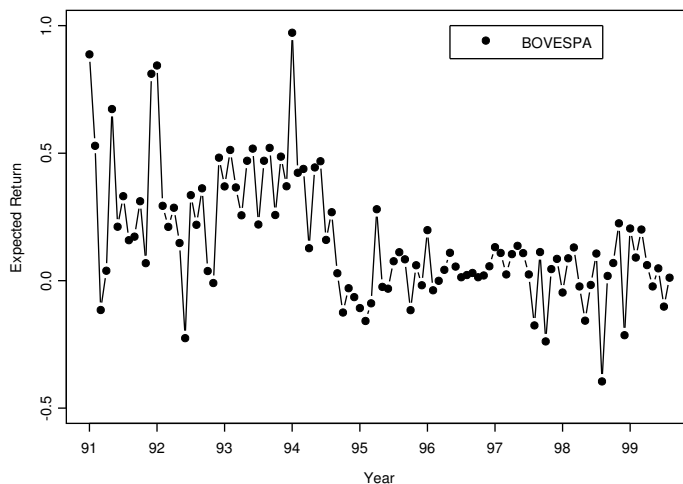
Fig. 1. BOVESPA series of returns.

The application focuses on an important index from the Brazilian stock market, the BOVESPA index, expressed in terms of monthly returns, from January, 1991, to August, 1999, as seen in Figure 1. As usual in finance, the return series is defined by the transformation $R_t = (P_t - P_{t-1})/P_{t-1}$, in which P_t is the closing price at the tth month.

It seems reasonable to assume that the observations are conditionally independent and distributed according to the normal distribution $\mathcal{N}(\mu_{[ij]}, \sigma^2_{[ij]})$. We shall assume the normal-inverted-gamma distribution, a natural conjugate prior distribution, successfully used to model stock market data (Hsu, 1982). For the present case we are adopting the following:

$$\mu_{[ij]}|\sigma^2_{[ij]} \sim \mathcal{N}(0.0, \sigma^2_{[ij]}), \text{ and } \sigma^2_{[ij]} \sim \mathcal{IG}(0.01/2, 4/2).$$
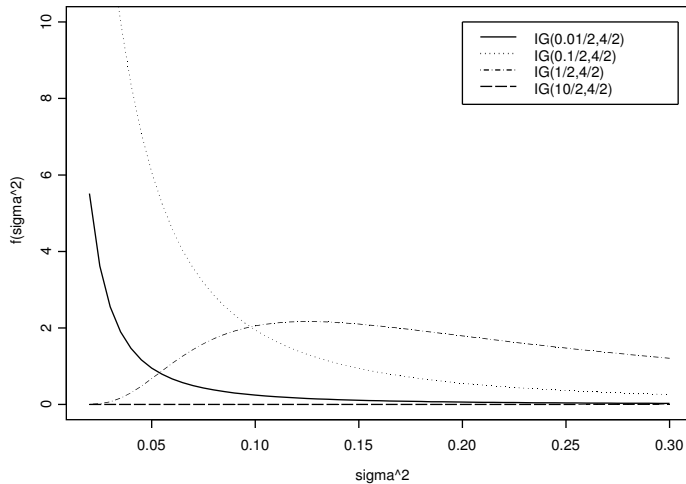
9

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

As we can see in Figure 2, the inverted-gamma distribution considered concentrates its mass in low values but presents some variability. These specifications were chosen because the Brazilian market is emerging and therefore more susceptible to the world political scenario than developed markets (Mendes, 2000). Different settings for $a/2$ and $d/2$ that could be adopted for different markets are seen in Figure 2.

In regard to the parameter p that indexes Yao's (1984) cohesions, a beta prior distribution was adopted. Several different prior specifications were considered, plotted in Figure 7. Firstly, it is possible to be completely non-informative about p, by assuming for instance a beta prior distribution such as $\mathcal{B}(1.1, 1.1)$. If we assumed prior distributions that concentrate most of its mass in small values, such as $\mathcal{B}(1, 50)$, we would mean that it is reasonable to expect a small number of changes. On the other hand, a prior distribution with a high average value, such as $\mathcal{B}(50, 50)$, would mean that small blocks of returns and a high number of change points may be expected a priori.

Figure 3 shows the product estimates of the expected returns and Figure 4, for the variance (or volatility), for all different prior specifications for p. It is no surprise that the product estimates were different for different prior distributions for p as earlier noticed (Loschi and Cruz, 2002). However, it is remarkable that for all of them change points were identified in the expected returns around September, 1994, April, 1995, and July, 1997. Additionally, change points were identified in the volatility around October, 1992, September, 1994, April, 1995, July, 1997, and January, 1999.

Figure 5 presents the posterior probability of each point to be a change point. It is remarkable that the estimates corresponding to the prior distributions
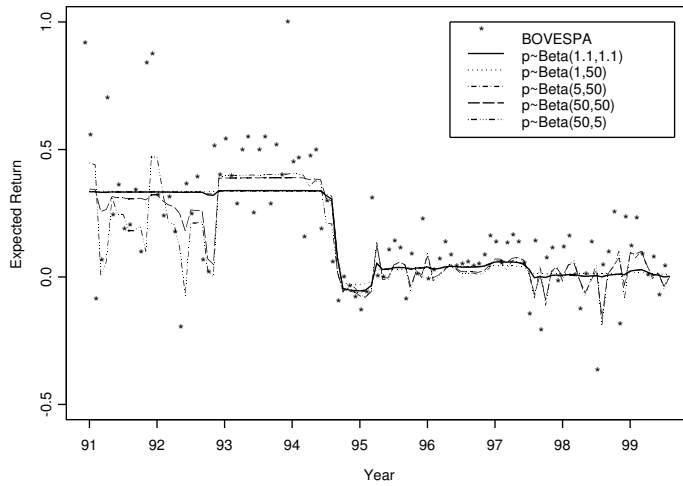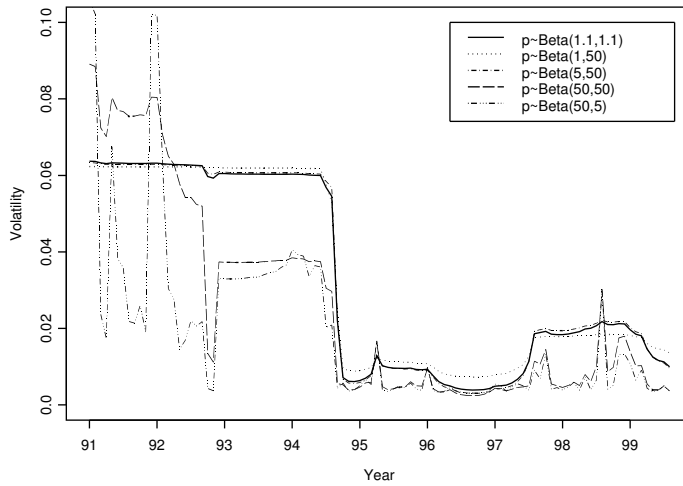
10

Fig. 3. Posterior estimates for μ.

Fig. 4. Posterior estimates for σ.

$\mathcal{B}(1.1, 1.1)$, $\mathcal{B}(1, 50)$, and $\mathcal{B}(5, 50)$ are almost identical. Based on these estimates it is easily identified three months with "high" probability to be a change point, namely September, 1994, with probability 64.7%, April, 1995, with probability 60.4%, and August, 1997, with probability 48.4%. The scenario described by priors $\mathcal{B}(50, 50)$ and $\mathcal{B}(50, 5)$, would be that of a high assurance of a highly unstable market with high probability of structural changes expected a priori. As a result, much more months would be identified as change points, specially after 1995. Since we do not think this was the case, we shall not go further on analyzing these priors.

Figure 6 depicts the estimates for the posterior probability that each instant

11
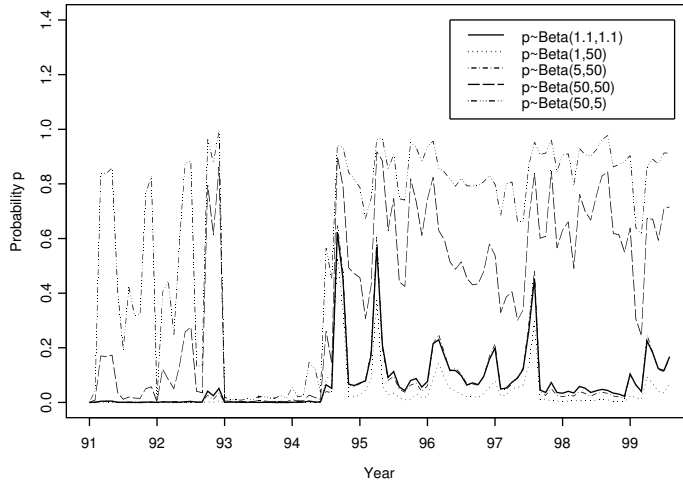
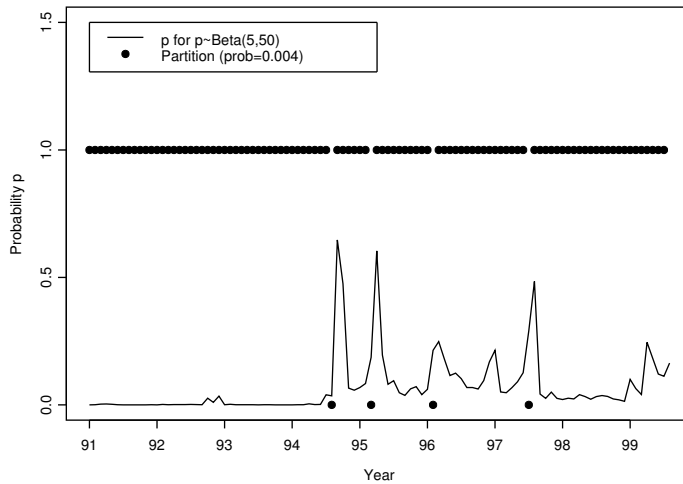Fig. 5. Posterior probability of change.

Fig. 6. Probabilities of change and the most probable partition.

is a change point, by Eq. (16), and the most probable partition, by Eq. (7), considering the prior we believe is the most appropriate for the Brazilian market, i.e., $\mathcal{B}(5, 50)$. The most probable posterior partition is $\rho = \{0, \text{Sep}/94, \text{Apr}/95, \text{Mar}/96, \text{Aug}/97\}$ which occurs with probability 0.004. Notice that this partition indicates that October, 1994, May, 1995 , April, 1996, and September, 1997, are change points. However, from Figure 5, a different partition may be formed, by considering those months whose probabilities to be change points are the highest, that is $\rho = \{0, \text{Aug}/94, \text{Mar}/95, \text{Jul}/97\}$. Although the estimated probability of the latter partition is smaller than 0.004 (≈ 0.00134), such a partition is more intuitive than the former. Thus, the prob-
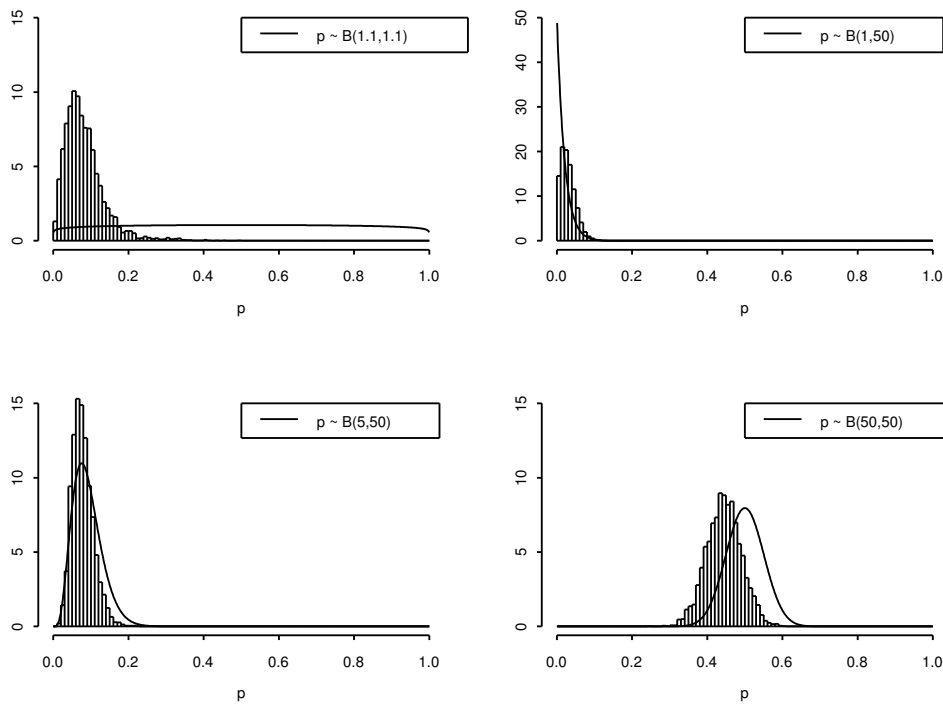
12

Fig. 7. Prior and Posterior Distributions of p.

Table 1
Descriptive statistics for the prior and posterior distributions of p.

	Prior Distribution		Posterior Distribution				
Prior	Mean	StDev	Mean	StDev	Q1	Median	Q3
$p \sim \mathcal{B}(1.1, 1.1)$	0.5000	0.2795	0.0823	0.0540	0.0464	0.0718	0.1047
$p \sim \mathcal{B}(1, 50)$	0.0196	0.0192	0.0302	0.0198	0.0149	0.0270	0.0414
$p \sim \mathcal{B}(5, 50)$	0.0909	0.0384	0.0781	0.0279	0.0583	0.0745	0.0945
$p \sim \mathcal{B}(50, 50)$	0.5000	0.0498	0.4442	0.0461	0.4138	0.4444	0.4749
$p \sim \mathcal{B}(50, 5)$	0.9091	0.0384	0.7273	0.0495	0.6956	0.7299	0.7620

abilities of change seems to be a more meaningful tool to identify probable partitions than by Eq. (7).

Figure 7 shows the prior and posterior distributions for p for some of the prior distributions considered. It is noticeable that the posterior distribution of p has a low valued mode (see also Table 1). Assuming that $p \sim \mathcal{B}(5, 50)$ and considering the square loss function, the Bayes estimate of p is 0.0909 which decreases for 0.0781 in the posterior evaluation. Similar conclusions can be
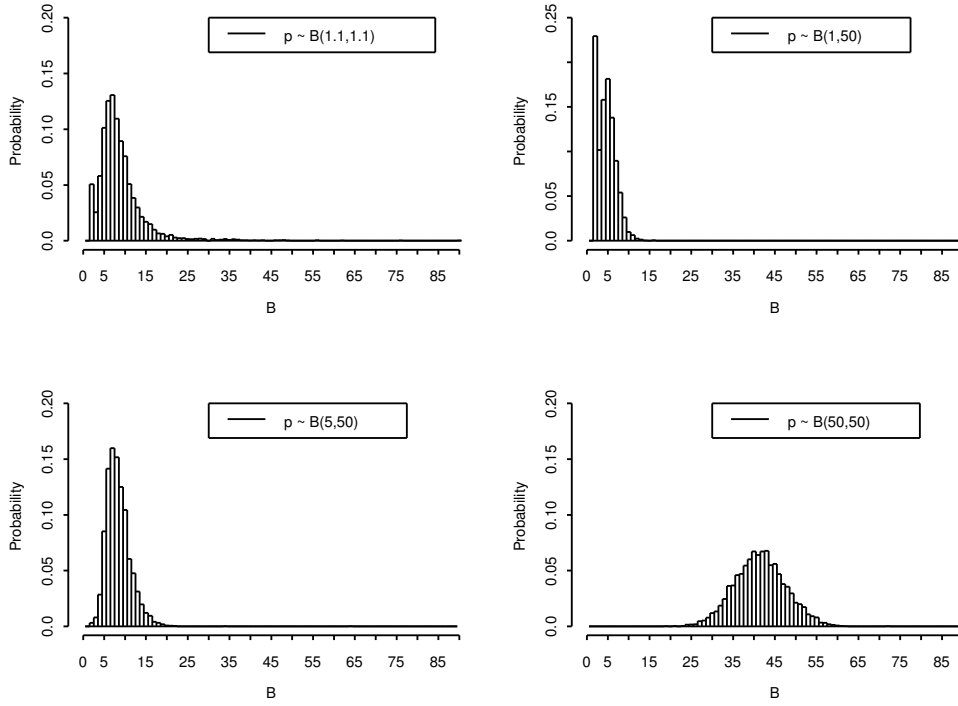
13

Fig. 8. Posterior distributions of B.

Table 2
Descriptive statistics for the prior and posterior distributions of B.

| Prior | Prior Distribution | | Posterior Distribution | | | | | |
	Mean	StDev	Mean	StDev	Mode	Q1	Median	Q3
$p \sim \mathcal{B}(1.1, 1.1)$	52.5	29.1	8.74	5.67	7	6	8	10
$p \sim \mathcal{B}(1, 50)$	3.02	2.42	4.68	2.37	2	3	5	6
$p \sim \mathcal{B}(5, 50)$	10.4	4.90	8.40	2.95	7	6	8	10
$p \sim \mathcal{B}(50, 50)$	52.5	7.19	41.5	6.16	43	37	41	45
$p \sim \mathcal{B}(50, 5)$	94.6	4.90	66.0	5.59	68	63	66	70

drawn for the other prior distributions. Observe also that the posterior estimates were generally more precise (lower standard deviation) than the prior estimates, exception made for prior $\mathcal{B}(50, 5)$ for which the standard deviation increased (0.0384 vs. 0.0495, as seen in Table 1).

Figure 8 shows the posterior distribution of the number of block B, Eq. (15), and Table 2 presents some descriptive statistics for the prior and posterior distributions of B. As earlier observed concerning p, the posterior distribution

of B has an unique low mode, for all prior specification. Notice also that the posterior distributions present lower standard deviation in comparison with the prior distribution, exception made for prior $\mathcal{B}(50, 5)$ (4.90 vs. 5.59, as seen in Table 2). In conclusion, in the posterior evaluation, the expected number of change points were lower and the posterior estimates were in general more accurate (lower standard deviation). In any case the summaries of ρ and B are sensitive to the prior. Thus, proper prior elicitation becomes crucial.

5 Summary and Conclusions

The classical product partition model (PPM) was considered to the identification of multiple change points in the means and variances of normal data sequences. This paper extends previous work by providing a full Bayesian analysis for the change point problem by means of both the PPM and Yao's (1984) cohesions, and by proposing a method to compute the probability that each instant of time is a change point.

Conjugate prior distributions were assumed for the means and variances and a beta prior distribution was considered to describe the prior behavior of the parameter p that indexes Yao's (1984) cohesions and represents the probability to have a change at a given instant of time. The PPM was tailored to provide new information, namely the posterior distribution of the probability to have a change in any instant of time and the posterior probability that each instant is a change point.

The methodology was applied to an important Brazilian stock market data. Several different prior specifications for p were considered. The results indicated that the new method is quite effective and may provide useful new information. Mainly, it can be concluded that the posterior probability that each instant to be a change point provides a better tool for decision-makers than the posterior distribution on the random partition formed by the instants when change points occurred.

15

References

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

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

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

Gamerman, D., 1997. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Chapman & Hall, London, UK.

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

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

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

Loschi, R. H., Cruz, F. R. B., 2002. 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.

Loschi, R. H., Cruz, F. R. B., Iglesias, P. L., Arellano-Valle, R. B., 2003. A Gibbs sampling scheme to the product partition model: An application to change point problems. Computers & Operations Research 30 (3), 463–482.

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

Quintana, F. A., Iglesias, P. L., 2003. Bayesian clustering and product partition models. Journal of the Royal Statistical Society B 65 (2), 557–574.

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