SOBRAPO

A FUZZY/BAYESIAN APPROACH FOR THE TIME SERIES CHANGE POINT DETECTION PROBLEM

Marcos Flávio S.V. D'Angelo[1*], Reinaldo M. Palhares[2], Ricardo H.C. Takahashi[3] and Rosangela H. Loschi[4]

ABSTRACT. This paper addresses the change point detection problem in time series. A methodology based on the Metropolis-Hastings algorithm applied to time series modeled as a process with Beta distribution is discussed. In order to make this methodology useful in practice, a fuzzy cluster technique is applied to the initial time series at first, generating a new data set with Beta distribution. Bayesian procedures are considered for inference and the Metropolis-Hastings algorithm is used to sample from the posteriors. In the clustering process, a Kohonen neural network is used having as objective to find the best centers of the time series to be used in the fuzzyfication process. Finally, it will be presented a simulation results in the series of the electric energy consumption in Brazil, between January of 1976 and December of 2000, five months before the blackout occurred in 2001. Such result illustrates the efficiency of the proposed methodology for change point detection in time series.

Keywords: change point, fuzzy clustering, Metropolis-Hastings.

1 INTRODUCTION

Traditional methods of describing a time series behavior are based on models which satisfactorily explain its dynamics. Some fundamental question must be answered, however, in order to obtaine a model that reasonably fits to data: Are there switching regimes in the series? Can a single model, all the time, fit the dynamics behavior? If there are significant changes for the underlying

*Corresponding author

[1]Universidade Estadual de Montes Claros – UNIMONTES, Departamento de Ciência da Computação, Av. Rui Braga, s/n, 39401-089 Montes Claros, MG, Brazil. E-mails: marcos.dangelo@unimontes.br / dangelo@cpdee.ufmg.br

[2]Universidade Federal de Minas Gerais – UFMG, Departamento de Engenharia Eletrônica, Av. Antônio Carlos, 6627, Pampulha, 31270-901 Belo Horizonte, MG, Brazil. E-mail: palhares@cpdee.ufmg.br

[3]Universidade Federal de Minas Gerais – UFMG, Departamento de Matemática, Av. Antônio Carlos, 6627, Pampulha, 31270-901 Belo Horizonte, MG, Brazil. E-mail: taka@mat.ufmg.br

[4]Universidade Federal de Minas Gerais – UFMG, Departamento de Estatística, Av. Antônio Carlos, 6627, Pampulha, 31270-901 Belo Horizonte, MG, Brazil. E-mail: loschi@est.ufmg.br

time series, it seems natural to find out those change points or change periods before modeling the whole process.

Change points detection is a problem of interest in many different areas and it has been used to analyze financial time series (Oh KJ, Roh TH & Moon MS, 2005), ecological time series (Beckage, Joseph, Belisle, Wolfson & Platt, 2007), number of crimes series (Loschi RH, Gonçalves FB & Cruz FBR, 2005), and many others. The main techniques for change points detection presented in the literature are based on sequential statistical tests and Bayesian analysis. The most common statistical test on the change point detection problem is the CUSUM (Cumulative Sum). The CUSUM test proposed by (Hinkey, 1971) is widely used in the change point detection, and applications of this method can be found in (Hadjiliadis & Moustakides, 2006), (Lee, Nishiyama & Yoshida, 2006), and (Lee S, Park S, Maekawa K & Kawai K, 2006), as well as its modifications and extensions. In the Bayesian approach, inference about the changes are made based on the posterior for the positions of the change point. In this setting, the MCMC (Markov chain Monte Carlo) methods are commonly used to approximate the posteriors. The Bayesian analysis context the reference work for multiple change points detection is the paper of (Barry & Hartigan, 1993) that introduced the product partition model (ppm) to identify change points in the mean of data having a Gaussian distribution. In (Loschi RH & Cruz FRB, 2005b) the posterior probability of any instant being a change point is used as an evidence measure that the behavior of a sequence of data changes at any moment. The effectiveness of this technique in identifying changes in a Poisson distribution when compared with the measure proposed by (Hartigan, 1990) is presented in (Loschi RH & Cruz FRB, 2005a).

However, all the previous approaches necessarily demand some type of prior knowledge about the time series, namely the type of distribution that models the data behavior. Thus, differently from former techniques, the method proposed in this paper does not require any prior knowledge about statistical properties of the time series before the application of the MCMC procedure. This is made by the introduction of a first step, in which a fuzzy set technique is applied in order to cluster and to transform the initial time series, about which there is no *a priori* knowledge of its distribution, into a time series which can be approximated by a Beta distribution. Specifically in this first step, the Kohonen network is used to find centers of the clusters, and in the sequel the fuzzy membership degree is computed for each point of the initial time series, generating a new time series with beta distribution. This new time series generated in the first step allows systematically to apply the same strategy to detect the change via Bayesian change point model with a fixed reference distribution: the Beta distribution. Thus the Metropolis-Hastings algorithm (Gamerman, 1997) can be used to sample from the posteriors including the one that indicates the position of the change point.

Assuming that p_1 is the first operation point, p_2 is the second operation point, $\varepsilon(t)$ is the noise signal with $\pi(\cdot)$ distribution and m is the position of the change point, throughout the paper, it is considered the following time series:

$$f(t) = \begin{cases} p_1 + 0.1 * \varepsilon(t) - 0.1 * \varepsilon(t-1), & \text{se } t \leq m ; \\ p_2 + 0.1 * \varepsilon(t) - 0.1 * \varepsilon(t-1), & \text{se } t > m . \end{cases} \tag{1}$$

The Figure 1 show the time series behavior with fixed $p_1 = 1$, $p_2 = 2$, $m = 30$ e $\varepsilon(t) \sim U(0, 1)$ and 60 samples of time series.
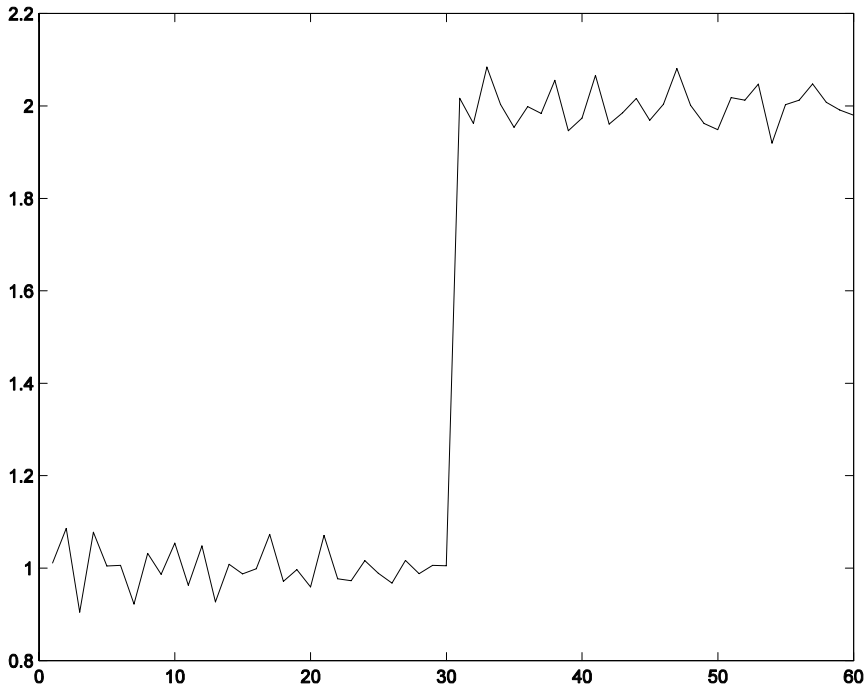

Figure 1 – Time series behavior with fixed $p_1 = 1$, $p_2 = 2$, $m = 30$ e $\varepsilon(t) \sim U(0, 1)$ and 60 samples of time series.

To establish notation, along this paper, $U(0, 1)$ denotes the uniform distribution in the interval $[0, 1]$, $N(0, 1)$ denotes the standard Normal distribution with zero mean and unit variance, $t(5)$ denotes the Students-t distribution with 5 degrees of freedom, $gamma(a, b)$ denotes the Gamma distribution with scale parameter a and shape parameter b and $\Gamma(a)$ denotes the gamma function evaluated in a. All the time series, $e.g.$, $y(t)$, denotes a discrete-time sequence.

2   TIME SERIES TRANSFORMATION BY FUZZY SETS

One of the main applications of the fuzzy sets theory (Zadeh, 1965) is for clustering detection, in which the problem of grouping the data into k separate categories is solved by defining different pertinences to each element belonging to category.

Thus, in this section, the first step of the proposed formulation to the change point detection problem is detailed, namely the fuzzy clustering of the initial time series.

Definition. *Let $y(t)$ be a time series, and consider a positive integer K. The set*

$$C = \big\{C_i \,|\, \min\{y(t)\} \leq C_i \leq \max\{y(t)\}, \; i = 1, \ldots, k\big\}$$

that solves the minimization problem

$$\min \sum_{i=1}^{k} \sum_{\mu_i(t) \in C_i} \|\mu_i(t) - C_i\|^2 \tag{2}$$

is called the fuzzy cluster centers *set for the time series* $y(t)$. *The* fuzzy membership degree *or fuzzy membership function value of* $y(t) \in C_i$ *(which means* $y(t)$ *belongs to the cluster* C_i) *is given by*

$$\mu_i(t) \triangleq 1 - (y(t) - C_i)^2 / \sum_{j=1}^{k} (y(t) - C_j)^2 \tag{3}$$

Given the set C of cluster centers, it is an easy task to measure the distance between each point in the time series $y(t)$ and each center C_i. The problem of finding the centers can be solved, for instance, via k-means (Kaufman & Rousseeuw, 1990), C-means (Bezdek, 1981), and Kohonen network (Kohonen, 2001). In this paper, the Kohonen network technique is used.

The proposed fuzzy clustering to transform a given time series into a new one is described below:

1. Input the time series $y(t)$;

2. Find the set of k centers $C = \{C_i | \min\{y(t)\} \le C_i \le \max\{y(t)\}, i = 1, 2\}$, that minimizes the Euclidean distance given in (2) and that is illustrated in Figure 2, considering, for example, the time series (1). For the proposal of this paper, it is considered $k = 2$ centers.

3. Compute the fuzzy membership degree given in (3), for each sample of the time series, $y(t)$, with respect to each center C_i (as illustrated in Figure 3 considering, for example, the time series (1)).

Since the idea is to find just one change point, only two centers are found. The function $\mu 1(t)$ defines a new time series, in which the change point is detected.

From Figure 3, it is noticeable that the fuzzy membership degree is concentrated around 1 (one) for the first 30 values observed before the change point (see also the histogram in Figure 4), and is concentrated around zero for the last 30 values observed after the change point (see also the histogram in Figure 5). Furthermore, the distribution of $\mu_1(t)$, defined according to (3), is confined into the interval [0, 1], as shown below:

if $y(t) \longrightarrow C1$ then $\mu_1(t) \longrightarrow 1^-$;

if $y(t) \longrightarrow C2$ then $\mu_1(t) \longrightarrow 0^+$;

if $C1 \longrightarrow C2$ then $\mu_1(t) \longrightarrow 1/2$;

if $y(t) \longrightarrow [C1, C2]$ then $\mu_1(t) \longrightarrow [0, 1]$.

Using the Kullback-Leibler divergence (Kullback & Leibler, 1951) one may conclude that the distributions of $\mu_1(t)$ before and after the change point can be approximated by Beta distributions with different input parameters, say, for $t \le m$, one assumes a *beta(a, b)* distribution, and for
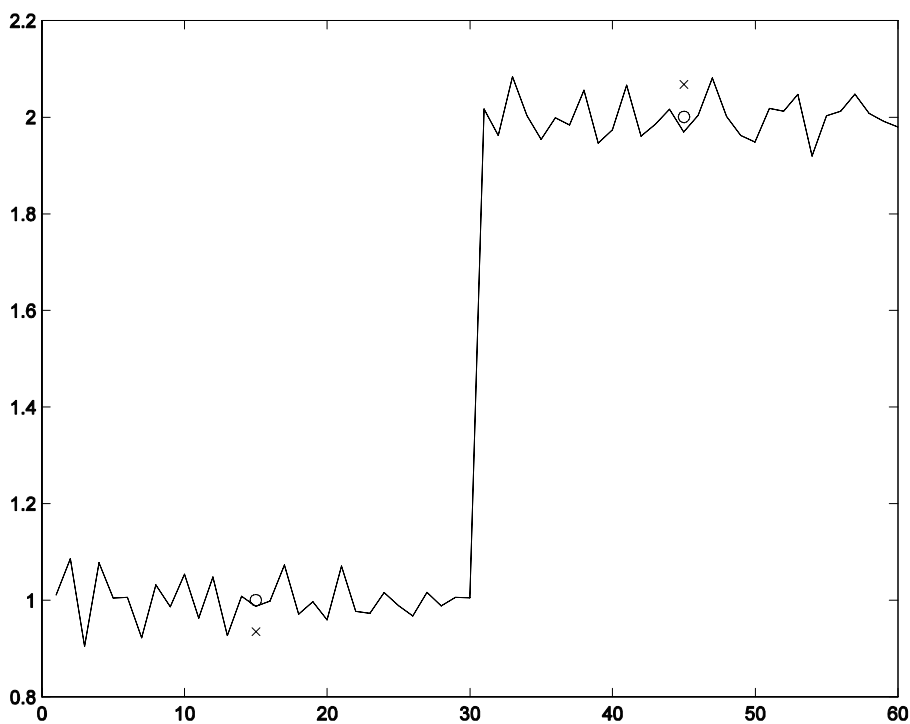
Figure 2 – Time series centers considering the Figure 1: × – before of Kohonen network learning and ○ – after of Kohonen network learning.
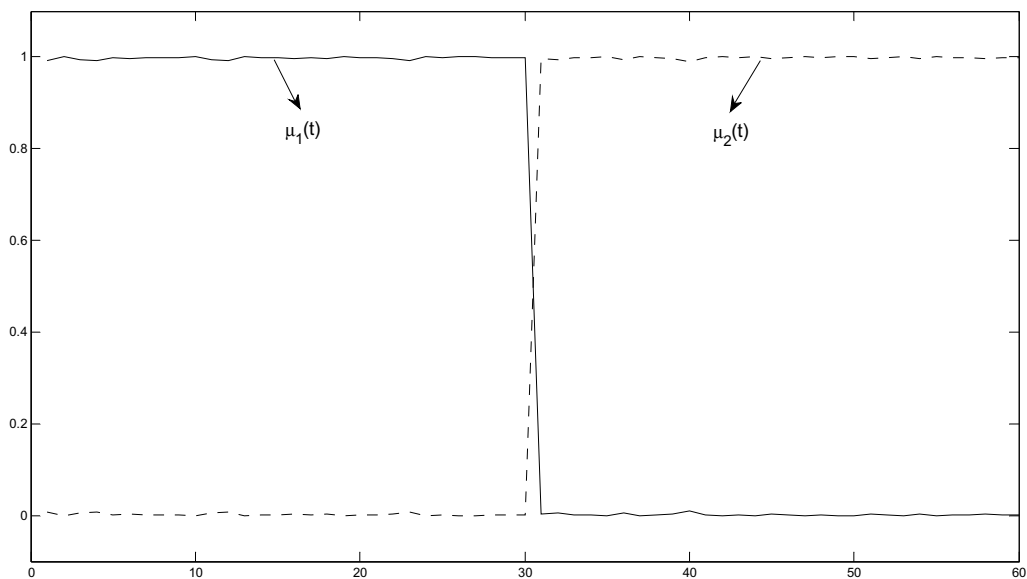

Figure 3 – $\mu_1(t)$ and $\mu_2(t)$.

$t > m$ it is assumed the $beta(c, d)$ distribution. For example, in the case in which there is a change point in the time series, the parameter a is greater than the parameter in the Beta distribution of $\mu_1(t)$ if $t \leq m$, and the parameter c is smaller than the parameter d in the Beta distribution of $\mu_1(t)$ if $t > m$. Figures 6-8 presents an illustration of the distributions of $\mu_1(t)$ for the time series in (1) with sample size $n = 200$. It is assumed that the change point took place at the position $m = 100$. This empirical test has been performed starting from several time series, with different probability distributions for $\varepsilon(t)$, say, $U(0, 1)$ in Figure 6, $N(0, 1)$ in Figure 7, and $t(5)$ in Figure 8. In all these cases, the behavior of the transformed time series, obtained by applying the clustering technique, was well described by Beta distributions with similar shapes.
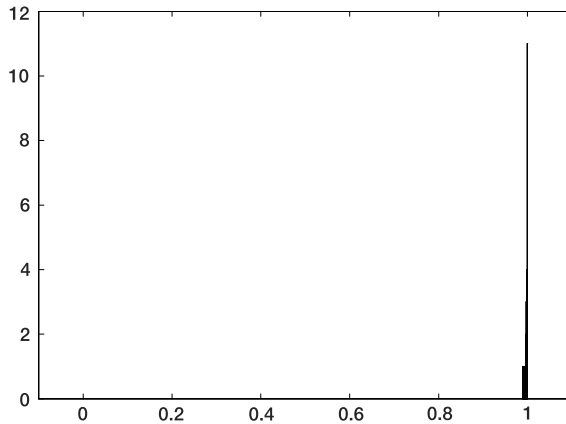

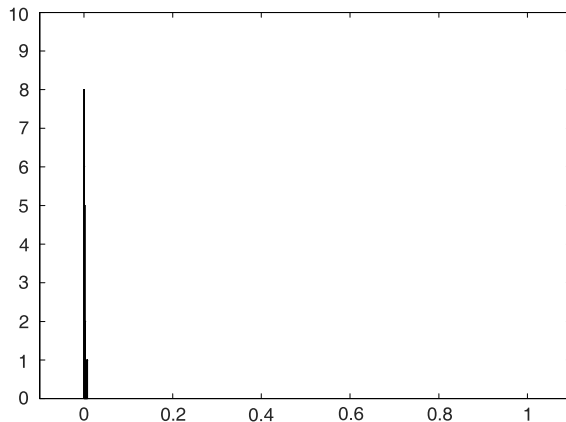**Figure 4** – Histogram of $\mu_1(t)$ for ≤ 30.

Figure 5 – Histogram of $\mu_1(t)$ for > 30.

Since the clustering technique assures the same type of probability distribution function for the new time series, $\mu_1(t)$ this fixed statistical model can be assumed in the Bayesian formulation to detect the change point in the transformed time series. In this paper, the Metropolis-Hastings algorithm is used to sample from the posteriors, since it is a powerful and simple strategy to do it.
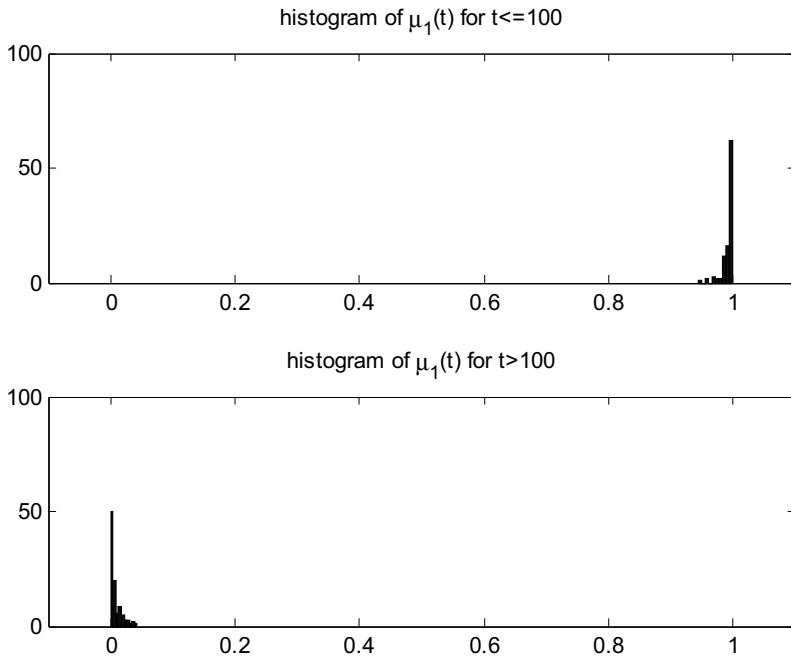
histogram of $\mu_1(t)$ for t<=100

histogram of $\mu_1(t)$ for t>100

Figure 6 – Histograms of μ_{1t} for $\varepsilon(t) \sim U(0, 1)$.

histogram of $\mu_1(t)$ for t<=100

histogram of $\mu_1(t)$ for t>100
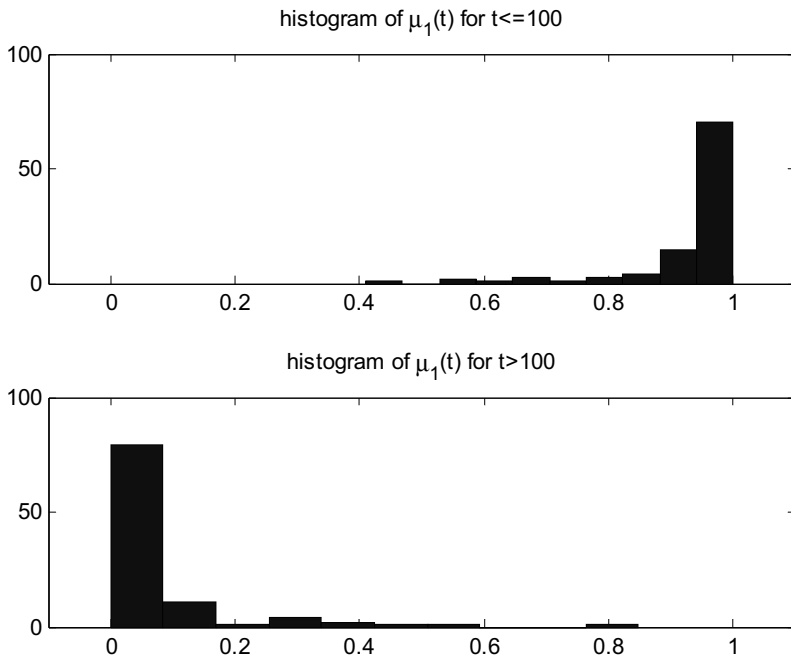
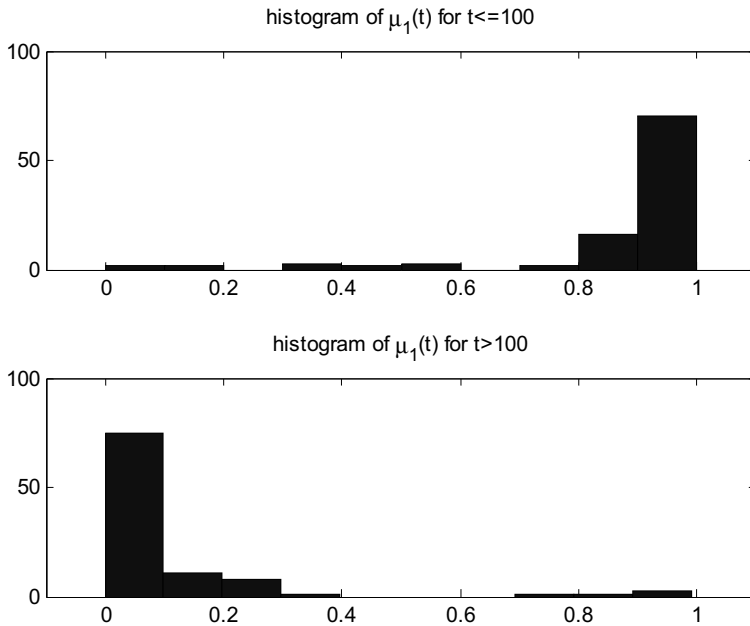**Figure 7** – Histograms of μ_{1t} for $\varepsilon(t) \sim N(0, 1)$.

Figure 8 – Histograms of μ_{1t} for $\varepsilon(t) \sim t(5)$.

3   METROPOLIS-HASTINGS FORMULATION

The goal of the Metropolis-Hastings algorithm (Gamerman, 1997) is to construct a Markov chain that has a specified equilibrium distribution π that in the paper is the posterior distribution of all parameters in the change point model.

Define a Markov chain as follows. If $X_{i-1} = x_{i-1}$, then generate a candidate value $x^* = y$ from a reference distribution with density $f_{X^*|X}(y) = q(x_{i-1}, x^*)$. The candidate value x^* is then accepted or rejected as a sample of π. The acceptance probability is

$$\alpha(x, y) = \min\left(1, \frac{\pi(x^*)}{\pi(x_{i-1})} \frac{q(x^*, x_{i-1})}{q(x_{i-1}, x^*)}\right) \tag{4}$$

which defines the transition kernel of the chain. If the candidate is accepted, then set $X_i = x^*$, otherwise, set $X_i = x_{i-1}$. It is possible to show that under general conditions the sequence $X_0, X_1, X_2, \ldots$ is a Markov chain with equilibrium distribution π.

In practical terms, the Metropolis-Hastings algorithm can be specified by the following steps:

1. Choose a starting value x_0, the number of iterations, R, and set the iteration counter $r = 1$;

2. Generate a candidate value $y \sim q(x_{i,.})$;

3. Calculate the acceptance probability (4) and generate $u \sim U(0, 1)$;

4. Compute the new value for the current state:

$$x^{t+1} = \begin{cases} y \,, & \text{if } \alpha(x, y) > u \\ x^t \,, & \text{otherwise} \end{cases}$$

5. If $r < R$, return to step 2. Otherwise stop.

As discussed in the previous section, the clustering technique generates a transformed time series with the following distribution

$$y(t) \sim beta(a, b) \,, \quad \text{for } t = 1, \dots, m$$
$$y(t) \sim beta(c, d) \,, \quad \text{for } t = m + 1, \dots, n$$

where a, b, c, d and the change point m are the parameters to be estimated. To complete the model specification, *prior distributions* are elicited for the parameters. As usually done whenever there is no prior information available, in this paper the following non-informative prior distributions are assumed:

$$a \sim gamma\,(0.1, 0.1)$$
$$b \sim gamma\,(0.1, 0.1)$$
$$c \sim gamma\,(0.1, 0.1)$$
$$d \sim gamma\,(0.1, 0.1)$$
$$m \sim U[1, \dots, n] \,, \text{ with } p(m) = 1/n \,.$$

These non-informative gamma prior distributions, with parameters 0.1, have been chosen with the purpose of spreading the whole parametric space.

Since the posteriors do not have a close form, the Metropolis-Hastings algorithm is used to sample from them. Its formulation is given by the pseudo-code in the Table 1, and the reference distributions

$$\frac{\pi(a^*)}{\pi(a_{i-1})}\,\frac{q(a^*, a_{i-1})}{q(a_{i-1}, a^*)}\,, \quad \frac{\pi(b^*)}{\pi(b_{i-1})}\,\frac{q(b^*, b_{i-1})}{q(b_{i-1}, b^*)}\,, \quad \frac{\pi(c^*)}{\pi(c_{i-1})}\,\frac{q(c^*, c_{i-1})}{q(c_{i-1}, c^*)}\,,$$

$$\frac{\pi(d^*)}{\pi(d_{i-1})}\,\frac{q(d^*, d_{i-1})}{q(d_{i-1}, d^*)} \quad \text{and} \quad \frac{\pi(m^*)}{\pi(m_{i-1})}\,\frac{q(m^*, m)}{q(m_{i-1}, m^*)}$$

are presented in the Appendix A. The change point m is obtained as the mode of its posterior distribution. It is approximated by the maximum value in the histogram for the R simulations of the pseudo-code in Table 1, with the exception of the border points of the distribution. The Figure 9 illustrates the convergence of parameter m for $p_1 = 1$, $p_2 = 2$, $\varepsilon(t) \sim U(0, 1)$, $R = 100$ and $m = 30$.

Table 2 shows the results for 100 replications of the time series assuming different values of the parameters p_2 and $\varepsilon(t)$. It is clear from Table 2 that the proposed methodology has presented

Table 1 – Pseudo-code of the Metropolis-Hastings algorithm.

Read number of simulations R
$a_0 \leftarrow Gamma(0.1, 0.1)$
$b_0 \leftarrow Gamma(0.1, 0.1)$
$c_0 \leftarrow Gamma(0.1, 0.1)$
$d_0 \leftarrow Gamma(0.1, 0.1)$
$m_0 \leftarrow U[1, \ldots, n]$
For $i = 1$ **to** R **Go**

$a^* \leftarrow Gamma(0.1, 0.1)$

$b^* \leftarrow Gamma(0.1, 0.1)$

$c^* \leftarrow Gamma(0.1, 0.1)$

$d^* \leftarrow Gamma(0.1, 0.1)$

$m^* \leftarrow U[1, \ldots, n]$

If $\min\left(1, \frac{\pi(a^*)}{\pi(a_{i-1})} \frac{q(a^*, a_{i-1})}{q(a_{i-1}, a^*)}\right) > U(0, 1)$

$a_i \leftarrow a^*$

Else

$a_i \leftarrow a_{i-1}$

End_If

If $\min\left(1, \frac{\pi(b^*)}{\pi(b_{i-1})} \frac{q(b^*, b_{i-1})}{q(b_{i-1}, b^*)}\right) > U(0, 1)$

$b_i \leftarrow b^*$

Else

$b_i \leftarrow b_{i-1}$

End_If

If $\min\left(1, \frac{\pi(c^*)}{\pi(c_{i-1})} \frac{q(c^*, c_{i-1})}{q(c_{i-1}, c^*)}\right) > U(0, 1)$

$c_i \leftarrow c^*$

Else

$c_i \leftarrow c_{i-1}$

End_If

If $\min\left(1, \frac{\pi(d^*)}{\pi(d_{i-1})} \frac{q(d^*, d_{i-1})}{q(d_{i-1}, d^*)}\right) > U(0, 1)$

$d_i \leftarrow d^*$

Else

$d_i \leftarrow d_{i-1}$

End_If

If $\min\left(1, \frac{\pi(m^*)}{\pi(m_{i-1})} \frac{q(m^*, m_{i-1})}{q(m_{i-1}, m^*)}\right) > U(0, 1)$

$m_i \leftarrow m^*$

Else

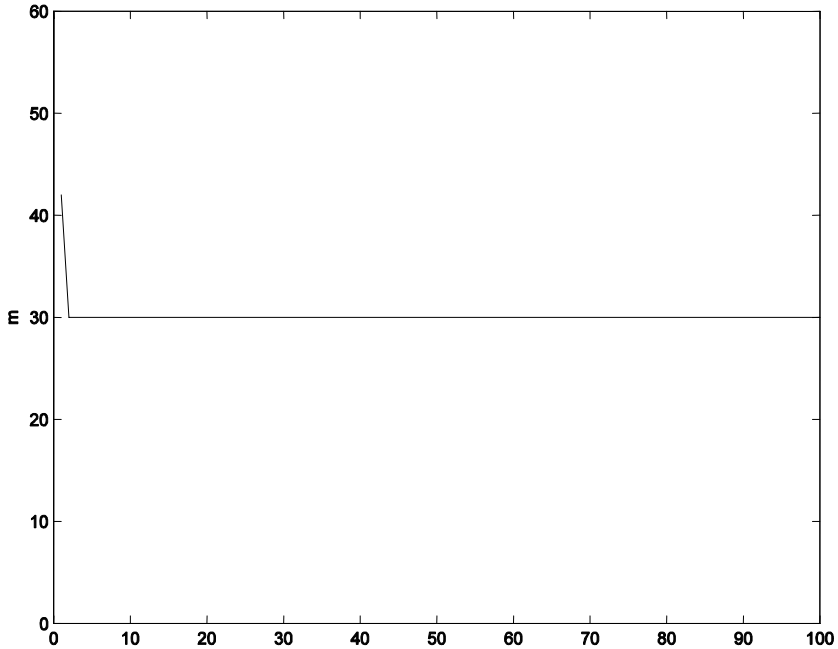$m_i \leftarrow m_{i-1}$

End_If

End_For

Figure 9 – Convergence of parameter m.

100% of correctness in detecting the change point in almost all situations. However, in the case when the time series changes from point $p_1 = 1$ to point $p_2 = 1.1$, the proposed approach has detected points in the neighborhood of $m = 30$, including it. In Table 2, the sign 'XX' indicates that the percentage of correctness was not conclusive. This fact is illustrated in the Figure 10, where a test was performed assuming $p_1 = 1$, $p_2 = 1.1$, $\varepsilon(t) \sim U(0, 1)$, $R = 1000$ and $m = 30$. This behavior is due to the fact that both the noise and the signal change have the same amplitude.

Table 2 – Results of change point detection.

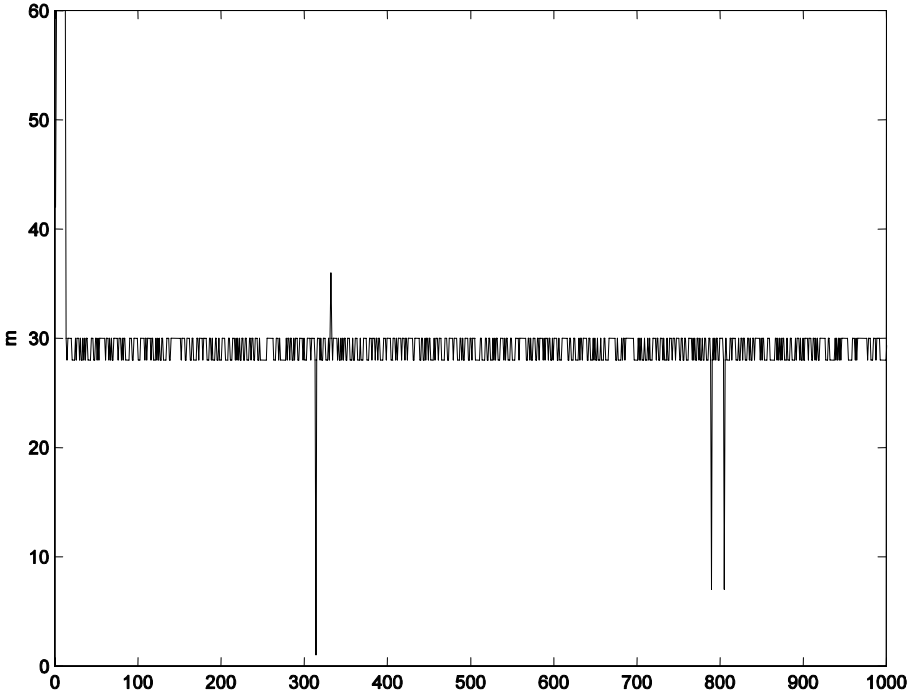| p_1	p_2	m	Noise size	% of correct detection $\varepsilon(t) \sim U(0, 1)$	% of correct detection $\varepsilon(t) \sim N(0, 1)$	% of correct detection $\varepsilon(t) \sim t(5)$
1	2	30	0.1	100	100	100
1	1.9	30	0.1	100	100	100
1	1.8	30	0.1	100	100	100
1	1.7	30	0.1	100	100	100
1	1.6	30	0.1	100	100	100
1	1.5	30	0.1	100	100	100
1	1.4	30	0.1	100	100	100
1	1.3	30	0.1	100	100	100
1	1.2	30	0.1	100	100	100
1	1.1	30	0.1	XX	XX	XX
1	1	30	0.1	100	100	100

Figure 10 – Convergence of parameter m.

4   A CASE STUDY

The energy crisis faced by Brazil in early 2001 triggered a process of energy rationing in the country. This phenomenon was responsible for modifying habits and behaviors taken so far as standards when it came to consumption of electricity in Brazil. Since this fact has been occurred more useful studies to predict the electricity sector make necessary. In this case study the proposed methodology is used to assess the change in the trend at a particular time series that characterizes the energy consumption only in the Southeast Brazil (industry trade) in the period January 1976 to December 2000 (five months before the blackout in 2001), as presented in Figure 11, available at (IPEA).

Since such time series can experienced a change in its trend in the analyzed period, it is important to find this change point before modeling more precisely the entire process as discussed in the introduction section. Before applying the proposed methodology, the series was decomposed into a time series trend and a time series of cycles, (Hodrick & Prescott, 1997). The trend time series is illustrated in Figure 12. One can notice changes in the rate of growth of consumption between points 200 and 250, and points 250 and 300. The process of detection of change points in the time series trend, in Figure 12, was found by the taking its time-derivative and is illustrated in Figure 13. The result of the proposed methodology for the change point detection in Figure 13 is illustrated in Figure 14, which shows the effectiveness of this technique. Notice that the results in Figure 14 have been obtained after 1000 simulations of the pseudo-code in Table 1.
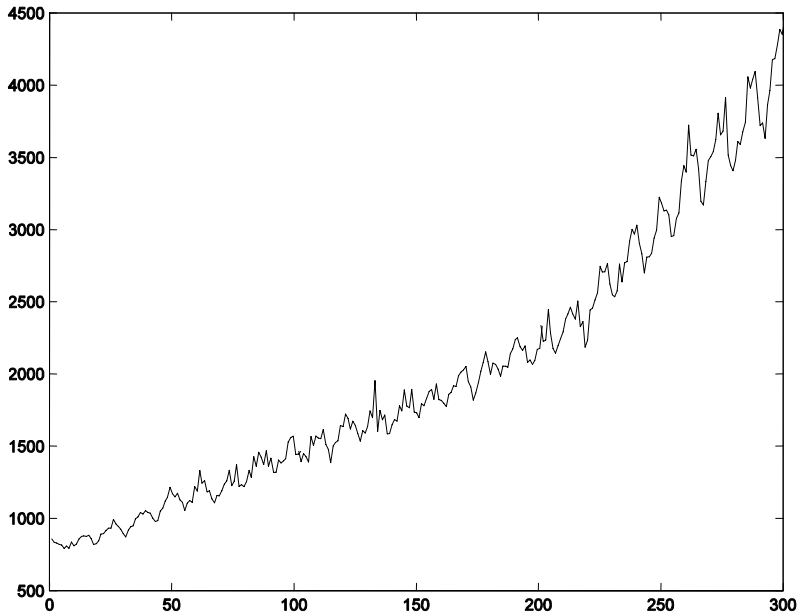
Figure 11 – Time series that characterizes the energy consumption only in the Southeast Brazil (industry trade) in the period January 1976 to December 2000.
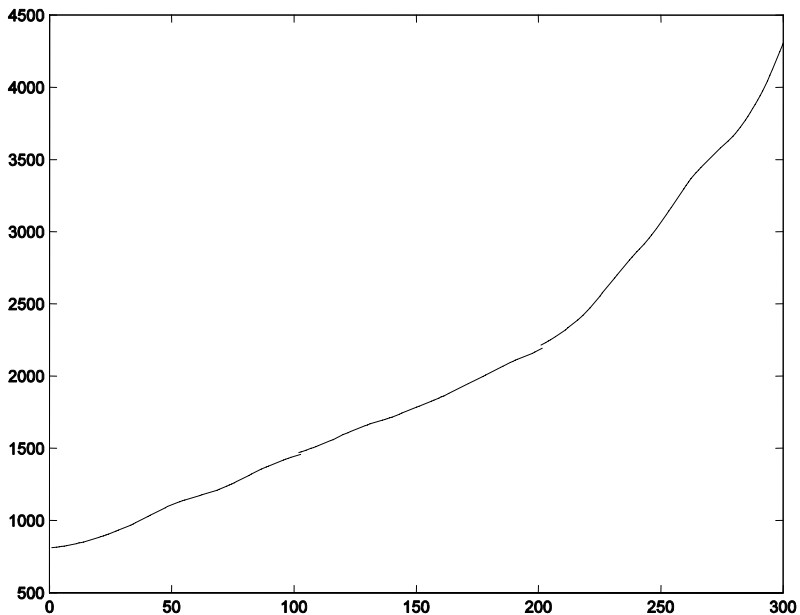

Figure 12 – Time series trend.

Further this methodology cannot be used as "blackout detector", but as an auxiliary tool in the forecasting process in this time series to be used in the planning of the expansion in the energy sector.
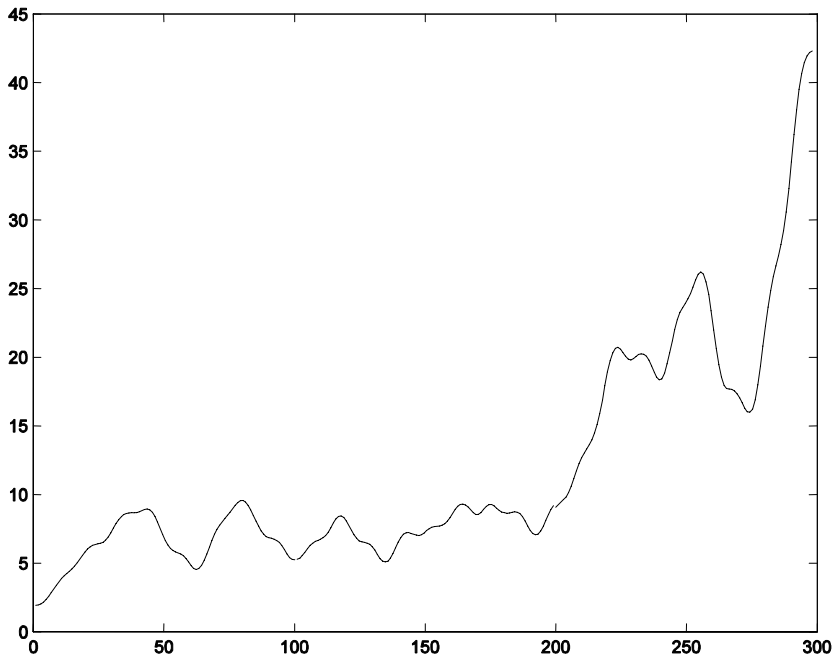
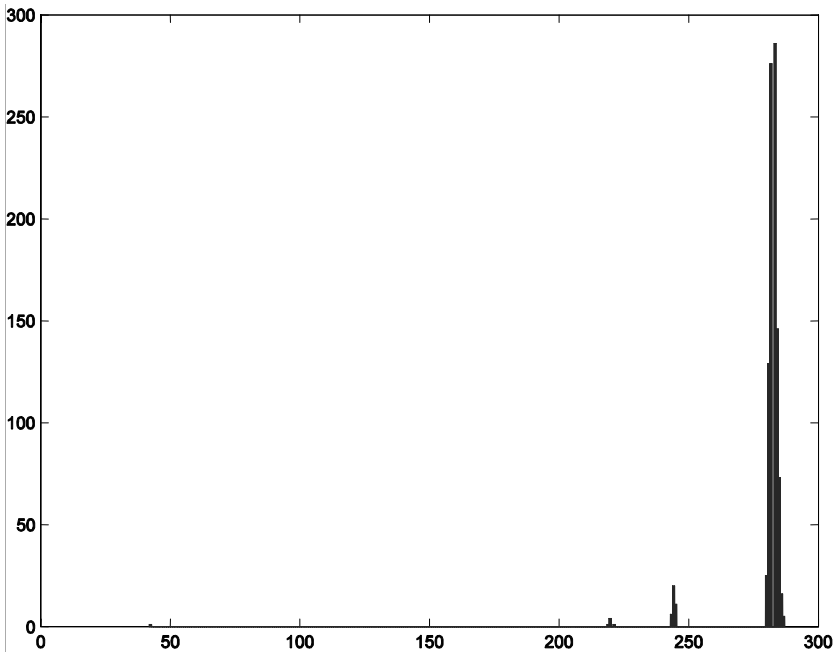**Figure 13** – Time-derivative of the time series trend.

Figure 14 – Response of the proposed methodology.

5   CONCLUSIONS

In this paper a novel fuzzy/Bayesian methodology for change point detection in time series was proposed. The methodology is based on a two step formulation. Firstly, a fuzzy clustering generates a transformed time series with Beta distribution to be used in the second step. In the second step, a Bayesian modeling is used for change point detection in the transformed time series. The Metropolis-Hastings algorithm is used to sample form the posteriors and the position of the change point is thus obtained taking the posterior mode as it estimation.

ACKNOWLEDGEMENTS

REFERENCES

[1]   BARRY D & HARTIGAN JA. 1993. A Bayesian Analysis for Change Point Problems, **88**(421): 309–319.

[2]   BECKAGE B, JOSEPH L, BELISLE P, WOLFSON DB & PLATT WJ. 2007. Bayesian change-point analyses in ecology. *New Phytologist*, **174**: 456–467.

[3]   BEZDEK JC. 1981. *Pattern recognition with fuzzy objective function algorithms*. Plenum Press.

[4]   GAMERMAN D. 1997. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman & Hall.

[5]   HADJILIADIS O & MOUSTAKIDES V. 2006. Optimal and Asymptotically Optimal CUSUM Rules for Change Point Detection in the Brownian Motion Model with Multiple Alternatives. *Theory of Probability and its Applications*, **50**(1): 75–85.

[6]   HARTIGAN JA. 1990. Partition Models, **19**(8): 2745–2756.

[7]   HINKEY DV. 1971. Inference About the Change Point from Cumulative Sum Test, **26**: 279–284.

[8]   HODRICK R & PRESCOTT EC. 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking*, **29**: 1–16.

[9]   IPEA – INSTITUTO DE PESQUISAS ECONÔMICA APLICADA. (n.d.). Retrieved may 10, 2008, from http://www.ipeadata.gov.br.

[10]   KAUFMAN L & ROUSSEEUW PJ. 1990. *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons.

[11]   KOHONEN T. 2001. *Self-organizing maps*. Springer.

[12]   KULLBACK S & LEIBLER RA. 1951. On Information and Sufficiency. *Annals of Mathematical Statistics*, **22**(1): 79–86.

[13]   LEE S, NISHIYAMA Y & YOSHIDA N. 2006. Test for Parameter Change in Diffusion Processes by Cusum Statistics Based on One-step Estimators. *Annals of the Institute of Statistical Mathematics*, **58**(2): 211–222.

[14]   LEE S, PARK S, MAEKAWA K & KAWAI K. 2006. Test for Parameter Change in ARIMA Models. *Communications in Statistics: Simulation and Computation*, **35**(2): 429–439.

[15] LOSCHI RH & CRUZ FRB. 2005. Bayesian Identification of Multiple Change Points in Poisson Data. *Advances in Complex Systems*, **8**: 465–482.

[16] LOSCHI RH & CRUZ FRB. 2005. Extension to the product partition model: computing the probability of a change. *Computational Statistics and Data Analysis*, **48**(2): 255–268.

[17] LOSCHI RH, GONÇALVES FB & CRUZ FBR. 2005. Avaliação de uma medida de evidência de um ponto de mudança e sua utilização na identificação de mudanças na taxa de criminalidade em Belo Horizonte. *Pesquisa Operacional*, **25**(3): 459–463.

[18] OH KJ, ROH TH & MOON MS. 2005. Developing time-based clustering neural networks to use change-point detection: Application to financial time series. *Asia-Pacific Journal of Operational Research*, **22**(1): 51–70.

[19] ZADEH LA. 1965. Fuzzy Sets. *Information and Control*, **8**(3): 338–353.

A   Appendix

In this appendix is presented the acceptance probability for the *posterior* distribution of the parameters a, b, c, d and m described in the Metropolis-Hastings algorithm. The reference distributions used in the work are the *prior* distributions *gamma*(0.1, 0.1).

The reference distribution in Metropolis-Hastings is computed by likelihood function:

$$f(y|a, b, c, d, m) \sim \prod_{i=1}^{m} \frac{\Gamma(a+b)}{\Gamma(a)+\Gamma(b)} y_i^{a-1}(1-y_i)^{b-1} \prod_{i=m+1}^{m} \frac{\Gamma(c+d)}{\Gamma(c)+\Gamma(d)} y_i^{c-1}(1-y_i)^{d-1}$$

1. For the parameter a:

$$\frac{\pi(a^*)}{\pi(a_{i-1})} \frac{q(a^*, a_{i-1})}{q(a_{i-1}, a^*)}$$

$$= \frac{\pi(a^*)}{\pi(a_{i-1})} \frac{f(y|a^*, b^{i-1}, c^{i-1}, d^{i-1}, m^{i-1})\pi(a^*)\pi(b^{i-1})\pi(c^{i-1})\pi(d^{i-1})\pi(m^{i-1})}{f(y|a^{i-1}, b^{i-1}, c^{i-1}, d^{i-1}, m^{i-1})\pi(a^{i-1})\pi(b^{i-1})\pi(c^{i-1})\pi(d^{i-1})\pi(m^{i-1})}$$

$$= \frac{\left[\frac{\Gamma(a^*+b^{i-1})}{\Gamma(a^*)\Gamma(b^{i-1})}\right]^{m^{i-1}} \left[\frac{\Gamma(c^{i-1}+d^{i-1})}{\Gamma(c^{i-1})\Gamma(d)}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}}\left[y_j^{a^*-1}(1-y_j)^{b^{i-1}-1}\right]}{\left[\frac{\Gamma(a^{i-1}+b^{i-1})}{\Gamma(a^{i-1})\Gamma(b^{i-1})}\right]^{m^{i-1}} \left[\frac{\Gamma(c^{i-1}+d^{i-1})}{\Gamma(c^{i-1})\Gamma(d^{i-1})}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}}\left[y_j^{a^{i-1}-1}(1-y_j)^{b^{i-1}-1}\right]}$$

$$\times \frac{\prod_{j=m^{i-1}+1}^{n}\left[y_j^{c^{i-1}-1}(1-y_j)^{d^{i-1}-1}\right]}{\prod_{j=m^{i-1}+1}^{n}\left[y_j^{c^{i-1}-1}(1-y_j)^{d^{i-1}-1}\right]} \frac{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}a^{*0.1-1}e^{-0.1a^*}\right]^2}{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}a^{i-1^{0.1-1}}e^{-0.1a^{i-1}}\right]^2}$$

$$= \frac{\left[\frac{\Gamma(a^*+b^{i-1})}{\Gamma(a^*)}\right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}} y_j^{a^*-1} \left\{\left[\frac{a^{i-1}}{a^*}\right]^{0.9} e^{-0.1(a^*-a^{i-1})}\right\}^2}{\left[\frac{\Gamma(a^{i-1}+b^{i-1})}{\Gamma(a^{i-1})}\right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}} y_j^{a^{i-1}-1}}$$

2. For the parameter b:

$$\frac{\pi(b^*)}{\pi(b_{i-1})} \frac{q(b^*, b_{i-1})}{q(b_{i-1}, b^*)}$$

$$= \frac{\pi(b^*)}{\pi(b_{i-1})} \frac{f(y|a^i, b^*, c^{i-1}, d^{i-1}, m^{i-1})\pi(a^i)\pi(b^*)\pi(c^{i-1})\pi(d^{i-1})\pi(m^{i-1})}{f(y|a^i, b^{i-1}, c^{i-1}, d^{i-1}, m^{i-1})\pi(a^i)\pi(b^{i-1})\pi(c^{i-1})\pi(d^{i-1})\pi(m^{i-1})}$$

$$= \frac{\left[\frac{\Gamma(a^i+b^*)}{\Gamma(a^i)\Gamma(b^*)}\right]^{m^{i-1}} \left[\frac{\Gamma(c^{i-1}+d^{i-1})}{\Gamma(c^{i-1})\Gamma(d)}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}}\left[y_j^{a^i-1}(1-y_j)^{b^*-1}\right]}{\left[\frac{\Gamma(a^i+b^{i-1})}{\Gamma(a^i)\Gamma(b^{i-1})}\right]^{m^{i-1}} \left[\frac{\Gamma(c^{i-1}+d^{i-1})}{\Gamma(c^{i-1})\Gamma(d^{i-1})}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}}\left[y_j^{a^i-1}(1-y_j)^{b^{i-1}-1}\right]}$$

$$\times \frac{\prod_{j=m^{i-1}+1}^{n}\left[y_j^{c^{i-1}-1}(1-y_j)^{d^{i-1}-1}\right]}{\prod_{j=m^{i-1}+1}^{n}\left[y_j^{c^{i-1}-1}(1-y_j)^{d^{i-1}-1}\right]} \frac{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}b^{*0.1-1}e^{-0.1b^*}\right]^2}{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}b^{i-1^{0.1-1}}e^{-0.1b^{i-1}}\right]^2}$$

$$= \frac{\left[\frac{\Gamma(a^i+b^*)}{\Gamma(b^*)}\right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}}(1-y_j)^{b^*-1} \left\{\left[\frac{b^{i-1}}{b^*}\right]^{0.9}e^{-0.1(b^*-b^{i-1})}\right\}^2}{\left[\frac{\Gamma(a^{i-1}+b^{i-1})}{\Gamma(b^{i-1})}\right]^{m^{i-1}} \prod_{j=1}^{m^{i-1}}(1-y_j)^{b^{i-1}-1}}$$

3. For the parameter c:

$$\frac{\pi(c^*)}{\pi(c_{i-1})} \frac{q(c^*, c_{i-1})}{q(c_{i-1}, c^*)}$$

$$= \frac{\pi(c^*)}{\pi(c_{i-1})} \frac{f(y|a^i, b^i, c^*, d^{i-1}, m^{i-1})\pi(a^i)\pi(b^i)\pi(c^*)\pi(d^{i-1})\pi(m^{i-1})}{f(y|a^i, b^i, c^{i-1}, d^{i-1}, m^{i-1})\pi(a^i)\pi(b^i)\pi(c^{i-1})\pi(d^{i-1})\pi(m^{i-1})}$$

$$= \frac{\left[\frac{\Gamma(a^i+b^i)}{\Gamma(a^i)\Gamma(b^i)}\right]^{m^{i-1}} \left[\frac{\Gamma(c^*+d^{i-1})}{\Gamma(c^*)\Gamma(d^{i-1})}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}}\left[y_j^{a^i-1}(1-y_j)^{b^i-1}\right]}{\left[\frac{\Gamma(a^i+b^i)}{\Gamma(a^i)\Gamma(b^i)}\right]^{m^{i-1}} \left[\frac{\Gamma(c^{i-1}+d^{i-1})}{\Gamma(c^{i-1})\Gamma(d^{i-1})}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}}\left[y_j^{a^i-1}(1-y_j)^{b^i-1}\right]}$$

$$\times \frac{\prod_{j=m^{i-1}+1}^{n}\left[y_j^{c^*-1}(1-y_j)^{d^{i-1}-1}\right]}{\prod_{j=m^{i-1}+1}^{n}\left[y_j^{c^{i-1}-1}(1-y_j)^{d^{i-1}-1}\right]} \frac{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}c^{*0.1-1}e^{-0.1c^*}\right]^2}{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}c^{i-1^{0.1-1}}e^{-0.1c^{i-1}}\right]^2}$$

$$= \frac{\left[\frac{\Gamma(c^*+d^{i-1})}{\Gamma(c^*)}\right]^{n-m^{i-1}} \prod_{j=m^{i-1}-1}^{n} y_j^{c^*-1} \left\{\left[\frac{c^{i-1}}{c^*}\right]^{0.9}e^{-0.1(c^*-c^{i-1})}\right\}^2}{\left[\frac{\Gamma(c^{i-1}+d^{i-1})}{\Gamma(d^{i-1})}\right]^{n-m^{i-1}} \prod_{j=m^{i-1}+1}^{n} y_j^{c^{i-1}-1}}$$

4. For the parameter d:

$$\frac{\pi(d^*)}{\pi(d_{i-1})} \frac{q(d^*, d_{i-1})}{q(d_{i-1}, d^*)}$$

$$= \frac{\pi(d^*)}{\pi(d_{i-1})} \frac{f(y|a^i, b^i, c^i, d^*, m^{i-1})\pi(a^i)\pi(b^i)\pi(c^i)\pi(d^*)\pi(m^{i-1})}{f(y|a^i, b^i, c^i, d^{i-1}, m^{i-1})\pi(a^i)\pi(b^i)\pi(c^i)\pi(d^{i-1})\pi(m^{i-1})}$$

$$= \frac{\left[\frac{\Gamma(a^i+b^i)}{\Gamma(a^i)\Gamma(b^i)}\right]^{m^{i-1}} \left[\frac{\Gamma(c^i+d^*)}{\Gamma(c^i)\Gamma(d^*)}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}} \left[y_j^{a^i-1}(1-y_j)^{b^i-1}\right]}{\left[\frac{\Gamma(a^i+b^i)}{\Gamma(a^i)\Gamma(b^i)}\right]^{m^{i-1}} \left[\frac{\Gamma(c^i+d^{i-1})}{\Gamma(c^i)\Gamma(d^{i-1})}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}} \left[y_j^{a^i-1}(1-y_j)^{b^i-1}\right]}$$

$$\times \frac{\prod_{j=m^{i-1}+1}^{n} \left[y_j^{c^i-1}(1-y_j)^{d^*-1}\right]}{\prod_{j=m^{i-1}+1}^{n} \left[y_j^{c^i-1}(1-y_j)^{d^{i-1}-1}\right]} \frac{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}d^{*0.1-1}e^{-0.1d^*}\right]^2}{\left[0.1^{0.1}[\Gamma(0.1)]^{-1}d^{i-1^{0.1-1}}e^{-0.1d^{i-1}}\right]^2}$$

$$= \frac{\left[\frac{\Gamma(c^*+d^{i-1})}{\Gamma(c^*)}\right]^{n-m^{i-1}} \prod_{j=m^{i-1}-1}^{n} \left[(1-y_j)^{d^*-1}\right]\left\{\left[\frac{d^{i-1}}{d^*}\right]^{0.9} e^{-0.1(d^*-d^{i-1})}\right\}^2}{\left[\frac{\Gamma(c^{i-1}+d^{i-1})}{\Gamma(d^{i-1})}\right]^{n-m^{i-1}} \prod_{j=m^{i-1}+1}^{n} \left[(1-y_j)^{d^{i-1}-1}\right]}$$

5. For the parameter m:

$$\frac{\pi(m^*)}{\pi(m_{i-1})} \frac{q(m^*, m)}{q(m_{i-1}, m^*)}$$

$$= \frac{\pi(m^*)}{\pi(m_{i-1})} \frac{f(y|a^i, b^i, c^i, d^i, m^*)\pi(a^i)\pi(b^i)\pi(c^i)\pi(d^i)\pi(m^*)}{f(y|a^i, b^i, c^i, d^i, m^{i-1})\pi(a^i)\pi(b^i)\pi(c^i)\pi(d^i)\pi(m^{i-1})}$$

Sendo $\pi(m^{(\cdot)}) \sim U(0, 1)$, então:

$$\frac{\pi(m^*)}{\pi(m_{i-1})} \frac{q(m^*, m)}{q(m_{i-1}, m^*)}$$

$$= \frac{f(y|a^i, b^i, c^i, d^i, m^*)}{f(y|a^i, b^i, c^i, d^i, m^{i-1})}$$

$$= \frac{\left[\frac{\Gamma(a^i+b^i)}{\Gamma(a^i)\Gamma(b^i)}\right]^{m^*} \left[\frac{\Gamma(c^i+d^i)}{\Gamma(c^i)\Gamma(d^i)}\right]^{n-m^*} \prod_{j=1}^{m^*} \left[y_j^{a^i-1}(1-y_j)^{b^i-1}\right]\prod_{j=m^*+1}^{n} \left[y_j^{c^i-1}(1-y_j)^{d^i-1}\right]}{\left[\frac{\Gamma(a^i+b^i)}{\Gamma(a^i)\Gamma(b^i)}\right]^{m^{i-1}} \left[\frac{\Gamma(c^i+d^i)}{\Gamma(c^i)\Gamma(d^i)}\right]^{n-m^{i-1}} \prod_{j=1}^{m^{i-1}} \left[y_j^{a^i-1}(1-y_j)^{b^i-1}\right]\prod_{j=m^{i-1}+1}^{n} \left[y_j^{c^i-1}(1-y_j)^{d^i-1}\right]}$$