



NEURO/FUZZY/BAYESIAN APPROACH FOR BEHAVIOR CHANGE DETECTION IN TIME SERIES RELATED TO DAM SAFETY MANAGEMENT

¹FERNANDO MUCIO BANDO, ²JAIR MENDES MARQUES,
³JOSIELE PATIAS, ⁴LUIZ ANTÔNIO TEIXEIRA JUNIOR

¹ Asstt. Prof., Eng. and Exact Sciences Center, State University of West Paraná, Foz do Iguaçu, Brazil

² Assoc. Prof., Department of Mathematic and Statistic, Federal University of Paraná, Curitiba-PR, Brazil

³ Civil Engineering and Architecture Division, Itaipu, Foz do Iguaçu, Brazil

⁴ Assoc. Prof., ILATIT, Federal University of Latin America Integration, Foz do Iguaçu-PR, Brazil

E-mail: ¹ fernandobando@gmail.com, ² jair.marques@utp.br, ³ jpatias@itaipu.br,

⁴ l Luiz.a.t.junior@gmail.com.

ABSTRACT

This article is aimed to present the efficiency of neuro/fuzzy/Bayesian approach in the detection of automatic behavior change in time series generated from monitoring instruments used in detection security management of a dam. The method for detecting change points in a time series is divided into three steps: clustering of the time series data using a self-organizing map of Kohonen; construction of a fuzzy set to transform the original time series with arbitrary distribution in a new series which distribution probability may be approximated by a beta distribution; applying a simulation of Monte Carlo via Markov chains for determining the switching point. The identified points are classified as levels of alerts that assist in a dam safety management. The efficiency of the method is validated through applying in a time series used in the dam safety management of Itaipu hydroelectric power plant.

Keywords: Dam Safety. Artificial Neural Network. Sets Fuzzy. Monte Carlo Simulation.

1. INTRODUCTION

The construction of a hydroelectric power plant is greatly important for the development of a country and shows the enormous engineering capability in which high qualified professionals are involved. However, this type of building requires extensive watching because the possibility of an unexpected behavior may result in undesirable consequences concerning economy, environment and safety of people [1]. The safety of Dams is a large discussed issue all over the world and Brazil shows great concern with the health of its dams.

On September 20 the law n°12.334 [2] came into effect, which establishes the National Policy of Safety of dams intended to be loaded with water for any use, temporary or permanent disposal of tailings and the accumulation of industry discards, in this context it was also created the National System of Dam Safety Information (NSDSI). In the article 2 item III of this law the dam safety is defined as a set of conditions aimed to preserve its

structural and operational integrity besides the preservation of life, health and environment.

Itaipu power plant is free from the obligations of this law due to its binationality, yet it adopts procedures internationally accepted and it was one of the supporters of the law n°. 12334 [3]. In this regard Itaipu currently has about 2792 installed instruments and 10 percent of them are equipped with an automatic system of data acquisition which aims to monitor the dam structural behavior through measurements related to stress, deformation, misplacement, pressure, water infiltration, downstream and upstream river levels, temperature and rainfall. These instruments enable the obtain scans at a controllable frequency from 5 to 30 minutes. The information is processed, stored and sent to a central station which is responsible for processing, analyzing and comparing data and trigger alarms [4]. Since the amount of data stored on a daily basis is very large it is needed to use mathematical methods for processing and analyzing of data, which enables the construction of models that describe the occurrences measured



by instruments and therefore, identify and predict possible behavior changes so to efficiently manage the dam safety.

Diagnosing a possible dam structural pattern change often occurs through identifying the behavior change of time series data collected by monitoring instruments. Many techniques for detecting changing points on time series are now presented in scientific articles [5], [6] and [7]. Monitoring systems used to detect failures are called Fault Detection and Isolation (FDI), such systems are meant to recognize abnormal behavior of components in a process. However the FDI techniques depend in general on previous knowledge of the behavior of time series such as deterministic and statistical models. Amongst these techniques it is possible to highlight Luenberger observer; Kalman filter; artificial neural networks; and neuro-fuzzy networks.

In some time series it is not possible to obtain the necessary previous knowledge to apply the mentioned techniques. A present study which seeks to avoid this situation is presented in [8] in which a neuro/fuzzy/bayesian formulation is suggested to identify a changing point in time series without having a previous knowledge of such time series besides enabling the detection of incipient changes, in other words changes which gradually affect the usual operation of a structure [9] which are difficult to detect through other approaches that only uncover changes with large irregularities in usual operation conditions, which are termed as abrupt changes.

Therefore, this article presents a variation of the neuro/fuzzy/bayesian formulation for detecting changing points in time series use on identifying data behavioral change collected by monitoring instruments of a dam, creating as a final result an auxiliary method for dam safety management.

The present work is organized in five sections Section two presents the importance of Dam safety study via instrumented monitoring. The neuro/fuzzy/bayesian method is explained in section three. The data processing as well as its main results is presented in section four and at last the research conclusion is given.

2. DAM SAFETY

Dam safety concept comprises many aspects such as structural, hydraulic, geotechnical and operational. Such aspects must be considered both

in the project and during the whole operation lifetime of a dam. A way of evaluating the behavior and integrity of a dam is through an instrument system able to monitor its structural and geotechnical condition [10].

According to [10] the main target of an apparatus plan is to ensure a proper safety level for a dam. Monitoring through instruments enables the collect of quantitative data which allows to obtain extremely important information such piezometric pressure, strain, stress and water level. However, a proper analysis of the obtained data is needed to identify a critical condition. An instrument system combined with data analysis techniques allows us to track the dam safety level through its operation lifetime and also checking if the premises set out in the project remain under control.

NBR 8681 – Actions and safety of structures [11] sets out the necessary verification related to stability analysis of a dam made of concrete as to evaluate movement safety such as sliding, tumbling down, floating, pressure at the base of foundation and structure, deformation and vibrations.

Physically the movements mentioned to safety monitor are explained due to the fact that the difference in water level creates a hydraulic gradient between upstream and downstream of a dam forcing the water from the reservoir go downstream so that the water level is balanced. For such, the water percolates mainly through the dam base solid. In the course of this process the water which penetrated the solid provoke upwards force under the dam; such force is called base suppression. In addition, the water of the reservoir also causes horizontal force which works from upstream to downstream over the dam, which is called hydrostatic pressure against the dam wall. These two main types of occurrence are considered destabilizing forces which combined may cause the tumbling or sliding of the dam wall both by directly efforts and by the relief of structure weight [12].

Loading conditions and material quality can often change in the course of a dam operation, hence the set of installed instruments in a dam is an important identifying tool of such changes.

A way of identifying possible behavior changes can be done through visual analysis of time series generated by monitor instruments spread throughout the structures and base according to project criteria. Nonetheless, a high quantity of

instruments, considering large dams increases remarkably the time of analysis to be done both manually and in in scale, this way a method development which seeks to automatically detect time series that demonstrate behavior change after any occurrence of incidents outside the structure is greatly important because the identification of sites which demand attention is quickly done, which helps the dam structural safety management.

3. METHODOLOGY

The detection of changes in time series is a very important method which is frequently used on safety of systems that have monitor devices and aim to maintain productivity standards, policies of correcting actions and safety management [13]. Change point analysis in a time series can be divided into tasks such as detection of change, pointing the moment of occurrence; classification of change, determining the necessary level of attention for the treatment of such change and identifying the cause of behavior change in the time series.

There are many strategies to treat problems of detection of changes in time series but most of the techniques presented need a previous knowledge of the behavior of data that describe a time series.

A technique which does not use the previous knowledge of the time series has been described in [8], [9] and [14] works. This methodology is based on the theory of fuzzy sets associated with the bayesian statistic and has as main contribution an approach which allows the detection of change points in a time series without the need of knowing previously the deterministic or probability models that are to describe the collection of data.

The method is divided into three steps (Figure 1):

1. Clustering of time series data through a self-organizing map of Kohonen;
2. Construction of a new time series by fuzzing the initial time series;
3. Use of Monte Carlo simulation via Metropolis-Hastings algorithm for determining the change point.

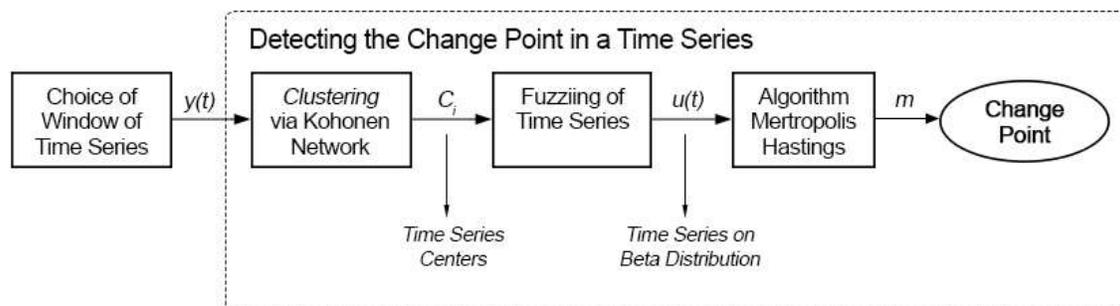


Figure 1: Steps of methodology for detection of change point of time series.

3.1. Clustering via Kohonen Network

The chosen method for clustering of data is based on a self-organized neural network of Kohonen [15]. The main reason for the choice of clustering of data via Kohonen network is because this method allows flexible setting, thus determining besides centers, a proper quantity of sets ignoring unnecessary sets.

As this study is to focus only in time series of a unique dimension only one door for the network is used. The number of neurons is defined in the beginning of the process and decreases during of iterations through a low performance elimination

criteria defined by the number of associations with elements. Also early in the process the maximum number of iterations as well as the learning rate is defined for the first stage of the methodology. The impact of neuro networks are equally spread between the maximum and minimum values of the time series. In the end of the process the neuro networks which remain are defined as being clusters centers.

Neuro network training is neither supervised nor competitive [16]. Only the value of the neuron which wins is adjusted, therefore the neighboring radius is always considered equal 1. The learning rate is updated every iteration and varies from 0,1



to 0,01 in a linear basis. In the end of the training the neurons which had a low performance, or rather, won few times are not considered. This process is repeated until the remaining neurons have a considerable performance. After processing all iterations the neurons which remain are defined as clusters centers for the next step.

3.2. Fuzzifying of Time Series

The theory of fuzzy set has been applied in this step as a method of clustering. The classic clustering methods set apart data in k categories, however in many cases some elements cannot belong to a specific category because they belong to two or more categories simultaneously. The use of clustering fuzzy methods is a good alternative to solve such problem since an element can belong to more than one category simultaneously [17].

The fuzzy clustering method had been used in this study to generate a new time series based on clusters centers defined by the Kohonen network. D’Angelo et al. [18] have shown in empirical way that even without any previous knowledge of the original time series, the new series generated through fuzzy clustering can be properly approximated to a series on beta distribution.

This way, given a time series $y(t)$ and, considering a integer positive k, the fuzzy cluster consists of defining a set $C = \{C_i; \min\{y(t)\} \leq C_i \leq \max\{y(t)\}, i = 1, 2, \dots, k\}$, which solved the problem of minimizing

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

where,

$$\mu_i(t) = \left[\frac{\sum_{j=1}^k \|y(t) - C_j\|^2}{\sum_{j=1}^k \|y(t) - C_j\|^2} \right]^{-1}, \tag{2}$$

Is the degree of fuzzy pertinence of $y(t)$ in relation to each center C_i .

The set $C = \{C_i, i = 1, 2, \dots, k\}$, which minimizes the equation (1), is called center set of time series $y(t)$ and in this article it was defined through clustering via Kohonen network already presented.

Since this works seeks to find only one change point at a time [9], just two centers are found every application of the method and therefore two

functions of pertinence are defined, $\mu_1(t)$ e $\mu_2(t)$. In the next step of the method it is needed only one of the pertinence functions, in this article, it is used the function $\mu(t) = \mu_1(t)$.

Through statistical testes it come to the conclusion that the function $\mu(t)$ can be approximated to a beta distribution function with diferente parameters of entrance or rather, a beta distribution (a, b) to $t \leq m$ and a beta distribution (c, d) to $t > m$. in which the point m is the time series change point.

3.3. Metropolis-Hastings Algorithm

Once the previous steps transformed the original series, with a distribution of any probability, on a new time series $\mu(t)$ with a function of beta distribution probability, so a new statistic set model is considered in a Bayesian formulation in order to estimate the beta distribution parameters which approximates to a new time series and therefore it is possible to estimate the parameter m which is the detection of a possible change point [8]. In this step the Metropolis-Hastings algorithm has been used to execute the Monte Carlo simulation via Markov chains with the aim of estimate parameters.

The definition of a Markov chains is given as: Given a random vector $\theta_i = (a_i, b_i, c_i, d_i, m_i)$, it is chosen a candidate value $\theta = (a, b, c, d, m)$ of a distribution with density $f_{\theta|\theta_i}(\mu) = q(\theta; \theta)$. The function q is known as chain transition nucleus. It is a function that depends on two variables the current estate of the chain θ_i and the candidate value θ .

The candidate value θ is accepted or rejected depending on the probability value of acceptance given by

$$\alpha(\theta_i, \theta) = \min \left(1, \frac{\pi(\theta)q(\theta; \theta_i)}{\pi(\theta_i)q(\theta_i; \theta)} \right). \tag{3}$$

If the candidate value is accepted so $\theta_{i+1} = \theta$, on the contrary $\theta_{i+1} = \theta_i$. This way, if the candidate value is rejected, the Markov chain has a sequence reiteration. Therefore, the sequence $\theta_0, \theta_1, \theta_2, \dots$ makes a Markov chain with balanced distribution π .

In practical terms the Metropolis-Hastings algorithm is specified on the fact that in the previous steps the processed time series follows the



given distribution: $\mu(t) \sim \text{beta}(a, b)$ to $t = 1, \dots, m$ and $\mu(t) \sim \text{beta}(c, d)$ to $t = m + 1, \dots, n$.

The parameters estimated by the algorithm are a, b, c, d and the change point m . In this case the choice of initial values is made using uninformative distribution, on our study: $a, b, c, d \sim U(0,1)$ e $m \sim U\{1, 2, \dots, n\}$.

For the balance function π it is used a density function of distribution $\text{Gama}(0,1,0,1)$ to calculate parameters a, b, c e d :

$$\pi(x) = \frac{0,1^{0,1} x^{0,1-1} e^{-0,1x}}{\Gamma(0,1)}, \quad (4)$$

and $U\{1, 2, \dots, n\}$ to calculate parameter m :

$$\pi(m) = 1/n. \quad (5)$$

The likelihood function of μ in relation to parameters a, b, c, d and m is given by

$$f(\mu|a, b, c, d, m) = \prod_{i=1}^m \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu(i)^{a-1} (1-\mu(i))^{b-1} \prod_{i=m+1}^n \frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} \mu(i)^{c-1} (1-\mu(i))^{d-1}, \quad (6)$$

be $\Gamma(x)$ the gama function.

Finally the estimated value of parameter m is regarded as being the changing point of a time series.

For better control the results the proposed methodology is applied to a series subsequence in size n . And to determine all change points or the subsequence the method generalization for detecting multiple change points in a time series is applied [14] hence, the change points are found and stored. Afterwards, the subsequence is slid through the whole time series in order to repeat the process.

4. APPLICATION: ITAIPU HYDROELECTRIC POWER PLANT

Itaipu is regarded as one of the greatest hydroelectric projects in the world. It is the result of the efforts and commitment of two neighbor countries Brazil and Paraguay. The dam is located in the Parana river where the countries are

bordered, 14 kilometers upstream from the international bridge which connects the city of Foz do Iguacu, in Brazil to Ciudad del Este, in Paraguay [19].

The Itaipu dam is a structure (concrete, rockfill and soil) which serves to dam the water and get a 120 meters gap (nominal steep fall) which allows the operation of turbines. On the top of the main dam there are inlets, structures with gates that allow the water go through them and reach the pipelines and reach the spiral structures that make the turbine spin.

The construction of Itaipu has been of great importance for the development of the country and showed the enormous Brazilian engineering capability. A far advanced work for the time which involved high quality professionals in its building. Itaipu dam is 7.919 meters length and 196 meters high. 12.3 million cubic meters of concrete have been used in its construction, such dimensions turned the dam into a concrete study reference and dam safety [4].

Due to its giant size, Itaipu hydroelectric power plant has kept since its project, constant attention regarding the structural safety of its dam. During its construction and later after its operation, a large number of instruments have been installed both in its structure and also on its foundation in order to monitor various parameters such underpressure on foundation and structure, seepage flow through the foundation, deformation, structural alteration, temperature, piezometric level amongst others.

Itaipu currently has about 2792 installed instruments and 10 percent of them are equipped with an automatic system of data acquisition which aims to monitor the dam structural behavior through measurements related to stress, deformation, misplacement, pressure, water infiltration, downstream and upstream river levels, temperature and rainfall. These instruments enable the obtain scans at a controllable frequency from 5 to 30 minutes. The information is processed, stored and sent to a central station which is responsible for processing, analyzing and comparing data and trigger alarms [4].

One of the targets of this work is to apply the proposed methodology in a time series using the dam safety management. The time series selected for analysis (Figure 2) was generated from about 1.900 weekly measurements, performed from 1980

to 2014, from a piezometer installed in the base of key block D38 of the Itaipu dam. This instrument is responsible for generating structural stability monitor data in terms of sliding, tumbling or floating directly affected by the level of piezometric pressure on the concrete–rock interaction. For a better understanding, the block

D38 is one of the blocks of foothills out of a set of 64 blocks which are located on the right margin between the spillway and the main dam of relieved gravity in which all blocks are identical in structure and profile and are 17 meters length in the shaft, thus with different height. [19].

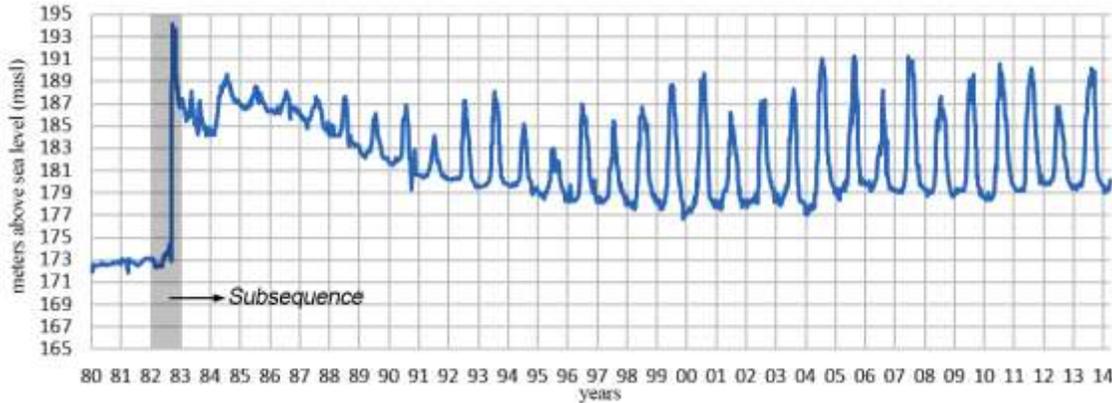


Figure 2: Time serie - Piezometer

5. RESULTS AND DISCUSSIONS

In order to detect change points of this time series, an algorithm has been developed in the software MATLAB R2011 based on the presented methodology, and applied to windows of time series formed by subsequences of 30 elements, one at a time. Finally, these windows are slid along the time series as to identify possible change points.

The subsequence, which was selected to illustrate the results, has 30 points between measurements on position 70 to 100, embracing the measure dates of the second semester in 1982, when the reservoir of Itaipu was filled with water which cause a remarkable increase in piezometric pressure levels in the concrete–rock interface. This behavior change is noticeably visible in the time series generated by the piezometer under analysis.

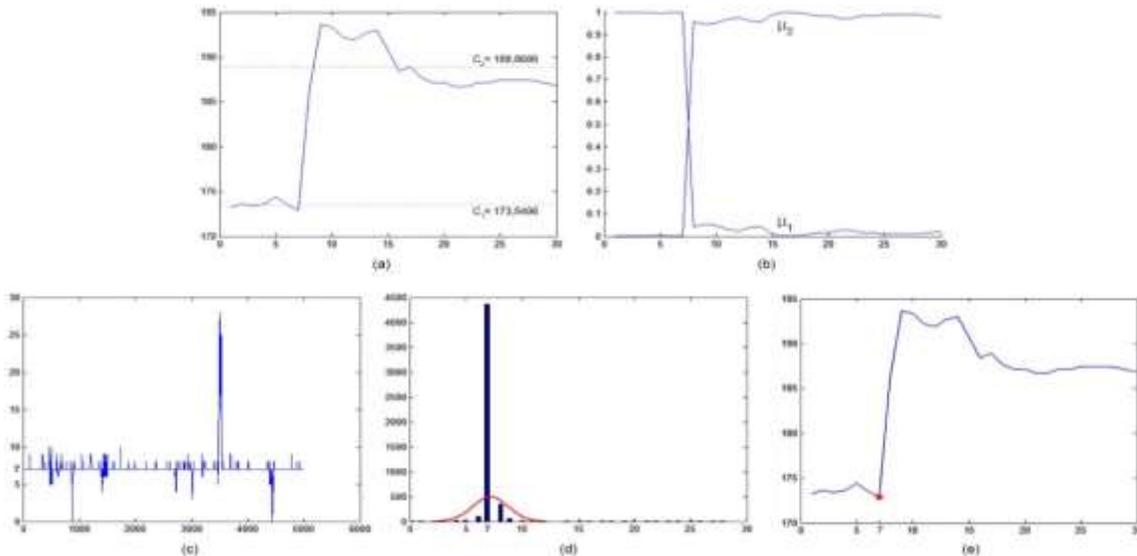


Figure 3: Results of methodology of change point detection applied to subsequence.

With the aim of showing that the methodology exposed in the article is potentially able to automatically detect the change points in a time series without the need of any previous statistic knowledge of the series under analysis, the method has been used on the chosen subsequence.

Initially the algorithm generated, from a complete time series, the subsequence of interest, in this case 30 points which vary between measurements 70 and 100 (Figure 3a) and thus in the new formed series, the position 1 represents the position 71 from the original series, the position 2 represents the position 72 and so on, as such the point found will be added to the value 70 from the original series to determine in which measurement it is possible to find the change point in the original time series. Afterwards a clustering via Kohonen network of the selected subsequence is made which allows us to obtain the centers $C1 = 173.5496$ e $C2 = 188.8686$ (Figure 3a). Once the centers are defined the fuzzy process is applied to the original series through the pertinence function defined in Equation (2). This step generates two new time

series μ_1 e μ_2 which can be approximated by a beta distribution (Figure 3b).

Once having the previous knowledge of probability distribution of the transformed series μ_1 , it is possible to develop a Monte Carlo simulation via Metropolis-Hastings algorithm in order to estimate the change point m . Figures 3c and 4d show results generated by 5000 simulations revealing that the change point likely occurs in the position $n.7$ (or position 77 of the complete time series). Therefore, it is possible to see in the Figure 3e that the position 7, automatically detected by the presented methodology, correspond perfectly to the intuitive knowledge about what is one of the behavior changes in values generated through measurements in some physical process.

At last the methodology has been applied to all sliding windows which, going along the complete time series, enables us to observe all change points of a time series as shown in Figure 4.

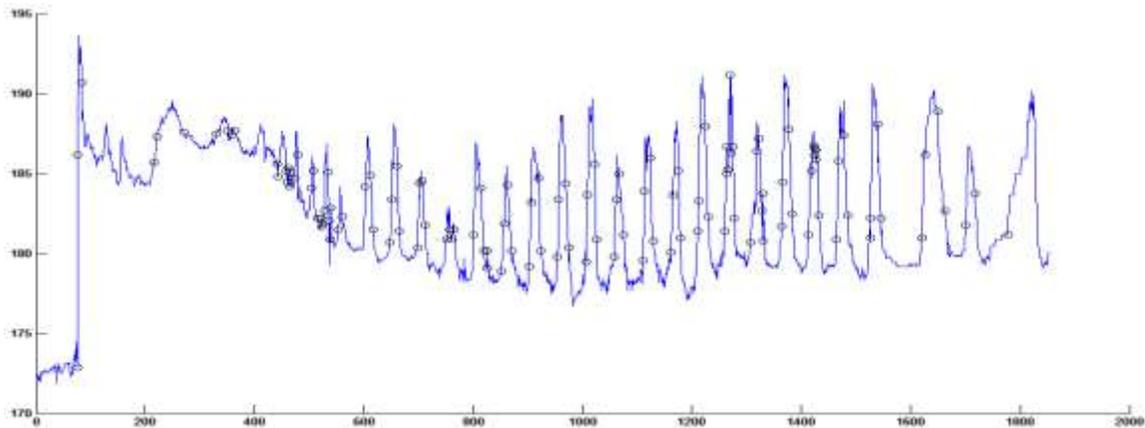


Figure 4: Change points of time series detected through methodology.

5. CONCLUSION

The safety management of a dam embraces many areas and one of them is the constant monitor of the involved structures, which most of the times is given by a high quantity of installed instruments in strategic places, hence a gigantic amount of data for analysis is generated. Therefore, analytical techniques which help the reading of data, are large interest to dam safety management.

With this in mind, it is possible to affirm that the methodology used in change point detection in a

time series presented in this article, is a relevant contribution to help dam safety management. Implementing it may accelerate the process of identifying past occurrences, giving knowledge and experience on future other cases.

Yet a lot can be enhanced, leading to different study possibilities. An intriguing theme for future studies is the classification of change points found through the Fuzzy theory and analysis of probability distribution of change points, enabling the analysis of a line of flows that describe the



relation of each instrument to possible block behavior change.

ACKNOWLEDGMENTS

The authors thank for the contributions of the following institutions:

- Ceasb - Itaipu Binacional: technical support and provision of data;

- Fundação Araucária: financial viability of this study;

- PPGMNE - UFPR: theoretical support and encourage the development of this work.

REFERENCES

- [1] R. N. C. dos Santos, *Enquadramento das Análises de Riscos em Geotecnia*. p 257. Dissertação (Mestrado em Geotecnia para Engenharia Civil) - Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Lisboa, 2006, p. 281.
- [2] Brasil, *Lei no. 12.334, de 20 de setembro 2010*.
- [3] C. Neumann Jr, C. Osako, J. Patias, e C. Porchetto, *Reflexos da Lei de Segurança de Barragens na Itaipu Binacional*. Rio de Janeiro - RJ: XXVIII Seminário Nacional de Grandes Barragens, p. 1-11, 2011.
- [4] Itaipu, *Usina Hidrelétrica de Itaipu: Aspectos Técnicos das Estruturas*. Foz do Iguaçu - PR, 2008.
- [5] D. V. Hinkey, "Inference About the Change From Cumulative Sum Test", *Biometria*, vol. 26, p. 279–284, 1971.
- [6] D. Barry e J. A. Hartigan, "A Bayesian Analysis for Chance Point Problems", *J. Am. Stat. Association*, vol. 88(421), p. 309–319, 1993.
- [7] R. H. Loschi e F. R. B. Cruz, "Extension to the Product Partition Model: Computing the Probability of a Change", *Comput. Stat. Data Anal.*, vol. 42(2), p. 255–268, 2005.
- [8] M. F. S. V. D'Angelo, R. M. Palhares, R. H. C. Takahashi, e R. H. Loschi, *Fuzzy/Bayesian change point detection approach to incipient fault detection*, vol. 5. IET Control Theory and Applications, v. 5, p. 539, 2011.
- [9] R. Souza, R. M. Palhares, M. F. S. V. D'Angelo, e W. Caminhas, "Atualização de Distribuições de Probabilidade e Taxas de Confiabilidade Utilizando Técnicas de Detecção de Falhas", in *XIX Congresso Brasileiro de Automática, CBA*, 2012, p. 3164–3169.
- [10] J. F. A. Silveira, *Instrumentação e Comportamento de Fundações de Barragens de Concreto*. São Paulo - SP: Oficina de Textos, 2003.
- [11] ABNT - Associação Brasileira de Normas Técnicas, *Ações e Seguranças nas Estruturas (NBR 8681)*. 2003.
- [12] R. Villwock, *Técnicas de Agrupamento e de Hierarquização no Contexto de KDD - Aplicação a Dados Temporais de Instrumentação Geotécnica - Estrutura da Usina Hidrelétrica de Itaipu*. p 124. Tese(Doutorado em Ciências) Programa de Pós-Graduação em Métodos Numéricos em Engenharia, UFPR, Curitiba - PR, 2009, p. 125.
- [13] R. Isermann e P. Balle, "Trends in the Application of Model-Based Fault Detection and Diagnosis of Technical Processes", *Control Eng. Pract.*, vol. 5 (5), p. 707–719, 1997.
- [14] M. F. S. V D'Angelo, R. M. Palhares, L. B. Cosme, L. a. Aguiar, F. S. Fonseca, e W. M. Caminhas, "Fault detection in dynamic systems by a Fuzzy/Bayesian network formulation", *Appl. Soft Comput. J.*, vol. 21, p. 647–653, 2014.



-
- [15] T. Kohonen, *The self-organizing map*, vol. 78, n° 9. Proceedings of the IEEE, v. 78, n. 9, p. 1464-1480, 1990.
- [16] T. Kohonen, *Self-organized formation of topologically correct feature maps*, vol. 43. Biological Cybernetics, v. 43, p. 59-69, 1982.
- [17] W. Pedrycz, "Fuzzy Clustering with Partial Supervision", *Pattern Recognition Lett.*, vol. 3 (1), p. 13–20, 1985.
- [18] M. F. S. V. D'Angelo, R. M. Palhares, R. H. C. Takahashi, R. H. Loschi, L. M. R. Baccharini, e W. M. Caminhas, "Incipient fault detection in induction machine stator-winding using a fuzzy-bayesian change point detection approach", *Appl. Soft Comput.*, vol. 11, p. 179–192, 2011.
- [19] Itaipu: Usina Hidrelétrica, *Projeto: Aspectos de Engenharia*. Foz do Iguaçu - PR, 2009.