Journal of Physics:
Conference Series

PURPOSE-LED PUBLISHING™

PAPER • **OPEN ACCESS**

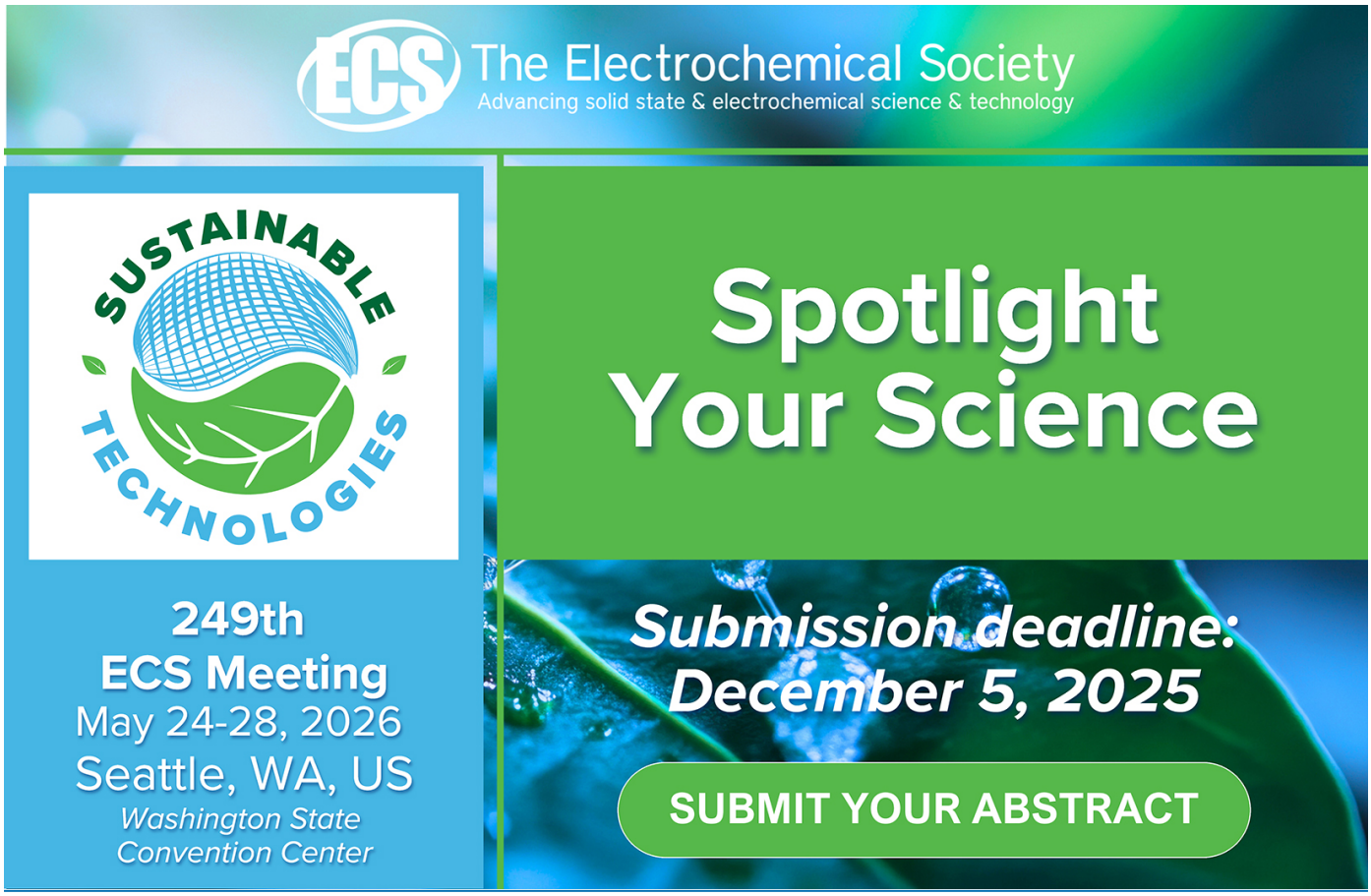
Change-point analysis of the demand for medicines

To cite this article: J C Valderrama Balaguera *et al* 2020 *J. Phys.: Conf. Ser.* **1513** 012018

View the article online for updates and enhancements.

You may also like

- Real-time residual life prediction based on kernel density estimation considering abrupt change point detection
Weizhen Zhang, Hui Shi, Jianchao Zeng et al.
- Detection and localization of change points in temporal networks with the aid of stochastic block models
Simon De Ridder, Benjamin Vandermarliere and Jan Ryckebusch
- Application of change-point analysis to the selection of representative data in creep experiments
Setareh Zomorodpoosh, Nicklas Volz, Steffen Neumeier et al.



This content was downloaded from IP address 8.243.64.131 on 06/10/2025 at 21:49

Change-point analysis of the demand for medicines

J C Valderrama Balaguera[1], **H F Castro Silva**[1], **and T Velásquez Pérez**[2]
[1] Grupo de Investigación Observatorio, Universidad Pedagógica y Tecnológica de Colombia, Sogamoso, Colombia
[2] Grupo de Investigación en Tecnología y Desarrollo en Ingeniería (GITYD), Universidad Francisco de Paula Santander, Seccional Ocaña, Colombia

E-mail: hugofernando.castro@uptc.edu.co, tvelasquezp@ufpso.edu.co

Abstract. During the last years, the model of change points has been widely used to predict data series and it can be demonstrated success cases of its application in diverse areas. However, the usefulness of the theory of the points of change for the prognosis of the demand for medicines has not been proven. In this research with a quantitative approach, the use of the theory of points of change in prognosis of the demand for medicines in Healthcare Providers in Colombia is analyzed. The forecasts obtained with the theory of change point are compared with the forecasts obtained with traditional method of Holt-Winters. As a result of this study, the "best" model is identified and recommendations for the implementation of the proposed model are presented. This study is useful for administrators of entities providing health services interested in optimizing budget allocation and improving the inventory planning strategy. The results presented in this article, will allow to incorporate the results of the methodology in the field of modeling of systems of inventories of medicines.

1. Introduction

The inadequate use of forecasting methodologies together with the inability to predict fluctuations that may occur in some areas, generate a decoupling between organizations so that the performance of some of them can affect the performance of others directly or indirectly [1]. The data recorded in the time of demand for medicines in health service providers in Colombia is a good example of the fluctuation in these factors. One of the reasons why this situation arises is due to the fact that in organizations the tools or concepts that are intended to be applied are not studied in depth, this causes that the results regarding an objective are not satisfactorily fulfilled.

Focusing on the forecasts according to Devarajan and Jayamohan [2], these occupy the most strategic position in the working capital structure of most companies and good management is a tool for the company to maintain a competitive advantage, even in an economy uncertain. Supervision, treatment and an adequate analysis of data and forecasts can become the key to a company's success, given that poor management and poor planning of activities directly and negatively affects the performance of organizations.

The model of change points has been widely used and it can be shown successful cases of its application in various areas [3], such as the studies by Costa and Gon [4], Killick, *et al.* [5], among others. In its basic foundation the modeling of the time series and its forecasts is carried out mainly by classical methods such as the exponential smoothing model, Kalfman filter, Holt, and Holt-Winters. The selection of methods depends on factors such as trend, cyclicity, random effect and seasonality.

On the other hand, the demand for medicines is a process that is affected by the implementation of

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

new policies in their use, in accordance with the current regulations established by the "Instituto Nacional de Vigilancia de Medicamentos y Alimentos (INVIMA)", these changes in the regulations translate into a change in the behavior of the series over time with a high impact on the value of the average and volatility of the series. Since for the data of the demand of the medicines in Colombia the changes or interventions prevail in the time and the value of the error is very high when predicting using classic methodologies, according to the initial analysis made to the series. It is proposed to investigate the utility of the application of a model with points of change to predict the demand for medicines in health service providers, analyzing the adverse effects of the non-incorporation of these changes in the modeling of the series.

2. Background

The analysis of the change points is a methodology that seeks the estimation of the points where changes occurred in the time series [6]. When a disturbance appears in the time series, it is said that the process is outside the statistical control, the moment in which the change occurs is called point of change (PC) and the point of change analysis (APC) is the tool that studies this phenomenon [7] The objective of the APC is to discover the abrupt changes of properties that are behind the data of the series of time, facilitating the analysis of the characteristics of the series before and after a PC is detected [8].

The statistical inference about the points of change has two aspects, the first is to detect if there is any change in the sequence of observed random variables. And the second is to estimate the number of changes and their corresponding locations. The methods for PC estimation are mainly the likelihood ratio, non-parametric and bayesian [3]. Some authors also considered the problem of the point of change in other configurations of the model, such as gamma and exponential.

The likelihood model is used for the location of the change point, as well as an information criterion developed, for the case of known variance [9]. "This model is a hypothesis test that helps choose the" best "model between two nested models. Understanding a nested model means that one is a special case of the other. The best model is one that makes the data more likely or maximizes the likelihood function. However, although the concept is relatively easy to understand, the calculations to find the entries for the procedure are not" [10]. The likelihood ratio tests use soft and monotonous functions like the logarithm, which allow the numerical convergence of the algorithms. Most statistical software packages have built-in functions to handle them.

According to Picard [11], those seeking a point of change primarily investigate a problem in two different ways: (a) detect a fault that does not affect the average of the observations but changes its covariance structure. When this happens, they recommend considering procedures that take into account mild or profound disturbances, performing a nonparametric type test such as the Kolmogorov-Smirnov test. And (b) detect a fault and estimate the change parameters for a failure that occurs at the same time in the mean and the covariance of an autoregressive process of bounded order. It is recommended to apply a likelihood ratio test and make the appropriate normalizations.

According to Chen and Gupta [3], in many situations the researcher faces the problem of detecting the number and location of PC in a series. This problem is evident in disciplines such as the stock market, quality control, geological data analysis, meteorology and genetic data analysis. In this context, the Bayesian model (MB) has been used effectively as it is confirmed in their studies Aronis, *et al.* [12], Barry and Hartigan [13], Cmejla, *et al* [14]. The application of the MB carried out by Parent, *et al.* [15], focuses on the analysis of hydro meteorological weather series and develops a comparison of the PC model with forecasts. Jeong and Kim [16], use the MB, to determine the PCs and the length of the segments between each of these by means of the application of a genetic algorithm (GA).

In the same way, Li and Lund [17], detect multiple PCs through metadata, for the estimation of multiple PCs is considered through a MB. Another method that is frequently used is linear regression, as do Yan, *et al.* [18], Eck [19]. For example, Lu, *et al.* [20], use the regression method together with a tool known as theoretical regression trees, propose a novel method for PC detection in machine

monitoring. The use of the likelihood method for estimating PC has become one of the most extensively used techniques, for example in the stock market due to the volume of data such as Zhitlukhin and Ziemba [21].

In recent years, studies on time series have increased and with this the need to be able to estimate efficiently and precisely the location of PCs, then a method used for their estimation is briefly presented. According to Killick, in his research an approach commonly used to identify multiple points of change is to minimize [22]. In Equation (1), C is a function that takes a value for a segment, for example, the negative logarithmic probability and $\beta f(m)$ is a penalty to protect against excessive adjustment. This method through minimization considers 2^{n-1} solutions reducing to $\binom{n-1}{m}$ if m is known.

$$\sum_{i=1}^{m+1}[C(Y\tau_{I-1}+1):_{\tau i})]+\beta f(m) \quad (1)$$

On the other hand, in the field of forecasts of time series, one of the most used methods is the Holt-Winters method. This is a triple softening exponent method and has the advantage of being easy to adapt as new real information is available. The Holt-Winters method is an extension of the Holt method that considers only two softening exponents. Holt-Winters considers level, trend and seasonal of a certain series of times. This method has two main models, depending on the type of seasonality, the seasonal multiplicative model and the seasonal additive model [23]. Pérez, *et al.* [22], assure that this method is frequently used for items with seasonal demand and requires start-up values, using historical data.

Forecasts provide a competitive advantage for companies, through these tool managers try to always make better estimates of what will happen in the future, the main objective of making forecasts is to make good estimates, a good forecast represents an essential part of the operations of efficient services and manufacturing [22].

3. Evaluation and diagnosis of the data

This research describes the impact of the estimation of the points of change in the prognosis of the demand for medicines in an entity that provides health services, by using historical numerical data, together with its analysis. The health provider has a systems area that is responsible for maintaining digital records of the sales of medicines made to their patients, likewise the pharmacy area, has the support of a software for the registration of the movements of medicines.

It is also worth mentioning that the organization does not have defined methods for calculating the demand forecasts for medicines. The medication requirements are the responsibility of the pharmacy manager, who based on his experience determines the amount necessary for the correct provision of the service. Due to the nature of the entity, and its principles, including the quality in the provision of the service, it is not possible that there are unmet demands. In the event that an unforeseen event occurs in the supply chain, the physicians of the entity are informed in a timely manner, they approve the absence of one medication with another, thus avoiding complications in the patients' health.

The consumption records were found in a database with the historical record of the demands of the health service provider.

For the study, an analysis is made of the information provided by the pharmacy of the case study entity, where a list of 315 medications was taken into account, and a consolidated demand for each one is made during a period of 313 weeks. Within the analysis of the information collected is shown in Table 1, the description and presentation of the six drugs that stand out from the total list, due to their high level of demand during this period of time. For the analysis of the demand by means of the PC theory, the data of Acetaminophen syrup 150 mg / 5 ml fco x 60 ml will be used, which is a medicine that is intended for children between 1 and 12 years old; whose characteristics are each 100 mL of syrup contains acetaminophen 3g, and its sanitary registration before INVIMA is 2017M-007901-R3 A 2006M-007901-R2.

Table 1. Consolidation of drugs with the highest demand.

	Name of the medication	Total demand for 314 weeks (un)	Average weekly demand (un)
1	Acetaminophen 500 mg tab	379901	1209
2	Acetaminophen Syrup 150 mg / 5 ml FCO x 60 ml	78320	249
3	Dipyrone magnesium 2 g / 5 ml amp	210921	671
4	Ranitidine 50 mg2 ml amp	169824	540
5	Tramadol 50 mg1 ml amp	114970	366
6	Metoclopramide hydrochloride 10 mg amphotericin	110574	352

4. Analysis and discussion

For the development of this research, R® was used. Figure 1 to Figure 4 summarize the sequence of data analysis. As can be seen in Figure 1, in some periods there are abrupt changes in the series, these are the points of change which are the objective of the investigation. In Figure 2, the components of the series are observed (trend, seasonality and randomness). For the series of Figure 2, changes in the mean and in the variance are observed, this happens because there was a change in the level and the variance was reduced, by decomposing the series we can analyze each of the patterns that the series takes over time to identify if there is growth (as seen from the beginning of 2012 until 2014), or a decrease (from 2014 to 2017), analyze the patterns of change that are recurrent (it is observed that the patterns in the seasonality are repetitive during each year), and interpret the fluctuations caused by unpredictable events (in mid-2013 the randomness of the series reaches its highest point and in mid of the year 2014 reaches its lowest point).

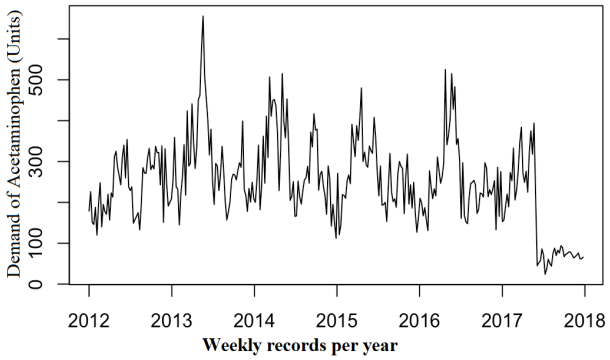


Figure 1. Weekly demand series acetaminophen syrup 150 mg / 5 ml x 60 ml fco.

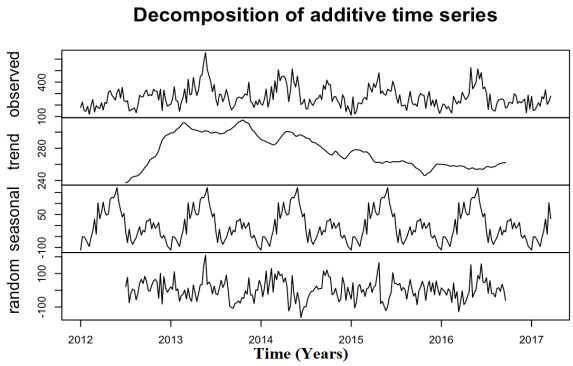


Figure 2. Components of the time series.

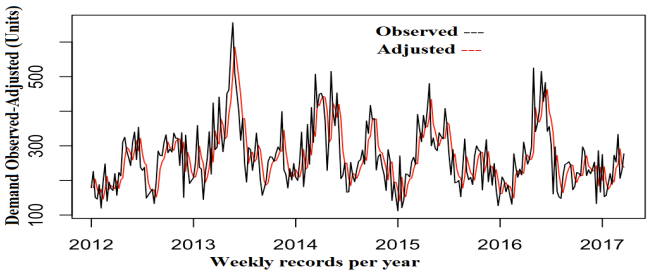


Figure 3. Holt-Winters smoothing of the observed weekly demand – adjusted.

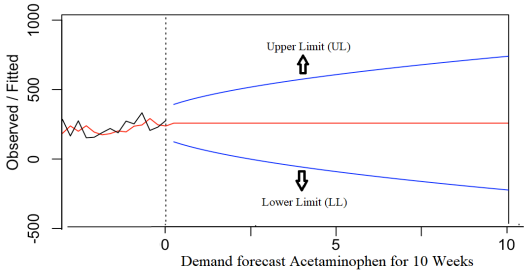


Figure 4. Demand forecast acetaminophen syrup 150 mg / 5 ml fco x 60 ml.

Figure 3 shows that the exponential smoothing adjusts the expected data to the observed data, within the smoothing parameters only α is assigned equal to 0.52. On the other hand, the sum of the square errors (SSE) of the prediction function is calculated, which gives a value of SSE = 1278134. In

Figure 4, the forecast is represented with the Holt-Winters method of 10 periods for the analyzed data series. The results of the forecast can be seen in Table 2, where the estimated demand during the next 10 weeks with its respective tolerances is evidenced.

Table 2. Consolidated demand forecast for acetaminophen syrup 150 mg / 5 ml fco x 60 ml for the 10 future periods.

Period	Estimated	Upper limit (UL)	Lower limit (LL)
1	258.47	393.32	123.63
2	258.47	411.00	105.94
3	258.47	426.84	90.11
4	258.47	441.31	75.64
5	258.47	454.72	62.23
6	258.47	467.26	49.68
7	258.47	479.10	37.85
8	258.47	490.33	26.62
9	258.47	501.04	15.91
10	258.47	511.30	5.65

Table 3. Consolidated demand forecast for acetaminophen syrup 150 mg / 5 ml fco x 60 ml for the 10 future periods in the new series.

Period	Estimated	Upper limit (UL)	Lower limit (LL)
1	148.29	282.68	13.90
2	148.29	301.05	-4.47
3	148.29	317.44	-20.86
4	148.29	332.38	-35.80
5	148.29	346.20	-49.61
6	148.29	359.11	-62.52
7	148.29	371.27	-74.69
8	148.29	382.80	-86.22
9	148.29	393.80	-97.22
10	148.29	404.32	-107.74

According to the results obtained in R, the calculation of the forecast using the Holt-Winters model, the demand for Acetaminophen syrup from period 1 will be ≈258 units of weekly medication.

After having made an analysis, it is useful to be able to trace the change points in the original data to visually inspect if the estimated change points are reasonable. For this purpose, a tracing method is included, for example, a change in the variance is denoted by a vertical line at the location of the change point, while a change in the mean is indicated by horizontal lines, which represent the mean value in different segments within the study series [22]. The results obtained from the estimation of the change points of the case study series show that the number of PCs when there are variations in the mean is 288, when there are changes in the variance 3 PCs appear whose location are in the data 69 (PC 1: corresponding to the date 2013-04-21, demand of 334 units), 76 (PC2: 2013-06-30, demand: 249) and 283 (PC3: 2017-05-28, demand: 215); and finally there are 7 PCs estimated when there are changes in the mean and variance (location: 43, 45, 182, 234, 283 and 310).

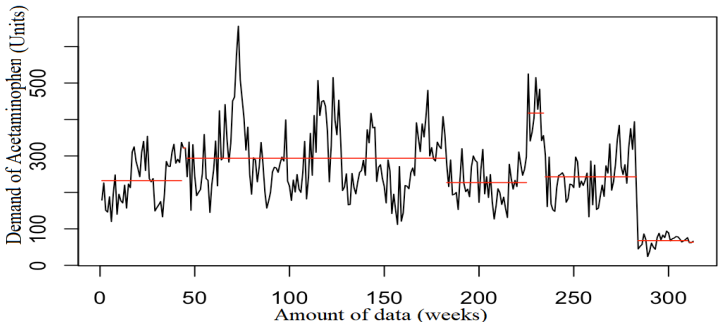


Figure 5. Estimation of change points in mean and variance.

The next step for the analysis of the data series consists in the estimation of the change points. It is important to mention that for the development of this research the estimated PCs were taken into account when there are changes in the variance, the other results will be the object of study for future research.

Therefore, once the PCs are estimated and located in the variance, the objective is to perform again the steps illustrated in Figure 1 to Figure 4 and compare the results when they are estimated, and the PCs are not estimated in the forecast of the demand.

As can be seen in Figure 5, in PCs in the variance the three CPs are clearly seen, for the comparison of the forecast the data that appear from PC 2 was taken into account (date: 2013-06-30, demand: 249). For Holt-Winters or for any strategy the results will be depending on the previous observations, what is done is that the new forecast takes into account the data that appear from the second PC, and once this PC occurs, the behavior of the new series is analyzed, with this it becomes steps 1 to 4 described above that is what you want to test, since in the history of the series is the first PC that does not get up again. Studying the new series, the results obtained are shown below in Table 3. The sum of the square errors of the prediction function that for this case gave a value of SSE = 1001440. And the calculation of the forecast by the Holt-Winters model of the demand for Acetaminophen for the new series during the 10 periods will be ≈ 148 units of weekly medication.

5. Conclusions

Comparing the results of the calculated forecasts with the original series and the new series (taking into account the estimation of the CPs in the variance) the SSE decreased considerably by 14%, and the forecast for the 10 future periods decreased by 110 units. For the use of this methodology in this class of studies it is important to have a large volume of data between the points, given that if there are consecutive perturbations in the series it will be difficult to estimate the forecasts.

For this study when PCs were estimated in the case of variance, it can be clearly seen that there is at the end a PC of change in variance (PC # 3) of the original series (see Figure 5), for this case if the data from this disturbance was estimated, and the size of the new series is very short, which made it difficult to estimate a forecast since there is no change in the period of the series.

With respect to considering the points of change in the mean, or in the variance, or in the mean and variance it is done through the decomposition of the series (see Figure 2), when there are changes in the trend and that change is not different over many observations in time, that kind of series for this type of studies would not work, because it means that the series varies in a very consecutive way as the case of a series of time in the price of the dollar because of the number of fluctuations that it has in real time, and not will be used to calculate long-term forecasts. Contrary to what happens in the series of this investigation of acetaminophen in syrup where clear intervention is noted that is why the PC is estimated and based on this estimate make the predictions allowing a better prediction.

For the purposes of this investigation, the new time series was taken from PC 2, until the last data corresponding to No. 313 (date: 2017-12-24, demand: 66) (Figure 5), rejecting PC 3 for the purpose of making several calculations. The calculation of the forecast was made with the data that appear after PC 3, and it was evidenced that the forecast could not be made since there is no change in the period of the series. Also to conclude a third calculation was made with a new series taking the data that is between PC 2 and PC 3; and in this way it was evidenced that the forecast can be further improved, so that the SSE took a much lower value than that previously calculated equal to 996046.5.

Through this research a theoretical contribution is made in the methodological development for the estimation of forecasts when applied in the analysis of the demand of medicines to a health entity.

Consequently, the application of the change points model serves as a tool to analyze the aspects, variables and processes that affect the forecast of the demand and recalculation of the inventory level for the company that wishes to study.

References

[1] R Sanchis, R Poler 2010 Estrategias de Gestión de los Procesos y Operaciones en Escenarios de Personalización en Masa *4th International Conference on Industrial Engineering and Industrial Management XIV Congreso de Ingeniería de Organización* (Donostia: Asociación para el Desarrollo de la Ingeniería de Organización)

[2] D Devarajan, M S Jayamohan 2016 Stock control in a chemical firm: combined FSN and XYZ analysis *Procedia Technology* **24** 562

[3] J Chen, K Gupta 2012 *Parametric Statistical Change Point Analysis* (Boston: Birkhäuser Basel)

[4] M Costa, A M Pereira, L Teixeira 2016 Change-point detection in environmental time series based on the informational approach *Electronic Journal of Applied Statistical Analysis* **9(2)** 267

[5] R Killick, I A Eckley, K Ewans, and P Jonathan 2010 Detection of changes in variance of oceanographic time-series using changepoint analysis *Ocean Engineering* **37(13)** 1120
[6] J A Garza Venegas, A E Cordero Franco, V G Tercero Gómez, J F Camacho Vallejo 2012 Estimación de múltiples puntos de cambio en series de tiempo con distribución normal utilizando un algoritmo heurístico de construcción *Congreso Latino-Iberoamericano de Investigación Operativa, Simpósio Brasileiro de Pesquisa Operaciona* (Río de Janeiro: Asociación Latino-Iberoamericana de Investigación Operativa-ALIO)
[7] S Pérez, E Cordero, V Tercero, A Chávez 2013 *Análisis de Punto de Cambio, Dentro del Control Estadístico de Procesos* (San Nicolás de los Garza: Universidad Autónoma de Nuevo León)
[8] S Liu, M Yamada, N Collier, M Sugiyama 2013 Change-point detection in time-series data by relative density-ratio estimation *Neural Network* **43** 72
[9] P J Plummer 2012 *Decting Change–Points in a Compound Poisson* (Kansas City: University of Central Missouri)
[10] E L Lehmann 2006 On likelihood ratio tests *Optimality* **49** 1
[11] D Picard 1985 Testing and estimating change-points in time series *Adv. Appl. Probab.* **17(4)** 841
[12] K P Aronis, I Magou, R Dekker, G Tagaras 2004 Inventory control of spare parts using a Bayesian approach: A case study *European Journal of Operational Research* **154(3)** 730
[13] D Barry, J A Hartigan 1993 A Bayesian analysis for change point problems *Journal of the American Statistical Association* **88(421)** 309
[14] R Cmejla, J Rusz, P Bergl, J Vokral 2013 Bayesian changepoint detection for the automatic assessment of fluency and articulatory disorders *Speech Communication* **55(1)** 178
[15] E Parent, L Perreault, J Bernier, B Bobe 2000 Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited *Journal of Hydrology* **235** 221
[16] C Jeong, J Kim 2013 Bayesian multiple structural change-points estimation in time series models with genetic algorithm *Journal of the Korean Statistical Society* **42(4)** 459
[17] Y Li, R Lund 2015 Multiple changepoint detection using metadata *Journal of Climate* **28(10)** 4199
[18] Q Yan, Z Sun, Q Gan, W L Jin 2018 Automatic identification of near-stationary traffic states based on the PELT changepoint detection *Transportation Research Part B: Methodological* **108** 39
[19] D J Eck 2018 Bootstrapping for multivariate linear regression models *Stat. Probab. Lett.* **134** 141
[20] G Lu, Y Zhou, C Lu, X Li 2017 A novel framework of change-point detection for machine monitoring *Mech. Signal Processing* **83** 533
[21] M V Zhitlukhin, W T Ziemba 2016 Exit strategies in bubble-like markets using a changepoint model *Journal Quantitative Finance Letters* **4(1)** 47
[22] R Killick 2014 Changepoint: An R package for changepoint analysis *J. Stat. Softw.* **58(3)** 1
[23] O M Rivero 2016 *El Método de Pronóstico Holt-Winters* (Perú: WordPress) Consulted on: https://administration21.files.wordpress.com/2017/01/pronc3b3sticos-holt-winters-omr-nov2016.pdf)