# Non-parametric Methods of Finding Changepoints in Multidimensional Time Series

Dmitriy Klyushin[1], Andrii Urazovskyi[1]
[1] *Taras Shevchenko National University of Kyiv, prospekt Glushkova, 4D, 03680, Kyiv, Ukraine*

**Abstract**
Modern companies contain a huge number of processes. The larger and more complex the business that one person or company is engaged in, the more impossible it becomes to keep track of all the processes in manual mode. Each element is at risk of failure, and in order to prevent or quickly respond to various breakdowns, you need to be able to recognize errors automatically so that in case of critical situations, seek help from a specialist who can fix everything. For modern systems, which can be important every second and which constantly process huge flows of information, a method is needed that will allow you to quickly recognize changes in state, for example, a deterioration in the patient's condition or server failure. To solve these and similar problems, a new method based on the use of Fisher's linear discriminant and Petunin statistics is proposed. To simulate the process, after turning on the sensors to capture data and a continuous flow of information, a multidimensional time series will be generated, after which the method will recognize changepoints that indicate that the object has changed state or something has happened to it. A clear probabilistic interpretation of the method underlying this classification greatly expands its capabilities within the framework of risk-informed systems.
**Keywords**
Time series, change point, nonparametric statistics, risk-informed systems.

## 1. Introduction

The use of automatic systems and artificial intelligence to recognize changepoints in multidimensional time series provides great opportunities for risk-informed systems. For example, in the fields of medicine, engineering, economics, cybersecurity, which may be narrowly focused or if the region lacks qualified employees. This will help to optimize the use of human resources, which can be directed to the management or solving critical issues, for example, those related to people's lives.

Nuclear power plants provide another important example of the responsible use of risk-informed systems. In the design, use, economics, and licensing of such energy sources, safety plays a key role. Since such facilities have been used for several decades, it is necessary to ensure the integrity and operability of vital elements of nuclear power plants in order to prevent, and otherwise reduce or mitigate the consequences of accidents that have occurred. Historically, plant designers have redesigned nuclear power plant systems to provide reliability in the form of redundant and varied safety features and to ensure that even in the event of abnormal and unplanned situations, the health and safety of workers and the public can be protected with a high degree of confidence.

For a method to be useful, it must have the following properties, namely:

1. High accuracy to minimize the possibility of false positive and false negative results
2. Stability so that single outliers or anomalies cannot severely corrupt the data series and create a false changepoint.

CITRisk'2022: 3rd International Workshop on Computational & Information Technologies for Risk-Informed Systems, January 12, 2023, Neubiberg, Germany
EMAIL: dokmed5@gmail.com (D. Klyushin); urazovskya@gmail.com (A. Urazovskyi)
ORCID: 0000-0003-4554-1049 (D. Klyushin); 0000-0002-7918-2876 (A. Urazovskyi)

3. Independence from the main distributions, so that the method is as versatile as possible and can be applied in different areas and for different situations, processes, objects.

4. Low cost of computing to be able to work online without using a lot of computing power and without overloading the server.

5. Balanced sensitivity, that is, not too high so as not to react if the patient simply turned over on the other side, but not too low so as not to miss the explosion of a reactor at a nuclear power plant. This article will talk about a new non-parametric method for detecting changepoints in multivariate time series based on the metric developed in (Klyushin and Petunin, 2003) and demonstrated advantages over Kolmogorov-Smirnov and Wilcoxon statistics (Klyushin and Urazovskyi, 2021), and to show its medicinal benefits.

In section 2.1, we describe the state of the art in the field of detection of changepoints in multivariate time series. In section 2.2 we will consider the algorithm for calculating the Petunin statistic and its properties. In section 2.3 we will consider the algorithm for constructing the Fisher linear discriminant. In section 3.1 we present the results of numerous numerical experiments with a wide range of distributions. Section 3.2 considers possible applications of the proposed algorithm. For example, we can investigate the presence of various diseases in a virtual patient by measuring its parameters, such as heart rate, blood oxygen saturation and body temperature, which need to be monitored.

## 2. Theoretical part

### 2.1. Literature review

When considering a task of searching and recognizing change points of random multidimensional time series there are many ideas for applying the results obtained. A change point of time series is a point that separates two pieces of a series into ones that have different distributions. The developed methods for solving the problem of finding a point of change in multidimensional time series are usually separated into algorithms for streaming (for data coming one by one) and already fully known (when the whole series is already given completely) data. Streaming data goes constantly in time, and already known ones are given in their entirety. A detailed review of the change point detection method in the already given time series was done in (Truong, Oudre and Vayatis, 2020). Since we want to consider an independent of initial distributions method for streaming data, we will consider the streaming algorithms discussed in articles over the past few years.
A common and complex problem that arises when considering our problem is the increase in dimension, which can add, for example, problems with the speed of calculations. In (Alippi et al, 2016), the problem of determining the point of change in the time series was discussed using Kullback–Leibler divergence and the log-likelihood function with different distributions. The authors showed that the more the data dimension grows, the worse we can notice changes in the set value. In the article (Wang and Zwetsloot, 2021), attention is paid to the problems associated with increasing the data dimension. The authors described a method for detecting a change point using control charts. Their algorithm makes it possible to detect sparse shifts of the mean vector.
Detection of a change point can be considered with different variations: simply look for the presence of a change point anywhere in the time series or localization of exactly the coordinates of the desired point. In (Jaehyeok, Ramdas, and Rinaldo, 2022), attention is paid to the presence of a change point somewhere, but the determination of the coordinate of the desired point does not have sufficient accuracy and localization. A Bayesian method proposed in (Sorba and Geissler, 2021) has lineax computational complexity with respect to the number of points, but puts the researcher in a tradeoff of speed and accuracy.

The method discussed in (Navarro, Allen, and Weylandt, 2021) showed excellent performance of the convex network clustering method, which unfortunately requires large computational costs. One of the main methods for finding change points, discussed in (Tickle, Eckley, and Fearnhead, 2021), like many of today's change point detection methods (used in this article for terrorist research), has the assumption that the data flow is independent of time. Although this criterion may be resistant to the violation of this condition, however power may decrease.

The proposed method is needed to process a continuous stream of data, can track outliers and does not rely on any specific assumptions and known facts about the distribution of data. Below we consider articles whose authors approached the problem from the same angle. In (Wendelberger et al., 2021) the extended Bayesian Online Changepoint Detection was developed. A method proposed in (Adams and MacKay, 2007) is dedicated to the exploration of geographical data. In (Cooney and White, 2021) the authors considered algorithms proposed exclusively for exponential models. To increase the accuracy of functioning, in (Castillo-Matteo, 2021) the authors made assumptions on data distribution. In (Hallgren, Heard, and Turcotte, 2021), data on the type of distribution helped to optimize the computational complexity of the. The paper (Fotoohinasab, Hocking, and Afghah, 2021) has the same drawback. It is required to make a priori assumptions about the data in order to find the changepoints in the model. To determine the points of change in a multivariate time series more precisely, it is often necessary to pre-process the data (Fearnhead and Rigaill , 2018). In (Harle et al., 2014), the authors reviewed the Bayesian method for segmenting multivariate time series using the MCMC method and Gibbs sampling. The authors have demonstrated that change points are stably detected and their coordinates localized by implicitly examining the dependency structure. Similar ideas were proposed in the article (Renz et al., 2021) for gesture recognition. Change point estimation using the Yule-Walker moment estimator (Gallagher et al., 2021) is unstable due to the large shifts in the means.

In (Wang et al, 2019), the authors consider an algorithm that is used to stream data using a huge matrix dependent on the dimension of the source data space. Similar methods discussed in (Romano et al., 2021) where the authors proposed a method called Functional Online CuSUM (FOCuS). The idea is rolling the window and running the previously developed methods in parallel for all window sizes. The efficiency and applicability of the algorithm was shown by detecting anomalies in the computer server data.

Analysis of papers on this topic shows that the most desirable qualities by searching for points of change in our problem are 1) stability, 2) high accuracy of calculations, 3) speed of work and 4) independence from basic distributions. Below, we describe such an algorithm based on so-called Petunin's statistics.

## 2.2. Petunin's statistics

The Petunin's statistic ($p$-statistic) is a measure of proximity between samples proposed by the Ukrainian mathematician Yuriy Petunin. It is used to test the hypothesis that the distribution functions of two samples are equal.

Let us consider two general populations $G$ and $G'$ and corresponding distribution functions $F_G$ and $F_{G'}$.

Let there be two samples $x = (x_1, x_2, \dots x_n) \in G$ and $x' = (x_1', x_2', \dots x_m') \in G'$, and $x_{(1)}, \leq x_{(2)} \leq x_{(3)} \dots \leq x_{(n)}$ and $x'_{(1)} \leq x'_{(2)} \leq x'_{(3)} \dots \leq x'_{(m)}$ - corresponding ordinal statistics and it is necessary to determine whether they belong to the same distributions. Suppose that $F_G(u) = F_{G'}(u)$, then

$$P\left(A_{ij}^{(k)}\right) = P\left(x'_k \in \left(x_{(i)}, x_{(j)}\right)\right) = p_{ij}^{(n)} = \frac{j-i}{n+1} \tag{1.1}$$

If we have a sample $x' \in \left(x'_{(1)}, x'_{(2)}, x'_{(3)}, \dots, x'_{(m)}\right)$, we can find the frequency $h_{ij}$ random event $A_{ij}$ and confidence intervals $\left(\Delta_{ij}^{(1)}, \Delta_{ij}^{(2)}\right)$ for probability $p_{ij}$ at a given level of significance $\beta$, i.e

$$B = \left\{p_{ij} \in \left(\Delta_{ij}^{(1)}, \Delta_{ij}^{(2)}\right)\right\}, p(B) = 1 - \beta \tag{1.2}$$

According to (Van der Waerden, 1969)

$$\Delta_{ij}^{(1)} = \frac{h_{ij}^{(n)} n + \frac{g^2}{2} - g\sqrt{h_{ij}^{(n)}\left(1 - h_{ij}^{(n)}\right) n + \frac{g^2}{4}}}{n + g^2} \tag{1.3}$$

$$\Delta_{ij}^{(2)} = \frac{h_{ij}^{(n)} n + \frac{g^2}{2} + g\sqrt{h_{ij}^{(n)}\left(1 - h_{ij}^{(n)}\right)n + \frac{g^2}{4}}}{n + g^2} \tag{1.4}$$

Where $g$ satisfies the condition $\phi(g) = 1 - \frac{\beta}{2}$ ($\phi(g)$ – the density of the normal distribution). Value $g$ determines the level of significance of the confidence interval $I_{ij}^{(n,m)} = (\Delta_{ij}^{(1)}, \Delta_{ij}^{(2)})$; According to rule $3\sigma$ (Petunin, Klyushin, Ganina, Borodai and Andrushkiv, 2001) at $g = 3$ the level of significance of this interval does not exceed 0.05. Let's denote by $N$ number of all the confidence intervals $I_{ij} = (\Delta_{\text{ij}}^{(1)}, \Delta_{\text{ij}}^{(2)})$. It is clear that $N = \frac{n(n-1)}{2}$. Denote by $L$ – the number of those intervals $I_{ij}$, which contain probability $p_{ij}^{(n)}$. Statistics $h^{(n)} = \frac{L}{N}$ we will call $p$-statistics and it will be a measure of closeness $\rho(x, x')$ between samples $x$ and $x'$. Let's substitute the obtained value $h$ in the formula for calculating confidence intervals, we will get a confidence interval $I = (\Delta^{(1)}, \Delta^{(2)})$ to test the hypothesis $H$ with a level of significance approximately equal to 0.05 (Klyushin and Petunin, 2003)

## 2.3. Fisher's linear discriminant

The terms Fisher's linear discriminant and LDA are often used interchangeably, although Fisher's original article (Fisher, 1936) actually describes a slightly different discriminant, which does not make some of the assumptions of LDA such as normally distributed classes or equal class covariances.

Suppose two classes of observations have means $\overline{\mu_0}$, $\overline{\mu_1}$ and covariances $\Sigma_0$, $\Sigma_1$. Then the linear combination of features $\overline{w} \cdot x$ will have means $\overline{w} \cdot \overline{\mu_\iota}$ and variances $\overline{w}^T \Sigma_i \overline{w}$ for $i = 0,1$. Fisher defined the separation between these two distributions to be the ratio of the variance between the classes to the variance within the classes:

$$S = \frac{\sigma_{between}^2}{\sigma_{within}^2} = \frac{(\vec{w} \cdot \overrightarrow{\mu_\iota} - \vec{w} \cdot \overrightarrow{\mu_0})^2}{\vec{w}^T \Sigma_1 \vec{w} + \vec{w}^T \Sigma_0 \vec{w}} = \frac{\left(\vec{w} \cdot (\overrightarrow{\mu_\iota} - \overrightarrow{\mu_0})\right)^2}{\vec{w}^T (\Sigma_0 + \Sigma_1) \vec{w}}$$

This measure is, in some sense, a measure of the signal-to-noise ratio for the class labelling. It can be shown that the maximum separation occurs when

$$\vec{w} \propto (\Sigma_0 + \Sigma_1)^{-1} (\overrightarrow{\mu_\iota} - \overrightarrow{\mu_0})$$

When the assumptions of LDA are satisfied, the above equation is equivalent to LDA.

Be sure to note that the vector $\vec{w}$ is the normal to the discriminant hyperplane. As an example, in a two dimensional problem, the line that best divides the two groups is perpendicular to $\vec{w}$.

Generally, the data points to be discriminated are projected onto $\vec{w}$; then the threshold that best separates the data is chosen from analysis of the one-dimensional distribution. There is no general rule for the threshold. However, if projections of points from both classes exhibit approximately the same distributions, a good choice would be the hyperplane between projections of the two means, $\vec{w} \cdot \overrightarrow{\mu_0}$ and $\vec{w} \cdot \overrightarrow{\mu_1}$. In this case the parameter $c$ in threshold condition $\vec{w} \cdot \vec{x} > c$ can be found explicitly:

$$c = \vec{w} \cdot \frac{1}{2}(\overrightarrow{\mu_0} + \overrightarrow{\mu_1}) = \frac{1}{2} \overrightarrow{\mu_1}^T \Sigma_1^{-1} \overrightarrow{\mu_1} - \frac{1}{2} \overrightarrow{\mu_0}^T \Sigma_0^{-1} \overrightarrow{\mu_0}$$

# 3. Practice part

## 3.1. Numerical experiments

The purpose of our experiments is to demonstrate the accuracy of the following algorithm for a stationary time series, which should find the first changepoint and test the homogeneity hypothesis.

At the beginning we take $width$ and designate the elements $x_1, \ldots, x_{width}$ - starting ones, with which we will continue to work using the sliding window method. When we have a sample $(x_{i+1}, x_{i+2}, \ldots, x_{i+width})$, we do the following with it:

- Building a linear Fisher discriminant for samples $(x_1, x_2, \ldots, x_{width})$ and $(x_{i+1}, x_{i+2}, \ldots, x_{i+width})$ and find the projections on the line.

- Rotate the resulting straight line so that only one coordinate remains, and make the rest the same. Getting projections $(p_1, p_2, \dots, p_{width})$ and $(p_{i+1}, p_{i+2}, \dots, p_{i+width})$
- Calculate the Petunin's statistics $p_{stat}$ for the resulting sets of projections
- If $p_{stat} \geq 0.95$, then we say that the new sample has the same distribution as the original one, otherwise we say that the other.
- Shifting the sample $(x_{i+1}, x_{i+2}, \dots, x_{i+width})$ one position to the right and start the algorithm from the beginning. We do this until all the data is gone.

If sample after element $x_n$ become inhomogeneous, then the point $x_{n+1}$ regarded as a changepoint.

To demonstrate how the algorithm works, we take a series of length $N = 400$ and divide it into 4 equal intervals with different distributions. Then we run our algorithm 100 times and average the values of Petunin's statistics (P statistics), after which we display the obtained values in two colors: blue is not less than 0.95, that is, for those samples that have the same distribution as the original and red less than 0.95 - having a different distribution.

For each experiment, we calculated five measures of error: mean absolute error (MAE), mean squared error (MSE), mean squared deviation (MSD), root mean squared error (RMSE), root mean squared error (RMSE), and root mean squared error (RMSE), and normalized root mean squared error (NRMSE). To demonstrate the effectiveness of the described algorithm, we will rely on the latter value. As is well known, if NRSMR > 0.5 the results can be considered as random. If a NRMSE is close to 0, then the results are considered good.

### 3.1.1. Almost non-overlapping uniform distributions with different means

Let's consider a saltatory time series, which is composed of uniform distributions that practically do not overlap. On this time series, we will be able to test the shift hypothesis.

**Table 1**

Time intervals and uniform distributions with different means

| Time interval | Distribution $T_1$ | Distribution $T_2$ | Distribution $T_3$ |
|---|---|---|---|
| 0-99 | U(65;75) | U(96.5;97.5) | U(36.4;36.7) |
| 100-199 | U(100;110) | U(97.0;99.0) | U(38.0;39.0) |
| 200-299 | U(65;75) | U(96.5, 97.5) | U(36.4;36.7) |
| 300-399 | U(70,90) | U(97.5;99.0) | U(37.0;37.5) |

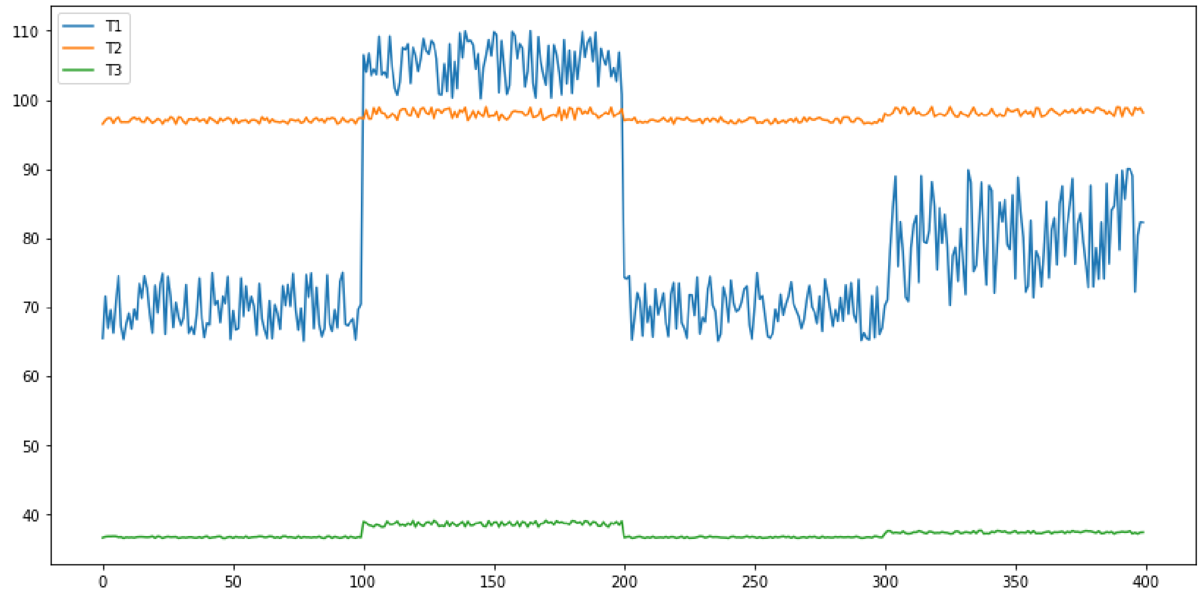


**Figure 1**: An example of time series consisting of samples from almost non-overlapped uniform distributions with different means and their change points

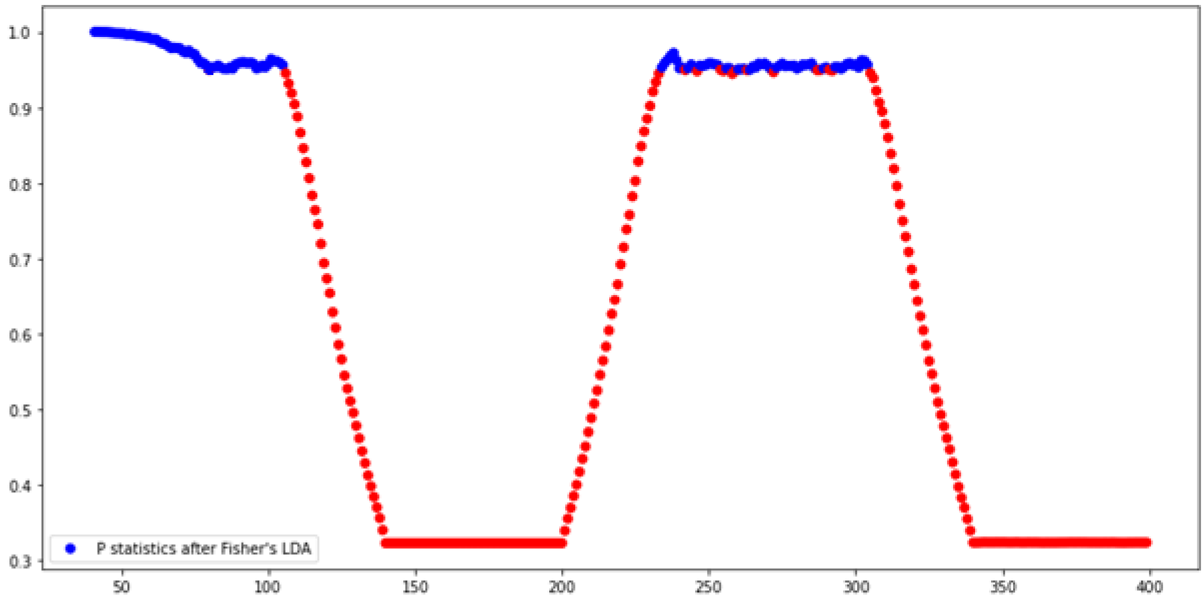


**Figure 2**: Values of the $P$ statistics for practically non-overlapping uniform distributions with different means. Blue dots indicate times when the values are not less than 0.95, and red dots indicate the opposite, when they are less

**Table 2**

Error measures for almost non-overlapping uniform distributions with different means

| Error measure | Value |
|---|---|
| MAE | 21.13 |
| MSE | 583.21 |
| MSD | 16.95 |
| RMSE | 21.13 |
| NRMSE | 0.21 |

As can be seen from Table 1 and Figure 1, the desired change point is 100. In Figure 2, we see that the p-statistic takes values greater than 0.95 only near intervals that have a distribution similar to the first one and the measures of error we can see in Table 2.

### 3.1.2. Uniform distributions with different means, which are initially strongly overlap, then slightly overlap, and finally no overlap

Let's consider a saltatory time series, which is composed of uniform distributions that are initially strongly overlap, then slightly overlap, and finally no overlap. On this time series, we will be able to test the shift hypothesis.

**Table 3**

Time intervals and uniform distributions with different means, which are initially strongly overlap, then slightly overlap, and finally no overlap

| Time interval | Distribution $T_1$ | Distribution $T_2$ | Distribution $T_3$ |
|---|---|---|---|
| 0-99 | U(60;70) | U(96.0;97.0) | U(36.4;36.7) |
| 100-199 | U(63;73) | U(96.3;97.3) | U(36.5;36.8) |
| 200-299 | U(70;80) | U(97.0, 98.0) | U(36.7;37.0) |
| 300-399 | U(85,95) | U(99.0;99.9) | U(37.5;37.8) |

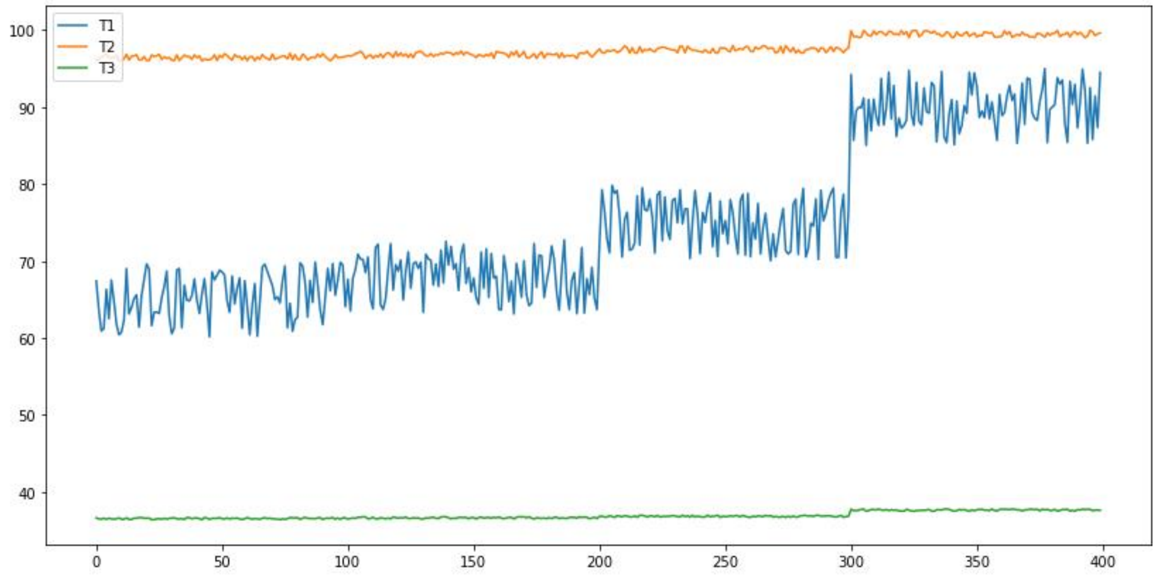


**Figure 3**: Time series consisting of samples from uniform distributions with different means, which are initially strongly overlap, then slightly overlap, and finally no overlap

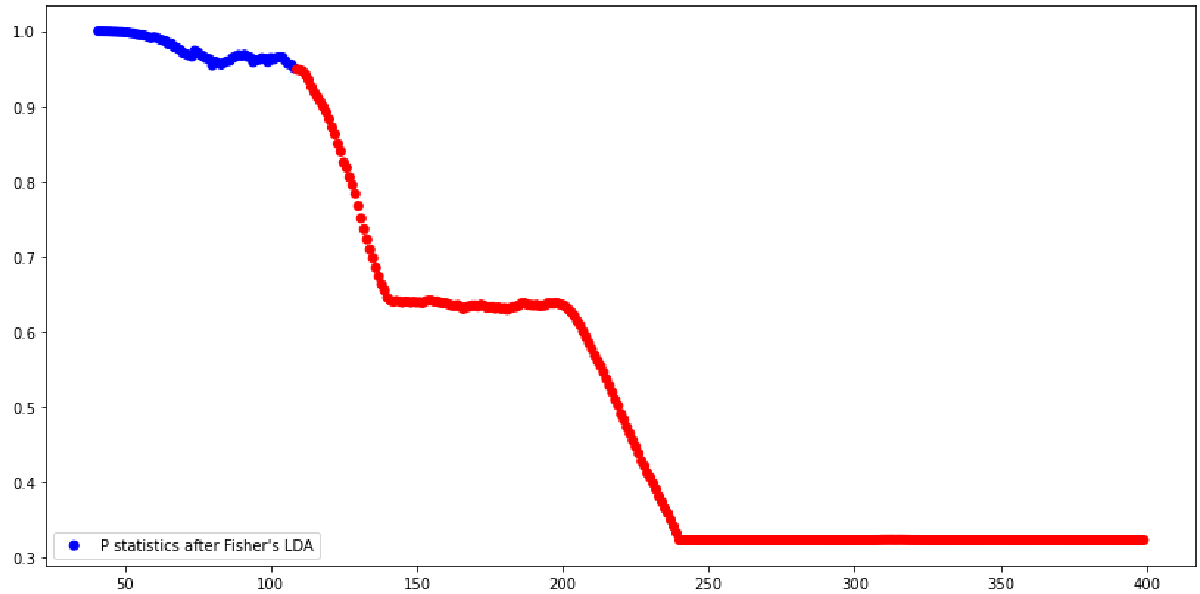


**Figure 4**: Values of the $P$ statistics for samples from uniform distributions with different means, which are initially strongly overlap, then slightly overlap, and finally no overlap. Blue dots indicate times when the values are not less than 0.95, and red dots indicate the opposite, when they are less

**Table 4**

Error measures for uniform distributions with different means, which are initially strongly overlap, then slightly overlap, and finally no overlap

| Error measure | Value |
|---|---|
| MAE | 23.21 |
| MSE | 653.63 |
| MSD | 18.01 |
| RMSE | 23.21 |
| NRMSE | 0.23 |

As can be seen from Table 3 and Figure 3, the desired change point is 100. In Figure 4, we see that the p-statistic takes values greater than 0.95 only in the first interval and the measures of error we can see in Table 4.

### 3.1.3. Normal distributions with different means that almost do not overlap

Let's consider a saltatory time series, which is composed of normal distributions with different means that almost do not overlap. On this time series, we will be able to test the shift hypothesis.

**Table 5**

Time intervals and normal distributions with different means that almost do not overlap

| Time interval | Distribution $T_1$ | Distribution $T_2$ | Distribution $T_3$ |
|---|---|---|---|
| 0-99 | N(70;2) | N(96.0;0.15) | N(36.5;0.05) |
| 100-199 | N(105;2) | N(96.3;0.33) | N(38.5;0.15) |
| 200-299 | N(70;2) | N(97.0, 0.15) | N(36.5;0.05) |
| 300-399 | N(80,4) | N(99.0;0.25) | N(37.3;0.98) |

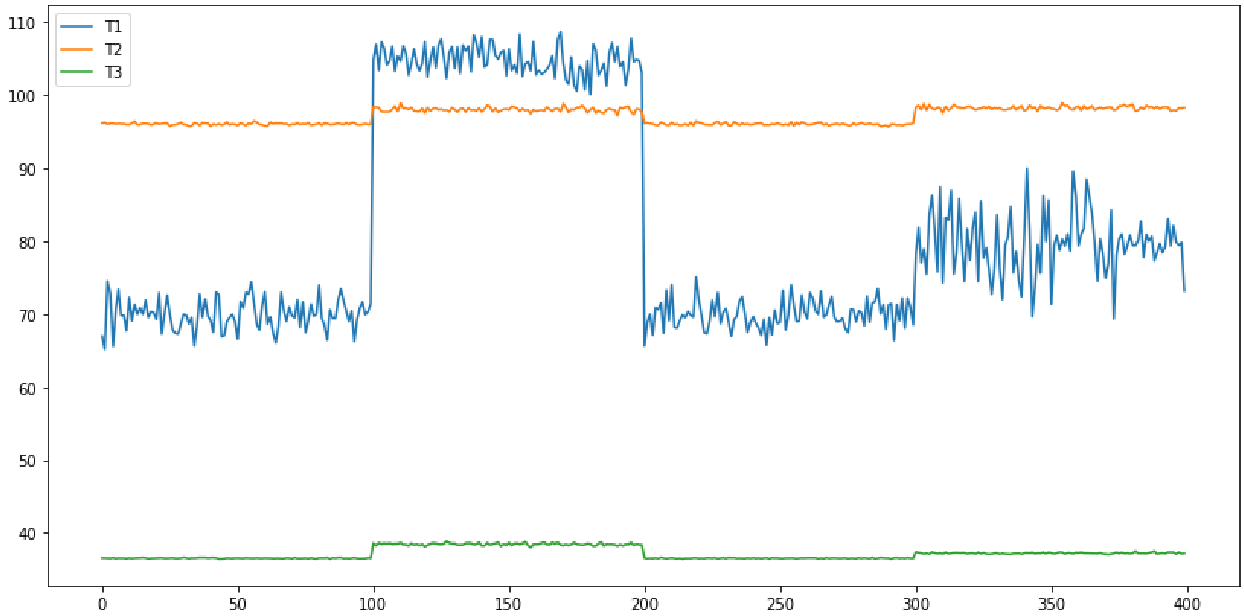


**Figure 5**: Time series consisting of samples from normal distributions with different means that almost do not overlap

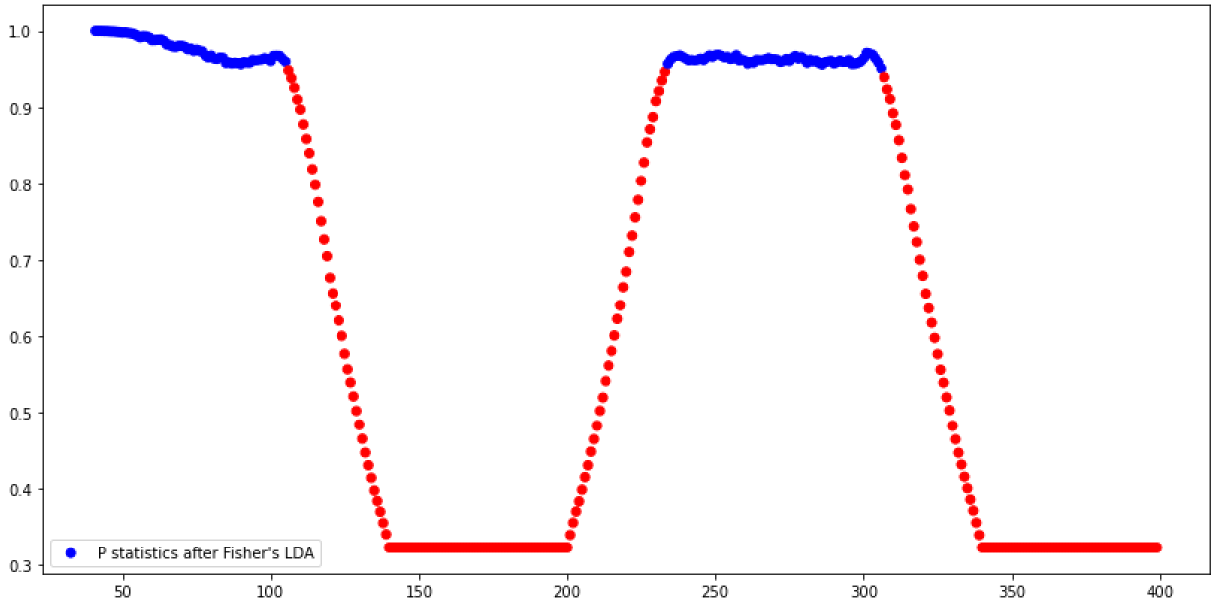


**Figure 6**: Values of the $P$ statistics for samples from normal distributions with different means that almost do not overlap. Blue dots indicate times when the values are not less than 0.95, and red dots indicate the opposite, when they are less

**Table 6**

Error measures for Normal distributions with different means that almost do not overlap

| Error measure | Value |
|---|---|
| MAE | 19.54 |
| MSE | 509.52 |
| MSD | 16.02 |
| RMSE | 19.54 |
| NRMSE | 0.19 |

As can be seen from Table 5 and Figure 5, the desired change point is 100. In Figure 6, we see that the p-statistic takes values greater than 0.95 only near intervals that have a distribution similar to the first one and the measures of error we can see in Table 6.

### 3.1.4. Normal distributions with the same means, but with variances that gradually begin to differ

Let's consider a saltatory time series, which is composed of normal distributions with the same means, but with variances that gradually begin to differ. On this time series, we will be able to test the scale hypothesis.

**Table 7**

Time intervals and normal distributions with the same means, but with variances that gradually begin to differ

| Time interval | Distribution $T_1$ | Distribution $T_2$ | Distribution $T_3$ |
|---|---|---|---|
| 0-99 | N(70;1) | N(97.0;0.10) | N(36.55;0.05) |
| 100-199 | N(70;2) | N(97.0;0.15) | N(36.55;0.10) |
| 200-299 | N(70;3) | N(97.0,0.20) | N(36.55;0.15) |
| 300-399 | N(70;5) | N(97.0,0.30) | N(36.55;0.20) |

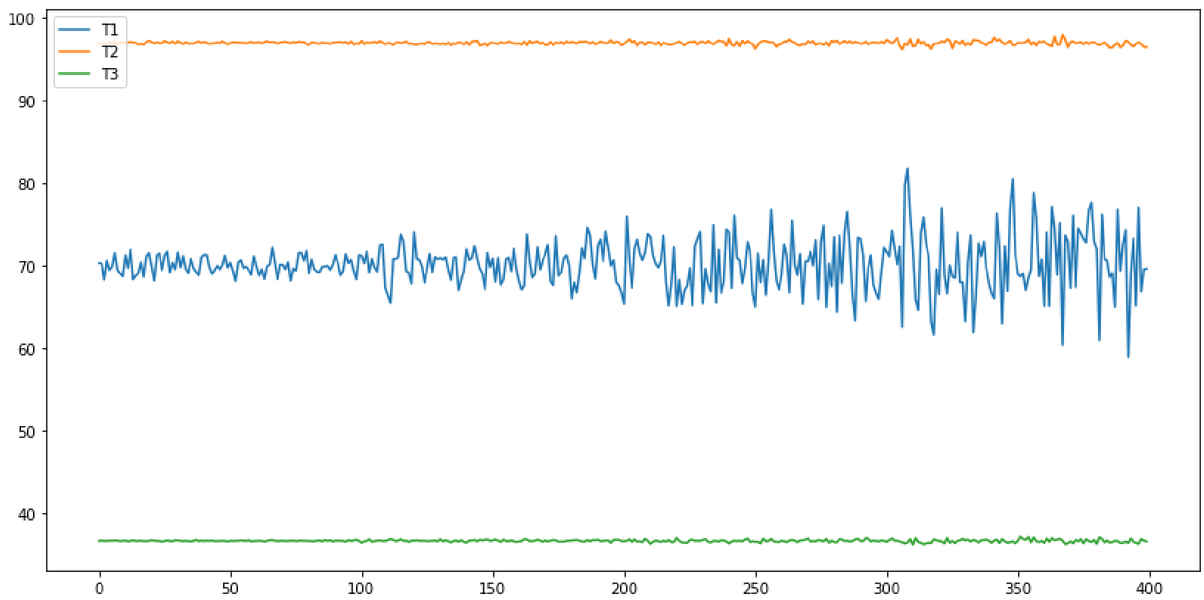


**Figure 7**: Time series consisting of samples from normal distributions with the same means, but with variances that gradually begin to differ

As can be seen from Table 7 and Figure 7, the desired change point is 100. In Figure 8, we see that the p-statistic takes values greater than 0.95 only in the first interval and the measures of error we can see in Table 8.

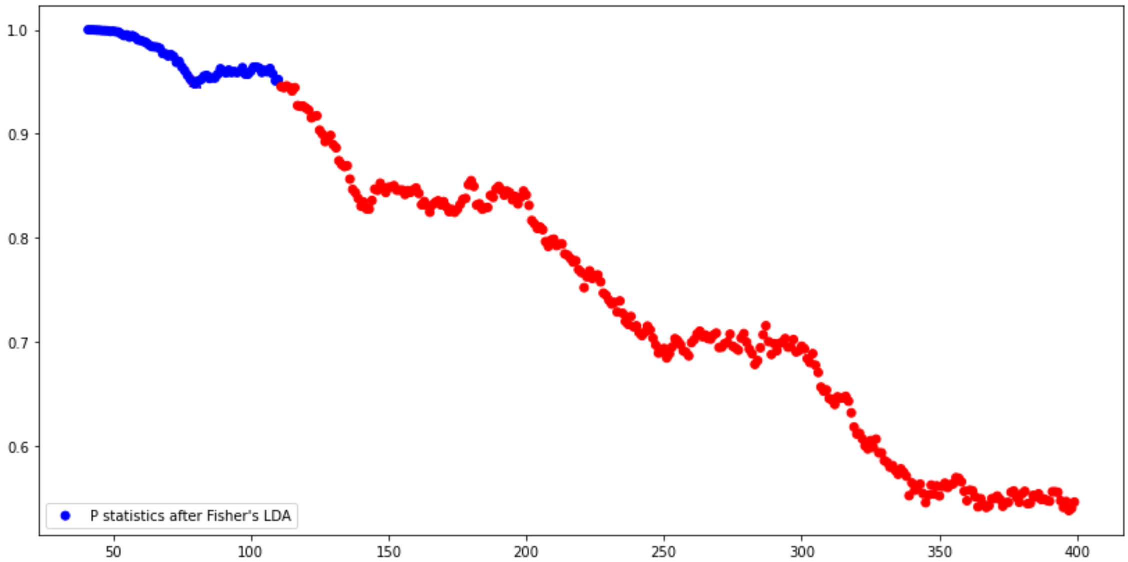

**Figure 8**: Values of the $P$ statistics for samples from normal distributions with the same means, but with variances that gradually begin to differ. Blue dots indicate times when the values are not less than 0.95, and red dots indicate the opposite, when they are less

**Table 8**

Error measures for normal distributions with the same means, but with variances that gradually begin to differ

| Error measure | Value |
|---|---|
| MAE | 21.74 |
| MSE | 594.18 |
| MSD | 12.76 |
| RMSE | 21.74 |
| NRMSE | 0.21 |

### 3.1.5. Normal distributions with the same means, but with variances that differ more strongly

Let's consider a saltatory time series, which is composed of normal distributions with the same means, but with variances that differ more strongly. On this time series, we will be able to test the scale hypothesis.

**Table 9**

Time intervals and normal distributions with the same means, but with variances that differ more strongly

| Time interval | Distribution $T_1$ | Distribution $T_2$ | Distribution $T_3$ |
|---|---|---|---|
| 0-99 | N(70;1) | N(97.0;0.10) | N(36.55;0.05) |
| 100-199 | N(70;5) | N(97.0;0.50) | N(36.55;0.25) |
| 200-299 | N(70;7) | N(97.0,1.00) | N(36.55;0.5) |
| 300-399 | N(70;10) | N(97.0,1.50) | N(36.55;0.75) |

As can be seen from Table 9 and Figure 9, the desired change point is 100. In Figure 10, we see that the p-statistic takes values greater than 0.95 only in the first interval and the measures of error we can see in Table 10.

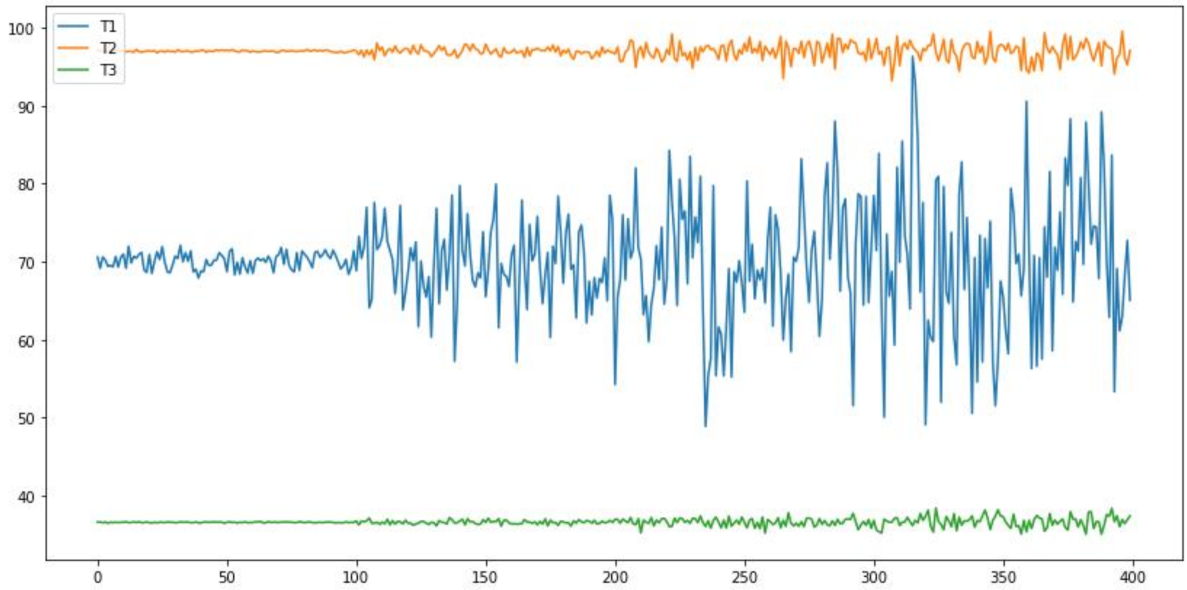


**Figure 9**: Time series consisting of samples from normal distributions with the same means, but with variances that differ more strongly

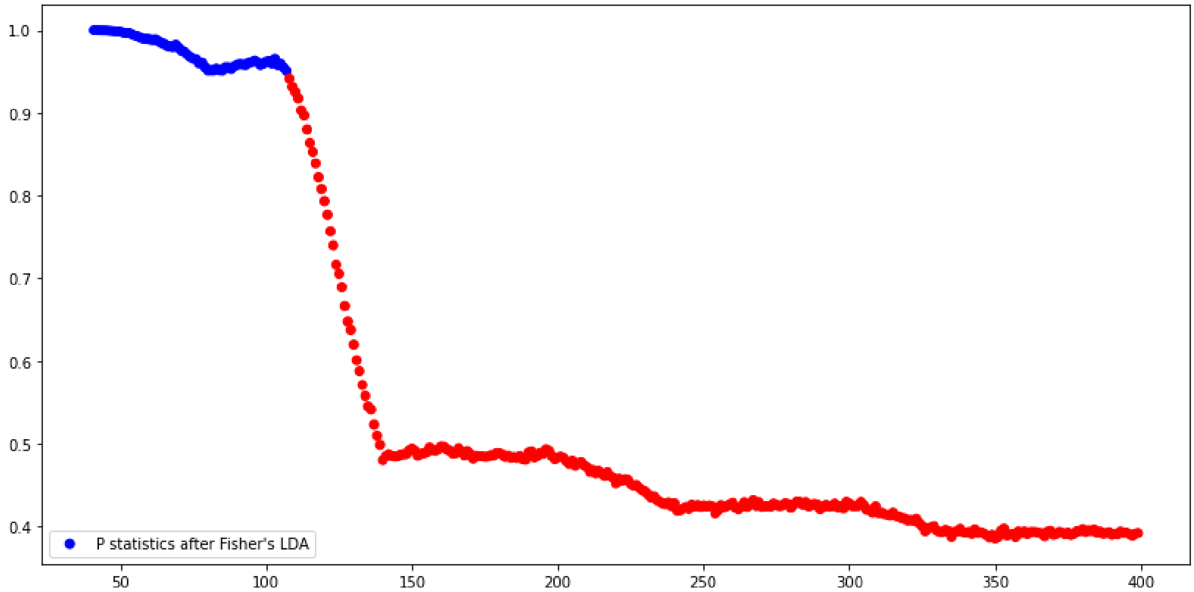


**Figure 10**: Values of the $P$ statistics for samples from normal distributions with the same means, but with variances that differ more strongly. Blue dots indicate times when the values are not less than 0.95, and red dots indicate the opposite, when they are less

**Table 10**

Error measures for normal distributions with the same means, but with variances that differ more strongly

| Error measure | Value |
|---|---|
| MAE | 20.03 |
| MSE | 537.41 |
| MSD | 13.53 |
| RMSE | 20.03 |
| NRMSE | 0.20 |

## 4. Conclusion

In this chapter, an algorithm for finding changepoints using Fisher's linear discriminant and Petunin's statistics was described. Experiments demonstrate fairly fast and accurate recognition when changing the distribution function for a wide range of distributions. This gives a clear presentation of the results, which means that this algorithm can be applied to risk-informed systems, in particular, to work in clinics, to monitor the condition of patients with coronavirus.

## References


[1] D.A. Klyushin, Y.I. Petunin, Nonparametric population equivalence test based on measure of closeness between samples, Ukrainian Mathematical Journal, 2nd. ed., 2003, pp. 147-163

[2] D.A. Klyushin, A.V. Urazovskyi, Nonparametric Test for Change-Point Detection of IoT Time-Series Data, Chapter in: Kumar P., Obaid A., Cengiz K., Balas A. (Eds.) A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems, Intelligent Systems Reference Library, volume 210, Springer, 2021, pp. 99-122

[3] R. A. Fisher, The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, volume 7, 2nd. ed., 1936, pp. 179–188

[4] B.L. Van der Waerden, Mathematische Statistic, Springer-Verlag, Berlin, 1957; English. transl. of 2nd. ed. (1965) Springer-Verlag, Berlin and New York, 1969

[5] Y. I. Petunin, D. A. Klyushin, K. P. Ganina, N. V. Borodai, R. I. Andrushkiv, Computer diagnosis of breast cancer, Bulletin of Kyiv University, Ser. cybernetics, 2, 2001, pp. 58-68

[6] S. Mika, Fisher Discriminant Analysis with Kernels, IEEE Conference on Neural Networks for Signal Processing IX, 1999, pp. 41–48. DOI:10.1109/NNSP.1999.788121

[7] C. Truong, L. Oudre, N. Vayatis, Selective review of offline change point detection methods, Signal Processing, volume 167, 2020, 107299. DOI:10.1016/j.sigpro.2019.107299.

[8] C. Alippi, G. Boracchi, D. Carrera M. Roveri, Change Detection in Multivariate Datastreams: Likelihood and Detectability Loss, Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16), 2016, pp. 1368–1374.DOI:10.48550/arXiv.1510.04850.

[9] Z. Wang, X. Lin, A. Mishra, R. Sriharsha, Online Changepoint Detection on a Budget, 2021 International Conference on Data Mining Workshops (ICDMW), 2021, pp. 414–420. DOI:10.1109/ICDMW53433.2021.00057.

[10] G. Romano, I. Eckley, P. Fearnhead, G. Rigaill, Fast Online Changepoint Detection via Functional Pruning CUSUM statistics, 2021. arXiv preprint arXiv:2110.08205v2. DOI:10.48550/arXiv.2110.08205.

[11] Z. Wang, I. M. Zwetsloot, A Change-Point Based Control Chart for Detecting Sparse Changes in High-Dimensional Heteroscedastic Data, 2021. arXiv preprint arXiv:2101.09424v1. DOI:10.48550/arXiv.2101.09424.

[12] S. Jaehyeok, A. Ramdas, and A. Rinaldo, E-detectors: a nonparametric framework for online changepoint detection, arXiv preprint arXiv:2203.03532v1, 2022.DOI:10.48550/ arXiv.2203.03532.

[13] O. Sorba, C. Geissler, Online Bayesian inference for multiple changepoints and risk assessment. arXiv preprint arXiv:2106.05834v1, 2021. DOI:10.48550/arXiv.2106.05834.

[14] L. Wendelberger, J. Gray, B. Reich, A. Wilson, Monitoring Deforestation Using Multivariate Bayesian Online Changepoint Detection with Outliers, 2021. arXiv preprint arXiv:2112.12899v2.

[15] P. Adams, D. Mackay, Bayesian Online Changepoint Detection, 2007. arXiv preprint arXiv:0710.3742v1. DOI:10.48550/arXiv.0710.3742

[16] P. Cooney, A. White, Change-point Detection for Piecewise Exponential Models, 2021. arXiv preprint arXiv:2112.03962v1. DOI:10.48550/arXiv.2112.03962

[17] J. Castillo-Mateo, Distribution-Free Changepoint Detection Tests Based on the Breaking of Records, 2021. arXiv preprint arXiv:2105.08186v1. DOI:10.48550/arXiv.2105.08186.

[18] K. L. Hallgren, N. A. Heard, M. J. M. Turcotte, Changepoint detection on a graph of time series, 2021. arXiv preprint arXiv:2102.04112v1. DOI:10.48550/arXiv.2102.04112
[19] A. Fotoohinasab, T. Hocking, F. Afghah, A Greedy Graph Search Algorithm Based on Changepoint Analysis for Automatic QRS Complex Detection, Computers in Biology and Medicine, 130, 2021, 104208. DOI:10.1016/j.compbiomed.2021.104208
[20] P. Fearnhead, G. Rigaill, Changepoint Detection in the Presence of Outliers, Journal of the American Statistical Association, 114, 2018, pp. 169–183. DOI:10.1080/01621459.2017.1385466
[21] F. Harlé, F. Chatelain, C. Gouy-Pailler, S. Achard, Rank-based multiple change-point detection in multivariate time series, 22nd European Signal Processing Conference (EUSIPCO), 2014, pp. 1337–1341. DOI:10.5281/zenodo.43927.
[22] K. Renz, N. C. Stache, N. Fox, G. Varol, S. Albanie, Sign Segmentation with Changepoint-Modulated Pseudo-Labelling, 2021. arXiv preprint arXiv:2104.13817v1. DOI:10.48550/arXiv.2104.13817
[23] C. Gallagher, R. Killick, R. Lund, X. Shi, Autocovariance Estimation in the Presence of Changepoints, 2021. arXiv preprint arXiv:2102.10669v2. DOI:10.48550/arXiv.2102.10669
[24] M. Navarro, G. I. Allen, M. Weylandt, Network Clustering for Latent State and Changepoint Detection, 2021. arXiv preprint arXiv:2111.01273v1. DOI:10.48550/arXiv.2111.01273.
[25] S. O. Tickle, I. A. Eckley, P. Fearnhead, A computationally efficient, high-dimensional multiple changepoint procedure with application to global terrorismincidence, 2020. arXiv preprint arXiv:2011.03599v2. DOI:10.1111/rssa.12695