Metadata of the chapter that will be visualized in SpringerLink

Book Title	Advanced Computing in Industrial Mathematics
Series Title	
Chapter Title	Change Point Analysis as a Tool to Detect Abrupt Cosmic Ray Muons Variations
Copyright Year	2019
Copyright HolderName	Springer Nature Switzerland AG

Corresponding Author	Family Name	**Tchorbadjieff**
	Particle	
	Given Name	**Assen**
	Prefix	
	Suffix	
	Role	
	Division	Institute of Mathematics and Informatics
	Organization	Bulgarian Academy of Science
	Address	Sofia, Bulgaria
	Division	
	Organization	South-West University
	Address	Blagoevgrad, Bulgaria
	Division	Institute for Nuclear Research and Nuclear Energy
	Organization	Bulgarian Academy of Science
	Address	Sofia, Bulgaria
	Email	atchorbadjieff@math.bas.bg
Author	Family Name	**Angelov**
	Particle	
	Given Name	**Ivo**
	Prefix	
	Suffix	
	Role	
	Division	Institute of Mathematics and Informatics
	Organization	Bulgarian Academy of Science
	Address	Sofia, Bulgaria
	Division	
	Organization	South-West University
	Address	Blagoevgrad, Bulgaria
	Division	Institute for Nuclear Research and Nuclear Energy
	Organization	Bulgarian Academy of Science
	Address	Sofia, Bulgaria
	Email	

Abstract

Recently, there have been an increasing number of studies using Big Data. They rely on large data sets of time series to detect artificial or natural patterns in processes of natural sciences and economy. The most possible outcome due to lack of rigid data processing is data contamination with abrupt drifts and regime shifts. They yield either inclusion of undetected errors or missed detection of important observations and events. Possible automatic tools for detection of regime shifts could be delivered from change point statistical methods. However, a major drawback for the most of the currently available change point (CP) methods is the challenge of complex temporal variations in non-stationary natural processes like cosmic rays observed at Earth. This kind of data analysis is applied to experimentally acquired time series from cosmic ray measurements. The observed parameters are muons produced in cosmic ray cascades in atmosphere and acquired in parallel with atmospheric and other meta-data. In this study, we test different approaches for change point detection in compound particle counting process.

Change Point Analysis as a Tool to Detect Abrupt Cosmic Ray Muons Variations

Assen Tchorbadjieff and Ivo Angelov

Abstract Recently, there have been an increasing number of studies using Big Data. They rely on large data sets of time series to detect artificial or natural patterns in processes of natural sciences and economy. The most possible outcome due to lack of rigid data processing is data contamination with abrupt drifts and regime shifts. They yield either inclusion of undetected errors or missed detection of important observations and events. Possible automatic tools for detection of regime shifts could be delivered from change point statistical methods. However, a major drawback for the most of the currently available change point (CP) methods is the challenge of complex temporal variations in non-stationary natural processes like cosmic rays observed at Earth. This kind of data analysis is applied to experimentally acquired time series from cosmic ray measurements. The observed parameters are muons produced in cosmic ray cascades in atmosphere and acquired in parallel with atmospheric and other meta-data. In this study, we test different approaches for change point detection in compound particle counting process.

1 Introduction

Identifying, quantifying, and understanding the nature of cosmic rays intensity variations at space and Earth atmosphere has been topic for many years in enormous number of different researches. Moreover, the possible consequent impact on climate and natural Earth radioactivity has been the focus of numerous recent research

A. Tchorbadjieff (✉) · I. Angelov
Institute of Mathematics and Informatics, Bulgarian Academy of Science, Sofia, Bulgaria
e-mail: atchorbadjieff@math.bas.bg

A. Tchorbadjieff · I. Angelov
South-West University, Blagoevgrad, Bulgaria

A. Tchorbadjieff · I. Angelov
Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Science, Sofia, Bulgaria

K. Georgiev et al. (eds.), *Advanced Computing in Industrial Mathematics*, Studies in Computational Intelligence 793,
https://doi.org/10.1007/978-3-319-97277-0_33

of cosmic rays, for more detailed example see the text of Dorman [1]. In most cases, the importance is derived from variations of Galactic Cosmic Rays (GCR) due to Solar perturbations and resulting of Space Weather variability in Space [2]. One of the observed effect on GCR flux are events known as Forbush decreases (FD). They are non-regular in time sudden decrease of particle fluxes which lasts from 3 to 5 days and differ in frequency and intensity [3].

Usually, the in-situ observations of CR are performed on decades-long time series data acquired by neutron and muon detectors, combined with relative atmosphere data. For their detection in large scale an automatic procedures are required for analysis of regime changes in CR variations. For this purpose, different implementations of a change point analysis were tested on available data. It is a statistical hypothesis testing for natural or artificial stochastic shifts in time series. It is mainly popular in financial mathematics, but it also gained a popularity in climatology and environmental science, or specially in detection abrupt changes in time series trend of observations in atmosphere science (see [4, 5]).

For the purpose of our research, variety different models are used for testing and detection of change points in time series with registered FD events. The used data records are acquired from the located at BEO Moussala (2925 m.a.s.l.) muon telescope [6]. The observed time series are intentionally chosen to include data that contain already detected and reported in the past FD events. This enables comparative analysis between independently acquired test results and actual situation. In addition, the data analysis must include related meteorological parameters, such as pressure, following the theoretical connection between intensity of secondary CR particle flux and atmosphere density [7]. Assuming the importance of this natural connection, two main different approaches for implementation are applied for change point analysis. The first one is direct change point analysis on modified univariate muon flux acquired after correction with ambient atmosphere pressure. The second approach is based on detection of regime changes in regression coefficients between muons and pressure. The explanation how theory of change point analysis is applied for both cases is shown in the next paragraph.

2 Change Point Analysis

The first method for regime change detection has been initially introduced as Cumulative Sum (CUSUM) sequential analysis technique by Page in 1954 [8]. The method is control chart scheme for identification of the subsamples and detection of the changes in the parameter value of sequential observations x_i with steps $i = 1, 2, \dots, n$. The method computes upper C_i^+ and lower C_i^- cumulative statistics with initial values equal to 0 such as:

$$\begin{aligned} C_0^+ &= C_0^- = 0 \\ C_i^+ &= \max(0, C_{i-1}^+ + x_i - k) \\ C_i^- &= \min(0, C_{i-1}^- + x_i - k). \end{aligned} \tag{1}$$

where k represents reference value. For example, when the shift of mean δ is known for i.i.d. and normally distributed x_i, k is equal to $\delta/2$. However, the model performs poorly for unknown δ with values significantly different than expected ones. An regime change is detected when C_i values reaches control limit h. It is a predefined value according average run length (ARL) and usually is proportional to the standard deviation.

2.1 Tests

However, despite the fact that the CUSUM sequence techniques is simple and easy to implement, usually more general statistical tool is required for multiple change point detection. Let $x_1, x_2, \ldots, x_n$ be a sequence of independent random vectors (variables) with any probability distribution functions $F_1, F_2, \ldots, F_n$. For detection of multiple change points the statistical test is run for the following alternative hypotheses H_0 vs H_A:

$$H_0 : F_1 = F_2 = \cdots = F_n \tag{2}$$

$$H_A : F_1 = \cdots = F_{k_1} \neq F_{k_1+1} = \cdots = F_{k_2} \neq F_{k_2+1} = \cdots = F_{k_{q+1}} = \cdots = F_n \tag{3}$$

where $1 < k_1 < k_2 < \cdots < k_q < n$ represents unknown number of changing points q with respective unknown positions $k_1, \ldots, k_q$. When $F_1, F_2, \ldots, F_n$ belongs to common parametric family $F(\theta)$, the null hypothesis is test about population parameters θ_i, $i = 1, \ldots, n$ and $\theta \in R^n$. Then the test is transformed to [9]:

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_n = \theta \quad \text{(unknown)} \tag{4}$$

$$H_A : \theta_1 = \cdots = \theta_{k_1} \neq \theta_{k_1+1} = \cdots = \theta_{k_2} \neq \theta_{k_2+1} = \cdots = \theta_{k_{q+1}} = \cdots = \theta_n \tag{5}$$

The computational approach to identify multiple change points is to compute [10]:

$$\min \left[\sum_{i=1}^{q+1} \left[C(x_{k_1,\ldots,k_q}) \right] + \beta f(k) \right] \tag{6}$$

where $C(x_{k_1}, x_{k_2}, \ldots, x_{k_q})$ is a cost function, usually twice negative log-likelihood. The additional part of $f(k)$ is a penalty for avoiding of over-fitting due to data size, number of change points or autocorrelation. There are many different types of proposed penalties. The most easy for implementation cost functions is Minimum AIC Estimate (MAICE) with penalty only on number of breaking points. But, because the penalty for data size is not considered, AIC based models show tendency to overfit. The criteria with implemented both of penalties on size and number of change points is Schwartz Information Criteria (BIC). Their formulation as selection among K models can be generalized by:

$$\mathrm{AIC}(k) = -2log(L(\theta_k)) + 2k, k = 1, 2, \ldots, K \tag{7}$$

$$\mathrm{SIC}(k) = -2log(L(\theta_k)) + klog(n), k = 1, 2, \ldots, K \tag{8}$$

For the computational work in this paper, the specially dedicated implementation is selected. It is the *changepoint* package [10] available for statistical computation with R [11]. The library delivers many features, which could be used directly for our research. Firstly, the implemented penalties are not constrained only to AIC and BIC, but there is an option for usage of user manually defined penalties. Secondly,the Gamma and Poisson distributions are implemented in addition to CUSUM and Normal. Another positive feature is the available variety of computational optimizations with implementations of splitting algorithms such as binary segmentation [12]; the Segment Neighbourhood [13]; and the PELT [14].

2.2 Using CP in Regression Models

In case of multivariate data, with vector of values y, the change point analysis is over regression model with a non-stochastic (p+1)-vector of predictors $\boldsymbol{x_i} = (1, x_{1i}, \ldots, x_{pi})$ is:

$$y_i = X\beta + \varepsilon_i, \ i = 1, \ldots, n, \tag{9}$$

where $\boldsymbol{\beta}'$ - is a $p + 1$ vector of unknown regression parameters and ε_i are random normally distributed errors with $N(0, \sigma^2)$. The change point analysis is observation about change of regression coefficients due to detected disconnection between the data before and after the point k. Then the statistical test is about the comparison between Null hypothesis of lack of differences in regression coefficients against the alternative of two different models with split point at k [9]:

$$H_0 : \quad \mu_{y_i} = x_i^{'}\beta \ , \ i = 1, \ldots, n \tag{10}$$

$$H_A : \begin{cases} \mu_{y_i} = x_i^{'}\beta_1, \ i = 1, \ldots, k \\ \mu_{y_i} = x_i^{'}\beta_2, \ i = k+1, \ldots, n \end{cases} \tag{11}$$

where $k = p + 1, \ldots, n - p - 1$ is the location of CP, where β, β_1, β_2 are unknown.

For computational work with regression models a specially dedicated package *strucchange* in R [15] is used. The estimated regression coefficients $\hat{\beta}$ are yielded from Ordinary least squares (OLS). The package provides computations at position k of residuals $\hat{u_i}$ and their recursive values $\tilde{u_i}$ with zero mean and σ^2 under H_0:

$$\hat{u}_i = y_i - x_i^{'}\hat{\beta}^n \tag{12}$$

$$\tilde{u}_i = \frac{y_i - x_i^{'}\hat{\beta}^{(i-1)}}{1 + x_i^T(X^{(i-1)'}X^{(i-1)})^{-1}x_i} \tag{13}$$

462814_1_En_33_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:2/8/2018 Pages: 13 Layout: T1-Standard

The decision for availability of CP is based on different implementations of CUSUM process of residuals. Then, the cumulative sum of standardized residuals $W_n(t)$, is defined as [15]:

$$W_n(t) = \frac{1}{\tilde{\sigma}\sqrt{\eta}} \sum_{i=k+1}^{k+|t\eta|} \tilde{u}_i \tag{14}$$

Thus, under H_0 $W_n \longrightarrow W$ as $n \to \infty$, which is standard Brownian Motion [15]. Similarly, the modified OLS-CUSUM process is equal to [15]

$$W_n^0(t) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=1}^{nt} \hat{u}_i \tag{15}$$

Then under $H_0 : W^0(t) = W(t) - tW(1)$, or the Brownian Motion starts at 0 and must finish there [15]. The MOSUM method is also implemented in the library, which is a moving sums of residuals.

3 Data

The used data for this research consist with datasets related to already reported and confirmed FD events. They include two periods - the first one is in the middle of February 2011, the second begins in March 1st 2012 and lasts until the end of May. We would refer to them as *Period 1* for year 2011 and *Period 2* in the following text. The time periods with their first and last dates are selected, as it is reported in [16, 17]. The two time series differ in size and number of FD evens which are included. The period of 2011 contains only single short FD event. Conversely, there are registered a chain of overleaped in time FD events and in addition to two other FD events during the *Period 2* (Table 1). Finally, an abrupt shift due to undocumented changes in the measurement procedures is available in data during the *Period 2*. This regime change is produced from higher values in the middle of May 2012. It was removed from the original report [17], but they are included in this work for more precise research. Finally, because the first detected FD during *Period 2* is on March 8th which is in only 7 rows after the first data on March 1st, the additional extended control dataset for daily data starting from Feb. 25th is assumed for verification that the short period before the event does not interfere the results.

The raw data consists of measured in 15 s vertical muon counts from 4 channels and pressure records in 10 min intervals. After both datasets are preprocessed for significant measurement errors they are synchronised in time and merged. For competitive data analysis three different datasets are assumed - two versions with hourly time resolution for periods 1 and 2 and one with a daily averages for events in 2012. The data for every test case are properly aggregated for the two different test scenar-

Table 1 Reported FD events

Start date	Until	Intensity (%)	Number of events	Reporting paper
2011–02–18	2011–02–20	4.5	1	[16][a]
2012–03–08	2012–03–17	6	3	[17]
2012–04–05	2012–04–07	2.5	1	[17]
2012–04–24	2012–04–27	3.5	1	[17]

[a]Note that the results are compared to more sensitive neutron flux, not muons

ios. This difference is required mainly because the very important physical negative dependence of CR particle flux on atmosphere density. For that reason two different type change point tests are assumed for every time period - CP regression models with implicitly included pressure as a predictor and modified muon time series with corrected pressure.

However, corrected data require some additional preparation. For correction are used β_P coefficients from regression models, OLS and Generalized Linear Models (GLM). Then the corrected values for time dependent flux $I(t)$ with averages of I_0 are equal to:

$$I_{corr} = \frac{I - I_0}{I_0} = \beta_P(P(t) - P_0) \tag{16}$$

where $P(t)$ are time relative pressure and P_0 is its average. Usually, the coefficients are computed from previously measured annual data without any significant regime shifts. However, for the current work are used coefficients computed from smaller time intervals occurred exactly before every observed period. The differences in β_P are negligible for the final outcome.

There are two important remarks that should be made for the observed data. Firstly, the overall distribution departure from the normality. The main reason are detected FD events, which skew density function of atmosphere corrected intensity with their lower values (Fig. 1). In general, the overall distribution of data with included FD events usually could not be generalized as Normal, Gamma or exponential for dif-

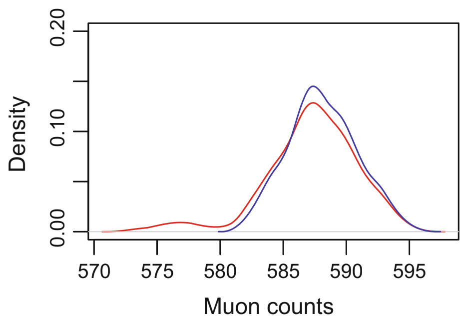

Fig. 1 A combined density plot of corrected with pressure muon data with 1 hour resolution. The selected period is the middle of February 2011, as it is shown in [16]. The density kernel of all data is with red line. The colour of data without reported FD decrease is blue

462814_1_En_33_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:2/8/2018 Pages: 13 Layout: T1-Standard

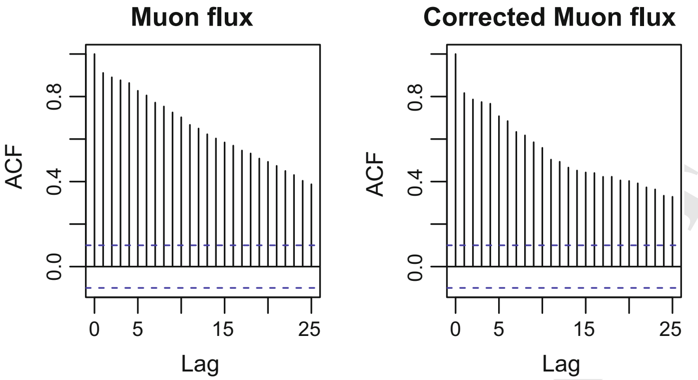

Fig. 2 Autocorrelation function for 1-h muon data (left) and for corrected with pressure values from the same data. The demonstrated period is mid-February 2011, as it shown in [16]

ferent cases. This is confirmed with the results for maximum-likelihood fittings with R function '*fitdistr*'.

Another important characteristic is the non-stationariness of aggregated data, a state common with many other natural processes [4]. The very possible explanation are complicated seasonality and trend with overlapped periods. In cosmic rays intensity, the seasonal periods are connected to rotations of Earth and Sun and daily, 27-daily, annual and 11 years cycles exist. As a result, the unit-root tests unsurprisingly deny stationary and autocorrelation decays very slowly for raw flux and pressure corrected data (see Fig. 2). To compute the optimal lags m it is used Augmented Dickey-Fuller test. All produced results are larger than 2 and the values of lags in every data set are shown in tables of results (Table 2).

4 Applications

All change point testing scenarios initially are run for tests with different penalties and segmentation algorithms without incorporation of autocorrelation properties. They are repeated for all aggregated versions of data. A brief description of implementation and most important results are explained in the following subsections.

4.1 Direct Approach

The function is **'efp'** from *structchange* library is the main method used for the change point analysis of regression coefficients. The models were build starting from the simplest regressive relation between muon counts and pressure. With assumption of periodicity, the regression model is extended with predictors of lagged data according expected periodicities. However, all test failed to produce correct results. Then, the already corrected muon data are run in regression models against their lagged values. The tests are run with pooling function **'breakpoints'**, which run all possible CP cases and select the optional model for minimal BIC or Residual Sum of Squares (RSS). With tests performed on data in 1 hour resolution using OLS-CUSUSM, the exact hour for the FD events with amplitude larger than >2.5% are detected. However, the numbers of breaking point is very strongly dependent on minimal segment size and usually differ from actual size. All other used variants failed to produce significant results. Some of the results are shown graphically in Fig. 3.

The change point analysis of corrected with pressure muon flux values are also computed with all possible functions **'cpt.mean', 'cpt.var'** and **'cpt.meanvar'** from *changepoint* library. Every distinct function is tested with all different options. Firstly, the used distributions were Normal and Gamma. They were run with all possible combinations of available segmentation methods and penalty functions. All results are either over-fitted or not complete, mainly in cases for complicated data from 2012. Some of the results are also shown graphically in Fig. 3.

In general, the proposed tests do not produce successful results. The main reason is the time dependence in measurement data, resulted with large autocorrelation lag for data either equal or shorter than 1 month [4]. Usually, this creates pattern in time series which could be easily confused with regime shift and the risk of false interpretation

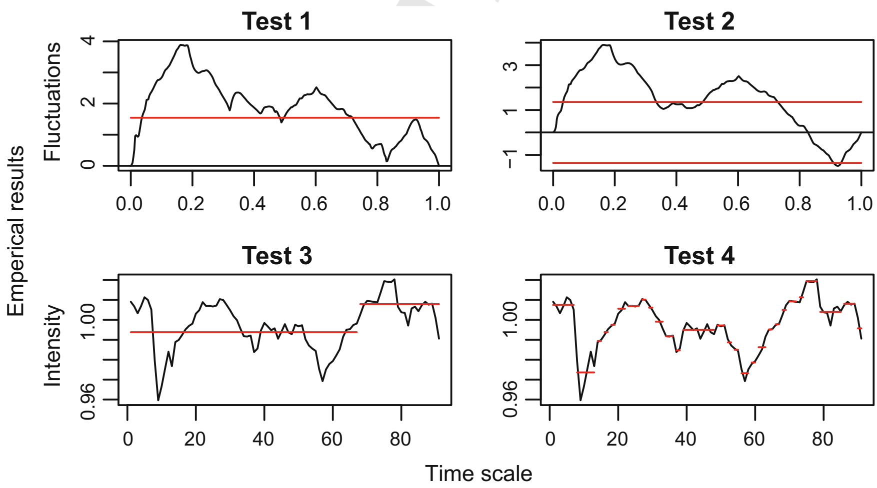

Fig. 3 Graphics show 4 tests detection of change points during the *Period 2*. Test 1 and Test 2 use regression with cumulative sums of standardized residuals for the first and OLS-CUSUM for the second test. Both of them are shown in Brownian motions scale. Tests 3 and 4 compute change point locations over pressure corrected muons reported in [17]. Test 3 and 4 are generated for hypothesis of Normal distribution with AIC penalty. The difference between those test is that for the last one is used PELT for computational acceleration

of ACT tends to increase [18]. However, the FD events represents abrupt shifts with purely random time of arrival, amplitude and the persistent time. Then, the solution for abrupt changes detection is to expand the penalty with incorporation of time dependency in the model.

4.2 Integrating the Autocorrelation

The initial steps for integration of autocorrelation are taken following the models proposed in similar works on environmental data (see [4, 5]). The daily data is computed with different versions of manually pre-computed SIC penalty with m-th order autocorrelation using **'cpt.meanvar'** function from *changepoint* library. The tests confirmed relatively precisely the period with extreme solar activity between March 8th and 19. The FD event in April 25 is also detected, but with a day earlier in comparison to the reported time in [17]. However the disturbances in beginning of April are not detected as separate event, but as a part of larger regime change until April 22. The results are shown in Fig. 4.

However, the tests were not successful in cases of data with hourly time resolution. It is a complex problem because we have data with compound periods incorporated and the proposed autocorrelation integration is on process with not fully removed time dependence. Moreover, the regime shifts of mean and variation usually correlates. Thus, a solution must implement approximation of well known tailed distribution with penalty correction representing the assumption of departure from the main distribution due to additional autocorrelation and regime shift. For this purpose we use a model with fixed penalty that enables the variance V to vary with the mean of distribution:

$$V(\mu) = \phi\mu^{s} \tag{17}$$

where ϕ is dispersion parameter. This constant ratio between mean and variation imply assumption for Tweedie distribution [19]. This is polymorphic distribution,

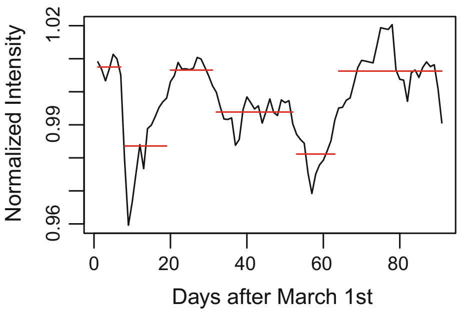

Fig. 4 CP analysis of *Period 2*. The two largest FD events are detected. The change point staring from April 5-th is detected, but the related FD is not distinguished

462814_1_En_33_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:2/8/2018 Pages: 13 Layout: T1-Standard

which with change of s, takes properties of well known distributions. For example, with $s = 2$ it is Gamma, but when $s = 0$ is equal to Norma distribution.

The newly modified model is based on data fitted to Gamma distribution (α, θ), where θ is rate ratio $(1/\mu)$, which inverse FD data on right-tailed. The CP analysis is performed by **'cpt.meanvar'** function with manually selected fixed penalty on variations equal to $\alpha Var(x)$. This, due to properties of Gamma distribution, is equal to $\alpha^2\theta^2 = \mu^2$, or just Tweedie fixed variation for Gamma distribution. Thus, any departure from Gamma may be penalized with $\alpha Var(x)^k$, where the power k enables correction on autocorrelation.The coefficient k is computed as $1 - 1/r$, where r is equal to:

$$r = log(n/m - 1)log(n - m - 2)\frac{1-\rho}{1+\rho}, \tag{18}$$

and n is a number of observed values, m is autocorrelation lag and $\rho < 1$. When the data is stationary, the value of k is equal to 1. When, the lag of autocorrelation is very large, $0 < s < 1$,the range where Tweedie distribution is not defined, thus k must be limited for values above 0.5. The value of $n - m - 2$ is yielded from reordering of the CP model following the assumption that number of CP must be less than n-m-2. The last part represents effective size correction as it is described in [5] with ρ equal to autocorrelation coefficient. Due to very large size n, the current formula is corrected in cases of hourly data as the first part is changed from $log(n/m - 1)$ to $log(n/m^2)$.

The results are obtained from computations performed over the very same datasets for Normally and Gamma distributed statistics with proposed fixed manual penalty. The tests for daily data in 2012, show that proposed model in [4] over-perform any tests with Gamma with fixed penalty. The main disadvantage in Gamma based models are their lesser sensitivity which leads to cut-off last 3 days of the biggest event in March 2012. However, both test with daily data missed the end of FD at April 5th 2012. The summarized results are shown in Table 2.

AQ3

However, according the tests with hour long time resolution, the models with proposed Gamma statistics with autoregression corrected penalty equal to r in Eq. 18 perform better than all other tests. It is important to be noticed that for events in 2012, the big disturbance in March is split on 2 parts, which exactly corresponds to report

Table 2 .

Dataset	n	m	k	Comparison with model of AR(q) correction in [4]
Hourly 2011[a]	382	7	0.63287	Better. Removes non-related CP a day before
Hourly 2012	2137	12	0.4953	Better. The large disturbance in March is slitted
Daily 2012	91	4	0.82696	Worse with March events shorter with 3 days

[a]Note that the data is about more sensitive neutron flux, not muons

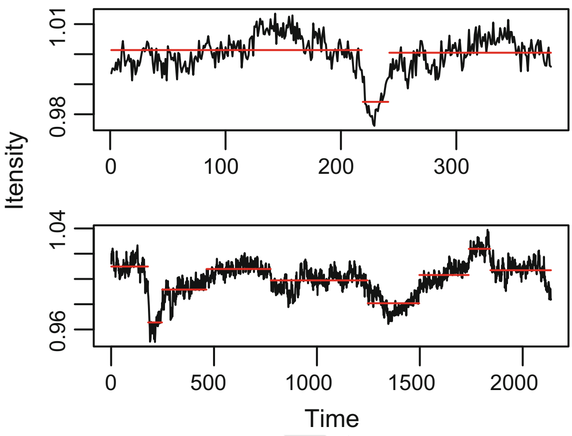

Fig. 5 Detected CP for data with 1 h resolution for both periods. The FD event in February 2012 is shown in figure above. The graphics below shows all events for observed period of 2012. The periods are hours after first hour

[17]. Secondly, the test also detected erroneous regime change in measurements after May 17th, which is shown as a increase of intensity. Thirdly, the begin times the FD events are also correctly detected. However, the sensitivity issue for the end of disturbances on April 5th 2012 remained (Fig. 5).

5 Conclusions

This paper presents first ever work on implementation of change point analysis of automatic detection of Forbush decrease with secondary CR muon flux. The work is tested on two independent in time events with two different approaches to the data. The used CP models, mainly split on dependence of relation between muons and pressure, performed with different success in comparison to already published reports. Most of tested models show weakness, that is partly solved with proposed test modification with incorporation of autoregression. However, some issues remained open without answer or for improvements.

Firstly, the sensitivity remained as issue on particular event on April 5th 2012. But the conclusion on effectiveness of CP analysis on CR muon data may not be drawn firmly. It is important to remember that the Solar disturbances and their impact on CR flux are not fully understood and they are still in research. Secondly, a possible connection with the theory of long memory processes is not investigated in this paper. Any extension to mixture model with implementation of Autoregressive fractionally integrated moving-average models (ARFIMA) could be investigated in future. Thirdly, the automatic procedure could be improved with incorporation of penalty intervals with CROPS option from ‘*changepoint*’ functionalities. Finally, all results and improvements must be done after more extended parallel research of Space Weather and CP analysis on cosmic rays data.

References

1. Dorman, I.: Cosmic rays and space weather: effects on global climate change. Ann. Geophys. **30**, 9–19 (2012)
2. Lilensten, J., et al.: What characterizes planetary space weather. Astron. Astrophys. Rev. 22–79 (2014)
3. Cane, H.: Coronal mass ejections and forbush decreases. Space Sci. Rev. **93**, 55–77 (2000)
4. Beaulieu, C., Chen, J., Sarmiento, J.: Change-point analysis as a tool to detect abrupt climate variations. Phil. Trans. R. Soc. A **370**, 1228–1249 (2012). https://doi.org/10.1098/rsta.2011.0383
5. Siedel, D.J., Lanzante, J.K.: An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes. J. Geophys. Res. **109** (2004). https://doi.org/10.1029/2003JD004414
6. Angelov, I., Malamova, E., Stamenov, J.: Muon telescopes at basic environmental observatory Moussala and South-West University Blagoevgrad. Sun Geosphere **3**(1), 20–25 (2008)
7. Dorman, L.I.: Cosmic Rays in the Earths Atmosphere and Underground. Springer Netherlands (2004) https://doi.org/10.1007/978-1-4020-2113-8
8. Page, E.S.: Continuous inspection schemes. Biometrika **41**(1–2), 100–114 (1954)
9. Chen, J., Gupta, A.K.: On change point detection and estimation. Commun. Stat. Simul. Comput. **30**(3), 665–697 (2001)
10. Killick, R., Eckley, I.: changepoint: an R package for changepoint analysis. J. Stat. Softw. **58**(3), 1 (2014)
11. R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2012)
12. Scott, A.J., Knott, M.: A cluster analysis method for grouping means in the analysis of variance. Biometrics **30**(3), 507–512 (1974)
13. Auger, I.E., Lawrence, C.E.: Algorithms for the optimal identification of segment neighbourhoods. Bull. Math. Biol. **51**(1), 39–54 (1989)
14. Killick, R., Fearnhead, P., Eckley, I.A.: Optimal detection of changepoints with a linear computational cost. JASA **107**(500), 1590–1598 (2012)
15. Zeileis, A.: strucchange: an R package for testing structural change in linear regression models. J. Stat. Softw. **7**, 1 (2002)
16. Abbrescia, M., et al.: Observation of the February 2011 Forbush decrease by the EEE telescopes. Eur. Phys. J. Plus (2011). https://doi.org/10.1140/epjp/i2011-11061-5
17. Tchorbadjieff, A.: Detection of Coronal Mass Ejections (CMEs) in the period of March–May 2012 at Moussala Peak. Proc. Bul. Acad. Sci. **66**(5), 659–666 (2013)
18. Norwood, B., Killick, R.: Long memory and changepoint models: a spectral classification procedure. Stat. Comput. **28**(2), 291–302 (2018)
19. Tweedie, M.C.K.: An index which distinguishes between some important exponential families. In: Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference. Indian Statistical Institute, Calcutta pp. 579–604 (1984)

Author Queries

Chapter 33

Query Refs.	Details Required	Author's response
AQ1	Please check and confirm if the authors and their respective affiliations have been correctly identified. Amend if necessary.	
AQ2	Please confirm if the inserted city and country name are correct. Amend if necessary.	
AQ3	Please provide captions for 'Table 2', as they are mandatory.	
AQ4	Please check and confirm if the inserted citation of 'Fig. 5' is correct. If not, please suggest an alternate citation.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	,
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	