Space weather study using change point analysis for in situ observations of cosmic rays muons

Cite as: AIP Conference Proceedings **2075**, 090011 (2019); https://doi.org/10.1063/1.5091225
Published Online: 26 February 2019

Assen Tchorbadjieff, and Ivo Angelov

Space Weather Study Using Change Point Analysis for in situ Observations of Cosmic Rays Muons

Assen Tchorbadjieff[1,a)] and Ivo Angelov[2,3,b)]

[1]*Institute of Mathematics and Informatics, Bulgarian Academy of Science, 8 Acad. G. Bonchev Str., Sofia 1113, Bulgaria.*
[2]*South-West University "Neofit Rilski", 66 Ivan Michailov Str., Blagoevgrad 2700, Bulgaria .*
[3]*Institute for Nuclear Research and Nuclear Energy, Bulgarian Academy of Science, 72 Tzarigradsho Shosse Blvd, Sofia 1748, Bulgaria.*

[a)]Corresponding author: atchorbadjieff@math.bas.bg
[b)]i_angelov@yahoo.com

Abstract. Coronal mass ejections (CMEs) are non-regular solar events, frequently connected with Space Weather disturbances, some of them with severe magnetic storms. The significant releases of plasma with CMEs impact modulations of Galactic Cosmic Rays (GCR) causing sudden and casual decrease of Cosmic Rays (CR), known as Forbush Decrease. For detection of such kind effects, uninterrupted in situ measurements of secondary CR muons may be used. To deal with the very fast growing data, an application based on Change Point (CP) analysis is used. It is an automatic statistical tool for sequential test analysis of abrupt regime changes in CR time series. For the study, results are reported with in-situ measured data at Moussala peak (2925 m.a.s.l.) for the period of 4 years - from 2014 to end of 2017 year and confirmed with satellite data for geoeffective solar activity connected with detected FD events.

INTRODUCTION

The CMEs are large scale plasma eruptions from the solar atmosphere into the interplanetary medium with particle speed in range of 20 – 2000 km s^{-1}, see Ref. [1]. Usually, the high speed CMEs are associated with flares, but conversely to the flare associated events that decelerate close to the Sun, the CMEs increase their speed. When the ejecta is with very high speed, the solar wind (SW) usually cause disturbances of geomagnetic activity. However, it is not very clear why some CMEs do not cause geomagnetic storms. For example, 132 Earthdirected CMEs were studied in Ref. [2], but only 45% of them caused severe geomagnetic storms with planetary index $Kp \geq 5$. From them, there are no connection between transit time and Kp until index values are not extended to above 7 ($Kp \geq 7$). However, despite that the properties of CMEs are very important driving force for severe magnetic storms, additional factors in coupling either or not with CMEs are important to magnetic activity - such as SW speed, pressure pulses or extended periods of southward IMF (negative Bz).

These irregular plasma ejections can be detected remotely by imaging and inference of different electromagnetic waves, such as X-ray, EUV or radio. The plasma, particles and magnetic materials also can be measured in Heliosphere by satellites, such as NASA SOHO. The methods for storm detections with estimation of particle fluxes are mainly based on identification of Forbush decreases (FDs) caused due to magnetic disturbances and Auroras or Ground Level Enhancements (GLE) detection for solar particle events. The first one, FD is abrupt decrease of Galactic cosmic rays (GCR) in their lower energy spectra due to particle deflection caused from disturbances in Interplanetary magnetic field (IMF). Firstly, it was detected by Forbush in 30-ties of 20th century. Then, with advancements in space study, the knowledge for FDs has been expanding. According the classification in Ref. [1] and [3], the largest FD events with usual decrease of particle flux above 5% are caused from CMEs and they start with sharp decrease in initial stage and relatively short period of restore. The recurrent events associated with high speed flows from coronal holes usually start with slow decrease of particle flux and reach minimum after a few days before slow restoration.

10th Jubilee International Conference of the Balkan Physical Union
AIP Conf. Proc. 2075, 090011-1–090011-6; https://doi.org/10.1063/1.5091225
Published by AIP Publishing. 978-0-7354-1803-5/$30.00

Less expensive methods to detect FDs are measurements with uninterrupted in-situ particle detectors for secondary neutrons and muons. Measuring the neutrons is a universal method to detect all possible solar particle events in addition to FDs. Conversely, the muon telescopes are not applicable to detection of Auroras and GLEs because the phenomenology of air showers' hard component. However, their sensitivity is enough for detection of FD events and their anisotropies. Moreover, their importance is growing with the need of more detailed impact estimations of secondary CR fluxes on advanced high energy physics experiments as CMS or neutrino studies. Hence, automatic tools for their appropriate, quick and full usefulness on purpose are required. For this task, a statistical tool based on change point analysis for detection of regime changes in muon flux time series is implemented. The results and analysis of muon data acquired form different large periods in between January 1st 2014 and December 31 2017 are compared and verified with publicly available data for solar activity from National Oceanic and Atmospheric Administration (NOAA) Space Weather Prediction Center (SWPC), Ref. [4].

DATA

The used in this research data are acquired from Muon telescope at Basic Environmental Observatory (BEO) at Moussala peak, the highest point in South-west Europe at about 2925m a.s.l. It consists of lead covered water Cherenkov detectors connecting 12 coincidence channels. The Cherenkov light is generated if the energy of traversing muon is high enough that its speed is greater than the speed of light in the water. Acquired data are time series of counting number of charged particles, mainly muons, from all 12 channels with time resolution of 15 seconds. The raw data are used for recomputing coincidence values for 5 different particle traverse directions - vertical, East-West, West-East, North-South and South-North. For more detailed description of construction and some already reported results see Ref. [5] and [6].

Data preprocessing

Despite that the telescope is in permanent operational regime, there are different interruptions in measurements. The main reasons for them are difficult conditions in high mountain stations. Some of them may be brief electricity interruptions, but others are more severe, such as equipment failure. These are the cases for both larger then month interruptions of the telescope operation in 2015 and 2017. However, there are other brief temporal and unpredicted technical conditions that could interfere in quality of measurements without abrupt interruptions. For their removal, all data are passed through local data filter with 3-σ confidence intervals over 60 minutes time window. The outcomes that remain outside from intervals are removed. Then, time dependent flux $I(t)$ with averages of I_0 is corrected for atmosphere density variations with pressure measured at Moussala with the standard procedure:

$$I_{\text{corr}} = \frac{I - I_0}{I_0} = \beta_P (P(t) - P_0), \tag{1}$$

where $P(t)$ is time relative pressure and P_0 is its average. The coefficients β_P are yielded from linear regression for preselected yearlong reference period. Finally, for our research, data are aggregated to daily and 1-hour averages. However, due to data quality reason, only periods with availability of at least 50% of valid data are included. The periods that do not fulfil this requirement are removed. In summary, the number of excluded values are less than 0.5% from every studied period.

TOOL FOR AUTOMATIC DETECTION

For detection of any possible regime change we consider identification of $\tau_{1:m}$ changepoint positions. Their detection is based on hypothesis tests with null hypothesis, H_0, corresponding to no changepoint ($m = 0$) and the alternative hypothesis, H_A for multiple m-number changepoints. The decision for validity is yielded after computation of [7]:

$$\min[\sum_{i=1}^{q+1} [C(x_{k_1,\dots,k_q})] + \beta f(k)], \tag{2}$$

where $C(x_{k_1}, x_{k_2}, ..., x_{k_q})$ is a cost function, mainly negative log-likelihood. The additional part of $f(k)$ is a penalty for avoiding of possible over-fitting due to data size, number of CP or autocorrelation. Most used and easy to implement cost functions are Minimum AIC Estimate (MAICE) with penalty on number of breaking points and Schwartz Information Criteria with penalty on size and number of CP simultaneously, see Ref. [7].

However, aggregated and corrected with pressure muon data usually are nonstationary time series (autocorrelation lag larger than 2). For this reason, we used specially modified CP analysis for mean and variance with integration of autocorrelation. It is implemented with modified penalty that enables the variance V to vary with the mean of distribution μ:

$$V(\mu) = \alpha\mu^k, \tag{3}$$

where α is dispersion parameter. The implementation used for aggregated muon flux is set to use option for manually defined penalty for *cpt.meanvar* function for CP computations from the package *changepoint*, designed for on statistical environment R [8] as it is described in Ref. [7]. The used probability distribution family is Gamma($\alpha, \frac{1}{\mu}$) and parameter k is equal to $1 - 1/r$, where values of r is computed from the following formula [9]:

$$r = \log(n/m^2)\log(n - m - 2)\frac{1 - \rho}{1 + \rho}, \tag{4}$$

n is total length of time series, m is autocorrelation lag, usually greater than 2, and ρ is the coefficient at lag m.

The proposed CP tool with daily data is used to estimate the strongest events. However, the storms with lower power could be missed due to many different reasons - overall intensity, many other events with the same low amplitude and etc. Moreover, in many cases geomagnetic activity shows multi-fractal scaling [10]. For this reason, the workflow is to continue change point analysis on predefined sub-periods. These new periods follows changing point split in previous CP analysis. Thus, the computational process repeats iteratively those in previous steps. However, because the sub-periods are usually very shorter than the previous one, a better time resolution is required - from daily to hour span. However, this resolution change require correction of coefficient r from formula 4. Its value must be corrected with substituting $\log(n/m^2)$ to $\log(n/m - 1)$, see Ref. [9].

RESULTS

The solar activity was very strong during year 2014, the first year of selected period. Then the trend was been decreasing until its lowest level at 2017. Because of this, the frequencies of severe storms are not regular and amplitude of FD during the first year overestimates those in the last. Because of this the change point analysis is performed separately for data from different years. Finally, the resulted data are normalized for every period separately. This is done intentionally, to achieve better sensitivity for periods with lower activity. But, it is important to note, that the scales of amplitudes of FD for different periods are not compatible.

The process is strictly straightforward - the CP software for 1 day annual data is run for detection of events with higher intensities for selected period. Then, the FDs with sudden sharp decrease are firstly classified. The identification of those FDs with slow and long decrease is next. Then identification of events related to other disturbances follows. After that, this process is repeated recursively for sub-events that have enough length. They are preselected sub-periods from previous iterations with higher level disturbances based on change point analysis. This loop could continue up to several times with only requirements for change of time resolution.

Then, the results are compared to NOAA reports. For event verifiers are used mainly the following data:

- the number of high intensity solar flares with up to 1 week time span of advance.
- the maximal reported speed of solar wind
- reports for fully or partially geoeffective CMEs and Coronal Hole High Speed Streams (CH HSS)

Data for IMF Bz component are also used in this analysis. However, to avoid extra additional explanations only Kp index is used in this report. Finally, the differences between the lowest value and last one before event beginning are used for measure of FS's amplitude. However, because data is normalized over its own period, this measure is not valuable for intra-year comparison.

Solar Events during 2014

The year 2014 is characterized with very active solar activity. The number of weeks with any kind of storms is larger than this with calm space weather. This complicates change point analysis and many weaker events remained unidentified with first iteration, shown in Figure 1. From there, 5 FDs events with step initial phase were identified, and all of them are occurred due to near and directly geoeffective CME episodes (Table 1). The most important conclusion from these results is that the higher intensity is connected to combined influence from CMEs and CH HSS on magnetic field. In addition to them, another 3 recurrent FD events on the beginning of May 22, September 27 and December 1st were detected. The first event was preceded by weak decrease starting May 17. Then during the event WS reached up to 500 km s^{-1}. The most probable reasons for the second symmetric FD are multiple southward Bz and current sheet crossings.

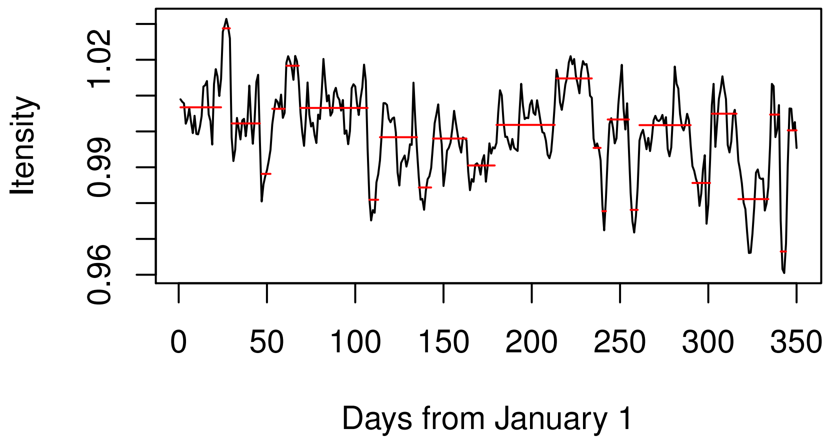

FIGURE 1. Results from Change Point analysis of detected muon flux during 2014. Data is atmosphere corrected and aggregated to 1 day resolution. Different regime changes are shown with changing up/down horizontal red lines.

TABLE 1. Brief description and time periods of large FD events caused by geoeffective CMEs detected during 2014.

	FD decrease max %	**CMEs number**	**Flares (class M/X) number** *	**Max SW Speed km s^{-1}**	**CH HSS available**	**Kp Index Max**
15-17 Feb	3.5	2	5(M)	487	yes	5
19-27 Feb	1.9	1	5(M)	743	no	6
19-25 April	3.9 †	2	1(M)		yes	5
12-14 Sep	2	2	3(M) and 1(X)	790	no	7
21-25 Dec	4.1	3	4(M) and 1(X)	450	yes	6

* Total number of reported flares during earthward CMEs eruptions.
† Interplanetary shock is available.

In addition, 4 additional sub-periods with different regimes during the initial CP analysis were identified. They last longer than discovered FDs. After CP analysis of their hourly data 2 smaller but significant FDs were discovered (see Table 2). Additionally to them, a weak FD is detected on around June 7th due to CH HSS and high speed SW. Unfortunately, the telescope were not operational for almost a week since May 29, a period that coincide with another strong solar activities. Finally, a large period of consecutive storms was observed in between July 7th and mid-August. But none of them were connected with any CME.

Solar Events during the rest of the period

The solar activity began to decrease in the following years. However, the data quality for periods closest to 2014 year was problematic. Firstly, a lot of data were removed due to data quality process for the period from January and May

TABLE 2. Brief description and time periods of weak FD events detected during 2014.

	FD decrease max %	CMEs number	Flares (class M/X) number *	Max SW Speed km s^{-1}	CH HSS available	Kp Index Max
27 Feb- Mar 1	1.5	1	3(M) and 1(X)	519	no	6
late 5-10 Apr	1.6	1	2(M)	525	no	4

* Total number of reported flares during earthward CMEs eruptions.

2015. Then, the telescope suffered serious technical damages after an accident at the end of May. The measurements were resumed after September 20. For this reason only sporadic periods were observed. The most important one is the period of March. For change point analysis data records with 1 hour resolution are used (Figure 2). The Sun was active during the period with a number of storms with effect on Earth. The event with highest intensity was arrival of CME on March 17th. The event persisted also on the next day. What is the most important for this event is that it was precedented by changing in time storm conditions since March 15. However, the results from CP computations split correctly CME arrival on March 18 from the previous disturbances.

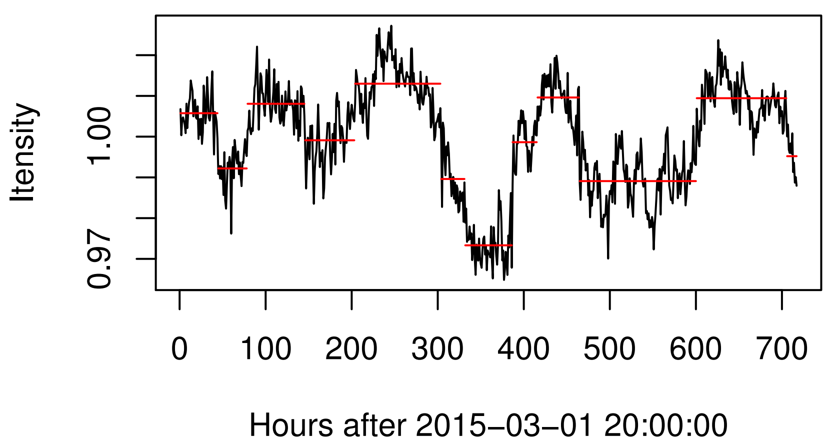

FIGURE 2. Results from Change Point analysis of detected muon flux during March 2015. Data is atmosphere corrected and aggregated to 1 hour time resolution. Different regime changes are shown with changing up/down horizontal red lines.

The muon telescope worked without interruptions during 2016. Solar activity was moderate through the year, but none earth-directed CME arrived. Meanwhile, 8 different periods with disturbances were detected. The main sources for these storms were strong solar wind and CH HS streams. However, the intensity of some decreases was moderate. For example, the strongest storms during 2016 with more than 3% decrease began on 14 June due to very strong wind with speed of about 772 km s^{-1} and CH HSS. However, it must be remarked that the amplitude of decrease is not compatible with FDs in 2014, because data for the period were normalized only over records from 2016 (Figure 3).

The solar activity during 2017 was also low. Unfortunately, the telescope was not operational during the summer, when some storms, such as this around August 17th, were reported. Because of these data gaps and possible data quality uncertainties, only data from August 1st until the end of year were studied. The period was calm, with only a single FD on September 8 connected to two Earth-directed CMEs, CH HSS and solar wind speed up to 680 km $s^{-}1$. There were also two additional periods with multiple magnetic disturbances. They were detected between 25 September-18 October and from 29 November to 16 December. During these periods the muon particle flux intensities decreased moderately for longer time mainly due to many CH HSS and high speed solar wind. But, there were not detected any Earth-directed CMEs.

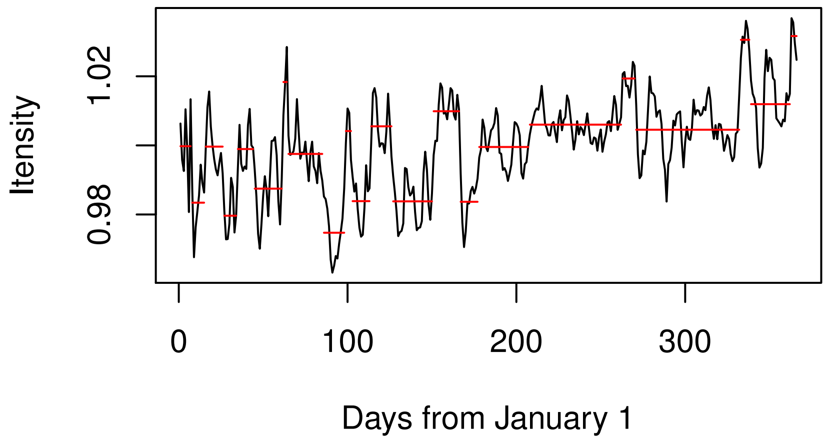

FIGURE 3. Results from Change Point analysis on detected muon flux during 2016. Data is atmosphere corrected and aggregated to 1 day span. Different regime changes are shown with changing up/down horizontal red lines.

CONCLUSIONS

The results shown in this report demonstrate efficacy of using changing point analysis of GCR particle flux for in situ detection of middle and strong geoeffective solar storms. The implemented automatic CP detection software was easy to run, fast and small resource consuming. Moreover, there are not detected neither false alarms nor missed serious or mid-size geoeffective storms during the observed period. Having this sensitivity, the further extension of number of detected events would expand the statistical significance for study of connection between solar activity, ejecta transport and geoeffectiveness. This could be proceeded with the following steps:

- After detection of the active solar periods, the data could be renormalized over the full period. It is easy to notice that the average particle flux in 2014 was about 2-4% lower than usual due to very strong solar activity.
- Similar CP analysis would be run for data acquired during the whole 24-th solar cycle. The obtained statistics, including active periods in 2011-13 years, would be representative for more detailed statistical analysis of connection between FDs and CMEs, CH HSS, wind speed and state of magnetic field. In addition, time delays between eruptions and FD detection could be estimated.

REFERENCES

[1] H. Cane, Space Science Review **93**, 55–77 (2000).
[2] Y. Wang and et al., Journal of Geophysical Research **93(A11)**, SSH 2–1–SSH 2–9 (2002).
[3] A. Belov and et al., Advances in Space Research **27(3)**, 625–630 (2001).
[4] N. S. W. P. Center, https://www.swpc.noaa.gov/ (2018).
[5] I.Angelov and et al, Sun and Geosphere **3(1)**, 20–25 (2008).
[6] A.Tchorbadjieff and et al, Bulg. J Astronomy **18(2)**, 74–83 (2012).
[7] R. Killick and I. Eckley, Journal of Statistical Software **58(3)**, 1–19 (2014).
[8] R Core Team, https://www.R-project.org/ (2018).
[9] A.Tchorbadjieff and I.Angelov, "Change point analysis as a tool to detect abrupt cosmic ray muons variations," in *Advanced Computing in Industrial Mathematics*, Studies in Computational Intelligence 793, edited by K. G. et al. (Springer Nature Switzerland, 2019), pp. 395–406.
[10] J. A. Wanliss and et al, Journal of Geophysical Research **110(A08214)**, 1–11 (2005).