Proceedings of the 30th International Cosmic Ray Conference
Rogelio Caballero, Juan Carlos D'Olivo, Gustavo Medina-Tanco,
Lukas Nellen, Federico A. Sánchez, José F. Valdés-Galicia (eds.)
Universidad Nacional Autónoma de México,
Mexico City, Mexico, 2008
Vol. 1 (SH), pages 241–244

30TH INTERNATIONAL COSMIC RAY CONFERENCE

Method of additional fluctuations for search of weak signals

SERGEY KARPOV[1], LEONTY MIROSHNICHENKO[2, 3]
[1]*Institute for Nuclear Research of RAS, Baksan Neutrino Observatory*
[2]*Instituto de Geofísica UNAM, México, D.F., México*
[3]*N.V. Pushkov Institute IZMIRAN, Troitsk, Moscow Region, Russia*
karpovsn@yandex.ru

Abstract: New statistical method for search of weak signals of various natures is suggested. This method is applied when average value of a signal does not give statistically significant excess over an average background of the device. The method uses a property of statistical distributions to increase number of the large fluctuations far from the mean value. Suggested method provides extraction of such deviations caused by a weak signal. The method of additional fluctuations is most useful in the search experiments working near to a limit of accuracy of the detector. Authors apply this method to interpret some peculiarities of muon bursts observed at the Baksan Underground Scintillation Telescope in close correlation with a number of GLE events of solar cosmic rays.

Introduction

In astrophysics there is a typical situation when some sought signal arises sporadically, exists for short time, and then disappears. At the same time, the time interval during which it is possible to expect to detect a signal, is often known from other (independent) observations. In such a case, there is a possibility to determine statistical distribution of the background fluctuations with necessary accuracy in absence of a sought signal. Then, it is necessary to find out changes of this distribution during the intervals when one can expect for the signal. Two characteristic examples of such signals are high-energy gamma-ray bursts [1] and muon bursts at the Baksan Underground Scintillation Telescope (BUST). The muon bursts being registered [2] in close correlation with solar flares and Ground Level Enhancements (GLE) of solar cosmic rays (SCR). In the first case the information about time of the event may be obtained from the satellite data on gamma-ray bursts of lower energy; in the second case – from ground-based and satellite observations of solar activity. In both cases, the events under study occur in the random moments of time. Between successive events there are long enough intervals for measurement of background fluctuations.

Statistical model

The following statistical model is put in a basis of the method. We assume that the background of detector in absence of a signal is measured, and its distribution of probability is known. The further consideration is proper for any distribution of background measured beforehand. But for a definiteness we consider Poisson distribution $P(n, \lambda)$ that takes place most frequently in nuclear physics, physics of particles and physics of cosmic rays. One example of behaviour of counting rate of detector with $\lambda = 100$ at measurement of a background is shown at the bottom panel of Fig.1 by a solid line. We also assume that during some time there is a signal (event) which can give additional counts of detector. However, the number of these additional counts is such that alteration of counting rate of detector during event is less or close to statistical accuracy of measurement of a background ($\sim 1\sigma$). If alteration of counting rate will be 3σ or more, then the signal is large enough, and use of the suggested method is not required. The example of small additional counting rate (a weak signal) is shown at the top panel of Fig.1. The dashed line on bottom panel shows the sum of counts of the background and weak signal.

At the first sight, an extraction of the signal from background and even ascertainment of the fact of its presence seems to be impossible. Nevertheless, if some conditions will be realized, then the suggested method allows making it. The main point is the fact that counting rate is stochastic variable liable to fluctuations. At large number of measurements, the background fluctuations may be also rather large (3-5σ and more).

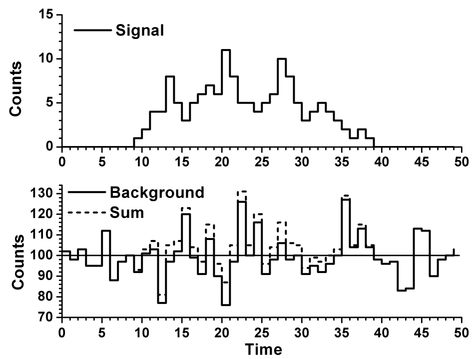

Figure 1: Examples of counting rate of the background distributed under Poisson law with λ = 100, weak signal and their sum.

If the fluctuation amplitude is small, for example, 1σ, then additional weak signal of ~ 0.5σ affects insignificantly a Poisson probability of such deviation from average background. So, at λ = 100 the probabilities to find out deviations $\geq 1\sigma$ and $\geq 1.5\sigma$ will be 0.17 and 0.076, respectively. If the fluctuation size is large enough, then even small additional counting rate reduces probability in many times. So for deviations $\geq 5\sigma$ and $\geq 5.5\sigma$ at λ= 100 the probabilities will be $1.9 \cdot 10^{-6}$ and $2.1 \cdot 10^{-7}$ respectively, i.e. difference makes up already almost one order of magnitude. As a result, the alteration of probability distribution of sum of the background and signal, in comparison with a pure background, will be very insignificant in the region close to the average counting rate, where large quantity of measurements of detector is concentrated.

Different situation will be observed in the region far from the average – in the tail area of Poisson distribution. Total number of large fluctuations is small, and noticeable part of them will have the changed amplitudes (see Fig.1). Therefore, the number of large fluctuations considerably grows even for a small additional signal. More correctly speaking, large fluctuations are shifted to the area of small probabilities in comparison with a background distribution. Just this change of the form of probability distribution in the region far from average is used in suggested method.

Description of the method

To find out a difference in the distributions in the region far from average, it is necessary to make rather large number of elementary measurements. As a rule, it is easily feasible for a background. For each physical event (signal) the number of measurements is usually limited by the ratio of the event duration to the duration of elementary measurement (a time step of measurements), by other parameters of the event and by the method of search for the signals. For example, in search for gamma-ray or BUST muon bursts, the number of elementary measurements is determined also by the size of sky area which is used in search for a signal and by step of division of this area onto smaller spatial cells. So, to enhance the method efficiency, it is necessary to use a series of similar events (bursts, solar flares, etc.). A set of events is often limited, too. Therefore, a basic idea of the method is following: to compare a signal with a background it is not required to obtain a full probability distribution during the event. It is enough to select by certain way only the biggest fluctuations for both the event and background.

The method of selection is following. Taking into account prospective properties of a sought signal and physics of the phenomenon, the duration of an interval of the event observation must be determined. It is a time window during which a search of a signal will be carried out. It is needed also to determine a position of observation interval relatively to a time reference point of the event. In the cases of gamma-ray or BUST muon bursts, respectively, the onset or maximum of low-energy burst or solar flare (in any kind of electromagnetic radiation) can be used as the time reference points. Additionally, the interval of observation can be divided onto a number of shorter time intervals (bins), if at the same time the detector allows obtaining acceptable statistical accuracy. The same is to do in relation to other properties of the event. For example, for gamma-ray and muon bursts the size and position of sky area for signal search are determined relatively to a spatial reference point (coordinates of gamma-ray burst and position of the Sun, respectively) that corresponds to prospective source. The search area also can be divided onto small spatial cells as the steps for sky scanning. Product of the numbers of temporal, spatial or other bins determines total number of elementary measurements (independent tests) N_{TST} during the event.

If all listed parameters are determined, then counting rate of each elementary measurement during the event should be analyzed and one maximal deviation from the average should be selected. It will be the maximal fluctuation with amplitude which could be increased due to additional counts of a signal. For this maximal burst the probability of its realization due to statistical fluctuations of the background should be calculated using probability distribution obtained earlier. For Poisson distribution the probability w of observation of n or more counts of detector in the given elementary measurement should be calculated by standard formula:

$$w = 1 - \sum_{i=0}^{n-1} \frac{e^{-\lambda} \cdot \lambda^{i}}{i\,!} \tag{1}$$

To calculate the probability of observation of such burst during all time of the event, it is necessary to take into account the number of elementary measurements N_{TST}. Product $w \cdot N_{TST}$ can be considered as average counting rate of bursts (new λ) with amplitude $\geq n$ during full interval of the observation. Then, the probability p_{ev} to find out one such burst or more during observation of the event, according to (1), will be equal to

$$p_{ev} = 1 - e^{-w \cdot N_{TST}} \tag{2}$$

Then the same procedure should be carried out for the control (background) intervals. The interval of observation, area of search, their dividing onto time bins and spatial cells should be the same, as at the analysis of the event. The moment of time should differ only. It is determined so that at this time there were no events which would be able to give additional counts, i.e. when the detector measures the background only. Depending on the properties of the detector, background and events, a control interval can be shifted onto minutes, hours or days forward or back relatively to the event under study. It is also necessary to take into account that the detector should be in the same conditions for measurements, as during the event. Observance of all these conditions provides an opportunity of direct comparison of the bursts found during the events and in control intervals. The described procedure of selection is repeatedly applied to all events of the series.

In presence of the signal, some part of the bursts will have the increased amplitude and smaller probability p_{ev} in comparison with the same bursts in the control intervals. To find out such difference, it is necessary to construct integral distribution of bursts k $(p_{ev} \leq p_i)$ depending on their probability. For a pure random process (background), such a distribution versus $1/p$ in double logarithmic scale should be a direct line with an inclination $\alpha = -1$. Really, according to "the law of big numbers", the probability p of some event is:

$$p = \lim_{N \to \infty} (k/N)\,, \tag{3}$$

N is a number of tests (in our case it is number of events of the investigated series), k – number of successes (number of events, where the bursts with $p_{ev} \leq p_i$ were fixed). At big enough N we can assign that $p = k / N$. Having expressed from here k through $(1/p)$, we get: $k = N \cdot (1/p)^{-1}$. Taking the logarithm of this equality we obtain the equation of a straight line. Thus, it is a consequence of "the law of big numbers" for pure random process. Two examples of distributions obtained in searching for extreme high-energy solar protons during GLEs [2] are represented on Fig.2. The distributions are constructed for the bursts selected both during the GLEs and during control intervals.

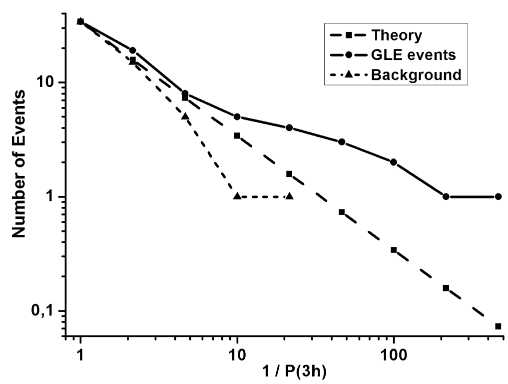

Figure 2: Comparison of integral distributions of bursts obtained at the analysis of GLE events and for control intervals.

If a background distribution do not described by a straight line with an inclination $\alpha = -1$, then the real probability distribution of the background differs from one taken in calculation. If there was a signal during the events, then integral distribution will have an absolute value of α less than 1. Corresponding straight line (or a curve) will pass some higher than theoretical and background straight lines. Just such behaviour is seen in Fig.2 for the bursts during the GLEs that testify to presence of additional flux of cosmic rays. The good enough agreement with the theory is observed for the control intervals.

Determination of signal properties

After the weak signal is found out, a question arises: can we tell something about properties of the signal? Consideration shows that the method cannot guarantee exact coincidence of properties of each found burst with properties of a signal. Moreover, the method does not guarantee that each burst contain a signal as the addition to large fluctuation. However, spatial and temporal burst distribution, as well as any other distribution obtained for series of the events, will correctly show on average real properties of a sought signal. Moreover, separate bursts of large amplitude with the properties analogous to the average properties of distribution will contain a sought signal with a high probability. The examples of properties studying of muon bursts during GLE events are submitted in works [2, 3].

Estimation of signal value

According to our statistical model, a sought signal makes up only a part of the amplitude of selected bursts. Generally speaking, it is impossible to separate quite correctly a signal from the background fluctuation. So, it is not realistic to obtain exact magnitude of the signal. One can expect only for a rather rough estimation by the order of magnitude. The bursts with maximal amplitude are best of all for estimation of signal magnitude. The magnitude of initial background fluctuation A_0 can be estimated with using following circumstance: if the initial background fluctuations will be submitted in the form of the integral distribution versus $1/p$ in double logarithmic scale, then they should lay on a straight line with the inclination $\alpha = -1$. This straight line sets theoretical probability, which should correspond to the background fluctuation. So, the magnitude of $1/p$ must be equal 34, in order the last right point in Fig.2 (GLE of 29 September 1989) would lay on theoretical straight line. In this case $p = 0.0294$. If we put p to (2), we shall find corresponding Poisson probability w_0, and from (1) we shall find n_0, and then A_0. For the GLE of 29 September 1989 we get $A_0 = 4.9\sigma$ in comparison with $A = 5.5\sigma$ for full amplitude. Thus, the additional counting rate of a signal will make up only 0.6σ or 11 % from full amplitude of the found burst. Therefore, the SCR flux obtained by suggested method of additional fluctuations is about 10 times less, than estimated earlier [3] with use of full amplitude of the burst.

Conclusions

Suggested method allows finding-out the weak signals with amplitudes less or close to accuracy of the background measurement. For this purpose, it is necessary to investigate a series of similar events, and then the signal from them can be found. Special selection procedure of the measurements, non-trivial probabilistic analysis and using of the control series of measurements – all of these techniques provide necessary reliability and sensitivity of the method of additional fluctuations. Its high efficiency is combined with visual presentation of obtained results. The method is universal enough and can be used in various experiments and observations. In particular, its application seems to help to resolve a long-standing problem of high-energy cutoff of the SCR spectrum [4].

References

[1] D.V. Smirnov, V.B. Petkov, and S.N. Karpov. Astron. Letters, 32, No.1, p.1-6, 2006.

[2] S.N. Karpov, L.I. Miroshnichenko, and E.V. Vashenyuk. Proc. 29th ICRC, Pune, 1, p.197-200, 2005.

[3] S.N. Karpov, L.I. Miroshnichenko, and E.V. Vashenyuk. Nuovo Cimento, 21C, p.551-573, 1998.

[4] L.I. Miroshnichenko. Izvestiya RAN, Ser. Phys., 67, No.4, p.463-465, 2003.