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. 3 (OG part 2), pages 1409–1412

30TH INTERNATIONAL COSMIC RAY CONFERENCE

Likelihood Method for 2-D Gamma-Ray Source Detection

G. H. SEMBROSKI[1], M. P. KERTZMAN[2], S. J. FEGAN[3]
[1]*Purdue University*
[2]*Depauw University*
[3]*University of California, Los Angeles*
sembroski@physics.purdue.edu

Abstract: In the analysis of Imaging Air Cerenkov Telescope (IACT) data, offset and/or extended source observations require more sophisticated schemes for determining the presence and statistical significance of any excess gamma-ray signal than the standard Li and Ma On-Off technique that has been conventionally used in the analysis of point sources. Although arrays of multiple telescopes have much less background contamination than a single IACT, they cannot eliminate all the background. In this paper we present a maximum-likelihood method for determining both the presence and statistical significance of a possible gamma-ray source in a field of non-zero background events. The method utilizes the ratio of the likelihood function under two hypotheses: first that the data can be explained by a pure background model and second that a source is present on top of the background, its strength given by maximization of the likelihood function. The method requires detailed knowledge of the detector response to background events, and the gamma-ray point-spread function in the detector. It produces a value for the signal strength along with the probability that the signal is inconsistent with being due to background alone.

Introduction

Imaging Atmospheric Cerenkov Telescopes (IACT) detect the Cherenkov light from both cosmic-ray and gamma-ray initiated air showers. The cosmic ray background is typically several orders of magnitude greater than the flux of gamma rays from a source. Many methods are employed to reduce the background (e.g. cuts on image parameters, cuts on reconstructed gamma-ray direction). However these methods do not completely eliminate the background of events from cosmic rays. Thus, we are always faced with extracting the source count rate and significance from a data set containing a large fraction of background events.

One common observation mode is the "on/off" mode, in which the telescope tracks the target object for some time interval and detects some number of events N_{on} in a small region of sky around the target location. The telescope is then moved to a location away from the object of interest and collects some number of events, N_{off}, in a region of the same size. Typically the off-source region is chosen such that the telescope traverses the same altitude and azimuth as the on-source run. The estimate of the number of gamma-ray events for the source region is therefore estimated by $N_{on} - \alpha N_{off}$, where α is the ratio of time spent in the on source region to the time spent on the off source region. Another method currently in use by several groups is the so-called "ring background" method, in which there is no separate off-source run. Instead, the background is estimated by forming an annulus around the target location, and using the events in this region to estimate the number of background events in the source region. In either method, the statistical significance of the result is generally evaluated by either equation 9 or equation 17 of Li and Ma [1].

These methods work reasonably well for pointed observations, that is, when the object of interest is well known. However, there are often interesting potential gamma-ray sources in which the emission is expected to be extended rather than point like, or for which we do not accurately know the position. In the case of a sky survey, both the existence and location of potential sources are unknown. In these cases the location of each event is reconstructed by some method (see [2] for a method applicable to observations with a single IACT), and a 2-D map of gamma-ray like events, which pass some gamma-ray like selection criteria, is created. The background estimate is more complicated in this case. Typically the camera response function is not uniform across the field of view. Also, a gamma-ray point source would not show up as a single point in the field of view. There is an intrinsic point-spread function associated with the reconstruction of the gamma-rays's origin from the shower image parameters (the shape of which is neglected by the typical on/off procedure described above). A typical procedure is to construct a 2-D off field map from the off field observation, apply a smoothing function to smooth out statistical fluctuations, apply the same smoothing procedure to the on-source map and then do a bin-by-bin on/off analysis[3]. This method assumes all the gamma-ray signal will be in one bin and thus ignores the intrinsic gamma-ray point-spread function and its shape, thus typically underestimating the statistical significance.

In order to overcome these shortcomings of the typical 2-D analysis for potential sources of unknown location, we have developed a maximum-likelihood method to search for gamma-ray signals in fields of non-zero background. This paper will describe the maximum-likelihood method as applied to the case of an unknown source location viewed by a single IACT. The method can easily be adapted for use with extended and point sources viewed by an array of IACTs (e.g. VERITAS [4]).

Method

We developed a maximum-likelihood method to test two hypotheses: (1) that our data is best fit by background and (2) that it is best fit by background plus a gamma-ray signal. This method requires models of the background and of the point-spread functions of gamma-ray events where the gamma-ray source is centered separately on each bin in the 2-D event map. The background model can be generated from Monte Carlo simulations. We have also generated background models based on data containing no gamma-ray signals (e.g. we have a plethora of "off-source" observations). The gamma-ray point-spread functions are generated by Monte Carlo. Here we describe the maximum-likelihood method, assuming we have background and gamma-ray models in hand.

The background model is a 2-D histogram which represents the expected outcome of a run which contains no sources of gamma rays. The background model is normalized such that the total count in the normalized 2-D histogram is one. The normalized values of each bin in the background histogram are called w_j. Note that by definition:

$$\sum_j w_j = 1 \tag{1}$$

Let us first consider the hypothesis that our data is best fit by a pure background model. The observed number of events in bin j of the histogram is n_j. If the data run contained no gamma-ray sources, the expected value of bin j would be:

$$\mu_j = Bw_j \tag{2}$$

where B is the total number of counts in the data histogram (assumed to be entirely due to background). We assume that the population in a bin should follow a Poisson distribution. In this case, μ_j is the mean of a Poisson distribution. Then, the probability of observing the value n_j in bin j is found from a Poisson distribution of mean μ_j:

$$P(n_j) = \frac{\mu_j^{\,n_j}}{n_j!} e^{-\mu_j} \ . \tag{3}$$

The probability that a particular set of data is fit by the background model is the product of these Poisson probabilities for each bin in the data histogram:

$$L_{background} = \prod_j \frac{\mu_j^{\,n_j}}{n_j!} e^{-\mu_j} = \prod_j \frac{(Bw_j)^{n_j}}{n_j!} e^{-Bw_j} \tag{4}$$

$L_{background}$ in equation 4 is called the likelihood function. The value of B in equation 2 is chosen to maximize the likelihood function and thus give the maximum probability that the data is consistent with the background model.

A similar approach is used for the hypothesis that the data is best fit by the background plus a signal. In this case, the expected value of bin *j* is:

$$\mu_j = Sa_j^k + Bw_j \tag{5}$$

Here *S* is the total number of counts due to signal, a_j^k is the normalized model signal contribution for bin *j* when the gamma-ray signal is centered in bin *k* and *B* is the total counts due to background. The likelihood function for the probability that the data is fit by a background plus signal, L_{signal}, is just equation 4, but with the μ given by equation 5 instead of equation 2. The best values of *B* and *S* are determined by maximizing this likelihood function. There is no analytical solution for equation 4 when μ_j is given by equation 5. In this case a grid search was used to find the best values for *S* and *B*.

To test the hypothesis that the data is best fit by the background model versus background plus data, we look at the ratio of the likelihoods: $L_{background}/L_{signal}$. The combination

$$R = -2\log\left(\frac{L_{background}}{L_{signal}}\right) \tag{6}$$

follows a χ^2 distribution with one degree of freedom. For each bin in the data histogram, the value of *R* was obtained assuming the signal is centered in that bin. The resulting 2-D map of χ^2 values was converted into equivalent probability values. These values give the bin-by-bin probability that the data histogram is best fit by the background model alone. Since *R* follows a χ^2 distribution with one degree of freedom, $\sqrt{R}$ is the equivalent standard deviation value of the probability.

Discussion

We applied this method to a 10 minute drift scan of Mrk421 taken by the Whipple 10m gamma-ray telescope in April 2004. The telescope was pointed 5 minutes ahead of Mrk421 and fixed at that altitude and elevation. Mrk421 then drifted across the center of the field of view of the camera. This run was conducted to test methods for reconstructing drift-scan events (see [5] for a discussion of the purpose and analysis methods of these drift-scan runs). The maximum-likelihood analysis of this Mrk 421 run yielded a gamma-ray rate of ~7 γ/min with a 6σ significance. This is consistent with results from normal runs taken immediately after the drift scan and with others taken the previous night. Four normal tracking runs taken on the previous night yielded a fairly steady rate of 8 γ/min with a significance of 14σ. A sequence of 28 minute runs taken immediately after the drift scan yielded rates 10.0 9.0, 8.0, 8.0, and 4.2 γ/min respectively, with significances ranging from 16σ to 9σ. The gamma-ray rate obtained by the maximum-likelihood analysis is consistent with the normal analysis. The lower significance obtained is reasonable because the source moves across the camera field of view which has an as non-uniform response.

Further testing of the method for single telescopes will be done by analyzing data collected on the Crab Nebula with the Whipple 10m telescope with the source offset varying amounts from the center of the field of view.

Although the method was developed for the situation of detecting a point source of gamma rays at an unknown location in a single camera field of view, it is applicable to other situations. For example, for an array of IACTs, a model would be needed for the response of the array, after all cuts and array processing, to background events. The gamma-ray models for this situation would depend on both the same cutting and array processing and on the nature of the astrophysical source (point source, extended source).

Acknowledgments

This research is supported by grants from the U.S. Department of Energy, the National Science Foundation, the Smithsonian Institution, by NSERC in Canada, by Science Foundation Ireland and by PPARC in the UK.

References

[1] Li and Ma, Astrophysical Journal **272**:317-324, 1983 September 1.
[2] Lessard, R. W., et al. 1999, in Proc. 26th Int. Cosmic Ray Conf. (Salt Lake City), 488.
[3] S. J. Fegan et.al. Astrophysical Journal **624**:638-655, 2005 May 10.
[4] Maier, G., et.al., The Status of VERITAS, In *International Cosmic Ray Conference*, 2007.
[5]Kertzman, M.P., et.al., The Whipple Strip Sky Survey, In *International Cosmic Ray Conference*, 2007.