29th International Cosmic Ray Conference Pune (2005) **00**, 101–104

Using of the analog complexing of parameters of space-temporal distribution of the EAS Cherenkov light for the analysis of the mass composition of cosmic rays

A. V. Yushkov, O. V. Zhurenkov
Altai State University, Lenin Ave. 61, Barnaul 656049, Russia
Presenter: A. V. Yushkov (yushkov@theory.dcn-asu.ru), rus-yushkov-AV-abs3-he11-poster

The space-temporal distribution of the EAS Cherenkov light initiated by p, O and Fe nuclei in the energy region 10–200 TeV per nucleus is described by 11 parameters. The ALTAI Monte-Carlo code has been used to perform simulations of the hadronic and electromagnetic interactions and detector's responses for real conditions. We applied analog complexing recognition method with optimization in relation to reference size for classification of primary nuclei with "event by event" method. Some results of these researches are given. Particularly, for some energies and distances the probability of true recognition of primary nuclei reaches value of 85%.

1. Introduction

At present, the elemental composition of primary cosmic rays above 10^{12} eV is studied by ground based experiments, generally equipped with a system of atmospheric Cherenkov telescopes (ACTs). Some parameters of the space-temporal distribution of Cherenkov light depend essentially on the atomic number of the primary nucleus [1]. The possibility of effective rejection of protons with the use of the same parameters set and χ^2 criterion also was shown in [1]. Such approach can be applied in gamma-astronomical experiments (see e.g. [2]), but for classification of three and more groups of nuclei it is not suitable.

Neural nets, evolution and genetic algorithms may be used for classification of complex multiparametric objects. In our work we applied one of the GMDH (Group Method of Data Handling) algorithms, namely, the analog complexing method.

2. Simulation

In our calculations we used the data on photoelectrons time registration, obtained with Monte-Carlo code ALTAI. The detailed description of the computing program can be found in [3].

For the data simulation we have chosen the installation geometry in the form of the square lattice (5×5) with the distance between adjacent telescopes equal to 50 m. The placement scheme of the telescopes is presented in Fig. 1. The position of the shower axis was simulated in the central quadrate of size 150 m $\times$ 150 m.

The radius of each telescope mirror is taken equal to

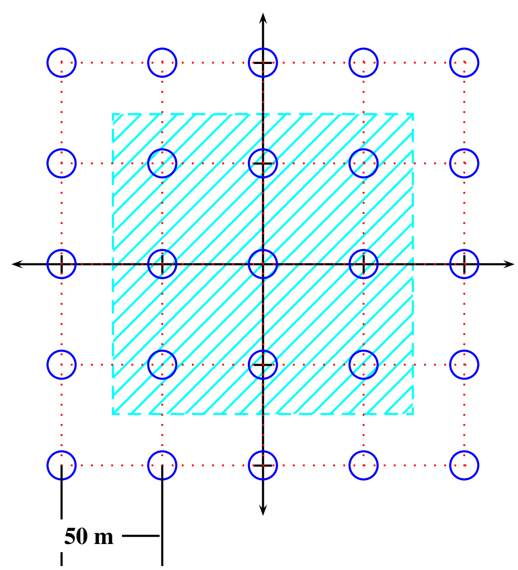

Figure 1. The installation geometry

Table 1. Structure of energy bins for different primary nuclei

Primary nucleus	E, TeV					
	1	**2**	**3**	**4**	**5**	**6**
p	10÷15	15÷20	20÷30	30÷50	50÷70	70÷100
O	15÷20	20÷30	30÷50	50÷70	70÷100	100÷150
Fe	20÷30	30÷50	50÷70	70÷100	100÷150	150÷200

1 m^2, focus distance $f = 1.5$ m, sight angle of a telescope $\alpha = 10°$. The threshold of the detector was set equal to $I_{\text{Tresh}} = 100$ ph. e. and for the trigger condition the 16 telescopes must be fired simultaneously. Only the vertical showers were simulated for the observation level of 2200 m a.s.l. (800 g/cm^2). It is necessary to note that the experimental setup, described above, does not represent a model of any real ACT array. However, for example AIROBICC [4], PACT [5] and TUNKA [6] installations have the similar technical characteristics and regular placement of ACTs.

We used three groups of primary nuclei: H, O and Fe. The sample sets were arranged by the following rules. EAS were simulated in energy range 10 ÷ 200 TeV per primary nucleus. This energy range was divided into 6 intervals, with numbers of simulated events for every particle type varying from 1500 to 100 for the lowest and the highest energy bins correspondingly. The initial energy was picked with the equal probability inside the energy bin. The bins borders for the different types of nuclei do not coincide (see Table 1). They were chosen from condition, that the primary nuclei from the bins with the same numbers must produce approximately the same average numbers of Cherenkov photons.

In the prospective experiment simulation the number of energy bin for particular nucleus was estimated from average number of Cherenkov γ-quants at the distance of ~ 75 m.

In the simulation of such experiment the additional errors are brought in the calculations not only by the evaluation of the incident particle energy, but also by the definition of the shower axis position. We used a simple method to define the coordinates (x, y) of the shower axis in the installation plane. For this purpose the coordinates of 9 telescopes with maximum amplitudes were used to get (x, y) in the following way:

$$x = \frac{\sum_{j=1}^{9} x_j I_j}{\sum_{j=1}^{9} I_j} \quad \text{and} \quad y = \frac{\sum_{j=1}^{9} y_j I_j}{\sum_{j=1}^{9} I_j},$$

where x_i, y_i — coordinates of the i-th telescope, and I_i — full number of the photoelectrons in the i-th telescope.

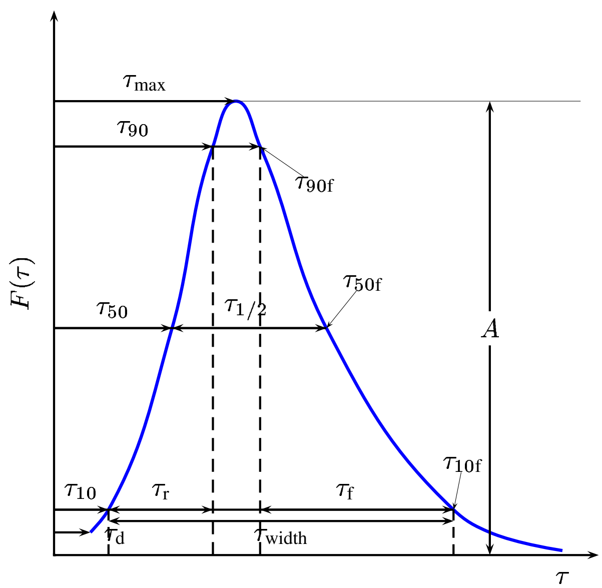

Figure 2. The basic parameters of the temporal distribution

The space-temporal distribution was reconstructed with time step $\Delta t = 0.25$ ns and step by distance from an axis $\Delta r = 25$ m. In Fig. 2 the form of temporal pulse and the basic parameters of temporal distribution are shown: τ_{d} — the arrival time of the first photon (in our calculations for all events is equal to 0); $\tau_{\max}$ — the position of the pulse maximum; τ_{10}, τ_{50},

τ_{90} — the arrival times of 10%, 50% and 90% of photons correspondingly; τ_{10f}, τ_{50f}, τ_{90f} — the arrival times of 10%, 50% and 90% of photons on the pulse falling side correspondingly; $\tau_{1/2}$ — the FWHM time; τ_r, τ_f — the pulse rise and fall times; τ_{width} — the width of the pulse at 10% level.

3. Analog complexing algorithm

The principles of GMDH were developed by A. G. Ivakhnenko in 1967 [7]. The general idea — selection — was borrowed from nature. There are many GMDH algorithms and one of them is analog complexing [8]. The aim of this algorithm is to choose from the basic sample the objects, lying at the closest distance in the multidimension space to the investigated unknown object (the basic sample consists of the great number of known objects). This method is very often used for the linked objects (e.g. consecutive in time) in a forecasting tasks. However, when the size of pattern equals to unity the data can not be linked and ordered. Then the task becomes the task of classification.

The characteristic property of the analog complexing is an absence of models, corresponding to every class of objects. This property is the advantage because of absence of the time waste for the building of models for objects (this is the most essential when the basic set frequently changes). On the other hand, this property is drawback because the full calculation procedure must be executed for the every checked object and the time of this calculations essentially depends on the size of the basic set (reference size) (see Fig. 3). We decided to reduce the influence of this drawback and therefore we had to solve the optimization task in relation to the basic set size. In the solving of this task not only the calculation time, but also the classification efficacy (see below) were taken into account. It is important, that the classification efficacy will be always higher for the p and Fe nuclei initiated EAS, than for EAS, initiated by oxygen nuclei.

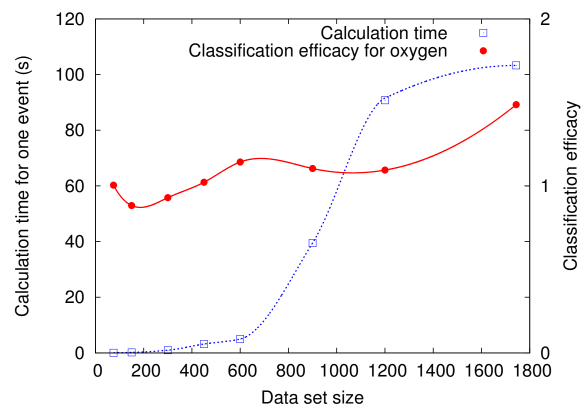

Figure 3. Dependence of the average calculation time for one event and of the classification efficacy for the primary oxygen nuclei on the basic set size

Hence, the classification efficacy for the oxygen nuclei initiated EAS was chosen as the main factor. In result of the optimization we determined the size of the basic set to be equal to 200 of the most characteristic events for every type of primary particle (i.e. 600 in total). The most characteristic events correspond to the least value of $R = \sum_{i=1}^{11} w_i(\tau_i - \bar{\tau}_i)^2$, where τ_i and $\bar{\tau}_i$ — the parameters of temporal distribution, described above (see Fig. 2), w_i — the weight of corresponding parameter. The values of the weights were taken from our results in [1] and were confirmed with the tests of the neural network model. If the number of the events (N_a) for some type of the primary particles (a) is less than 200 and less than for the other types of the primary particles, we take (N_a) events for all particle types.

4. Results of the classification

For the basic set construction the temporal parameters (type and energy of a primary particle and coordinates of a shower axis) for known events were used. We have performed the classification of the air showers gen-

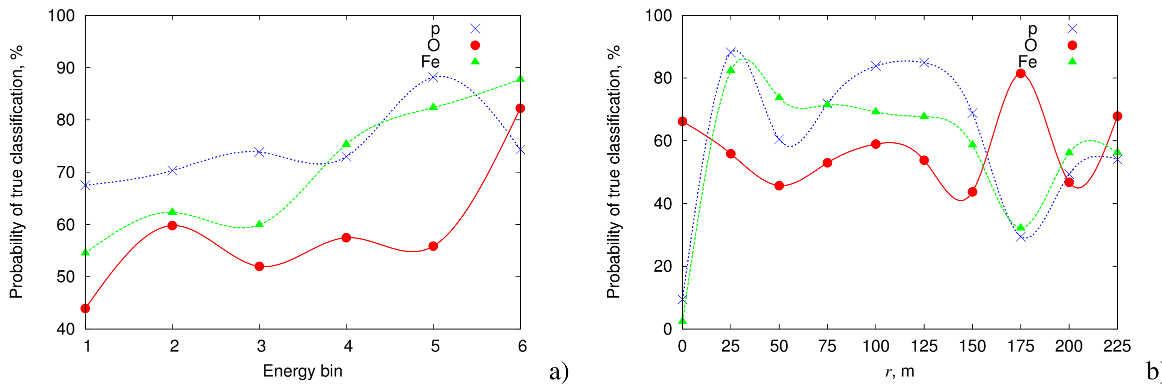

Figure 4. Probability of true classification: a) — distance from shower axis; b) — energy bin

erated by the different nuclei with the use of the analog complexing, as described above. Some results of the classification are given in Fig. 4.

Probabilities of true classification for every primary nucleus type for $r \sim 100$ are represented in Fig. 4 a); for the fifth bin energies ($E_p = 50 \div 70$ TeV, $E_{\rm O} = 70 \div 100$ TeV, $E_{\rm Fe} = 100 \div 150$ TeV) are given in Fig. 4 b). Recognition at the edges ($r \sim 0 \div 25$ m and $r \sim 200 \div 225$ m) is not provided with enough statistics and can not be used for any conclusions in our simulation.

We have introduced classification efficacy $\varkappa_{\rm nucl}$ for a primary type nucl as $\varkappa_{\rm nucl} = \dfrac{P_{\rm nucl}}{1 - P_{\rm nucl}}$, where $P_{\rm nucl}$ — the probability of true classification. So, the classification efficacy for the 5-th energy bin at 100 m from the axis is $\varkappa_p \simeq 5.2$, $\varkappa_{\rm O} \simeq 1.4$, $\varkappa_{\rm Fe} \simeq 2.2$. This approach may be recommended for reliable classification for distances $r \sim 75 \div 125$ m. The average values of the efficacy for these distances for energies 10-200 TeV are $\varkappa_p \sim 1.7 \div 7$, $\varkappa_{\rm O} \sim 1 \div 1.3$, $\varkappa_{\rm Fe} \sim 1.2 \div 3.3$. The classification efficacy slowly increases with the energy. This behavior agrees with results of [9], obtained with neural networks on the basis of data on N_μ^{tr} and N_e for PeV-region.

We recommend to use the given "event by event" method of primary particle type determination in analysis of space-temporal characteristics of Cherenkov light from EAS.

References

[1] O. V. Zhurenkov and A. V. Plyasheshnikov, Nuclear Physics B 75A, 296 (1999).
[2] V. R. Chitnis and P. N. Bhat, 26-th ICRC, Salt Lake City (1999), 5, 251.
[3] A. K. Konopelko and A. V. Plyasheshnikov, J. Phys. G: Nucl. Part. Phys. 26, 183 (2000).
[4] V. Fonseca, F. Arqueros, S. Bradbury et al., 24-th ICRC, Roma (1995), 1, 470.
[5] P. N. Bhat, B. S. Acharya, V. R. Chitnis et al., 26-th ICRC, Salt Lake City (1999), 5, 191.
[6] N. Budnev, D. Chernov, V. Galkin et al., 27-th ICRC, Hamburg (2001), 1, 581.
[7] A. G. Ivakhnenko, IEEE Transactions on Systems, Man, and Cybernetics (1971), SMC-1, 4, 364.
[8] Johann-Adolf Mueller and Frank Lemke, Self-Organising Data Mining. An Intelligent Approach To Extract Knowledge From Data (Berlin, Dresden, 1999), 1. Edition.
[9] M. Roth, T. Antoni, W.D. Apel et al., 27-th ICRC, Hamburg (2001), 1, 88.