Experiments with Mononuclear Cosmic Ray Beams

A. A. Vardanyan for the KASCADE Collaboration

Cosmic Ray Division, Yerevan Physics Institute, Alikhanyan Brothers' st. 2,Yerevan 375036, Armenia

Abstract. We discuss the capability and accuracy of non-parametric Neural Net method of classification and parameter estimation for the application to the analysis of the multi-dimensional experimental information provided by the Extensive Air Shower (EAS) observations in the KASCADE experiment. The methodical approach allows an event-by-event analysis of EAS measurements for a nonparametric estimate of the energy spectra of three different mass classes of the primary cosmic ray flux. Special emphasis is put on the possibility to obtain almost pure (mononuclear) samples of EAS for the two different mass groups. The implications and the potential bias for the estimate of the resulting distributions are discussed and displayed with the observables measured by the KASCADE central detector.

1 Introduction

Above a primary energy of a few hundred TeV the direct measurement of energy and mass of individual cosmic ray nuclei is unfeasible due to the drastic decrease of intensity with increase of energy. One has therefore to resort to the measurement of extensive air showers (EASs) which are produced when high energy cosmic ray particles enter into the earth atmosphere. To determine primary energy and mass from EAS observables has been tried since many years but has proved to be a very tough problem.

The idea to use advanced statistical techniques of multivariate analyses for isolating certain classes of EAS stems from an early proposal of A. Chilingarian and H. Zazyan (Chilingarian A. A., Zazyan H. Z., 1991a,b). The technique to prepare quasi-mononuclear beams by mass discriminative analyses of event-by-event EAS observations, was planned for the ANI experiment (ANI Collaboration, 1992; Chilingarian A. A., 1998). The realization of this proposal has become realistic by the recent results of the multi-detector experiment KASCADE (Kampert K-H. et al., 1999) which provides an accurate experimental basis for the event-by-event data analysis using measurements of several EAS characteristics.

This approach appears to be very promising with the aspect of refined tests of sophisticated EAS simulation codes (Heck D., et al., 1998) with alternative hadronic interaction models at extremely high energies (Drescher H. J. et al., 1999). Still the results of the KASCADE experiment concerning the energy spectrum and mass composition of primary cosmic rays are considerably affected by a model error, estimated for the energy slope to be 10 times larger than the statistical uncertainty (Chilingarian A. A. et al., 1999).

The present report introduces in recent results on energy dependence of mass composition and energy spectra of the primary flux different species. First attempt to analyze "mononuclear beams" interactions with atmosphere using KASCADE Central Detector (CD) information is made.

It is necessary to mention that results obtained within event-by-event analysis approach are conditional on the particular strong interaction model used for simulation of the primary passage through the atmosphere. The deconvolution of the triple uncertainty (primary mass, primary energy and strong interaction model) and obtaining of "pure" mononuclear beams even within one prechosen model will provide hints to understand direction in which family of strong interaction models have to evolve to meet experimental consistence criteria.

2 Primary Energy Estimation and Primary Mass Determination

The multi-layered perceptron (MLP) algorithm is used to determine the mass composition and the energy spectrum of the primary cosmic rays (PCR) in the knee region. This method provides opportunity of primary energy estimation as well as primary mass classification into multiple categories. The basics of Bayesian and neural regression and classification are described in (Chilingaryan A. A., 1994, 1995; Chilingarian

Correspondence to: A. A. Vardanyan (aro@crdlx5.yerphi.am)

A. A. et al., 1997).

For the estimation and classification tasks the same EAS observables are used. These are observables of the electromagnetic and muonic components measured by the KASCADE filed array detector installation:

• N_e: number of electrons in the EAS and
• N_μ^{tr}: truncated number of muons ($= 2\pi \int_{40\mathrm{m}}^{200\mathrm{m}} \rho_\mu(r) r dr$)

Restricting these observables is justified by following reasons:

• It is assumed that the electromagnetic and muonic component of EAS are described by the MC models with sufficient accuracy (partially insufficient knowledge about the hadronic component is illustrated in (Antoni T. et al., 1999; Roth M., 1999).
• Due to the larger statistical accuracy the uncertainties caused by strong EAS fluctuations are eliminated as compared with hadronic information of EAS.
• The KASCADE CD information can be used independently, after obtaining "mononuclear beams".

The results on the energy estimates demonstrate rather high accuracy ($\sim 25\%$ relative error) and almost unbiased estimation of energy (in whole energy interval except at the lowest and highest energies). Therefore we use a wider energy interval for simulated events to avoid over- and underestimation of primary energies at the boundaries.

Table 1. *Purity of classified events. Used observables: N_e and N_μ^{tr}*

$P_{\mathrm{i}\longleftarrow\mathrm{j}}$	j=p [%]	j=O [%]	j=Fe [%]
p	80	18	2
O	19	58	23
Fe	2	23	75

After estimating the primary energy each EAS event is classified as being induced by light (H,He), intermediate (CNO) or heavy (Si-Fe) nuclei (we will refer to these groups as "proton", "oxygen", and "iron"). The parameters of experimental events, classified as initiated from protons and iron nuclei, afterwards are compared with those of simulated ones. The results are given in table 1, which shows the contamination of misclassified events in each group of nuclei. One can see the rather high proportion of true classified events.

3 Energy Spectra of the Three Different Species of Primary Flux

Demonstrated in section 2 unbiased energy estimation and rather high percentage of the correct decisions in 3-way classification allow to use KASCADE array data for the physical analysis such as the determination of energy dependence of the mass composition and energy spectra of three species of the primary CR flux in the energy range $10^{15} - 2 \cdot 10^{16}$ eV. Figure 1 displays the differential energy spectra of three group of nuclei obtained by an event-by-event analysis (each shower was classified as originated by one of three types of the primary nuclei and the energy of that nucleus was estimated).

The knee feature is clearly seen for the all-particle and light nuclei spectra. For the spectrum of intermediate nuclei group the difference of the slopes is negligible and there is no evidence for a change in the spectral indices. For the heavy group of nuclei the decrease of the flux intensity and smaller spectral index at high energies is not observed. The fits of the energy spectra were done by a method described in ref. (Sokhoyan S. H. et al., 1998).

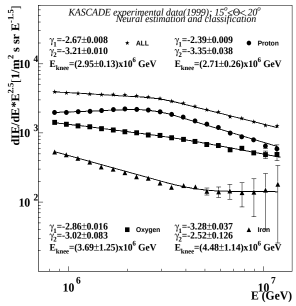

Fig. 1. *Differential energy spectra of three mass groups and the all-particle spectrum.*

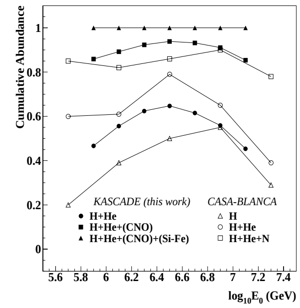

Fig. 2. *The cumulative abundances of different groups of nuclei*

In figure 2 one can see the comparison of the cumulative abundances of different groups of nuclei with recent results from CASA-BLANCA experiment (Fowler J. W. et al., 2000) using another experimental technique (measuring the lateral distribution of Cherenkov light in addition to the charged component of EAS) and completely different method of experimental data analysis. Similar trends and overall agreement in abundances of different nuclear groups are apparent.

4 The "Purification" Procedure

After estimating the misclassification rates, the possibility to select pure[1] nuclear beams was investigated. The developed neural information technique (Chilingarian A. A., 1998) allows to decrease the contamination of misclassified events in each class of nuclei. Of course, at the same time efficiency[2] of classification is reduced. The purification was done in the following way: the NN performs a nonlinear mapping of EAS multidimensional characteristics to the real number interval $[0, 1]$. Particular class assignments for three way classification are subintervals $[0. - 0.33]$, $[0.33 - 0.66]$ and $[0.66 - 1.]$ for first, second and third class respectively. The misclassification matrix for this mentioned decision intervals are given in Table 1. If the NN is trained well enough to have generalization capabilities NN output distributions for different classes are overlapping at subinterval boundaries. Therefore by the shrinking of the subintervals one can remove a large proportion of misclassified events, though simultaneously loosing some part of true classified events. From figure 3, where the purity versus efficiency is plotted,

[1] purity: fraction of true classified events in actual number of events assigned to a given class

[2] efficiency: fraction of true classified events in total number of events of a given class

one can see that the purity of proton and iron nuclei is larger than 90%, when the efficiency is still remaining not less than 50%. Purity estimates were obtained classifying 4000 control events (not used for the training) per class. For a given purity value the efficiency of proton events classification is always larger than the efficiency of iron event classification. Thus, the purification of proton events turns to be easier, than the purification of iron events.

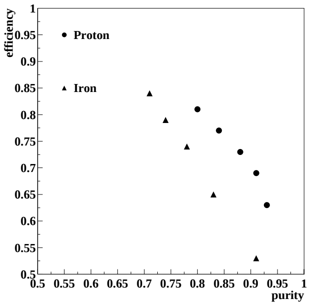

Fig. 3. *Event selection efficiency vs purity for proton and iron events (obtained by classification of the control samples).*

To prepare the "mononuclear" CR beams for investigation of the different nuclei interaction with the target (atmosphere) we have to check the purification procedure. First of all we have to investigate how the shrinking of the decision interval for different classes (on the NN output) affects on the distributions of EAS parameters (NN input). For this purposes we investigate and compare the "purified" and "rejected" EAS characteristics.

To check that no systematic distortions are introduced in EAS parameters by the purification procedure the one dimensional statistical tests (for detailed description on statistical tests used see (Chilingarian A. A., 1998)) are performed for initial and purified samples. The Table 2 displays the results of different tests for initial and purified proton and iron samples. Presented values are the probabilities of accepting the null hypothesis, that the samples are from one and the same population. If the tests produce small values of the probabilities one can reject the null hypothesis, i.e. there exists big difference between two examined samples.

Table 2. *The probability values of different tests for initial and purified proton and iron samples (t- Student, D- Kolmogorov-Smirnov, U- Mann-Whitnay)*

	Initial and pure protons			Initial and pure irons		
	t	D	U	t	D	U
N_μ^{CD}	0.11	0.36	0.13	0.18	0.44	0.12
N_h	0.36	0.99	0.49	0.35	0.99	0.28
E_h	0.36	0.77	0.40	0.38	0.72	0.23
E_h^{max}	0.30	0.81	0.29	0.40	0.84	0.26
E_{tot}	0.38	0.89	0.46	0.36	0.74	0.25

Table 2 clearly demonstrates that the probabilities are rather high and one can not reject the null hypothesis. So, the initial and purified proton and iron samples belong to the same population, which demonstrates that the purification does not introduce systematic distortions.

5 Experiments with "Pure" (mononuclear) CR

In previous sections the possibility of making precise ($\sim 25\%$ relative error) and unbiased estimation of primary energy, and accurate classification of primary particles into 3 categories in the range of $10^{15} - 2 \cdot 10^{16}$ eV was discussed (see also (Vardanyan A. A. et al., 1999)).

The purification technique described in sections 4, was used for "constructing" the mononuclear beams from KASCADE raw data. It is worth mentioning that achieved purity of proton and iron beams ($\sim 90\%$) exceeds the estimate obtained in (Chilingarian A. A., Zazyan H. Z., 1991a) ($70 - 80\%$). The relative error of primary energy estimation by the KASCADE array ($\sim 25\%$) almost coincide with the expected one.

The obtained beams purity and accuracy of energy estimation open possibility to consider the hadron component of EAS, detected by the CD. The comparisons of the QGSJET model (Kalmykov N. N. et al., 1997) and KASCADE data are depicted in figures 4a),b),c),d). The event selection procedure is equivalently done for the simulated and experimental data samples.

Due to the large statistical accuracy the experimental distributions of the hadronic parameters of the showers originating from the primary protons demonstrate rather smooth variation increasing with the energy. On the other hand the corresponding distributions, originating from the primary iron nuclei agree less. However, in general, the overall dependences are in agreement with QGSJet simulations.

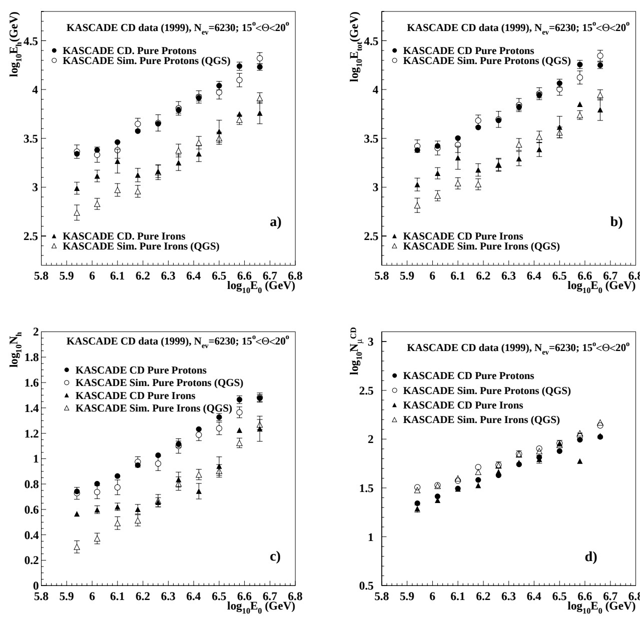

Fig. 4. *The Energy dependence of the Hadron Energy - a), Total Energy - b), Number of Hadrons - c) and Number of Muons - d) for the Proton and Iron Primaries*

It is worth to note that contamination of both proton and iron induced events by the intermediate nuclei has been ignored. The results in Figures 4c),d) can be compared with ref. (Antoni T. et al., 1999). But in the present case the primary energy is determined event-by-event.

If we take into account the limited efficiency ($\sim 90\%$) of the muon detecting facility, the agreement of the experimental data with the predictions is rather remarkable. It is interesting to point to the approximately primary invariant form of the muon number energy dependence, as demonstrated in Figure 4d).

It is also necessary to mention that the difference between figures presented in this paper and in ref. (Vardanyan A. A. et al., 1999) is caused by use of different data samples. In mentioned paper to improve statistics the array data were used with some cuts requiring at least 1 hadron and 5 muons in KASCADE CD, and the number of events survived cuts were compared with corresponding one in CD data sample where events in $0-30^{\circ}$ angular range were included. So, the real number of events which are the same in both array and CD samples is lesser in fact, since the array sample contains events only in $15-20^{\circ}$ zenith angle range. Therefore, it is easy to see that the statistics was not doubled but enlarged dramatically, which in its turn means that the applied cuts were very weak. Such cuts will allow to contribute events with shower core too far from CD and to introduce uncertanties in terms of shower core distance, primary energy, true number of hadrons and number of hadrons registered by the CD.

6 Concluding Remarks

Modern arrays of particle detectors covering large area are measuring different parameters of numerous secondary products of the primary cosmic ray interactions with the atmosphere. Only a simultaneous measurement of a large number of independent parameters in each individual Extensive Air Shower can yield reliable information to reconstruct the primary particle mass and its energy as well as the characteristics of strong interaction with atmosphere nuclei.

Current investigations are the first attempts to obtain energy spectra of 3 species of primary flux and to use ”mononuclear beams” for addressing the CR interaction problem on the *event-by-event* basis.

Used strong interaction model in general adequately explains allmost all EAS parameters. If, as we demonstrate, surface array parameters are described with sufficient accuracy, the hadron information demonstrates slightly more discrepancy with model predictions that could be caused by very poor statistics of CD events. Of course, better statistics for simulation and experimental data are required, other strong interaction models have to be tested as well. Nevertheless, we emphasize that the advocated approach is the only one which takes into account the shower fluctuations properly and is able to specify in a transparent way quantitative difference between model predictions and experimental data for different species of primary flux and various EAS parameters in wide range of the primary energy.

Acknowledgements. We would like to thank Dr. S. Ostapchenko for his interest to new data analysis methods and useful discussions.

References

ANI Collaboration. The ANI experiment on the investigation of the interactions of hadrons and nuclei in energy range 1000 - 100000 TeV. *NIM*, **A 323**:104, 1992.

Antoni T. et al. Test of high energy interaction models using the hadronic core of EAS. *J. Phys. G*, **25**:2161, 1999.

Chilingarian A. A. Analysis and Nonparametric Inference in High Energy Physics and Astroparticle Physics, 1998. Program Package ANI, (Users Manual, unpublished) http://crdlx5.yerphi.am/proj/ani.

Chilingarian A. A. et al. Estimation of the energy spectrum in the knee region by the KASCADE experiment. In *Proc. 26th ICRC*, vol. 1, p. 226. Salt Lake City, Utah, 1999.

Chilingarian A. A., Ter-Antonyan S., Vardanyan A. A. The comparison of Bayesian and Neural techniques in problem of classification to multiple categories. *NIM*, **A 389**:230, 1997.

Chilingarian A. A., Zazyan H. Z. Experiments with particle bundles in cosmic ray physics. Determination of strong interaction parameters by a pattern recognition method. *J. of Nuclear Phys.*, **54**:128, 1991a. (russian).

Chilingarian A. A., Zazyan H. Z. On the possibility of investigation of the mass composition and energy spectra of PCR in the energy range $10^{15}-10^{17}$ eV using EAS data. *IL Nuovo Cimento*, **14C**(6):555, 1991b.

Chilingaryan A. A. Neural classification technique for background rejection in high energy physics experiments. *Neorocomputing*, **6**:497, 1994.

Chilingaryan A. A. Detection of weak signals against background using neural network classifiers. *Pattern Recognition Letters*, **16**:333, 1995.

Drescher H. J. et al. NEXUS: The New Strong Interaction Model. 1999. (preprint hep-ph/9903296).

Fowler J. W., Forston L. F., Juni C.C.H., Kieda D. B., Ong R.A., Pryke C. L., Dommers P. A Measurement of the Cosmic Ray Spectrum and Composition at the Knee. 2000. (preprint astro-ph/00003190).

Heck D., et al. The Karlsruhe Exensive Air Shower Simulation Code CORSIKA. FZKA-Report 6019, Forschungszentrum Karlsruhe, Forshungszentrum Karlsruhe, Germany, 1998.

Kalmykov N. N., Ostapchenko S. S., Pavlov A. I. Quark-Gluon-String model and EAS simulation problem at ultra-high energies. *Nuclear Phys. B*, **52B**:17, 1997.

Kampert K-H. et al. The KASCADE experiment: Composition analysis and energy spectrum. In *Proc. 26 ICRC*, vol. 3, p. 159. Salt Lake City, Utah, 1999.

Roth M. *Analysis of air shower measurements of the KASCADE experiment by applying statistical classification methods for the determination of the eneregy and type of the primary cosmic radiation*. FZKA 6262, Forschungszentrum Karlsruhe, IK3, Forshungszentrum Karlsruhe, Germany, 1999.

Sokhoyan S. H. for the ANI Collaboration. Investigation of the integral EAS size spectra in the knee region. In *Proc. ANI 98 Workshop*, vol. FZK Internal report 6215, p. 55. Nor-Amberd, 1998.

Vardanyan A. A., Chilingarian A. A. and Roth M. for the KASCADE collaboration. On the Possibility of Selecting Pure Nuclear Beams from Measurements of the KASCADE Experiment. In *Proc. of the Workshop ANI 99*, vol. FZK Internal report 6472, p. 23. Nor-Amberd, 1999.