Proceedings of ICRC 2001: 93 © Copernicus Gesellschaft 2001

ICRC 2001

Determination of primary energy spectra from MAKET ANI data

A. A. Chilingarian, G. G. Hovsepyan, G. V. Gharagyozyan, S. S. Kazaryan, L. G. Melkumyan, S. O. Sokhoyan, and A. A. Vardanyan

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

Abstract. The unfolding of the primary energy spectra from size spectra measured by MAKET ANI installation is performed. The nonparametric regression method was used for estimation of energy of each detected shower. Simple method of the unfolding of size spectra was introduced as robust alternative to event-by-event analysis of EAS data. Both methods agree within experimental and methodical errors. The ways to utilize *a priori* knowledge for physical inference are discussed.

1 Introduction

The "all charged particle" spectrum of Cosmic Rays (CR) measured by surface arrays represents integrated information on both primary CR flux in vicinities of Earth - information for solving long standing problem of CR origin and acceleration - and on strong interaction of primary hadron with atmosphere nucleus - at energies ant secondary phase space that cannot be obtained on the modern colliders.

The aim of modern installations measuring Extensive air Showers (EAS) such as KASCADE (Klages H. O., et al., 1998) and ANI (ANI Collaboration, 1992) is to solve mentioned astrophysical and particle physics problems simultaneously. This ambitious goal requires measurement of the different EAS components, sophisticated simulation programs and adequate methods of multivariate statistical analysis.

We develop package of different multidimensional statistical methods (Chilingarian A. A., 1998), especially for solving inverse problems in astroparticle data analysis. In present paper we'll demonstrate usage of one of the modes of ANI package, namely adaptive nonparametric regression, for primary particle energy estimation. Simple unfolding method is introduced for reconstruction of the energy spectra from measured size spectra for different angles of incidence.

Correspondence to: A. A. Chilingarian (chili@yerphi.am)

2 Nonparametric Adaptive Regression

Usually, for experimental physics data analysis, the Likelihood Function cannot be written explicitly, and we deal with implicit, nonparametric models, for which no parametric form of underlying distribution is known, or can be assumed.

Nonparametric methods use much less stringent assumptions about population than those made in parametric statistics. Usually the underlying population distribution is assumed to be continuous only. Of course this assumption is rather mild comparing with the very specific assumptions made in parametric case.

Let us consider the stochastic mechanism that generates the observations $\mathbf{v}_i, i = 1, M_{exp}$ in a multivariate feature space - $\mathcal{V}$, $\mathbf{v}_i$ is one of M_{exp} d-dimensional vector of EAS parameters measured experimentally.

We assume that observations are random and can be described by some conditional probability density function depending on the primary particle type and its energy. The feature space $\mathcal{V}$ covers possible acceptable values of EAS parameters including cuts on shower age and size,...

We don't know full statistical description of how nature produces EAS from incident particles, nor we have the possibility to use particle beams outside of the atmosphere to calibrate the installations. That is why the total Monte-Carlo simulation of the EAS development in the atmosphere and in detectors is performed, including experimental data registration and handling for different primaries.

Usually the following parameters are used as inputs for a Monte-Carlo simulation program:

- the primary type;
- the primary energy;
- the angle of incidence;
- the strong interaction mode (one of possible alternatives).

The set of corresponding d-dimensional vectors $\mathbf{u}_i, i = 1, M_{tr}$ (training sets) obtained from simulations is an analog of the experimentally measured values of $\mathbf{v}_i$, but, unlike the experimental data we exactly know the energy and type of primary used in the simulations.

These labeled events include *a priori* information about dynamics of the EAS development and registration with inherent fluctuations, as well as the information on the primaries energy spectra. All statistical variability of events is given in a nonparametric form of simulation trials.

The problem is how to assign the probability measure in the primary particle parameters space. Which type of primary mass composition and energy spectra use in simulations? The same question has to be answered concerning the strong interaction model. The common answer to such a question is to use statistical procedures, not to depend crucially on the choice. Robust statistical procedures will depend strongly on the experimental data and not very strongly on particular type of the model, used for simulation.

Let's assume that training sample $(\mathbf{u}_i,\ E_j/A),\ i = 1, M_{tr}$, is generated according to the chosen energy spectrum, $A-$ is one of 5 primary nucleus used in simulation. The regression model usually is presented in the following form:

$$E_i = f_{model}(\mathbf{u}_i) + \epsilon_i \qquad i = 1, M_{tr} \qquad (1)$$

where f_{model} represents the stochastic algorithm used in simulation and ϵ_i are identically and independent distributed random numbers with zero mean and finite variance.

We are interested in estimation of energy of the experimental events:

$$\hat{E}_j = f_{exp}(\mathbf{v}_j) + \varepsilon_j \qquad j = 1, M_{exp} \qquad (2)$$

where f_{exp} represents the way nature used to generate EAS. Consequently ε_j has unknown distribution.

After validation of the model, when we prove that at least $f_{model} \sim f_{exp}$, we can construct energy estimator exploiting the *a priori* information contained in the training sample:

$$\hat{E_j} \ = \ \frac{\sum_{i=1}^{M_{tr}} E_i \ K(R_{i,j}^{mah}, h, W_i)}{\sum_{i=1}^{M_{tr}} K(R_{i,j}^{mah}, h, W_i)} \qquad j = 1, M_{exp} \qquad (3)$$

where $\hat{E}_j-$ is the energy estimate of the j-th experimental event, E_i- is the energy of the i-th event of training sample, W_i is the event weight, obtained according the particular simulation scheme, h is the kernel width (parameter controlling the degree of the "smoothness" of an estimate), and the Mahalonobis distance is the measure of the "closeness" of experimental and simulated events in the measurement metric space:

$$R_{i,j}^{mah} = \sqrt{(\mathbf{v_j} - \mathbf{u_i})^T \Sigma_i^{-1} (\mathbf{v_j} - \mathbf{u_i})}, \qquad (4)$$

where Σ_i is the sampling covariance matrix of the class to which $\mathbf{u}_i$ belongs (different primaries will produce different covariance matrices).

In present study we use the following form of the kernel function:

$$K(R_{i,j}^{mah}, h, W_i) \ = \ W_i \cdot e^{-\frac{R_{i,j}^{mah}}{h}}, i = 1, M_{tr}, j = 1, M_{exp}. \qquad (5)$$

As usual, we made the adaptation of estimator (Chilingarian A. A., 1989),calculating the energy for different values of smoothing parameter h simultaneously, and taking the median of this sequence as final estimate.

3 Size Spectra Unfolding

We use very simple algorithm of unfolding that proceeds in the following steps: [1]

1. Construct experimental integral size spectra in 5 angular bins (we divide experimental data to 5 independent pieces to check that same size spectra intensities, measured at different zenith angles, are corresponding to the one and the same primary energy, for details see figure on pg. 46 of(Chilingarian A. A., Sokhoyan S.O. et.al., 1999));

2. Perform full simulation of EAS and store shower parameters for each registered EAS;

3. Select 20 values of spectra intensities and perform constant intensity cuts (CIC, (Nagano M. et al., 1984)), in points $(I_{N_e}^i, i = 1, 20)$;

4. Determine 20 consequent values of shower sizes (for each angular bin), corresponding to the each intensity cut $(N_e^i, i = 1, 20)$;

5. For each obtained value of N_e^i consider all simulated events with $N_e^j > N_e^i$, select corresponding values of E_j and form 20 energy distributions;

6. Calculate the mode of each of this 20 distributions E_i^{mod};

7. Produce the integral energy spectra (E_i^{mod}, I_E^i), proceeding from assumption $(I_{N_e}^i \equiv I_E^i)$ for each of 5 angular intervals.

The simulated data used for the reconstruction of primary energy spectrum were generated by the CORSIKA (Heck D., et al., 1998) code and QGSJet (Kalmykov N. N., Ostapchenko S. S., Pavlov A. I., 1997) strong interaction model. Approximately 27000 events, initiated by different primaries were generated in $2 \cdot 10^{14} - 2 \cdot 10^{16}$ eV energy and $0 - 45^\circ$ zenith angle intervals. By changing the weights of each simulated event it is possible to obtain reference data correspondent to various assumptions about mass composition and partial energy spectra.

[1] In description of the algorithm we fix some parameters (number of angular intervals -5, and number of constant intensity cuts - 20). Of course, it is arbitrary choice and another selection is possible depending only on number of experimental and simulated events available

In Figure 1 the integral spectra for 3 angular intervals are presented. We reconstruct spectra separately for 3 independent data files, corresponding to the different angles of incidence. First of all, it is the check of uniformity of installation operation and angular resolution. Reconstructed energy spectra in contrast to size spectra must be isotropy, in a way we can see in the upper picture. In the bottom picture, where 3 spectra are artificially shifted from each other, we can detect the similarity of spectra shapes and good agreement in fitted values of spectra slopes and knee positions.

In Figure 2 the comparison of the energy spectra obtained by methods is presented. Obtained spectra slopes before and after knee and knee position could be considered as preliminary. The multiple comparisons of experimental and alternative simulation data, along with different statistical analysis methods are now performing for producing final estimates.

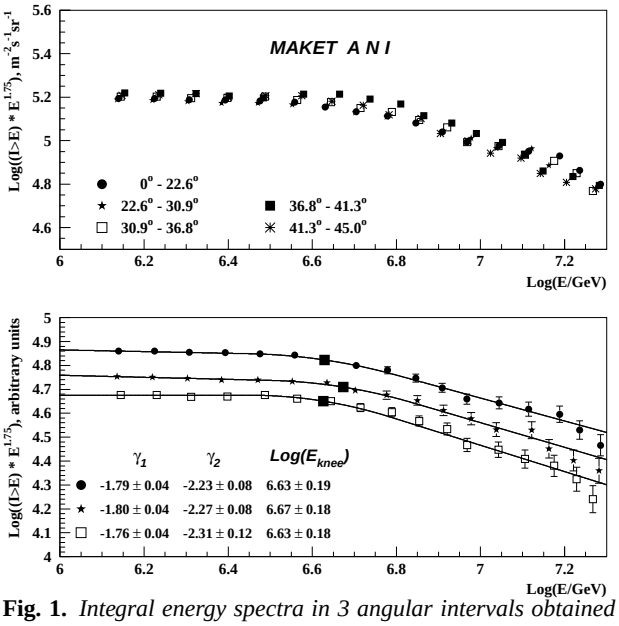

Fig. 1. *Integral energy spectra in 3 angular intervals obtained by unfolding of the size spectra*

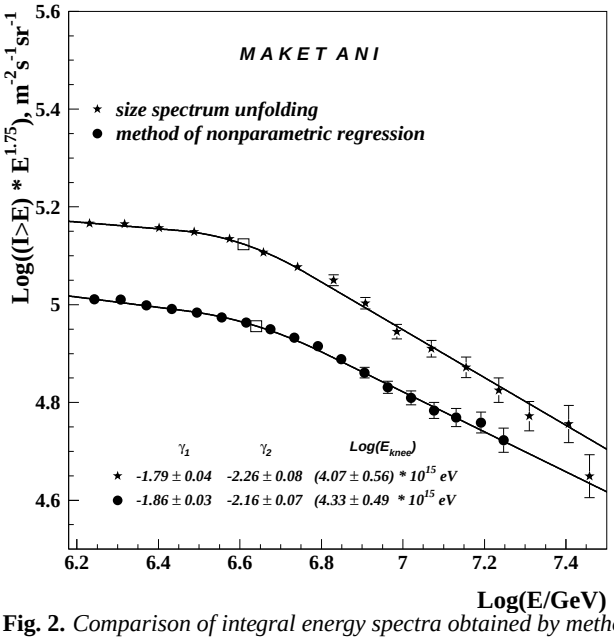

Fig. 2. *Comparison of integral energy spectra obtained by methods of nonparametric regression and unfolding of size spectra*

4 Discussion

For the first time we present the integral energy spectra obtained with MAKET ANI data (for details on the installation layout and data analysis methodology see ANI workshop proceedings (Chilingarian A.A.,Rebel H., Roth M., Zazyan M.Z., 1998; Chilingarian A.A., Haungs A., Rebel H., Sanosyan Kh.N, 2000) and rappporteur talk (Stanev T., 1999)).

We use the method of estimation of the individual energy of each registered shower by kernel regression function, and - unfolding of the size spectra, to produce the energy spectra. As one can see in Figure 2, the results of both methods are in good agreement with each other. and with energy spectra. Obtained values of spectra slopes and knee position also agree with values reported by KASCADE collaboration in Salt Lake City (Chilingarian A. A., Roth M., Vardanyan A. A. et al., 1999).

For both methods the simulations were used for weighted averaging of simulated energies in "nearest neighborhood" of experimental event (according to equations 3-5 of section 2), and for the calculation of the energy distributions modes corresponding to the showe "cutted" sizes (steps 5 and 6 in algorithm description in section 3).

Unfortunately, the realistic simulation of detector response is still underway (we plan to check the response function of installation registration channels by a special calibration experiment). Therefore, we are obliged to increase the detector threshold up to $N_e = 2 \cdot 10^5$, corresponding to $E_0 \sim 2 \cdot 10^{14}$. For this threshold the shower registration efficiency is approaching 100%. Consequently for producing energy spectra we use only $9 \cdot 10^5$ events instead of $2.3 \cdot 10^6$ available. Also we didn't use the shower age parameter, that proved to be very useful both for energy and primary type estimation on mountain altitudes (Chilingarian A. A., Ter-Antonyan S., Vardanyan A. A., et al., 1998).

The main goal of the development of the new methods of energy spectra reconstruction is to prove that the *a priory* assumptions on the partial energy spectra of different primaries used in simulation didn't influence the estimated spectra characteristics (slopes and knee position) crucially. I.e. we have to allow the experimental data to "speak for themselves", and avoid the "dictate" of *a priory* knowledge accumulated in the theoretical models. That is why we apply the robust methods of spectra unfolding along with the more precise and simulation dependent method of nonparametric regression. Good agreement of spectra characteristics, obtained by both methods prove the soundness of our approach. In addition we plan to perform multiple calculations of spectra using different theoretical models, changing the reference partial spectra and mass composition used in simulations. We also plan to use alternative methods of spectra unfolding, based on calculation of the multidimensional "er-

ror matrix" (Vardanyan A.A, 2000) and Bayesian adaptation (D'Agostini G., 1999).

Acknowledgements. The work has been partly supported by the research grant N 00-784 of the Armenian government, NATO NIG-975436 and CLG-975959 grants and ISTC A216 grant.

This publication is based on experimental results of ANI collaboration. The MAKET ANI installation has been set up as collaborative project of the Yerevan Physics Institute, Armenia and the Lebedev Institute, Russia. The continuous contribution of the Russian colleagues in the data analysis is thankfully acknowledged. This paper expresses the point of view of the given group of co-authors and isn't official paper of ANI collaboration. The Monte Carlo simulations used in this report were done by S.V.Ter-Antonyan.

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.

Chilingarian A. A. Statistical decisions under nonparametric a priori information. *Computer Physics Comunication*, **54**:381, 1989.

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., Roth M., Vardanyan 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., Sokhoyan S.O. et.al. Determination of the EAS Attenuation Length from Data of the ANI experiment. In *Proc. of the Workshop ANI 99*, vol. FZK Internal report 6472, p. 10. Nor-Amberd, 1999.

Chilingarian A. A., Ter-Antonyan S., Vardanyan A. A., et al. Energy Spectra and Elemental Composition Determination on Mountain Altitude and Sea Level. *Nuclear Phys. B*, **60B**:117, 1998.

Chilingarian A.A., Haungs A., Rebel H., Sanosyan Kh.N. Current Topics of Cosmic Ray Research with EAS Observations. In *Proc. of the Workshop ANI 99*, vol. FZK Internal report 6472, p. 139. Nor-Amberd, 2000.

Chilingarian A.A.,Rebel H., Roth M., Zazyan M.Z. Current Topics of Cosmic Ray Research with EAS Observations. In *Proc. of the Workshop ANI 99*, vol. FZK Internal report 6215, p. 132. Nor-Amberd, 1998.

D'Agostini G. Bayesian Resoning in High-Energy Physics:Principles and Applications. 1999. (preprint CERN 99-03).

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.

Klages H. O., et al. The KASCADE Experiment. *Nucl. Phys. B*, **60B**:151, 1998.

Nagano M. et al. Energy Spectrum of Primary Cosmic Rays between $10^{14.5} - 10^{18}$ eV. *Journal Phys. G: Nucl. Phys.*, **10**:235, 1984.

Stanev T. High Energy Interactions and Extensive Air Showers. In *26th International Cosmic Ray Conference*, vol. AIP 516, p. 407. Salt Lake City, 1999.

Vardanyan A.A. Private Communications. 1999.