ICRR

ICRR-Report -381-97-4

**Contribution to IX International Symposium
on Very High Energy Cosmic Ray Interactions
(18-24 Aug. 1996, Karlsruhe)**

ed. by A. Ohsawa

(January, 1997)

SW9717

Chacaltaya and Pamir Collaborations

A trial of classification of proton- and heavy-ion-induced families using neural networks

M.Tamada ^a

^aFaculty of Science and Technology, Kinki University, Osaka, Japan

A feed-forward neural network trained using backpropagation is applied to classify proton- and heavy-ion-induced cosmic-ray families. Fifteen input variables which characterize three-dimensional behaviour of the families are chosen. The network successfully classifies the events with classification efficiency $\sim 85\%$. The trained neural network is applied for estimating the fraction of heavy-induced events in the simulated families and also in the experimental families observed in the Pamir chambers.

1. INTRODUCTION

It is well known that the intensity of atmospheric families detected by emulsion chambers at high mountains is smaller than expected [1]. There are two possible explanations for the small intensity of families. Chacaltaya and Pamir groups [2] insist that it is due to changes of characteristics of nuclear interaction in high energy region of $E_0 \geq 10^{15} \text{eV}$, on the basis of their findings of exotic-type events called Centauro-species. Fuji-Kambala group [3], on the other hand, insists that the chemical composition of primary cosmic rays is changed and heavy primaries dominate over protons and light nuclei in those high energy region.

The emulsion chamber detects most efficiently atmospheric families induced by proton primaries. The fraction of heavy-induced families detected in emulsion chamber is expected to be small because of severe energy dissipation during passage through the atmosphere. However, if the chemical composition of primary cosmic-rays is changed drastically, e.g., the fraction of heavy primaries becomes more than 50% in $E_0 \geq 10^{15} \text{eV}$, as is claimed by Fuji and Kanbala experiment [3], we should detect heavy induced events in considerable fraction among observed families. In the present paper we show how to classify proton-induced families and heavy-induced ones using a neural network classifier [8], and estimate fraction of heavy-induced families in the observed families in the Pamir experiment.

2. ATMOSPHERIC INTERACTIONS

We assume simple geometrical model for hadron-nucleus interaction. Some details of the model are described in [4]. The number of intra-nuclear sub-collisions is sampled from Poisson distribution. In each sub-interaction secondary particles are generated by using UA5 algorithm [5]. In nucleus-nucleus interactions, the wounded-nucleon model is assumed. The number of wounded nucleons is taken from the formula found by T. Shibata et al. [6].

We calculated developments of nuclear and electromagnetic cascades in the air initiated from nuclear interaction by primary cosmic-rays and we have obtained 1,710 events of proton-induced families and 1,581 iron-induced families, at the Pamir altitude. Here families are selected by the criteria of visible energy sum $100 \leq \Sigma E(\gamma) < 1,000 \text{TeV}$ and shower multiplicity $N \geq 10$. The detection energy threshold is set to be 4 TeV both for gamma-rays and hadrons.

3. NEURAL NETWORKS

3.1. Network training

The simulated families are divided into two data sets. One is used for training the network, i.e., training data set which consists of 600 proton-induced events and 600 iron-induced events, and the other for testing the ability of the network to classify the events, i.e., test data set which consists of 870 proton-induced events and

Table 1

Fraction of the events with network output larger or less than 0.5.

	network output	
	$y_p < 0.5$	$y_p > 0.5$
proton	0.83	0.17
He	0.72	0.28
CNO	0.44	0.56
heavy	0.13	0.87
proton-dominant	0.79	0.21
heavy-dominant	0.71	0.29
experiment	0.84	0.16

870 iron-induced events.

We choose following 15 variables which characterize three-dimensional behaviours of families: i.e., (1) $\Sigma(E_\gamma + E_h^{(\gamma)})$, (2) ΣE_γ , (3) N_γ , (4) $< E_\gamma >$, (5) $\Sigma E_h^{(\gamma)}$, (6) N_h , (7) $< E_h >$, (8) $Q_h \equiv \Sigma E_h^{(\gamma)} / \Sigma(E_\gamma + E_h^{(\gamma)})$, (9) $< R_\gamma >$, (10) $< E_\gamma R_\gamma >$, (11) $< R_h >$, (12) $< E_h^{(\gamma)} R_h >$, (13) $< R >_{\gamma+h}$, (14) $< ER >_{\gamma+h}$, (15) $R_E \equiv \Sigma ER / \Sigma E$. Each input variable is normalized by the average value of that of proton-induced families, i.e., $I_i \equiv x_i / < x_i >$, where x_i is the above variable and $< x_i >$ is average of x_i in the proton-induced families.

The network topology here we use is 15-10-1 net, that is, 15 input nodes, 10 hidden nodes and 1 output node. The backpropagation algorithm[9] is used for network training. An input pattern, composed from the above 15 variables, in the training data set is presented to the feed-forward network and the corresponding output value, y_p is calculated. The connectivity weight w_{ij} between nodes of two successive layers are updated for every 20 pattern (10 protons and 10 heavies) presentations to the network so as to give the desired output value, $t_p = 0$ for proton-induced events and $t_p = 1$ for heavy-induced ones.

The initial connectivity weights are selected at random between $[-0.1, 0.1]$. The learning rate η is set to 0.0001 and momentum parameter α is set to 0.5. The sigmoid function $g(x) = 1/(1 + e^{-x})$ is used for activation function. We stop network

training at the training epoch becomes 10,000 and fix the network connectivity weight w_{ij} . Then proton-induced events and heavy-induced events in the test data set are presented to the above trained network. We show in Fig.1 the distribution of the network output y_p of those events in the test data set. Around 83% of proton-induced events give network output $y_p < 0.5$ and $\sim 87\%$ of heavy-induced events give output $y_p > 0.5$.

Figure 1. Distribution of the network output y_p for the events which are not used in network training.

Although the training data set contains only proton- and heavy-induced events, and network has never seen He-induced events and CNO-induced events, we study how the network works for He-induced events and CNO-induced events. In the figure we show also network output distribution for He-induced events and CNO-induced events. In Table 1 we summarize fraction of events which have network output larger or less than 0.5.

3.2. Fraction of heavy-induced events in simulated and experimental families

Two different chemical composition of primary cosmic-rays are assumed in simulation calculations. One is proton dominant composition (PD) and we use the spectrum formulated by

Nicolsky[7]. The other is heavy dominant composition (HD) and we use the spectrum formulated by Fuji-Kanbala group[3]. The fraction of protons and heavy primaries is respectively assumed to be 0.14 and 0.61 at $E_0 = 10^{16} \text{ eV}$ in HD case, in contrast to that it is assumed to be 0.42 and 0.31 in PD case. We have calculated nuclear cascade in the air and obtained 1703 atmospheric families, which satisfy the selection criteria, in PD case and 1,650 events in HD case. Those families are presented to the trained network described in section 3.1. Fig.2 shows distributions of network output for those two cases of chemical composition. The network output has a main peak at network output value around zero in both cases as is naturally expected. The number of events in the region of network output $y_p \geq 0.5$ which corresponds to heavy-induced families is apparently larger in HD case than in PD case. It is considered that the distribution actually reflect chemical composition of observed families in the two cases. The ratio between the number of events which have network output $y_p < 0.5$ and those which have network output $y_p > 0.5$ is summarized in Table 1.

216 experimental families which satisfy the same selection criteria, observed in the Pamir carbon chambers[2] are also presented to the trained network. In fig.2 network output distribution of these experimental families are also shown together with those of simulated families. As is seen in the figure, the experimental distribution is rather close to simulated one of PD case. The fraction of the events with network output less than 0.5 is ~ 0.84 , almost same to that of proton-induced events.

4. SUMMARY AND DISCUSSIONS

The neural network is applied for classification of proton- and heavy- induced cosmic-ray families. It was shown that the classification efficiency is $\sim 85\%$ both for proton-induced and heavy-induced events.

The neural-net classifier was also applied to the cosmic-ray families observed in the Pamir chambers. The network output distribution is almost same to that of proton- induced families and we

Figure 2. Distribution of the network output y_p for the families

can conclude that the fraction of heavy-induced events in the observed families is at most 2 – 3% of the events.

REFERENCES

1. Pamir Collaboration, Mt.Fuji Collaboration, Chacaltaya Collaboration (S.G.Bayburina et al.), *Nucl. Phys.* **B391** (1981) 1.
2. Chacaltaya and Pamir Collaboration (L.T.Baradzei et al.), *Nucl. Phys.* **B370** (1992) 365.
3. J.R.Ren et al., *Phys. Rev.* **D38** (1988) 1404.
4. M.Tamada, *J. Phys. G : Nucl. Part. Phys.* **21**(1995) 1387.
5. UA5 Collaboration (G.L.Alner et al.), *Nucl. Phys.* **B291** (1987) 445.
6. Y.Niihori, T.Shibata, I.M.Martin, E.H.Shibuya, A.Turtelli Jr., *Phys. Rev.* **D36** (1987) 783.
7. S.I.Nicolsky, *Proc. Int. 3rd Symp. on Cosmic Rays and Particle Physics* (Inst. for Cosmic Ray Research, Univ. of Tokyo, Tokyo, 1984) 507.
8. L.Lonnblad, C.Peterson and T.Rognvaldsson, *Comp. Phys. Comm.* **81** (1994) 185.
9. D.E.Rumelhart, G.E.Hinton, R.J.Williams, *Parallel Distributed Processing* (MIT Press, Cambridge, 1986) Vol.1 p.318

Simultaneous observation of families and accompanied air showers at Mt. Chacaltaya I

N. Kawasumi^a, I. Tsushima^a, K. Honda^a, K. Hashimoto^a, T. Matano^b, N. Inoue^c, K. Mori^c, A. Ohsawa^d, M. Tamada^e, N. Ohmori^f, N. Martinic^g, R. Ticona^g, F. Osco^g and C. Aguirre^g

^aFaculty of Education, Yamanashi University, Kofu, Yamanashi, 400 Japan

^bDepartment of Applied Physics and Chemistry, Fukui University of Technology, Fukui, 910 Japan

^cFaculty of Science, Saitama University, Urawa, Saitama, 388 Japan

^dInstitute for Cosmic Ray Research, University of Tokyo, Tanashi, Tokyo, 188 Japan

^eFaculty of Science and Technology, Kinki University, Higasi-Osaka, Osaka, 577 Japan

^fFaculty of Science, Kochi University, Kochi, 780 Japan

^gInstituto de Investigaciones Fisicas, Universidad Mayor de San Andress, La Paz, Bolivia

Simultaneous observation of families and accompanied air showers is carried out at Mt. Chacaltaya with emulsion chambers and a air shower array. The present analysis is based on the data of experiment from 1969 to 1986. 47 families with $\Sigma E_{\gamma} \geq 10$ TeV are coupled to the accompanied air showers of size 10^5 to 10^8 . Comparison of the results with the simulation calculation, there is discrepancy between the experimental data and the simulation calculation, which indicate a need of more rapid energy dissipating process in the atmosphere in high energy region of $E_0 > 10^{16}$ eV.

1. Introduction

We have observed the families in the air shower cores with emulsion chamber (32 blocks, total area 8 m²) installed at the center of air shower array at Mt. Chacaltaya since 1979[1].

The air shower array observes the Ne, while the emulsion chamber observes the high energy components in the core of air shower. Simultaneously observation of the family and the air shower is essentially important for the investigation of characteristics of nuclear interaction at the energy range greater than 10^{15} eV.

2. Experiment

The air shower array covers an area about 50 m radius by 35 plastic scintillation detectors. 5 plastic scintillation detectors are located in pyramid shape in the center of array to measure the arrival direction of the air showers. Core location, Ne and age parameter of individual air shower are determined by fitting the density distribution to NKG function.

Thirty-two blocks (0.25 m² of each area) of emulsion chamber are installed in the center of air shower array. The emulsion chamber consists of lead plate of 15 cm thick and 14 sensitive layers of two sheets of X-ray films which are inserted at every 1 cm lead plate.

Burst detectors are located just below each emulsion chamber unit. Large burst signals in lead plate arrival by the hadronic component