LABORATORI NAZIONALI DI FRASCATI
SIS – Pubblicazioni

LNF-95/046 (P)
31 Agosto 1995

Identification of Cosmic Ray Electrons and Positrons by Neural Network

F. Aversa[3], G. Barbiellini[3], G. Basini[5], R. Bellotti[1*], V. Bidoli[2], M. Bocciolini[4], U. Bravar[3], M. Boezio[3], F. Cafagna[1], M. Candusso[2], M. Casolino[2], M.[3] Castellano[1*], M. Circella[1], A. Colavita[3], G. De Cataldo[1], C. De Marzo[1], M.P. De Pascale[2], N. Finetti[6], F. Fratnik[3], N. Giglietto[1], R.L. Golden[7†], C. Grimani[6], M. Hof[10], B. Marangelli[1], F. Massimo Brancaccio[4], W. Menn[10], J.W. Mitchell[9], A. Morselli[2], P. Papini[4], A. Perego[4], S. Piccardi[4], P. Picozza[2], A. Rainò[1], M. Ricci[5], P. Schiavon[3] M. Simon[10], R. Sparvoli[2], P. Spillantini[4], P. Spinelli[1], S.A. Stephens[8], S.J. Stochaj[7], R.E. Streitmatter[9], A. Vacchi[3] and N. Zampa[3]

[1]Università di Bari and INFN, Bari, Italy
[2]Università di Roma Tor Vergata and INFN, Rome, Italy
[3]Università di Trieste and INFN, Trieste, Italy
[4]Università di Firenze and INFN, Firenze, Italy
[5]INFN Laboratori Nazionali di Frascati, Italy
[6]Università di Perugia and INFN, Perugia, Italy
[7]New Mexico State University, Las Cruces, USA
[8]Tata Institute of Fundamental Research, Bombay, India
[9]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[10]Siegen University, Siegen, Germany
[†]Deceased

Abstract

A data analysis based on artificial neural network classifiers has been done to identify cosmic ray electrons and positrons detected with the balloon-borne NMSU/Wizard-TS93 experiment. The information is provided by two ancillary and independent particle detectors: a transition radiation detector and a silicon-tungsten imaging calorimeter. Electrons and positrons measured during the flight have been identified with background rejection factors of 80±3 and 500±37 at signal efficiencies of 72±3% and 86±2% for the transition radiation detector and the silicon-tungsten imaging calorimeter, respectively. The ability of the artificial neural network classifiers to perform a careful multidimensional analysis surpasses the results achieved by conventional methods.

PACS.: 29.85.+c; 96.40.De

Submitted to Astroparticle Physics

* Corresponding authors: R. Bellotti and M. Castellano,
Dipartimento di Fisica, Via Amendola 173, 70126 Bari, Italy

1 Introduction

The investigation of cosmic ray positrons and electrons is of special interest to study of cosmic ray propagation and production. Cosmic ray e^- are mainly of primary origin whereas e^+ are produced in nuclear interactions of the primary cosmic rays with the interstellar medium (ISM), via the $\pi^+ \rightarrow \mu^+ \rightarrow e^+$ decay chain [1]. In addition $e^\pm$ may also be produced by mechanisms such as pair production in strong magnetic fields near pulsars or the annihilation of heavy dark matter particles, such as massive neutrinos or supersymmetric particles, in the galactic halo [2]. Because of their low mass, $e^\pm$ lose significant amounts of energy by interacting with electromagnetic fields and photons during the propagation in ISM. Measurements of the equilibrium energy spectra of these particles and comparison with those calculated by the predictions of model can provide a better understanding of the origin of cosmic ray $e^\pm$ and of their propagation in the Galaxy.

Observation of cosmic ray positrons is a very difficult task because of the high proton background. Their detection requires complex instrumentations and involves stringent cuts on the data. From the instrumental point of view the main problems to be solved in balloon-borne experiments are the short exposure factors available. The harsh conditions experienced during a balloon flight also introduce some systematic changes of the detector performances with time. In order to identify the $e^\pm$ signal against the large background sophisticated analysis techniques are required.

Nowadays many procedures currently used in high energy physics - from real time pattern recognition (triggering) [3] to off-line data analysis [4] - are performed by the application of neural network (NN) techniques. NNs are particularly apt to classify complex phenomena and provide robust and reliable methods to design efficient and fast particle identification systems.

In this paper we show that classifiers based on feed-forward neural network are able to effectively identify electrons and positrons measured by two independent and ancillary particle detectors employed in the NMSU / Wizard-TS93 apparatus [5]. Thus, the neural network based analysis can be used for a careful evaluation of the energy spectra of cosmic ray positrons and electrons. The reliability of the results achieved is confirmed by means of the statistical compatibility, based on the Kolmogorov test, between the neural feature space of the simulated and real data.

2 The TS93 flight instrument

The TS93 experiment is devoted to the measurement of the energy spectra of cosmic ray positrons and electrons in the energy range from 5 to 50 GeV. The detector system employed is shown in fig. 1. It consists of: (1) a superconducting magnet, equipped with multiwire proportional chambers and drift chambers, used as spectrometer; (2) a set of plastic scintillator providing trigger, time-of-flight and absolute charge measurements; (3) a transition radiation detector, for identification of electrons and positrons with energy above 3 GeV and (4) a silicon-tungsten imaging calorimeter to identify different particles according to the topological and energetic

pattern released in the detector (i.e. straight tracks for muons and non-interacting protons; hadronic showers for interacting protons; electromagnetic showers for electrons and positrons).

The transition radiation detector [6], positioned at the top of the balloon payload, is composed of ten modules, each one made of a carbon fiber radiator followed by a $Xe - CH_4$ filled multiwire proportional chamber (MWPC). Each radiator consists of an alluminium frame containing four bags, each filled with 6 mm long, 7 μm thick and 1.76 g/cm^3 dense carbon fiber segments. The MWPCs are equipped with 256, 3 mm-spaced, wires and they have an active area of 76 x 80 cm^2. The MWPC signal processing is based on the cluster counting method [7].

The calorimeter [8], positioned at the bottom of the balloon payload, is composed of 5 silicon planes, sensitive both in the X and Y coordinates, interleaved with one radiation length (=3.5 mm) of tungsten, for a total calorimeter thickness of four radiation lengths. The sampling layer of the calorimeter is an array of 8 x 8 pairs (X-Y) of detectors (6 x 6 cm^2, divided in 16 strips, each 3.6 mm wide). Each sampling layer consists of two arrays having 128+128 readout channels. A built-in system equipped with ADCs and digital processors accomplishes the data acquisition.

3 The neural network classifier

A classification system transforms an information space (pattern space) in to a discriminating variable space (feature space) [9]. The study of the human decision-making process has led the investigators to formulate a formal paradigm in order to gain knowledge by examples [10]. Neural networks are very popular paradigms to build up example-based classifiers. In the present work the "three-layered feed-forward" Neural Network is considered [11]. It consists of units (formal neurons) arranged in contiguous layers, as shown in fig. 2. Each neuron k, $(k = 1, \ldots, N_h)$ belonging to the hidden layer, receives as input the output X_l $(l = 1, \ldots, N_i)$ of all the neurons l of the input layer, to which it is connected by $w_{lk}^{(1)}$ synaptic strengths. On the other hand, the neuron k is also connected, with strengths $w_{ki}^{(2)}$, to the neuron of the output layer. The transfer function of the neurons is the sigmoid funtion:

$$g(x, \theta) = \frac{1}{1 + \exp(-\beta(x - \theta))} \tag{1}$$

where θ and β are the neuron threshold and the gain factor, respectively.

The synaptic matrix $W^{(old)}$ is trained by showing to the network a data set of examples (training data set) and updating the weights according to the delta rule [11], which minimizes the mean squared error function E of the classification system. After each training epoch, the quality of the new model $W^{(new)}$ is estimated by means of an independent data set (test data set). The learning session is stopped when the error function E, evaluated on the test set, reaches its lowest value. The training and stopping procedures are unbiased general criteria to build up a classification system in an accurate way. A further method to check the reliability of the classification

Table 1. Characteristics of the neural network used to classify the TRD data.

Neurons in the input layer	10
Neurons in the hidden layer	5
Neurons in the output layer	1
β gain factor	1
θ neuron threshold	0.5
α momentum term	0.9
η learning rate	1
Learning algorithm	Backpropagation
Learning mode	Incremental
Number of Epochs	3500
Training Set Cardinality	12000
Test Set Cardinality	12000
Validation Set Cardinality	$\sim$ 10000
Test Classification Error	$\sim$ 6%
Kolmogorov Test	$\geq$10%

Table 2. Characteristics of the neural network used to classify the calorimeter data.

Neurons in the input layer	10
Neurons in the hidden layer	15
Neurons in the output layer	1
β gain factor	1
θ neuron threshold	• 0.5
α momentum term	0.9
η learning rate	1
Learning algorithm	Backpropagation
Learning mode	Incremental
Number of Epochs	500
Training Set Cardinality	3000
Test Set Cardinality	3000
Validation Set Cardinality	$\sim$ 100000
Test Classification Error	$\sim$ 1%
Kolmogorov Test	$\geq$10%

is to compare the output spaces of the test data set and the real data set using the Kolmogorov test [12]. Tables 1 and 2 summarize both the parameters and the Kolmogorov indicator for the NN classifiers used in this paper.

From the point of view of data analysis each input neuron is associated with a physical variable. Therefore NNs enable one to explore a multidimensional input space, by taking into account several complex information about the physical event. The neural network architecture described above has been skilled in order to improve the signal/background discrimination of the real data measured by the TRD and the calorimeter during the flight. In both cases three different kinds of data sets have been used for the analysis: (i) *training data*, for the synaptic weights formation; (ii) *test data*, for the error evaluation of the trained NN; (iii) *application data*, to validate the classication model and then evaluate the detector performances during the flight. Training and test data have been obtained through a simulation; application data is a sample of real events selected with high-purity, in a low acceptance region, by means of stringent one-dimensional cuts on the whole experimental data set. In particular the TRD (calorimeter) results have been carried out using application data selected by the time-of-flight, the spectrometer and the calorimeter (TRD).

4 Experimental results

The data set analyzed here was collected during the balloon flight of the NMSU / Wizard-TS93 instrument performed at 4.5 GV/c geomagnetic cutoff. The flight took place on September 8, 1993 from Fort Sumner, New Mexico (USA). Only particles fullfilling the following pre-selection criteria were taken into account in this analysis: (a) momentum p larger than 5 GeV/c and smaller than 50 GeV/c; (b) both ${\chi_x}^2$ and ${\chi_y}^2 \leq 5$ for the reconstructed track; (c) deflection uncertainty $\sigma_\eta \leq 0.02\ GV^{-1}$; (d) absolute electric charge $\mid Z \mid = 1$; (e) events moving downward in the detectors system (e.g. no albedo particles); (f) at least 6 TRD modules crossed by the particle [13, 14, 15].

4.1 TRD data analysis

In some previous work evaluations of particle discrimination capability of a TRD prototype were done on different classical sets of variables in a standard way[6, 16]. A good degree of accuracy was obtained by taking into account macroscopic variables such as the *number of planes fired in the track* n_{planes}, the *number of hits fired in the track* n_{hits}, or some suitable combination of them, such as the *geometrical mean* indicator $GM = (n_{planes} \cdot n_{hits})^{1/2}$. It is worth pointing out that these indicators have inherent drawbacks because they neglect several important physical and instrumental effects, such as, in particular:

(1) the detection efficiencies of each MWPC plane;

(2) the topological distribution of the hits in the spatial window which defines the physical track;

(3) the weak dependence of the yields of close TRD's modules, due to X-ray escaping detection in the relative MWPC and their conversion in the following one.

A more careful analysis has been carried out by using a neural network. In order to overcome the above-mentioned drawbacks, the pattern space used for the classification of the TRD data consists of the total number of hits detected by each TRD planes in a small spatial window (10 planes · 5 wires) around the physical track, as reconstructed using a cross-correlation algorithm [17], and the information from the magnetic spectrometer. Results obtained using a TRD prototype with test beam data are shown in [18, 19].

Neural network responses (feature spaces) for flight data are shown in fig 3. The background rejection factor at various signal efficiency is shown in fig. 4, together with the rejection achieved by means of one-dimensional cuts on the *geometrical mean* indicator [13, 15]. The reliability of the results obtained by the present analysis has been checked by performing a comparison between feature spaces for simulated and application data. It must be pointed out that the events used for this comparison were not used to train the network. The agreement between the two statistical distribution, evaluated by means of the Kolmogorov test, is positive at 10% significance level.

4.2 Calorimeter data analysis

Lepton/hadron discrimination is achieved with the calorimeter by exploiting the different longitudinal and lateral energy deposit profiles of e.m. and hadronic showers. To this end, some discriminating variables are derived from the patterns in the detector, describing the energy, the number of hits and the shower aggregation of each event. These discriminating variables are built as being maximally invariant with respect to position within the detector, while retaining the largest amount of information contained in the pattern.

In order to exploit completely the weak correlations among the variables, these latter are fed as input to the artificial neural network described in table 2, together with the rigidity of the event measured by the spectrometer. Moreover this enables to correlate automatically the rigidity of the particle with the different behaviour of the discriminating variables, in different parts of the energy spectrum. Results using a calorimeter prototype with test beam and simulated data are shown in ref. [20]

In the present work the discriminating variables considered are selected from the set of 12 variables used in conventional analysis [13, 14]. They include the total energy released and number of hits fired in the calorimeter, as well as the fraction of energy relased in clusters close to the track path. In order to judge the performance of the classifier the Kolmogorov test has been used, as described in section 3. As a result, a set of 10 physical variables has been selected and fed as input to the network.

The neural network responses for application data are shown in fig. 5. Figure 6 shows the background rejection factor of the calorimeter at various signal efficiency, together with the result achieved by conventional analysis [13, 14].

5 Conclusions

In this paper two different classification systems, for TRD and calorimeter respectively, are proposed to effectively identify electrons and positrons detected by the NMSU / Wizard-TS93 cosmic ray space experiment. Classifiers are based on a neural network computational architecture in order to carry out a careful multidimensional analysis. In both cases the results show that the neural network classifiers are capable of discriminating between positrons and protons with higher efficiencies (for a fixed contamination level) than those achieved by conventional analyses. Specifically, the data analysis methods described here permit to identify electrons and positrons with a background rejection of 80±3(stat) and 500±37(stat) at the enhanced efficiency of 72±3(stat)% and 86±2(stat)%, for the TRD and calorimeter, respectively. These results are particularly useful in the context of balloon-borne experiments where, in searching rare events with the constraint of short exposure times, a signal efficiency as large as possible is required. The neural network based analysis can be used to improve the evaluation of the energy spectra of cosmic ray positrons and electrons.

References

[1] R.J. Protheroe, Astrophys. J. **254** (1982) 391.

[2] M.S. Turner and F. Wilczek, Phys. Rev D **42** (1990) 1001.

[3] C.S. Lindsey et al., Nucl. Instr. and Meth. A **317** (1992) 346.

C. Kiesling et al., "Concept for a fast second level trigger using a neural network architecture for the H1 experiment at HERA", in *New Computing techniques in Physics research III*, K.H. Becks and D. Perret-Gallix Editors, (1993), World Scientific.

C. Baldanza et al., "Results from an on-line neural trigger within a fixed target experiment for the production of beauty particles", in *New Computing techniques in Physics research III*, K.H. Becks and D. Perret-Gallix Editors, (1993), World Scientific.

B. Denby et al., Nucl. Instr. and Meth. A. **356** (1995) 485.

[4] Delphi Coll., Phys. Lett. B **295** (1992) 383.

L3 Coll., Phys. Lett. B **307** (1993) 237.

Aleph Coll., Phys. Lett. B **313** (1993) 549.

Wizard-TS93 Coll., Proc. of 24th International Cosmic Ray Conference, Argalia Ed., Vol. 3 (1995) 730.

[5] R.L. Golden et al., Proposal to NASA NRA-92-OSSA-10.

[6] E. Barbarito et al., Nucl. Instr. and Meth. A **357** (1995) 588.

[7] C. W. Fabjan et al., Nucl. Instr. and Meth. **181** (1981) 119.

[8] F. Aversa et al., Nucl. Instr. and Meth. A **360** (1995) 17.

[9] R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis Wiley-Interscience, (1973).

[10] Yoh-Han Pao, Adaptive Pattern Recognition and Neural Networks Addison-Wesley, (1989).

[11] J. Hertz, A. Krogh and R.G. Palmer, Introduction to the Theory of the Neural Computation, Addison-Wesley, (1991).

[12] Eadie et al., Statistical Methods in Experimental Physics, North-Holland, (1971).

[13] F. Aversa et al., "Measurement of positron to electron ratio in cosmic rays above 5 GeV", submitted to Astrophys. J.

[14] F. Aversa et al. (Wizard-TS93 Coll.), Proc. of 24th International Cosmic Ray Conference, Argalia Ed., Vol. 3 (1995) 9.

[15] F. Aversa et al. (Wizard-TS93 Coll.), Proc. of 24th International Cosmic Ray Conference, Argalia Ed., Vol. 3 (1995) 714.

[16] Nucl. Instr. and Meth. A **313** (1992) 295.

[17] A. Rosenfeld and A. C. Kak, Digital Picture Processing Accademic Press, (1989).

[18] R. Bellotti, M. Castellano, C. De Marzo, N. Giglietto, G. Pasquariello and P. Spinelli, Comp. Phys. Comm. **78** (1993) 17.

[19] R. Bellotti, M. Castellano, C. De Marzo, G. Pasquariello G. Satalino and P. Spinelli, Nucl. Instr. and Meth. A **350** (1994) 556.

[20] M. Candusso, M. Casolino, M.P. De Pascale, A. Morselli, P. Picozza and R. Sparvoli, Nucl. Instr. and Meth. A **360** (1995) 371.

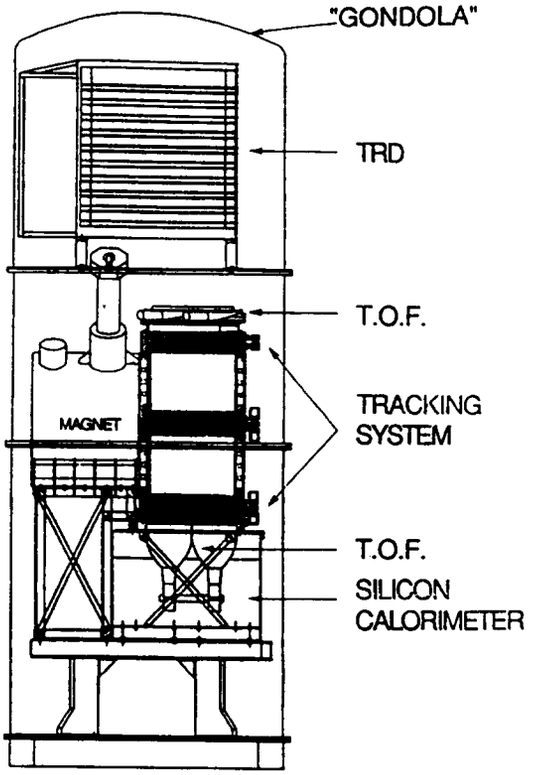

Fig. 1 – The New Mexico State University balloon-borne magnet facility equipped with Wizard TS93 apparatus used for the 1993 flight to identify cosmic ray electrons and positrons in the energy range from 5 to 50 GeV.

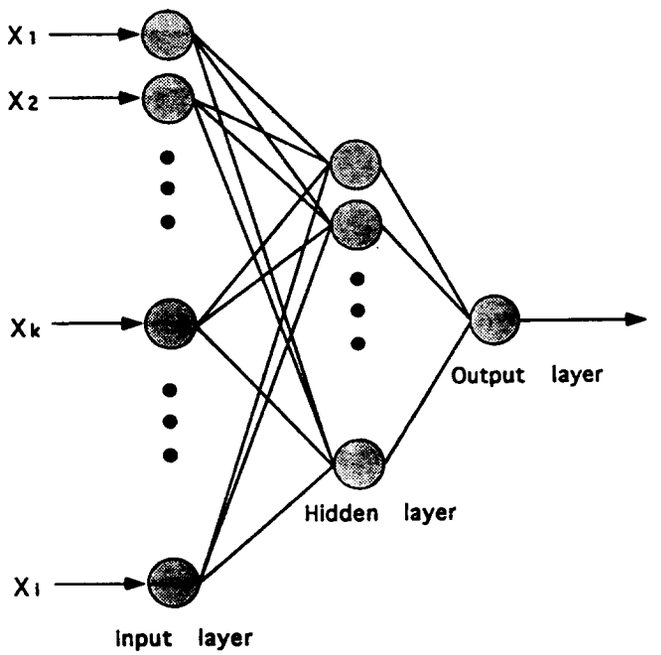

Fig. 2 – The three-layered feed-forward neural network classifier used to select electrons and positrons detected with the TRD and Calorimeter.

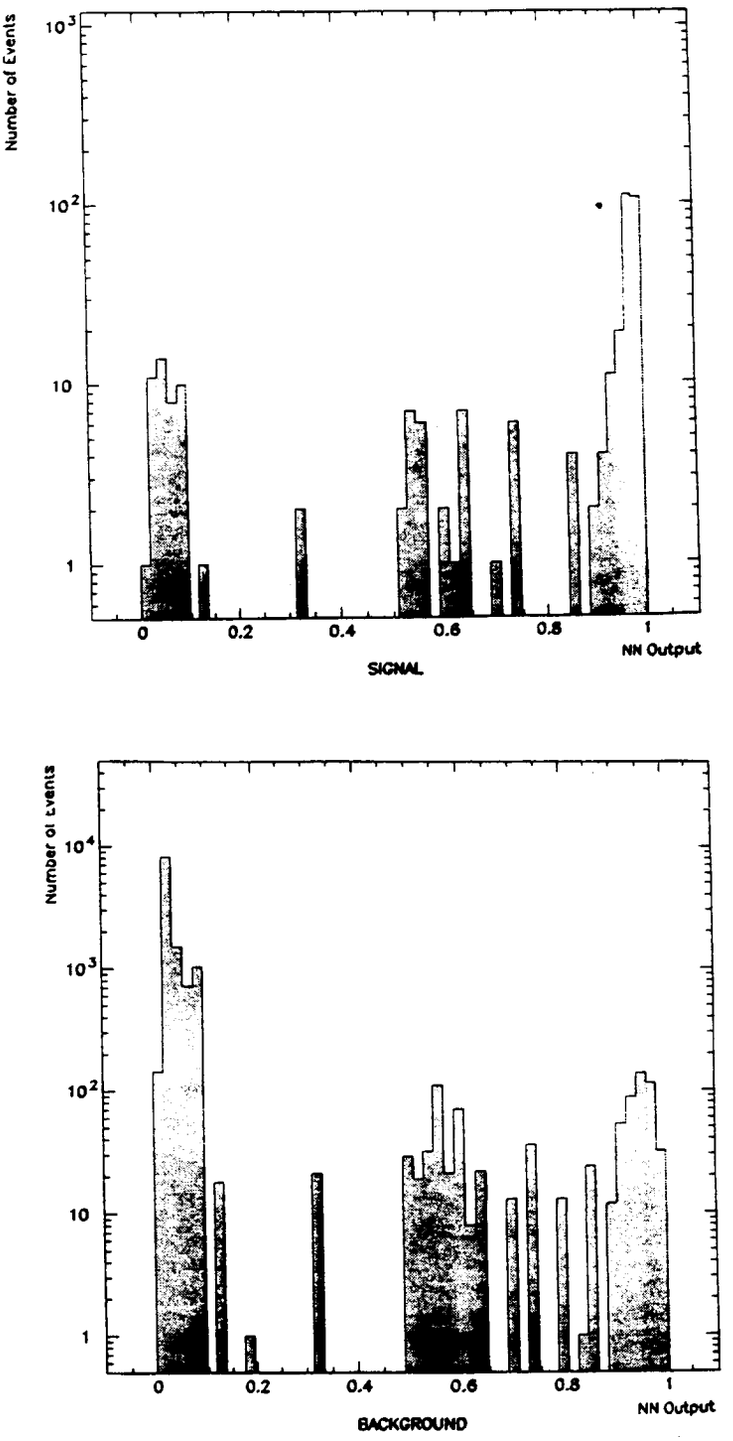

Fig. 3 – The neural network output for TRD (a) signal and (b) background events selected with high purity by means of the time-of-flight, spectrometer and calorimeter.

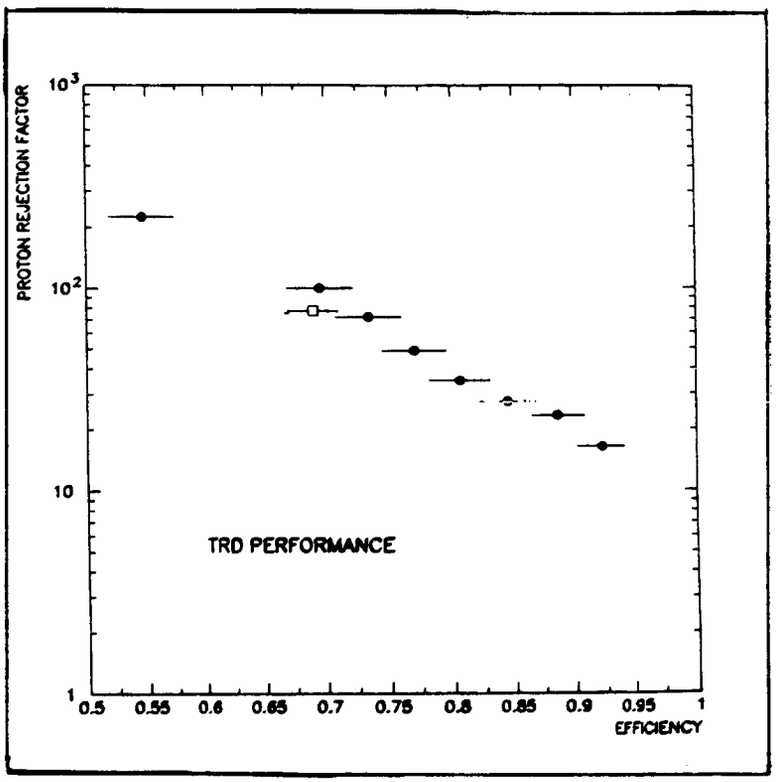

Fig. 4 – Proton rejection factor versus signal efficiency of the TRD on flight data (black circles); the result of the conventional analysis (ref. [13,15]) is also drawn (white box).

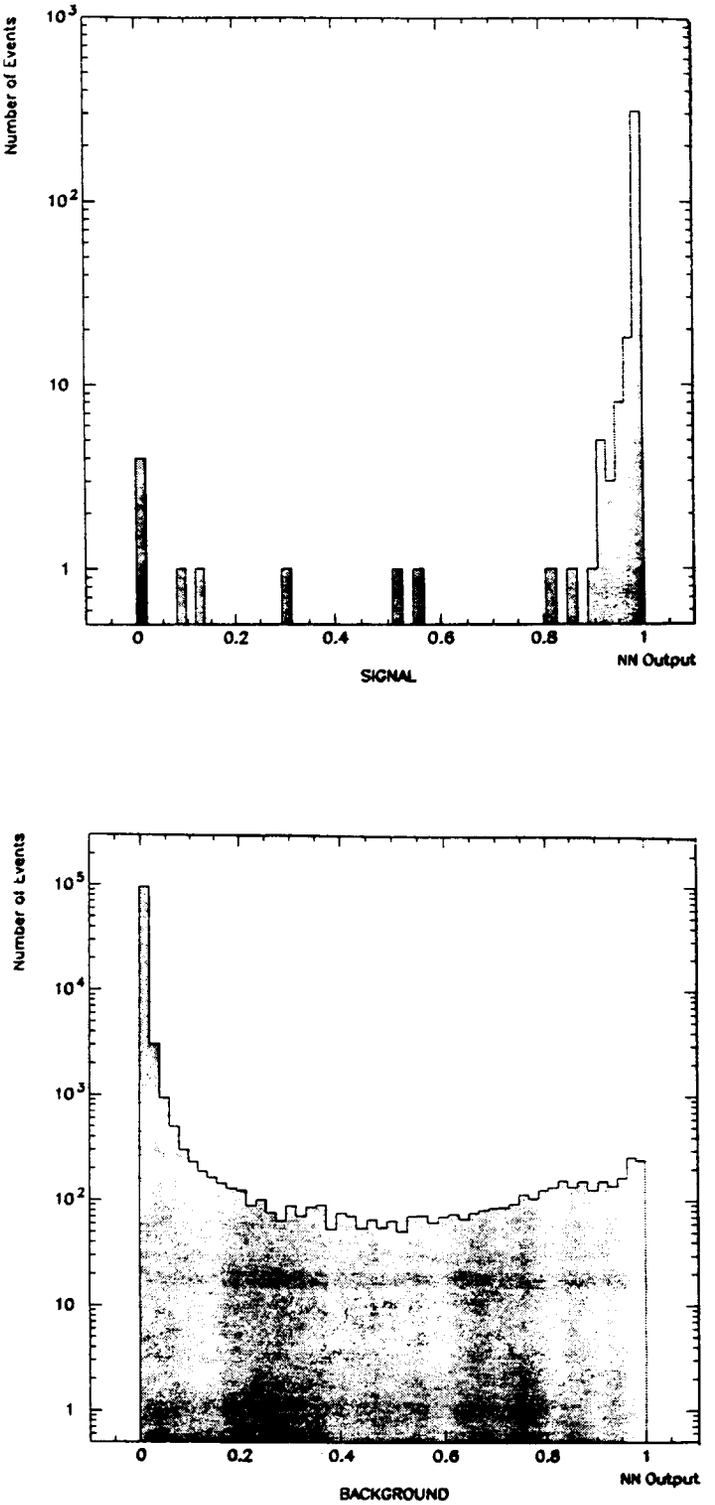

Fig. 5 – The neural network output for calorimeter (a) signal and (b) background events selected with high purity by means of the time-of-flight, spectrometer and TRD.

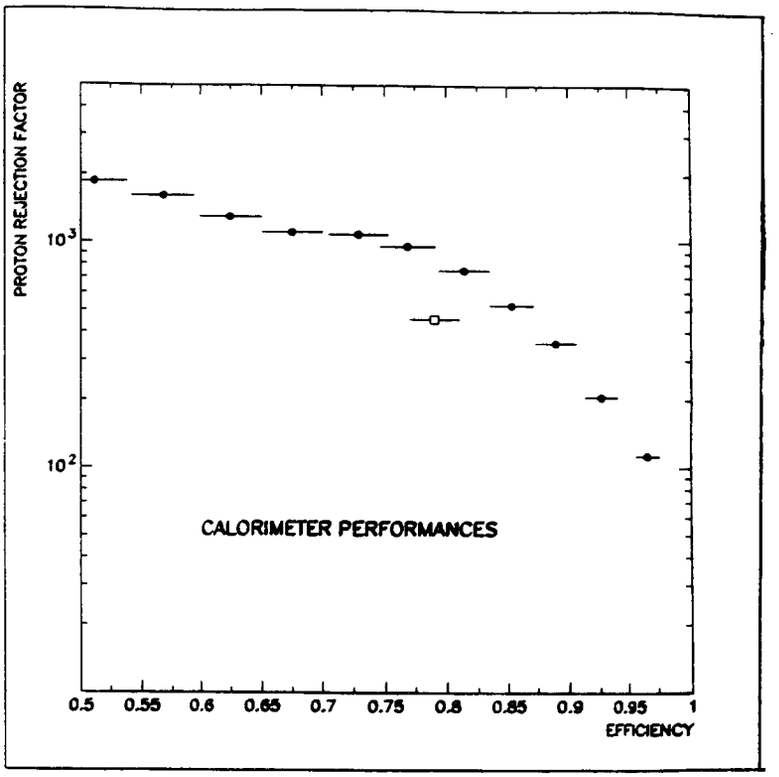

Fig. 6 – Proton rejection factor versus signal efficiency of the Calorimeter on flight data (black circles); the result of the conventional analysis (ref. [13,14]) is also drawn (white box).