

The use of Artificial Neural Network for dead channel data recovery in the DELPHI STIC calorimeter

M. Bonesini, M. Paganoni¹, F. Terranova

Dipartimento di Fisica, Università di Milano and Sezione INFN, Milano, Italy

S. Gumenyuk, L. Petrovykh ²

Sezione INFN, Milano, Italy and IHEP, Protvino, Russia

Abstract

Artificial Neural Networks (ANN) are a powerful tool widely used in High Energy Physics to solve track finding and particle identification problems. A entirely new class of application is related to the problem of recovering the information lost during data taking or signal transmission. Good performance can be reached by ANN when the events are described by quite regular patterns. Such a method was used for the DELPHI luminosity monitor (STIC) to recover calorimeter dead channels. A comparison with more traditional techniques is also given.

¹now at CERN, PPE division, Geneva

²Work partially supported by INTAS, contract INTAS-93-3602.

1 Introduction

Artificial Neural Networks (ANN) are powerful nonlinear tools for data analysis often used in High Energy Physics to solve track finding and particle classification problems [1, 2, 3]. More detector oriented issues are mainly limited to trigger and cluster reconstruction [4, 5]. A new field where ANN could be also fruitfully applied is related to possible loss of information during data taking. Nowadays, most detectors in particle physics consist of many independent channels. Each of them provides partial information about the current event and, sometimes, in case of faulty channels the global response of the detector can be reconstructed from the remaining active channels. ANN can perform this operation and give an estimate of the lost information. This technique is used for the DELPHI luminosity monitor (STIC) and a comparison with more traditional methods is given.

2 The STIC electromagnetic calorimeter.

The Small angle Tile Calorimeter (STIC) [6] is a lead - scintillator sampling calorimeter used to measure the LEP luminosity for the DELPHI detector [7, 8]. The STIC is composed of two cylinders ("arms") located ± 220 cm from the LEP interaction point; each arm of the STIC consists of 49 sampling layers made of 3.4 mm lead plates and 3.0 mm thick scintillator tiles. The tiles are arranged into 16 azimuthal sectors and 10 radial rings giving 160 towers projective to the interaction point (see fig. 1). The scintillation light is read out by WLS fibers running perpendicular to the planes with a density of about 0.8 fibers/cm² and is collected by Hamamatsu 1" R2149 phototetrodes. A precision machined tungsten "mask" in front of one arm (arm "C") defines the inner acceptance for the luminosity measurement.

The reconstruction of electromagnetic showers is performed by a clustering procedure which collects the energy deposited in the channels around the tower with the maximum signal. Each cluster is described by its total energy (E_{tot}) and its position at the STIC front face (radius R and azimuth φ). The main parameters of the STIC calorimeter are summarized in Table 1.

3 Influence of dead channels on the luminosity measurement and review of the methods used

The STIC luminometer has provided the absolute luminosity (\mathcal{L}) for DELPHI since May 1994 with a precision better than 0.1%. During the first 25 days of 1994 data taking (which correspond approximately to an integrated luminosity of 5 pb^{-1}) three channels did not work (period "A" of the STIC data taking). They were repaired and they were working properly for the rest of data taking (period "B" of the STIC data taking). Two of them are located in the outer rings of the calorimeter and, due to the angular dependence of the Bhabha cross section, the number of lost events is negligible for the luminosity measurement. The third channel ³ (number 154) is located in the fourth ring and therefore in that case a correction for the loss of events is needed. Without a specific procedure to recover dead channels a contribution to the systematic error on luminosity not higher than 0.03% is then expected ⁴. Dead channels distort both the energy and the position of the reconstructed electromagnetic clusters. Therefore the total number of selected Bhabhas (see Appendix 1) can be modified as well as the luminosity estimation. Since this error could give a contribution to the overall systematics (see Table 2), we have been forced to look for different methods to recover the dead channel information.

As mentioned above, each cluster reconstructed by the STIC calorimeter is described by three main parameters (E_{tot}, R and φ).

The cluster total energy is obviously $E_{tot} = \sum_{i=1}^n E_i$, where E_i is the energy detected by channel i and n is the total number of towers in a given cluster. The other parameters (R and φ) are calculated using the set of all E_i values by means of several estimators. In case of a dead channel it is necessary to estimate the value of the dead channel energy (E_{dead}) starting from the existing data. In principle, it is possible to reconstruct both R and φ without information on E_{dead} , using only the existing energy deposition

³In this paper this channel will be mentioned as C_{154} and E_{154} will be the energy deposited in it.

⁴Due to the $1/\theta^3$ behaviour of the Bhabha cross section and the small number of channels affected by this problem outside inner rings of the STIC calorimeter, the number of affected events is limited.

pattern (see Appendix 2).

Two different approaches for estimating E_{dead} were used in the analysis:

- E_{dead} parameterization as a function of the measured parameters of the reconstructed cluster;
- E_{dead} estimation using the Artificial Neural Network technique (see Section 5).

4 Dead channel energy recovery using a cluster parameterization

Different models were used to find correlations between the dead channel energy and the global cluster information. The sample of events from period "B" of data taking, where channel C_{154} was active, was used to set up and test the methods.

Good precision in the reconstruction of the dead channel energy (E_{154}) can be obtained by exploiting the characteristic signature of a Bhabha event, i.e. two back to back electromagnetic showers with energy around 45 GeV. A function which provides the energy in C_{154} can be built using as variables both the radius of the opposite cluster (R_c) and the total energy deposited in the 8 towers surrounding the dead one (E_{tot}^{surr}). This function is well-defined only for a limited sample of events (a slice of $f(R_c, E_{tot}^{surr})$ is shown in fig. 2). The cuts to define the sample are the following:

- Arm "C" (no dead channels), for the cluster with maximum energy:
 - $|\varphi_{max} - \varphi_{C_{154}}| \leq 22.5^\circ$
 - $14 \text{ cm} \leq R_{max} \leq 21 \text{ cm}$
 - $E_{max} \geq 29 \text{ GeV}$
- Arm "A" (dead channel 154):
 - at least 7 hits around C_{154}

The original and corrected energy spectra of Bhabha events are shown in fig. 3 and 4 (see also Table 3). The fraction of lost Bhabhas recovered with

this method is 83%. Attempts to increase this fraction by using more refined tuning of the same method were not successful.

The contribution to the overall systematic error in the luminosity measurement due to the statistical precision of the used method for dead channel correction is of the order of 0.01%.

5 Dead channel energy reconstruction with an Artificial Neural Network

For the problem discussed here the best version of the ANN is the "function fitting net" with back propagation updating ⁵. In this case the data processing is performed in two steps:

- **Network learning.** At this step the calorimeter data from the period "B" of the STIC data taking were used to train the ANN with the maximum statistics available.
- **Dead channel data reconstruction.** Afterwards, data from both periods "A" and "B" were used. A fraction of the period "B" data (which was never seen by the network at the learning stage) were used for learning tests and systematic error estimation. The data in period "A" were used for physics analysis.

Two input data streams were implemented for the ANN learning:

- the vector of 48 values of energy detected by the channels surrounding C_{154} ; this information is used to activate the top level of the network and does not include channel 154 itself.
- the energy detected by channel 154 which is used as the tagged value⁶.

The network has 48 nodes in the input layer, two hidden layers and one single node for the output. The architecture of the network is summarized in Table 4 (the terminology of reference [9] is used).

⁵The JETNET [9] ANN application package was used for the present analysis.

⁶After some investigations the best representation of the tagged value was expressed as: $TagValue = \ln\left(\frac{E_{154}}{E_{beam}} + 1.124\right)$ where E_{beam} is the energy of the LEP beams.

The learning required not more than 10000 events from period "B" . The network learning efficiency, which can be estimated in terms of mean error of the last epoch, is quite satisfactory, shows a stable behaviour and decreases with small fluctuations (see fig. 5). The input data were given to the network in a random order to make the learning procedure more efficient and all values were normalized in the range from 0 to 1. Since the radial density of events varies significantly, some additional efforts were made to obtain a more uniform learning and compensate for the $1/\theta^3$ behaviour of the Bhabha cross section. The patterns which are rare in the input data sample were used 2-3 times more often for learning and on the contrary, lower weights were given to the patterns which could have saturated the network.

The ANN data reconstruction quality was tested with the same sample of events used in Section 4 (20155 of these events were Bhabhas near the dead channel C_{154}).

Fig. 6 shows the distribution of $\Delta E_{154} = E_{154}^{orig} - E_{154}^{rec}$ where E_{154}^{orig} is the original energy detected by channel C_{154} and E_{154}^{rec} is the reconstructed one. The individual spectra of E_{154}^{orig} and E_{154}^{rec} are shown in fig. 7a and 7b. One can see a good agreement between the two distributions for $0 \leq E_{154}^{orig} \leq 20$ GeV. For bigger values of E_{154}^{orig} the response of the ANN is systematically smaller, but it does not affect the final selection of the Bhabha events used for the luminosity calculation. The total energy spectrum of the Bhabha clusters recovered by the ANN is shown in fig. 8. The difference between the original energy distribution and the distribution modified by the dead channel is clearly not negligible (fig. 8a); but the original and reconstructed spectra (fig. 8b) are in good agreement.

The procedure to get the cluster position (R and φ) used for the STIC data analysis needs information about the energy deposited in the channels around the barycenter of the cluster. The presence of the dead channel affects both values (see fig. 9a and 10a). The C_{154} energy recovery done by the ANN leads to a significant improvement in the cluster coordinate measurement (see fig. 9b and 10b). In particular, both plots show no systematic shift in the coordinate reconstruction. It has to be noted that the best precision for the cluster radial reconstruction in case of dead channel recovery is reached in the center of C_{154} ($\sigma \approx 38\mu m$) and it is worse near its border ($\sigma \approx 79\mu m$). This behaviour is opposite to the STIC radial resolution: all estimators in use show better performance near the borders of the tiles. The studies to

estimate the possibility of direct reconstruction of the cluster radial position for events with dead channels are discussed in Appendix 2.

The presence of the dead channel affects the number of reconstructed Bhabha events (some Bhabhas are recognised as non-Bhabhas and vice versa). Using the test statistics these smearing effects can be estimated. In particular, we observed no difference between the real and the reconstructed event distribution when applying radial and acoplanarity cuts (see Appendix 1). In fact a distortion in the number of selected Bhabhas after the energy cut was observed: 6 Bhabhas were recognized as non-Bhabhas and 27 non-Bhabhas recognised as Bhabhas. The contribution to the systematic error in the luminosity measurement due to the statistical precision of the method used to recover dead channels is of the order of 0.005%, negligible compared to the other sources of error shown in Table 2 and smaller than the one obtained with standard methods (see Section 4) ⁷.

The energy and impact cluster radius (R) are also reconstructed better with ANN methods than with traditional ones, as can be seen from Table 5.

6 Events classification with the ANN

ANN can be also used to classify Bhabha events directly without previous dead channel recovery. The parameters of this network are summarized in Table 6. The input layer of the ANN has a more sophisticated structure to take into account information from both calorimeter arms. This layer consists of 97 nodes. The first 48 values are reserved for the energy depositions detected by the channels surrounding C_{154} (the dead channel itself is not included). The second group of 49 nodes is used to accept the energy of the channels in arm "C" which are opposite to channel C_{154} . The output layer has one node with possible values in the range from 0 to 1. At the learning stage ⁸ the tagged value is 1 to mark Bhabha events and 0 for non-Bhabha events. At the testing stage the value of this node (V_{answer}) is used for event classification.

The ANN learning was done with 20000 events (around 70% of these

⁷The presence of dead channels can also introduce additional systematic errors as for example an inefficiency at trigger level ($\sim 0.015\%$ of the events can be lost in the case considered).

⁸Langevin learning with a Gaussian noise term is used.

events were Bhabhas selected by the standard STIC criteria) and 6951 events were used at the network testing stage (the total data sample was 40958 events).

Fig. 11 shows the network response for Bhabha (fig. 11a) and for non-Bhabha (fig. 11b) events. The following cuts are used for event classification:

- the event is a candidate Bhabha if $V_{answer} \geq 0.5$
- the event is a non-Bhabha candidate if $V_{answer} < 0.5$

The results of the ANN data analysis are summarized in the Table 7. A conservative estimate of the systematic error of this method can be calculated from the total number of wrongly classified events in the total data sample used for tests, giving a systematic error on luminosity smaller than 0.02%.

Conclusions

Two different approaches for dead channel recovery were developed and tested for the STIC luminosity monitor of DELPHI. The Artificial Neural Networks approach shows the best results in both energy and coordinate reconstruction (see Table 5). An error as low as 0.005% can be obtained on the luminosity measurement from the statistical precision of dead channel correction. Among the ANN models, the network described in Section 5 gives the minimum contribution to the total systematic error in luminosity and was finally used for physics.

Appendix 1

At e^+e^- colliders, the luminosity is measured by counting the rate of Bhabha events at small angles. The continuous progress in theoretical calculation of the Bhabha cross section and Monte Carlo simulation have reached an accuracy of $\simeq 0.2\%$ which has stimulated the construction of better luminosity monitors for high precision Z^0 physics at LEP.

In STIC luminosity analysis an event is classified as Bhabha if:

- there is at least one cluster in each arm with energy greater than 65 % of the beam energy ($E_{beam} = 45.638$ GeV in LEP during 1994 data taking)

- the radial position of the clusters with the maximum energy in the arms is: $R_C < 25.0$ cm for arm "C" and $8.2 < R_A < 28.0$ cm for arm "A"
- the absolute value of the acoplanarity (e.g. the value of $|\varphi_A - \varphi_C|$) of two maximum clusters in arms "A" and "C" is less than 20°

Appendix 2

Reconstruction of cluster radius can be done directly by the ANN without the intermediate step of E_{154} recovery. In this case the network structure was slightly modified (28 and 6 neurons in the hidden layers). The output layer provides a function of the cluster radius and the training procedure was set with 151907 events selected with the criteria described in Section 5).

In the training step, the original radius was computed by the standard STIC radial estimators ⁹ using also the information of C_{154} (available during period "B").

The testing step made use of a similar sample of events also from period "B" of the STIC data taking. However, events with the maximum energy deposit in rings number 1 or 7-10 were rejected since the performance of the ANN is poor and the contribution of C_{154} is negligible. These choices greatly improve the reconstruction performance but the whole technique remains less efficient than the method which employs the ANN for E_{154} reconstruction and recovers R analytically by means of STIC radial estimators. In particular, the long tails in the distribution of fig. 12 are essentially due to an implicit misinterpretation of E_{154} .

References

- [1] C. Bortolotto et al., Neural Networks in Experimental High Energy Physics, Int. Journal of Modern Physics **C3**, 733 (1992).
- [2] C.Peterson, Track Finding with Neural Networks, Nucl. Instr. Methods **A279**, 537 (1989).

⁹Essentially, they are functions of $\log \frac{E_1}{E_2}$ where E_1 and E_2 are the energy deposited above and below the tower with maximum signal.

- [3] W.S.Babbage and L.F.Thompson, The Use of Neural Networks in $\gamma - \pi^0$ Discrimination, Nucl. Instrum. Methods **A330**, 482 (1993).
- [4] H.Kolanoski, Application of Artificial Neural Networks in Particle Physics, DESY preprint DESY 95-061 (1993).
- [5] V.Breton et al., Application of Neural Networks and cellular automata to calorimetric problems, submitted to Nucl. Instrum. Methods.
- [6] A.C.Benvenuti et al. The DELPHI Small Angle Tile Calorimeter, DELPHI 94-157 CAL 120, 1994. Contribution to IEEE NSS Symposium, 1994.
- [7] P.Aarnio et al. The DELPHI detector at LEP, Nucl. Instr. Methods **A303**, 233 (1991).
- [8] S.J.Alsvaag et al. Performance of the new high precision luminosity monitor of DELPHI at LEP, DELPHI 95-68 PHYS 503, 1995. Contribution to EPS 95 Conference, Bruxelles.
- [9] L.Lönnblad, C.Peterson and T.Rögnvaldsson, JETNET 3.0 - A Versatile Artificial Neural Network Package, Lund Preprint LU TP 93-29 (1993).

Energy resolution for 45 GeV electrons	2.5 %
Energy linearity (from 10 to 100 GeV)	within ± 1.0 %
Radial resolution for different regions	250 μm at the border of tiles
	1.2 mm at the tower center
Resolution in the φ direction for different regions	1.0 mm at the border of sectors
	3.5 mm at the sector center
Number of hits for 45 GeV electron clusters	18.9 ± 2.2

Table 1: Main parameters of the STIC calorimeter

Source	Contribution to $\frac{\Delta\mathcal{L}}{\mathcal{L}}$ ^a
Monte Carlo statistics	0.2
Interaction point position	0.6
Precision of tungsten mask	0.4
Outer radius cut	0.2
Inner radius cut	0.2
Acoplanarity cut	0.1
Energy cut	0.3
Background subtraction	0.2
Trigger inefficiency	0.2
Total experimental	0.9
Total theoretical	1.6

^afraction of *per mill*.

Table 2: Sources of systematic error in luminosity

	Method of Sect. 4	ANN method
Total number of Bhabha events	118761	
Bhabhas lost if no correction done	990	
Bhabhas not reconstructed	164	6
Non-Bhabhas reconstructed as Bhabhas	4	27

Table 3: Results from Bhabha reconstruction by the traditional (Section 4) and ANN methods

Total number of layers used	4
Number of neurons in the 1-st hidden layer	25
Number of neurons in the 2-nd hidden layer	5
Number of patterns per update	10
Number of updates per epoch	100
Network activation function	$g(x) = x$
Learning rate (η)	0.2
Momentum parameter (α)	0.5
Number of epochs used for the ANN learning	60

Table 4: Characteristics of the ANN for the dead channel energy reconstruction

	Method of Sect. 4	ANN method
$\langle E_{rec} - E_{orig} \rangle$ (GeV)	0.110	0.051
$\sigma(E_{rec} - E_{orig})$ (GeV)	1.594	0.114
$\langle R_{rec} - R_{orig} \rangle$ (cm)	0.009	0.0034
$\sigma(R_{rec} - R_{orig})$ (cm)	0.11	0.0094

Table 5: Results from energy and radius reconstruction using the traditional and ANN methods.

Total number of layers used	3
Number of neurons in the hidden layer	40
Number of patterns per update	10
Number of updates per epoch	100
Network activation function	$g(x) = \frac{1}{1+\exp(-2x)}$
Learning rate (η)	1.0
Momentum parameter (α)	0.5
Number of epochs used for the ANN learning	70
Width of Gaussian noise in Langevin updating (σ)	0.01
Decrease in learning rate	0.999
Decrease in Langevin noise	0.99

Table 6: Characteristics of the ANN for the event classification

Events type	Number of events	ANN answer	
		Bhabha	non-Bhabha
Bhabha	4894	4888	6
non-Bhabha	2057	19	2038

Table 7: Results of the event classification with the help of the ANN for events near channel C_{154} , out of the total sample of 40958 events.

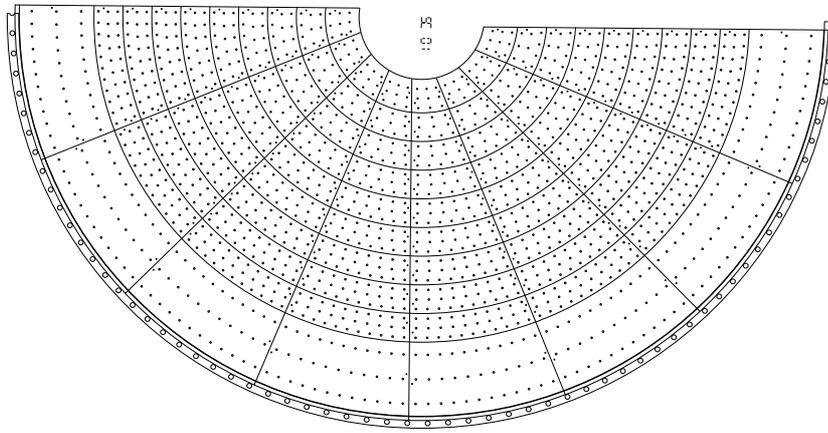


Figure 1: One plane of the calorimeter sandwich corresponding to one half arm.

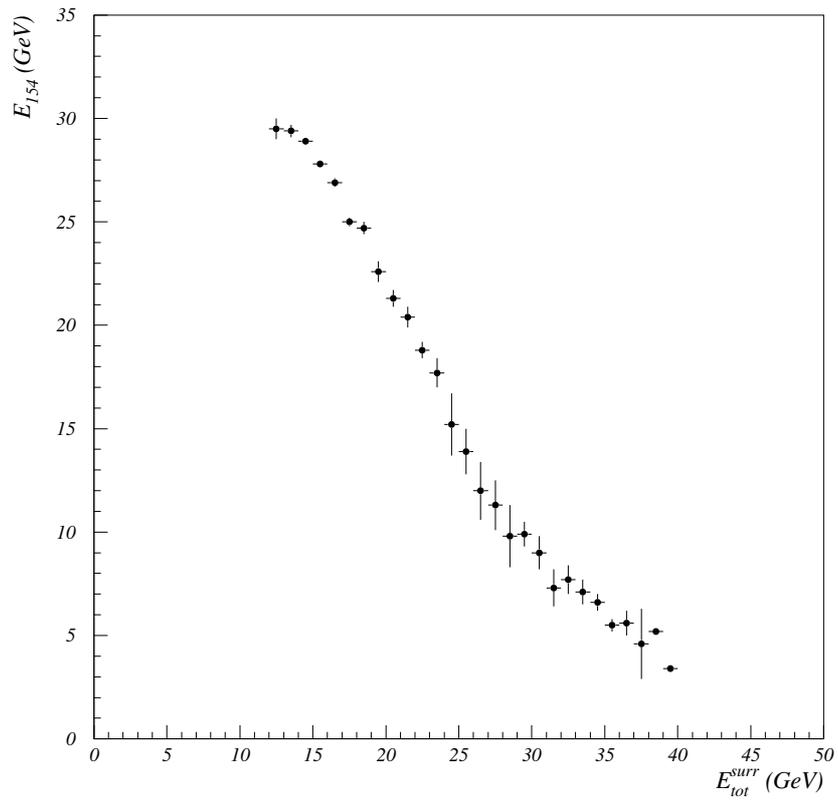


Figure 2: E_{154} as a function of E_{tot}^{surr} for $16 \leq R \leq 17$ cm.

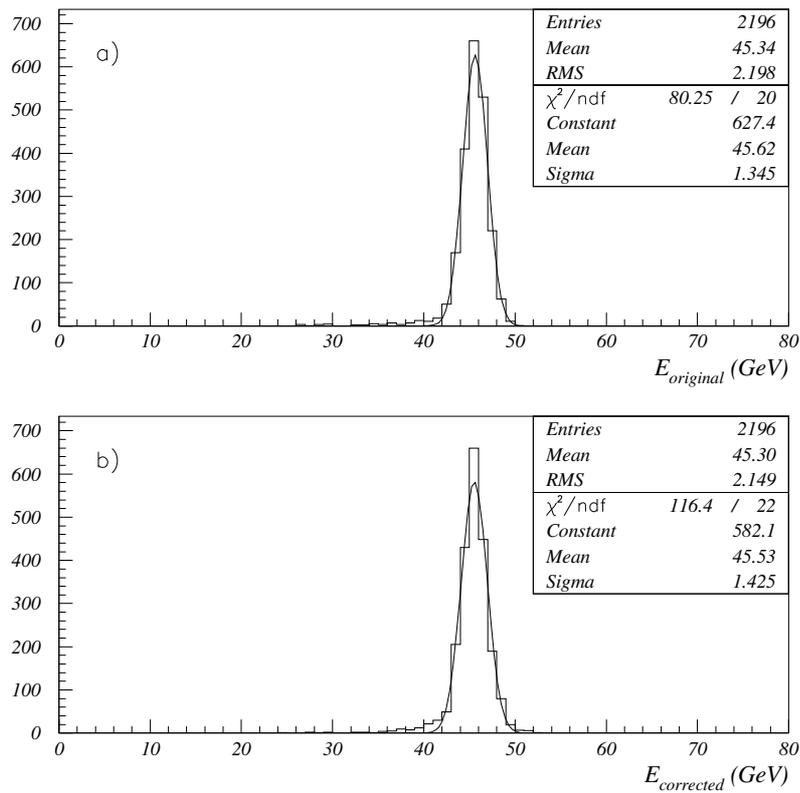


Figure 3: The total cluster energy distributions for the original data (a) and for the data corrected using the cluster energy parameterization (b).

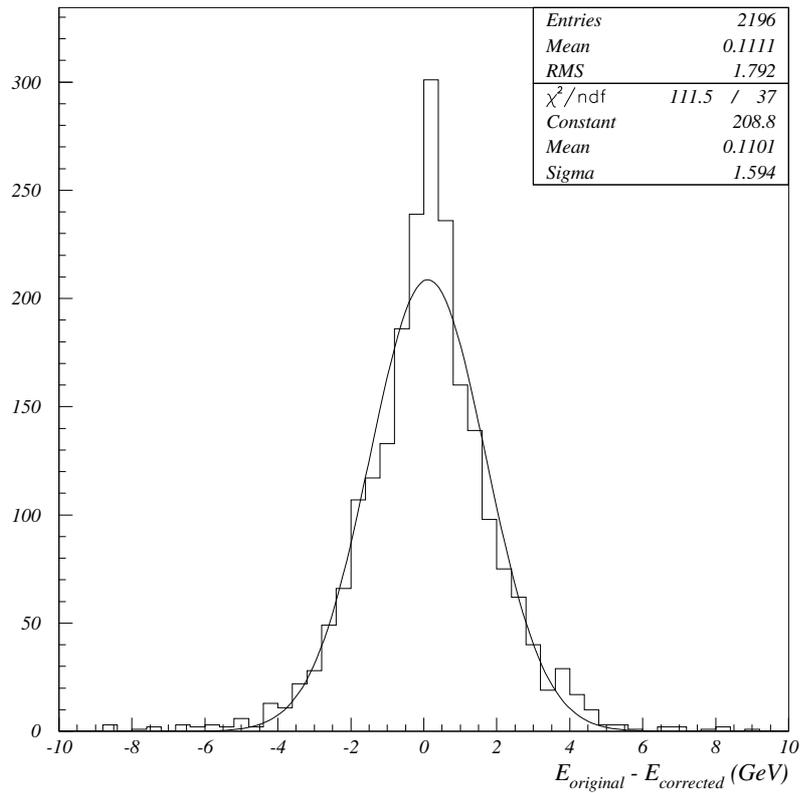


Figure 4: The difference between the original energy and the energy reconstructed using the cluster energy parameterization.

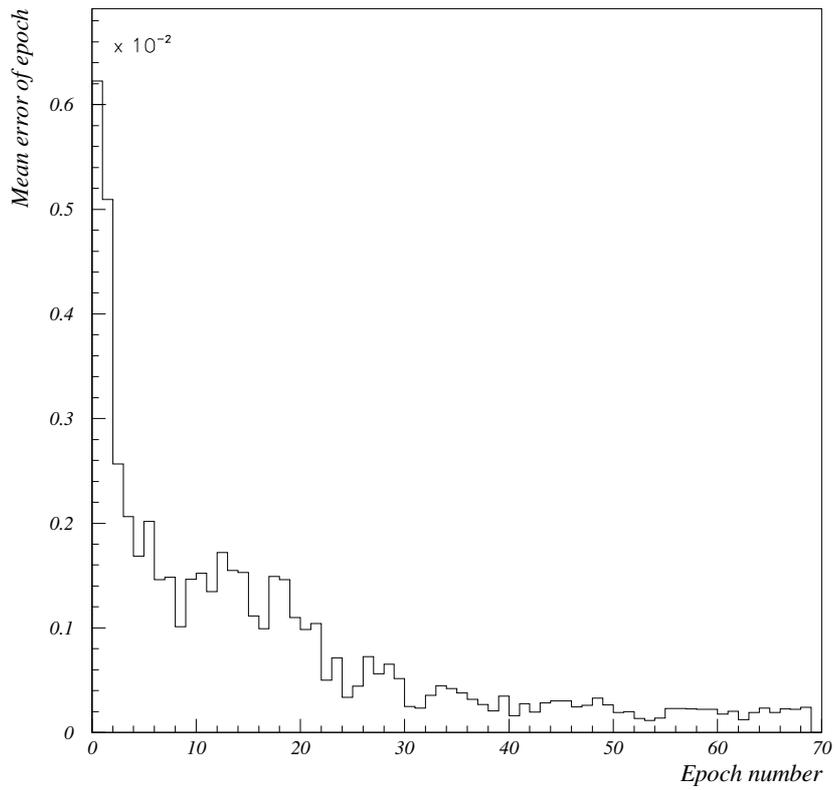


Figure 5: The network mean error as a function of the epoch number.

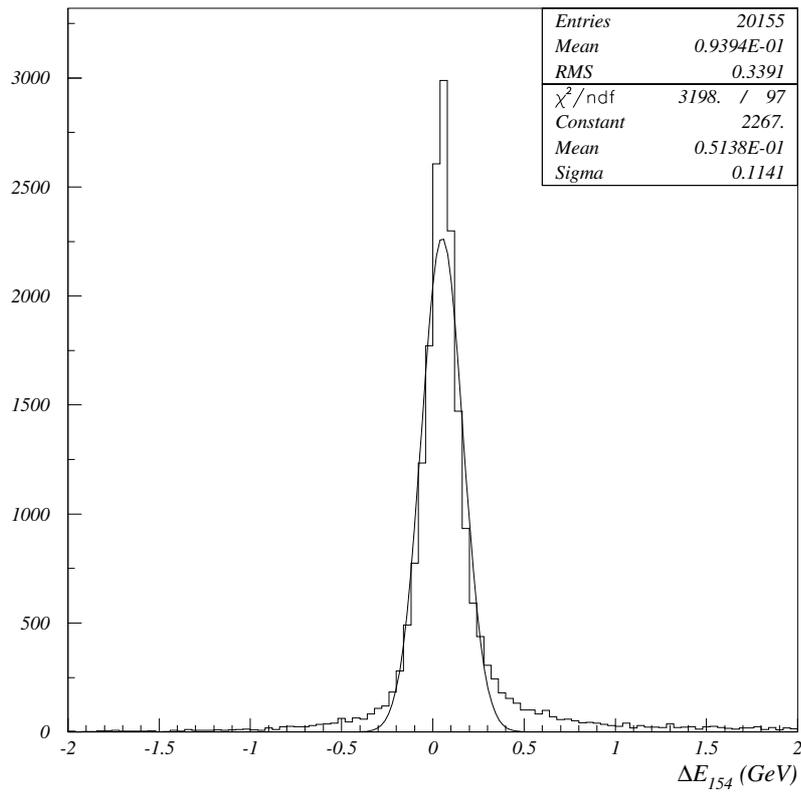


Figure 6: The difference between the original and the reconstructed energy of the counter C_{154} .

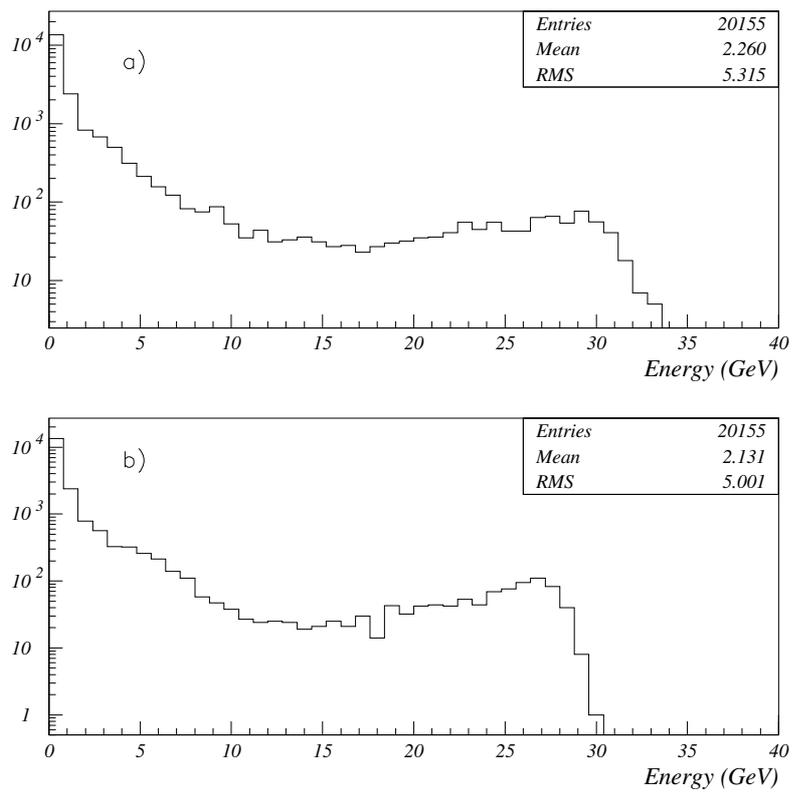


Figure 7: The original (a) and the reconstructed (b) energy deposition of the counter C_{154} .

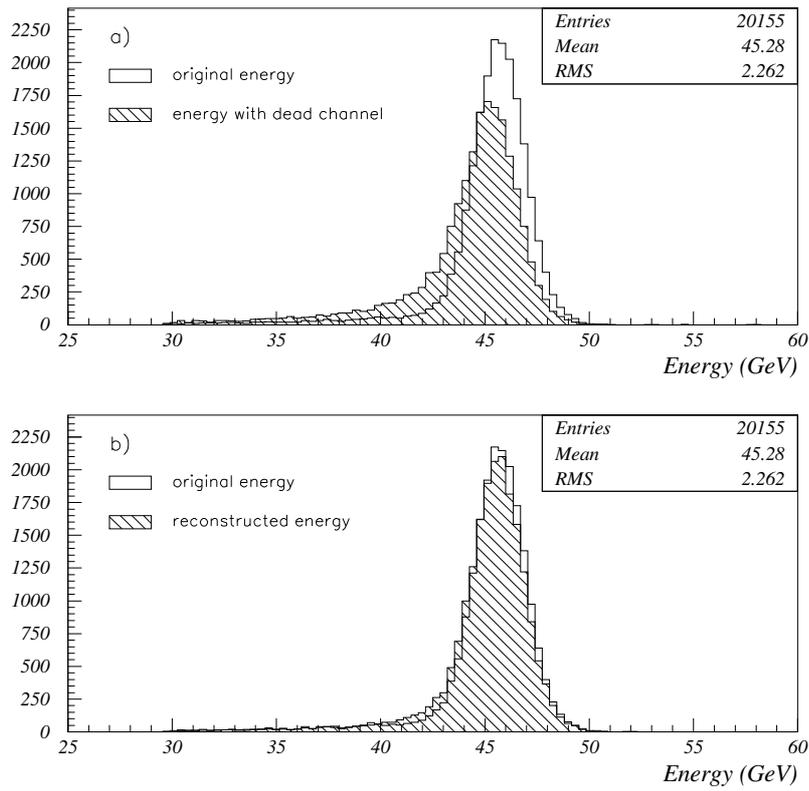


Figure 8: The total energy distributions for the Bhabha events without correction for dead channels (a) and with the dead channel recovery by ANN (b).

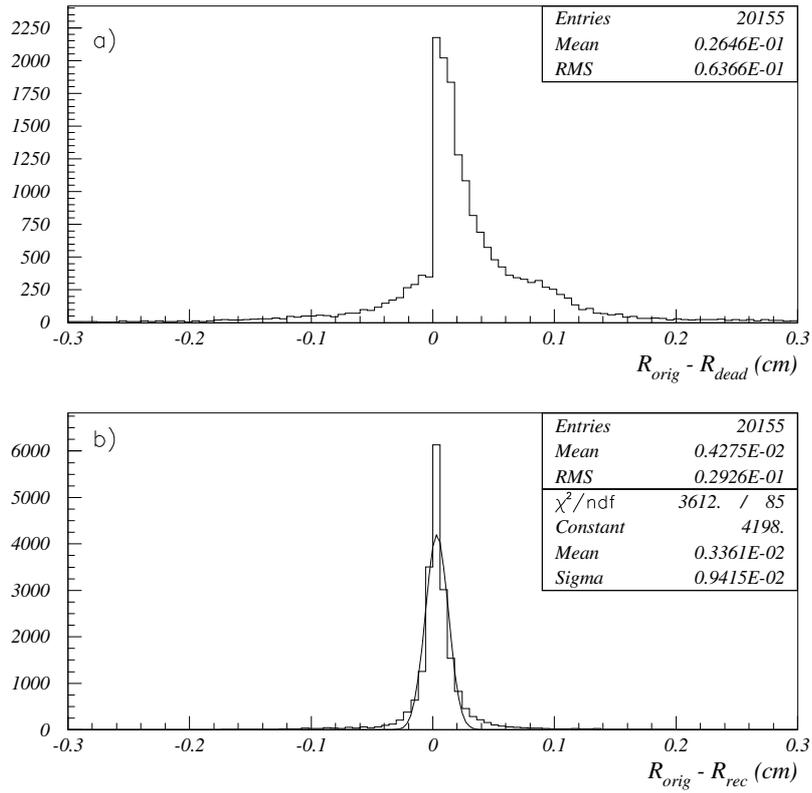


Figure 9: The difference between the original radial reconstruction of the cluster position and the data with the dead counter (a); the difference between the original radial reconstruction of the cluster position and the data recovered by the ANN (b).

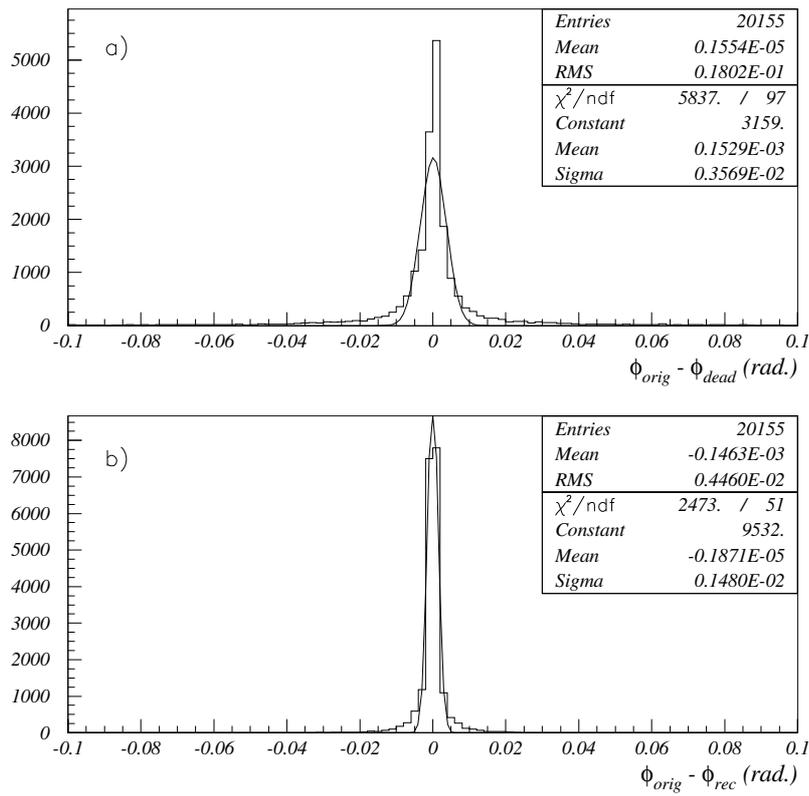


Figure 10: The difference between the original φ reconstruction of the cluster position and the data with the dead counter (a); the difference between the original φ reconstruction of the cluster position and the data recovered by the ANN (b).

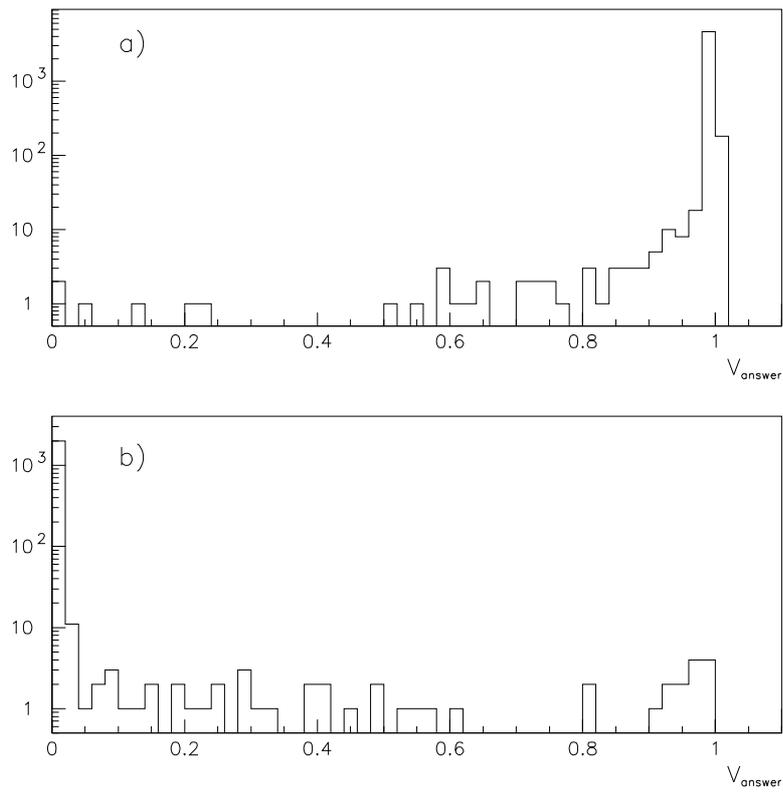


Figure 11: The classification ability of ANN for Bhabha (a) and non-Bhabha (b) events.

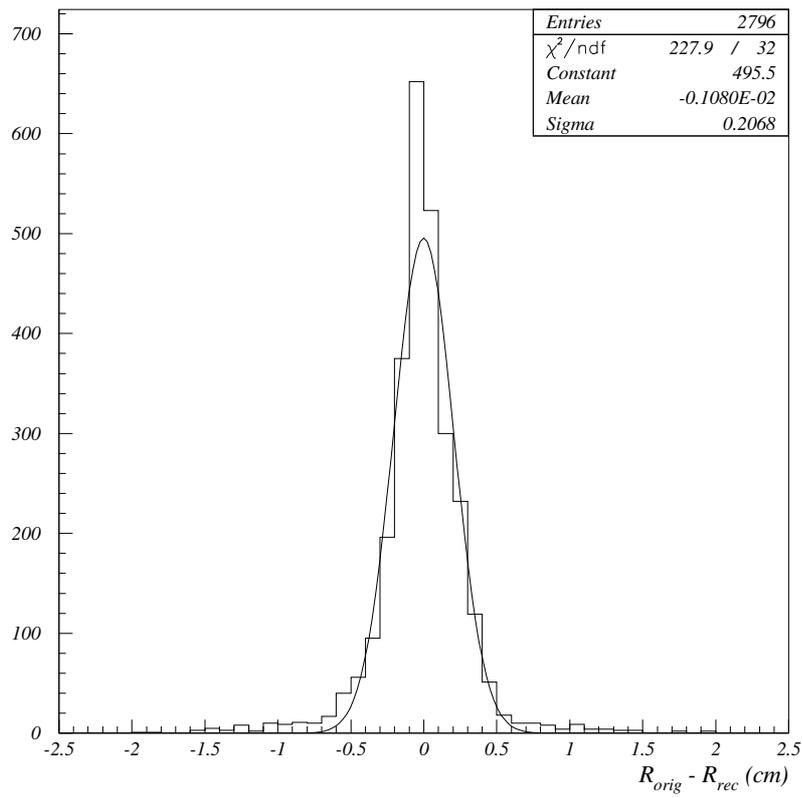


Figure 12: The difference between the original radial reconstruction of the cluster position and the cluster position after the direct radial reconstruction by the ANN.