

**Progress report on ν_τ CC - NC hadron
scattering event classification (preliminary)
& Phase II event classification**

N. Saoulidou & G. Tzanakos 24 - 10 - 2002

Outline

- **ANN for NC scattering - ν_τ CC separation**
- **ANN results on 38 kinks from Phase I & II**
- **ANN results on 4 ν_τ CC events of Phase I**
- **ANN for neutrino event classification**
- **Conclusion & On-going Work**

ANN for tau CC - NC scattering

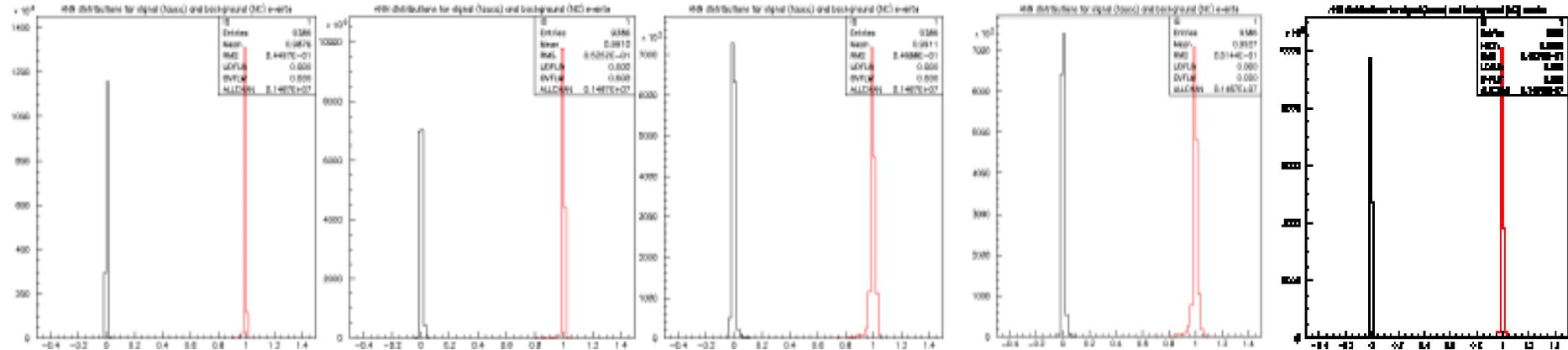
- **Goal** : To separate tau CC interactions from NC scattering interactions with the use of ANNs
- **Input Variables** :
 - **Daughter Momentum**
 - **Parent flight**
 - **Parent angle**
 - **Daughter angle**
 - **$\Delta\phi$ angle (between the parent and all the other primary tracks)**
 - **Daughter P_T**
- **Training Set** :
 - 20000 Tau CC interactions
 - 20000 NC scattering interactions

ANN construction

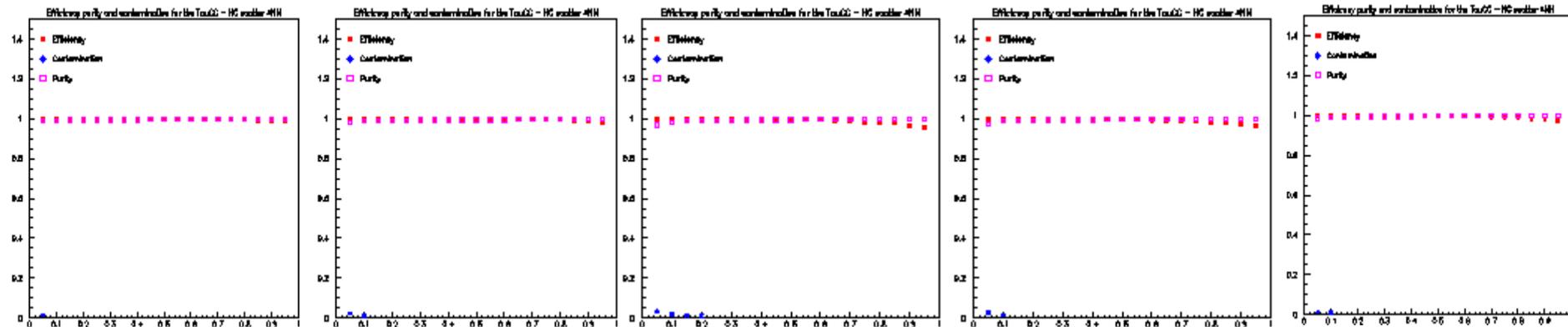
- We have constructed **five different** neural networks with the following characteristics:
 - **ANN1** : Four input variables (parent & daughter angle (θ_p, θ_d), decay length and daughter momentum (L_d, P_d))
 - **ANN2**: Four input variables (parent & daughter angle (θ_p, θ_d), decay length and daughter momentum (L_d, P_d)), with the daughter momentum **smearred by**
 - $\frac{\Delta p}{p} = 30\%$
 - **ANN3**: Five input variables (parent & daughter angle (θ_p, θ_d), decay length and daughter momentum, (L_d, P_d) and daughter P_T)
 - **ANN4**: Five input variables (parent & daughter angle (θ_p, θ_d), decay length and daughter momentum, (L_d, P_d) and $\Delta\phi$)
 - **ANN5**: Six input variables (parent & daughter angle (θ_p, θ_d), decay length and daughter momentum, (L_d, P_d) $\Delta\phi$ and P_T)

ANNs Performance

Output function on **Signal** and **Background** events



Efficiency Purity & Contamination



- *The performance of all five ANNs is very similar.*

ANNs Weights

ANN	ANN relative (%) weights for the Input Variables					
	P_d	θ_d	θ_p	l_d	$\Delta\phi$	P_T
1	42.6	24.4	16.1	16.9	—	—
2	23.6	32.6	29.0	14.7	—	—
3	25.7	22.2	25.3	9.1	—	17.6
4	26.9	35.1	18.5	13.6	5.9	—
5	21.7	25.2	23.3	9.1	7.0	13.7

- Since the ANN relative weights for the various input variables are different, the individual event probabilities will also differ.

ANN Results on Phase I & II kinks

- We computed the ANNs output on 38 kinks (supplied to us by Nonaka) and on the 4 tau CC events. 9/38 are Phase I kinks and the remaining Phase II.
- We repeated this procedure for the central value of the momentum and for 1 , 1.5, 2.0, 2.5 and 3 sigma below.

Kinks with ANN 1 , 2 &3 probability > 0.5

6

4

3

3

1

1

Momentum

p_{meas}

$p_{\text{meas}} -1\sigma$

$p_{\text{meas}} -1.5\sigma$

$p_{\text{meas}} -2\sigma$

$p_{\text{meas}} -2.5\sigma$

$p_{\text{meas}} -3\sigma$

Kinks with ANN 4 & 5 probability > 0.5

4

2

2

1

1

0

Momentum

p_{meas}

$p_{\text{meas}} -1\sigma$

$p_{\text{meas}} -1.5\sigma$

$p_{\text{meas}} -2\sigma$

$p_{\text{meas}} -2.5\sigma$

$p_{\text{meas}} -3\sigma$

ANN Results on Phase I 4 ν_τ events

- We computed the ANNs output on 38 kinks (supplied to us by Nonaka) and on the 4 tau CC events. 9/38 are Phase I kinks and the remaining Phase II.
- We repeated this procedure for the central value of the momentum and for 1 , 1.5, 2.0, 2.5 and 3 sigma below.

Kinks with ANN 1 , 2 &3 probability > 0.5

4

4

4

3

3

3

Momentum

$p_{\text{meas.}}$

$p_{\text{meas}} -1\sigma$

$p_{\text{meas}} -1.5\sigma$

$p_{\text{meas}} -2\sigma$

$p_{\text{meas}} -2.5\sigma$

$p_{\text{meas}} -3\sigma$

Kinks with ANN 4 & 5 ($\Delta\phi$ included) probability > 0.5

4

3

3

3

3

3

Momentum

$p_{\text{meas.}}$

$p_{\text{meas}} -1\sigma$

$p_{\text{meas}} -1.5\sigma$

$p_{\text{meas}} -2\sigma$

$p_{\text{meas}} -2.5\sigma$

$p_{\text{meas}} -3\sigma$

ANN Results on Phase I 4 ν_τ events (numbers)

P

Event	Run	P_d	θ_d	P_T	L_d	θ_p	$\Delta\phi$	ANN1	ANN2	ANN3	ANN4	ANN5
3263	25102	2.0	0.129	0.258	1815	0.190	0.176	1.0	1.0	0.9	0.7	0.7
3024	30175	2.9	0.094	0.271	4504	0.028	1.027	1.0	1.0	1.0	1.0	1.0
3039	1910	4.6	0.089	0.412	276	0.065	2.684	1.0	1.0	1.0	1.0	1.0
3333	17665	21.4	0.005	0.111	564	0.015	2.806	1.0	1.0	1.0	1.0	1.0

P - 1σ

Event	Run	P_d	θ_d	P_T	L_d	θ_p	$\Delta\phi$	ANN1	ANN2	ANN3	ANN4	ANN5
3263	25102	1.3	0.129	0.168	1815	0.190	0.176	0.9	0.9	0.8	0.4	0.4
3024	30175	2.1	0.094	0.197	4504	0.028	1.027	1.0	1.0	1.0	1.0	1.0
3039	1910	3.7	0.089	0.331	276	0.065	2.684	1.0	1.0	1.0	1.0	1.0
3333	17665	15.4	0.005	0.080	564	0.015	2.806	1.0	1.0	1.0	1.0	1.0

ANN Results on Phase II Kink events (numbers)

P

Event	Run	P_d	θ_d	P_T	L_d	θ_p	$\Delta\phi$	ANN1	ANN2	ANN3	ANN4	ANN5
3130	28864	0.8	0.097	0.078	3774	0.041	1.619	1.0	1.0	0.8	0.8	0.9
3111	6248	2.9	0.025	0.072	2780	0.055	0.021	0.9	1.0	0.8	0.6	0.6
3193	1361	3.1	0.019	0.058	1863	0.084	2.341	0.8	0.8	0.6	0.6	0.6
2809	11206	2.3	0.021	0.049	2023	0.058	0.245	0.8	0.9	0.6	0.4	0.4
3020	24151	8.8	0.005	0.044	2676	0.074	0.141	0.6	0.7	0.7	0.3	0.4
3182	10057	3.7	0.021	0.077	25	0.149	0.000*	1.0	1.0	0.9	0.9	0.7

P - 1 σ

Event	Run	P_d	θ_d	P_T	L_d	θ_p	$\Delta\phi$	ANN1	ANN2	ANN3	ANN4	ANN5
3130	28864	0.8	0.097	0.078	3774	0.041	1.619	0.9	0.9	0.8	0.7	0.7
3111	6248	2.9	0.025	0.072	2780	0.055	0.021	0.9	0.9	0.7	0.4	0.4
3193	1361	3.1	0.019	0.058	1863	0.084	2.341	0.3	0.5	0.4	0.4	0.3
2809	11206	2.3	0.021	0.049	2023	0.058	0.245	0.6	0.7	0.5	0.3	0.2
3020	24151	8.8	0.005	0.044	2676	0.074	0.141	0.2	0.3	0.4	0.1	0.2
3182	10057	3.7	0.021	0.077	25	0.149	0.000*	1.0	1.0	0.8	0.8	0.6

ANN Summary

- The performance of all ANNs (1-5) is quite satisfactory. The results are very similar as far as event classification is concerned.
- The first three ANNs, that do not use $\Delta\phi$ in the input variable set, give higher event probabilities both to the 4 Phase I tau events and to the Phase II kink events.
- The last two ANNs give lower event probabilities to one of the 4 tau events of Phase I (3263_25102) and, in general, to the Phase II kink events.
- There are six events, from the 38 Phase II kinks, with no identified lepton from the interaction vertex, that are assigned relatively high probabilities of being ν_τ CC events and could therefore be considered as interesting.
- The most dramatic feature of the **3 ν_τ CC events** of Phase I is the fact that their probabilities remain very high (~ 0.9 to 1.0) even when the daughter momentum is decreased by 2.5 sigma. This same characteristic exists in only **one Phase II kink event : 3182_10057** which is the only one that survives in this value of the momentum.

Event location status & Event characterization

- For the Phase II events we found that out of 227, 50 are not located by Bruce or us. We cross checked the results with Byron's list and the final number of not located events is close to 50 (or higher according to Byron).
- The Phase II events in Byron's list are accompanied by event characterization (numu CC, nue CC). So we compared this results with the ANN event characterization.

Method

- **Method :**

- Construct **two sequential** Neural Networks (ANN1 & ANN2) that will be **applied in the whole data set** :

a) The first to distinguish $v_{\mu} \text{ CC}$ from $v_e \& v_{\tau} \text{ CC} + \text{NC}$

b) The second to distinguish NC from $v_e \text{ CC} \& v_{\tau} \text{ CC}$

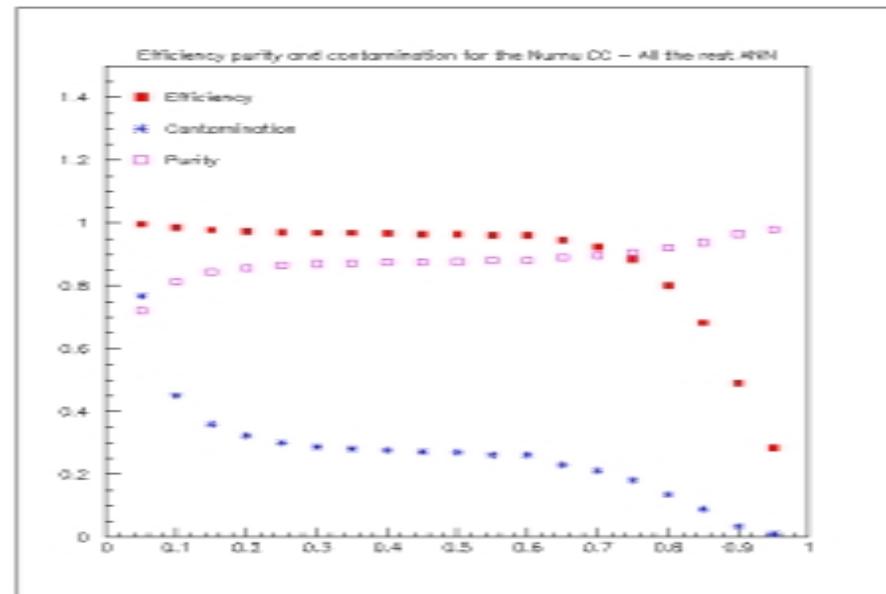
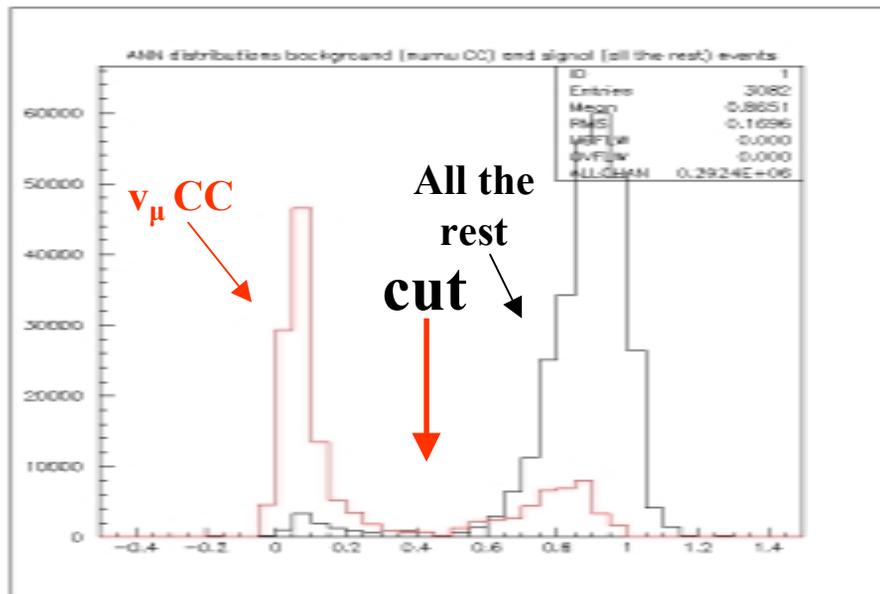
Training Set & Input Variables

- For **every period** we construct a **separate set of (2) ANN's** since every period has **different target configuration** and thus **different event characteristics**.
- For every period we use **5000 MC** events as a **training set**.

INPUT VARIABLES

- HITS** Total number of DC hits
(Total number of MID hits in the Central tubes)
- EMCAL** Total energy deposition
Number of clusters
Average Cluster energy
Mean value of the Clusters angle from the vertex with respect to the z - axis
Standard deviation of the Clusters angle
Mean Absolute deviation of the of the Clusters angle
Higher Moments of the Clusters angle : a) Skewness b) Curtosis
(Percentage of tracks with $E/P < 0.3$ (Muons))
- TRACKS** Number of final tracks
Number of DC tracks
(Number of tracks that have more than 3 hits in the MID system (Muons))
- OTHER** Total Pulse Height in the SF system

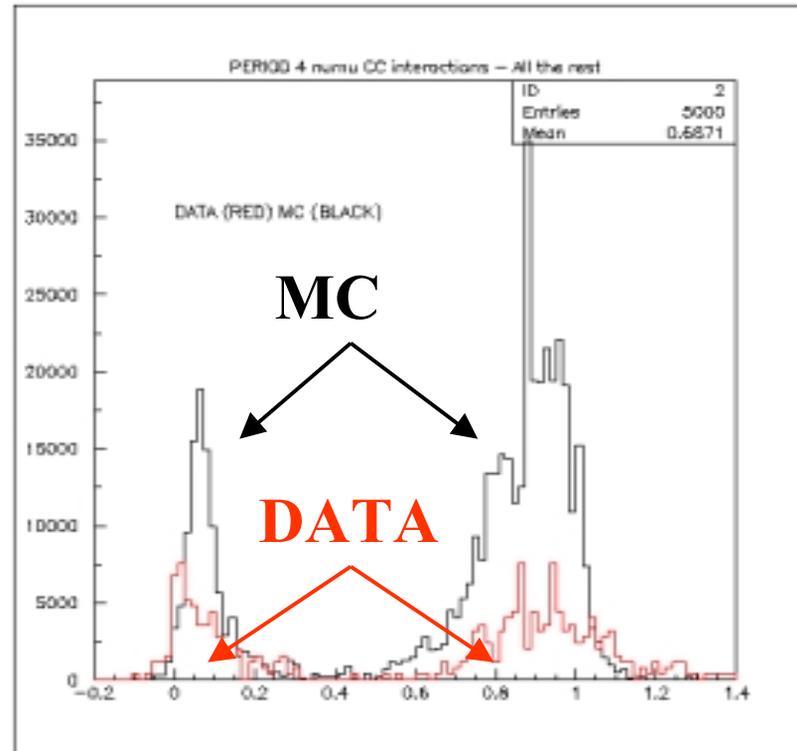
-Output of ANN1 (ν_μ CC - All the rest)-



- The **performance** of that network is **satisfactory**.
- With a **cut @ 0.5** in the network output function we select “**signal**” events and on the same time “**background**” events with :

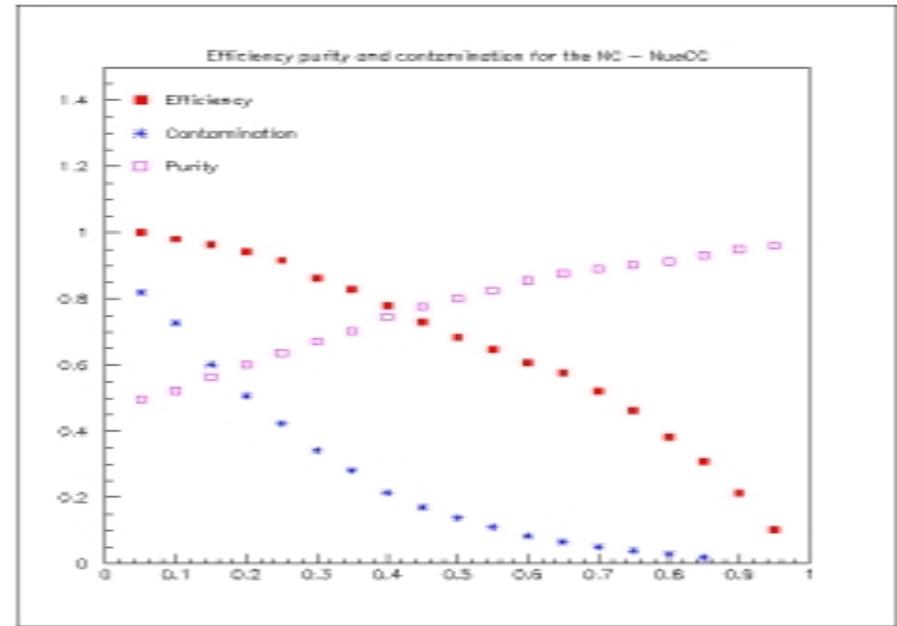
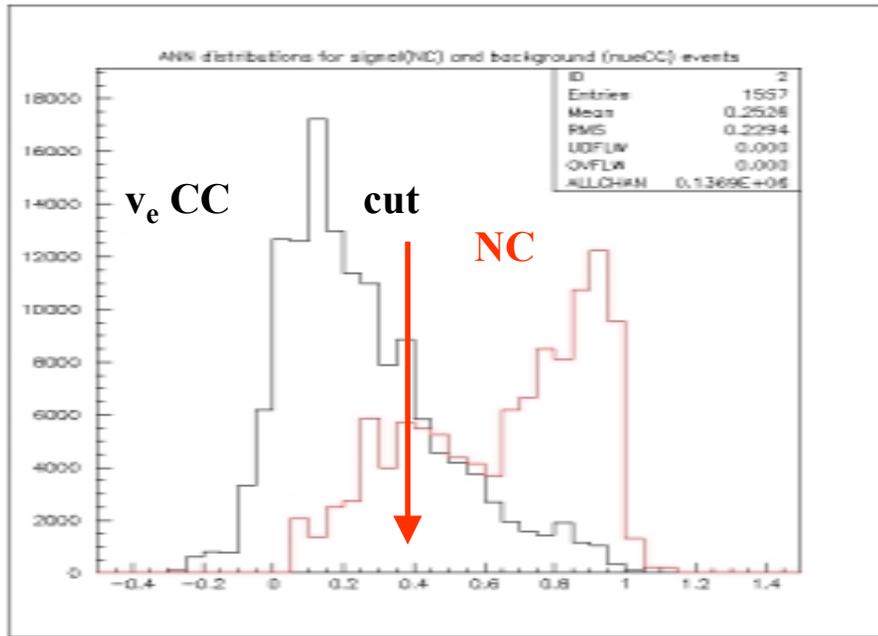
All the rest	efficiency 96 % - purity 88 %
ν_μ CC	efficiency 73 % - purity 96 %

ANN1 (ν_μ CC - All the rest) performance on MC & Real Data



- The performance of the **output function** of **ANN1** in **MC** events and in the **Real data set** is **very similar**.
- That indicates that the **results from ANN1** implementation in the **experimental data set** are **quite reliable**.

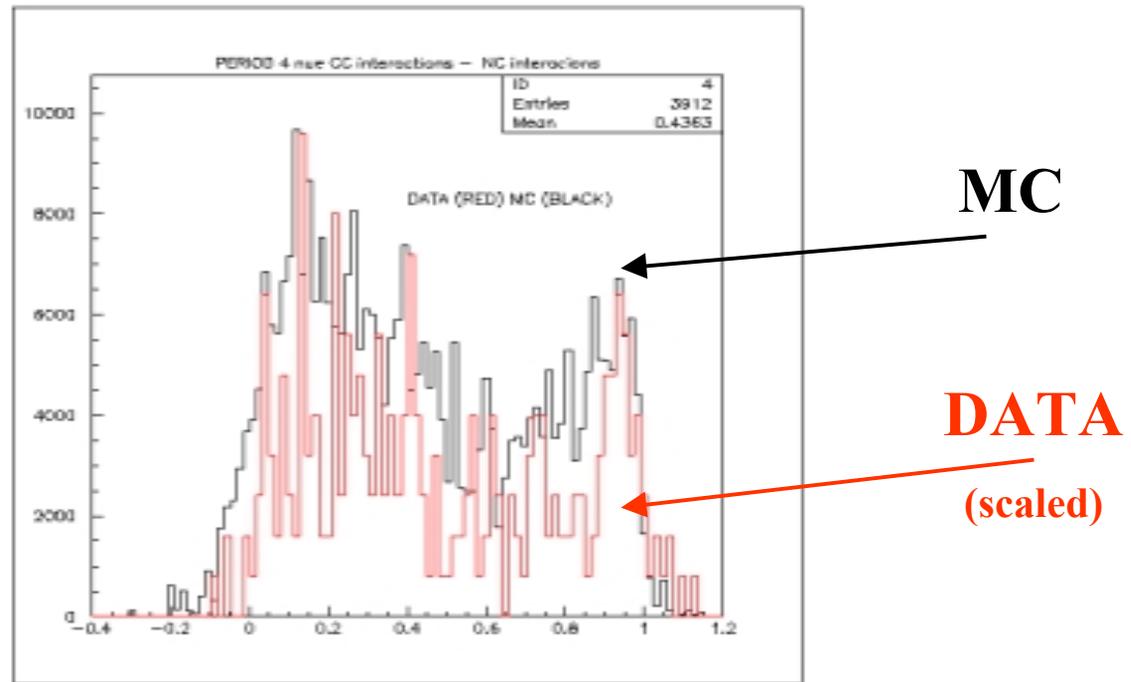
Output of ANN2 (NC - ν_e CC)



- This network shows a quite good behavior and by choosing a **cut @ 0.5** we select **signal (NC)** and at the same time **background** events (**ν_e CC**) with :

NC	efficiency 68 % - purity 80 %
ν_e CC	efficiency 86% - purity 76 %

ANN2 (NC - ν_e CC) performance on MC & Real Data



- The **performance** of the **output function** of **ANN2** in **MC** and in the **Experimental data** set is **very similar**.
- That permits us to consider the **results** of **ANN2** quite **reliable**.

ANN1 & ANN2 Results on Phase II events.

	Byron	ANN (agreement)	
ν_μ CC	100	92	→ 8 : 4 NC & 4 nue CC
ν_e CC	39	34	→ 5 : 2 NC & 3 numu CC

	ANN (total)
ν_μ CC	99
ν_e CC	56
NC	58

On going work

- We are continuing to study the tau CC - NC hadron scattering classification with the ANN approach.
- We will view all Phase II events in the event display in order to examine the validity of the ANN characterization. (There were some cases where some variables, like the number of final tracks with more than 3 hits in the MID, had the “wrong” value due to “old” spectrometer file.)
- We will finish the attempts for event location in the 14 events that come from our re-predictions and send the results to Nonaka.