Primary Data Processing Algorithms for Neutron Monitors

P Paschalis, C Sarlanis, H Mavromichalaki

Nuclear and Particle Physics Section, Physics Department, National and Kapodistrian University of Athens, Zografos, 15783 Athens, Greece

E-mail: emavromi@phys.uoa.gr

Abstract. The primary data processing of the neutron monitors is a necessary procedure in order to provide the worldwide network of neutron monitors with high quality data. The procedure should be performed in a real time code which means that it should be fast and make use only of the past measurements of a neutron monitor. In general, the data correction algorithms are based on the comparison among the different channels of the detectors. Such methods, which are used currently by the Athens neutron monitor station as well as by many other stations are the Median Editor and the Super Editor. In this work, two new algorithms that are currently being developed in the Athens Station are presented. The first one is based on an Artificial Neural Network model, while the second one is based on a pure statistical model.

1. Introduction

A neutron monitor consists of a number of independent counters (channels) each one of which measures the flux of the cosmic ray neutrons. The total count rate of the neutron monitor is based on the sum of the measurements of each channel. Unfortunately, in some cases the counting rates of one or more channels are distorted by an instrument variation [1]. In that case, the measurements of these channels are not related to the real flux of neutrons and the measurements should be rejected or corrected.

The primary processing algorithms aim to the correction or rejection of the channels that are distorted by instrument variations [2]. The correction should be performed in a real time basis, since the data have to be sent to NMDB whose data are used by online applications such as the GLE Alert. In the real time process, only the past measurements of the channels are known and for this reason all the primary data processing algorithms are based on the comparison of the different channels of the neutron monitor [3].

In general, a primary data processing algorithm should have three main characteristics. It should (a) be fast, (b) filter effectively the instrument variations and (c) leave unchanged the channels that are not distorted. The (a) and (b) are fulfilled by all the algorithms. As for the third characteristic, it is known that the algorithms in general compress the standard deviation of the no erroneous data. In this work, two new algorithms are presented based on Artificial Neural Networks (referred to as ANN from now on) and a pure statistical model.

2. Artificial Neural Network Algorithm

The ANN is a well known computational tool that can be used in many applications in a variety of fields [4]. It is composed of two or more layers and each one of them is composed of nodes named

"neurons". The nodes of each layer are connected with the nodes of the next layer through connections named "synapses". Each synapse is related to a weight factor which acts as a multiplier factor when a value is transferred through it. At each neuron, the input values are summarized and the result is processed by an activation function (usually a sigmoid function). The output of the neuron is transferred to the next layer and the process continues until the last – output layer. A general structure of an ANN is given in Fig. 1.
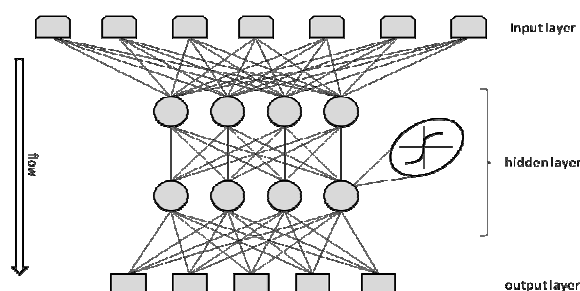

Figure 1. General structure of an Artificial Neural Network

When the ANN is firstly created, synapses are assigned with default or random values. In order for the network to produce a correct output for a specific input, synapses should have correct values. This is achieved through a learning procedure, during which the network is fed with training data (usually simulation data) and is forced to output the desired result. The network compares the actual output with the desired one and adapts the weights in small steps. After the training the ANN is ready to be used and gives an immediate response for any input.

The implementation of the algorithm for the case of the primary data processing was made in C++. The ANN used was the MLP class from ROOT data analysis framework, developed in CERN [5]. The training sample, used in the training of the ANN, was simulation data that were generated after a thorough analysis of the neutron monitor's statistics. It consisted of 10000 samples that represented some possible measurement cases of the neutron monitor. Each sample was distorted by random variations which represented the possible instrument variations. The parameters of the ANN are shown in Table 1.

ANN parameters	
Architecture	Input (6) : Hidden (30) : Output (6)
Activation function	Sigmoid
Learning method	Stochastic
Eta parameter	0.1
Training Sample	10000
Test Sample	2000
Weights Initialization	No randomization
Inputs	Normalized
Epochs	1000
Training time	$\approx$ 10 minutes @ Intel i7

Table 1: ANN parameters used in this work

After the training, the network was used for the correction of data in a real time procedure. The results are given in the result section.

3. Edge Editor
The second algorithm is based on a pure statistical model and is currently in an optimization phase. The general principale of the algorithm is shown in Figure 2. For this method it is supposed that the minutely data follow a distribution. At this early stage of the algorithm, a Gaussian distribution is

considered and its sigma is calculated after a thorough statistical analysis on the data of the neutron monitor.
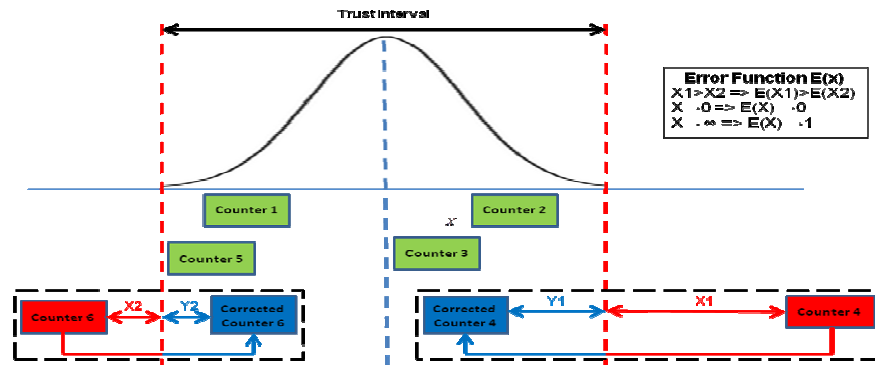

Figure 2. Operation diagram of Edge Editor

In the real time procedure of the algorithm, according to the measurements of the different counters of the neutron monitor, the most probable average value of the data is calculated. After that, the counters with their measurement within the trust interval (3 sigma or 4 sigma) remain unchanged. The counters that are outside the trust interval are corrected according to the following logic. The farther the measurement is from the edge of the trust interval, the closest to the mean value the corrected value is positioned. The logic is that a value that is farther from the trust interval is more possible to be an instrument variation so it has to be rejected. A value that is closer to the trust interval is more possible to be a statistical variation, so it does not have to be changed much. For this task, an error function is used which gives values between 0 and 1 and measures the level of error of the values that are outside the trust interval. The results of the method are also given in the result section.

4. Results

In this section, the correction using the two new algorithms is presented. The correction using the Median editor [6] is presented as well for comparison reasons. The results are given in Figure 3.
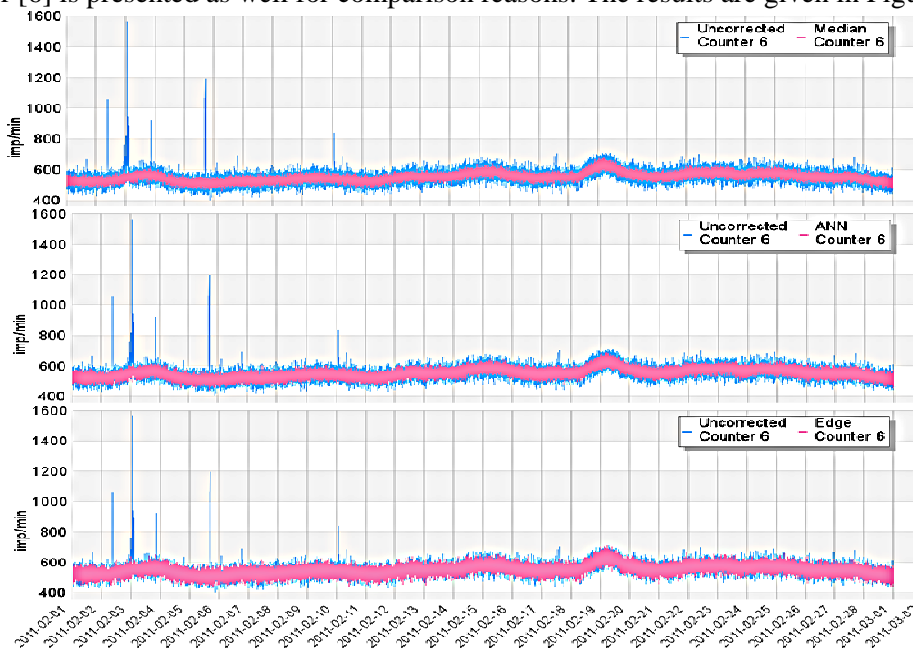

Figure 3. Uncorrected vs Corrected data of the counter 6 of Athens Neutron Monitor during the February 2011, using Median Editor (Upper), ANN (Middle), Edge Editor (Bottom).

All the algorithms successfully filter the instrument variations. The known behavior of the Median Editor that compresses the standard deviation of the data is presented. This behavior is also presented in the ANN algorithm but in a less significant level. In the Edge Editor, this behavior is almost absent.

In order to determine the affection on the non erroneous data, a quiet period of measurements where the counting rate is almost constant and without any instrument variation was selected. Such a period, was the one between the 17th and the 21nd of August 2011. For this time period, we calculated the mean value and the standard deviation of the uncorrected data and the corrected with the Median Editor, ANN method and Edge Editor data. The results are given in Table 2.

	Uncorrected Data		Median Editor Data		ANN Data		Edge Editor Data	
	mean value	standard deviation	mean value	standard deviation	mean value	standard deviation	mean value	standard deviation
Counter 1	598.69	31.43	598.8	18.01	598.21	23.65	598.34	28.55
Counter 2	609.28	31.92	610.49	18.36	608.52	25	609.16	29.65
Counter 3	533.83	30.51	532.59	16.02	532.08	22.11	533.46	26.77
Counter 4	529.65	29.26	527.72	15.87	531.13	23.18	529.12	26.11
Counter 5	507.33	29.59	505.33	15.2	505.75	20.96	506.93	25.74
Counter 6	534.48	30.02	536.49	16.14	534.09	21.52	534.28	26.44

Table 2: Statistics of Athens cosmic ray data for the quiet period of August 17-21, 2011

The mean value of all the correction algorithms is very close to the uncorrected data. However, the mean value of the Edge Editor is closer than the others. Moreover, as it was visually concluded from the figure 3, the Median Editor compresses the standard deviation. The ANN has a better behavior while the standard deviation of the Edge Editor is very close to the standard deviation of the raw data.

5. Conclusion

The two new algorithms, the ANN algorithm and the Edge Editor, can successfully filter the instrument variation of a neutron monitor. Compared to the Median Editor that is currently used in the Athens station, they present a better behaviour considering the compression of the standard deviation. Especially the Edge Editor, despite the fact that it is in an optimization phase, seems to work very effectively. In order to evaluate the stability of the algorithms, an additional application that corrects the data in a real time basis with all the presented algorithms, has been created. The corrected data are stored in a local database in Athens station, to be evaluated in time.

References
[1]   Chilingarian A, Hovhannisyan A, Mailyan B 2009 *Proc 31st ICRC (Lodz)* icrc0677
[2]   Hovhannisyan A, Chilingarian A 2011 *Adv Space Res* **47** 1544
[3]   Belov A V, Blokh Ya L, Klepach E G, Yanke V G 1988 *Kosmicheskie Luchi Moscow Nauka (in Russian)* **25** 113.
[4]   "Artificial Neural Networks", http://en.wikipedia.org/wiki/Artificial_neural_network
[5]   ROOT MLP Class, http://root.cern.ch/root/html/TMultiLayerPerceptron.html
[6]   Yanke V, Belov A, Klepach E, Eroshenko E, Nikolaevsky N, Kryakunova O, Sarlanis H, Mavromichalaki E, Gerontidou M 2011 *Proc 32nd ICRC (Beijing)* icrc0599