Nuclear Instruments and Methods in Physics Research A 714 (2013) 38–47

Contents lists available at SciVerse ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Optimization of neutron monitor data correction algorithms

P. Paschalis, H. Mavromichalaki*

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

ARTICLE INFO

Article history:
Received 5 January 2013
Received in revised form
4 February 2013
Accepted 25 February 2013
Available online 4 March 2013

Keywords:
Cosmic rays
Neutron monitors
Data correction

ABSTRACT

Nowadays, several neutron monitor stations worldwide, broadcast their cosmic ray data in real time, in order for the scientific community to be able to use these measurements immediately. In parallel, the development of the Neutron Monitor Database (NMDB; http://www.nmdb.eu) which collects all the high resolution real time measurements, allows the implementation of various applications and services by using these data instantly. Therefore, it is obvious that the need for high quality real time data is imperative. The quality of the data is handled by different correction algorithms that filter the real time measurements for undesired instrumental variations. In this work, an optimization of the Median Editor that is currently mainly applied to the neutron monitor data and the recently proposed ANN algorithm based on neural networks is presented. This optimization leads to the implementation of the Median Editor Plus and the ANN Plus algorithms. A direct comparison of these algorithms with the newly appeared Edge Editor is performed and the results are presented.

1. Introduction

The neutron monitors are the ground based detectors that measure the secondary cosmic ray flux [19]. The first neutron monitor stations have been in operation for more than 60 years, while new stations are still being established. The measurements of the neutron monitors are of great importance for the scientific community and play a key role as a research tool in the field of space physics, solar–terrestrial relations, and space weather applications. For this reason, nowadays, a great number of neutron monitor stations broadcast the measured cosmic ray intensity in real time. Recently, a European project (http://www.nmdb.eu/) has developed a database to which the neutron monitor stations may send their one minute resolution data. This database also contains the data archives of the neutron monitor stations. The final aim is the gathering of all the neutron monitor measurements in real time, if it is technically possible, and in a common format, in order for them to be instantly used from the scientific community.

The fact that the neutron monitor stations are spread worldwide, in locations with different rigidities and that their measurements may be available in real time, gives the opportunity for widespread usage [20,11]. The measurements are used by web users, mostly scientists, by applications and by online services and for tasks such as the prediction of the space weather [8] or the Ground Level Enhancement (GLE) Alerts [21,10,5,18]. These kinds of uses, apart from the real time measurements, require data of good quality. In order to establish the data quality, a neutron monitor station should verify the validity of the measurements and apply the necessary corrections, in order to exclude the parameters that affect or distort the data. The challenging aspect of this task is that these corrections should be performed in real time, while the data are transferred from the neutron monitor registration system to the Neutron Monitor Database.

Referring to the neutron monitor data, the meteorological and physical parameters, such as the atmospheric pressure, the snow that may cover the station and the very low temperatures, have a great effect on the measurements. These effects should be excluded from the measurements, since they cause changes that are not related to the variation of the cosmic rays. The correction of the data for the pressure is a straightforward procedure that requires only an accurate calculation of the barometric coefficient [13]. Moreover, the correction for the snow effect or for the very low temperatures that are met at some stations is performed by using specific models. However, apart from these effects, the measurements of the neutron monitors are in some cases distorted by unpredictable instrument variations. These variations are related to sporadic problems of the electronics and can be categorized in abrupt spikes, slow drifts and abrupt changes of the mean counting rate with or without recovery [1,3,6]. The correction of these variations is not a straightforward procedure. The reason for this is that by principal, the measurements of a neutron monitor have statistical variations and a distinction between them and the instrument variations is not always obvious. The task for the correction of the instrument variations

---

* Corresponding author. Tel.: +30 2107276890.
E-mail address: emavromi@phys.uoa.gr (H. Mavromichalaki).

0168-9002/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.nima.2013.02.031

is handled by correction algorithms that filter the data in real time, while they are transferred to the Neutron Monitor Database.

In order for a filtering algorithm to be effective, it should have three characteristics. It should be fast, so that it can be applied in real time, it should filter effectively all the instrument variations and finally it should leave the rest of the data unaffected. A filtering algorithm takes advantage of the fact that a neutron monitor consists of a number of identical channels [2]. Using this fact, the detection of an instrument variation on a channel can be performed by comparing its measurement with the measurements of the other channels. Based on this principle, a number of filtering algorithms have been implemented in the past. The most known algorithm, used currently by many neutron monitor stations such as the Athens station [9], is the Median Editor [1,3,6,22]. The Median Editor is a very fast algorithm, which filters all the instrument variations and has shown a very stable behavior. However, it has the disadvantage of distorting noticeably the data, even in the cases where no instrument variations are observed, as it is shown in Fig. 1. In order to overcome this issue, the Median Editor Plus concept has been introduced, according to which, the algorithm is applied only to the cases where an instrument variation is detected [22]. Another filtering algorithm that has recently been presented is the ANN algorithm, which makes use of an artificial neural network [14]. The ANN algorithm has shown a better behavior compared to the Median Editor, however the distortion of the non-erroneous data is still present. Finally, the Edge editor is another filtering algorithm that has recently been presented as well [12]. This algorithm hosts the Median Editor Plus concept and uses a validation criterion in order to distinguish the erroneous channels and apply corrections only to them. The Edge editor has shown a great behavior since the distortion of the non-erroneous data is almost unnoticeable. However, a direct comparison with the Median Editor and the ANN algorithm cannot be performed, since the latter does not use a validation criterion and is applied to all the measurements.

In this work, an optimization of the Median Editor and the ANN algorithm is performed. This optimization refers to the combination of the Edge editor's validation criterion with these algorithms in order to implement the "plus" version of the algorithms. Then, a direct comparison among the Edge editor, the Median Editor Plus and the ANN Plus algorithm is performed using the Athens neutron monitor data [9]. The comparison framework and the results are presented in the last sections of the manuscript.

2. Validation criterion of measurements

Referring to the characteristics that a real time filtering algorithm should have, the practice has shown that it is rather simple to implement an algorithm that is fast and that filters effectively all the instrument variations. The challenging point is to combine these characteristics with a behavior that does not affect the non-erroneous measurements. By principle, this is almost impossible since the processing of the measurements by an algorithm will cause little or great changes on them. The only way to protect the non-erroneous measurements from such effects is by not applying the algorithm on them. This issue leads to the conclusion that an optimized filtering algorithm should act in two steps, firstly towards the determination of the erroneous channels and secondly towards the application of a correction procedure only on them.

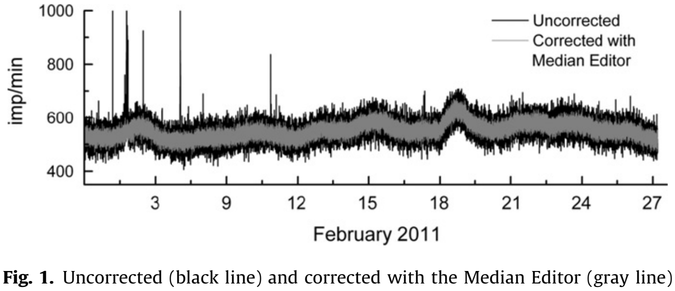

Fig. 1. Uncorrected (black line) and corrected with the Median Editor (gray line) data of the Athens NM's channel 6 for February 2011. The narrower variation of the corrected data implies a distortion of the original data.

The separation of the erroneous and the non-erroneous channels in the real time procedure is performed by using validation criteria. This kind of criteria can be constructed by performing a thorough data analysis on the past neutron monitor data that aims to the definition of the non-erroneous measurements pattern. Having defined this pattern, the real time measurements that follow it are considered as non-erroneous, while the ones that deviate from it are considered as erroneous. The pattern refers to the calculation of the physical statistical variations that each channel of the neutron monitor presents [7,15,12]. In this work, the Edge Editor's validation criterion is used for the determination of the erroneous measurements. The validation criterion of the Edge Editor [12], for the case of a neutron monitor with six channels as the Athens NM, is presented in Fig. 2. The validation criterion is separated into an offline analysis in order to calculate the necessary parameters and into an online application on the real time measurements.

Referring to the offline analysis, it is well known that each channel of the neutron monitor may measure a slightly different counting rate compared to the others, due to normalization factors. These factors correspond to the position and the characteristics of each tube (e.g. BF_3 density) and to the electronic modules. In order to perform an accurate analysis, it is required to normalize the measurements to the level of a selected reference channel 'j'. The reference channel can be any channel of the neutron monitor, independent of whether it is a channel that presents many or few instrument variations. The only use of the reference channel is the transformation of each channel's measurement into a common level by excluding the possible normalization factors. Therefore, the first step is the calculation of the normalization factors $R_{i,j} = (N_i/N_j)$ which refer to the ratio of the channels counting rates over the respective counting rate of the reference channel 'j', based on the historical data. The next step is the normalization of the historical data to the level that the reference channel 'j' measures by computing the variable $N_i^j = (N_i/R_{i,j})$. A statistical analysis is performed on the normalized measurements that results to the determination of the statistical variation σ_i^j for each channel. It has been shown in the past [12] that the statistical variations of the neutron monitor channels increase as the mean counting rate of the neutron monitor increases. In order to calculate the statistical variations σ_i^j, the mean counting rate of the neutron monitor (n_j) is calculated as the average of N_i^j for each minute. Then the N_i^j measurements are grouped by the n_j and the sigma of N_i^j is calculated in respect to n_j. Finally, a linear regression of σ^j with n_j gives the $\sigma_i^j = f(n_j)$ function. This procedure is performed for each channel. The offline analysis is taking place once and there is not any need for recalculation as long as the operational conditions of the neutron monitor are the same.

On the real time part, the measurements are normalized to the reference channel 'j' level and an estimation of the mean value is performed. On the contrary with the offline analysis, the estimation of the n_j cannot be done by simply averaging the normalized measurements, since one or more of them may contain an instrument variation. The weighted mean algorithm, that makes use of weight factors, is used for this task. Having estimated the n_j, the validation criterion calculates the estimated σ_i^j. Finally, a