Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 53 (2014) 71–76

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

Data reduction and correction algorithm for digital real-time processing of cosmic ray measurements: NM64 monitoring at Dourbes

Danislav Sapundjiev*, Myriam Nemry, Stanimir Stankov, Jean-Claude Jodogne

Royal Meteorological Institute of Belgium, Avenue Circulaire 3, B-1180 Brussels, Belgium

Received 29 May 2013; received in revised form 12 September 2013; accepted 28 September 2013
Available online 8 October 2013

Abstract

We present a data correction algorithm for real-time data processing for the NM64 galactic cosmic ray neutron monitor at the Royal Meteorological Institute (RMI) in Dourbes, Belgium. The correction is based on three main tests: a continuity test, tube ratios test and a derivative test. The continuity test works as a high pass filter with a threshold based on the entire recorded dataset. Additionally, it monitors whether the logging takes place at regular intervals (continuously). The ratios test identifies noisy sections and the final derivative test criterion will identify single or double spikes by testing them against the median increase of the intensities. Using these criteria, all data from the cosmic ray station at Dourbes is corrected in real time. Test results have been compared with data from verified neutron monitor stations with a similar geomagnetic cutoff rigidity.

Keywords: Cosmic rays; Data correction; Time series

1. Introduction

For many years, Galactic Cosmic Rays (GCR) measurements have been used to indirectly estimate the solar activity. More recently, real time data from a large number of stations have found various applications related to space weather. In the current era of telecommunications and satellite based technologies, the existing neutron monitors are organised into a world wide network which regained its importance – the great technological advance scored in the last 100 years have also increased the susceptibility of our everyday life to events from the near Earth atmosphere which were not of much concern to the previous generations. Possible space weather effects and impact on society have been the subject of various recent studies (NAS, 2008). The message conveyed by these reports is that space weather effects may cause significant damage to the material infrastructure and therewith affect the society. It becomes important to effectively predict at least those events with magnitudes posing hazards to the existing material and human health. Because GCR are partly modulated by the magnetic fields resulting from solar events, a pre-decrease or pre-increase in the intensity of the CR have been identified as precursors of possible space weather events and solar storms. For a large number of events, the lead-time of such precursors measured by neutron monitors is reported to be in the order of 4 h (Belov et al., 2001; Munakata et al., 2000). Cosmic ray anisotropy observed by ground level neutron monitor network may be efficiently able to forecast the arrival of an interplanetary shock directed towards the Earth (Leerungnavarat et al., 2003) resulting from loss cone and tail-in propagation (Nagashima et al., 1991). This shows that continuous cosmic ray monitoring from several conveniently positioned neutron monitoring stations offers a convenient way for space weather forecast (Kudela and Storini, 2006) with very attractive lead times.

The Royal Meteorological Institute of Belgium (RMI) has a long tradition and experience in observing the activity

---

* Corresponding author. Tel.: +32 060395446.
E-mail address: danislav.sapundjiev@meteo.be (D. Sapundjiev).

0273-1177/$36.00 © 2013 COSPAR. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.asr.2013.09.037

of the galactic cosmic rays. It is among the first institutes that started a continuous monitoring of the cosmic rays intensities. Data was published in the RMI publications until April 2004 and also sent to the World Data Center. At present, the neutron monitor at the RMI Geophysical Centre in Dourbes (50.1°N, 4.6°E) is used for detection of disturbances in the near Earth atmosphere caused by solar events. The data is used for development and testing of algorithms to forecast in real-time high intensity space weather effects. For this purpose a new data correction algorithm has been developed and implemented to answer the needs of a real time monitor of the GCR, a correct and reliable data has to be provided by the measurement station. In this article we describe the methods used for data control and correction at the neutron monitor station at Dourbes.

The primary GCR consists of about 90% protons and α-particles, 9%. The remaining components are heavier nuclei and exotic particles (positrons and anti-protons) (Grieder, 2001). On passing through the atmosphere and interacting with its components, the GCR produce secondary particles among which – neutrons. These neutrons are partly detected by a proportional tube counter filled with a suitable gas. With the addition of a heavy nuclear target and a moderator the intensity of the primary component can be monitored. This comprises what is called a super monitor NM-64 (Hatton and Carimichael, 1964).

The station started its operation in January 1965 with 12 IGY tubes. In July 1966 the station was equipped with 9-NM-64 tubes, which are in operation until today. The first records from the Dourbes station were produced by an IGY neutron monitor and from October 1966 by the 9-NM-64. Until 11 September 2003, the 9 tubes were combined into 3 sections of 3 tubes. The intensities were recorded every 15 min. Since September 2003, the intensities of the 9 tubes were recorded separately and the time resolution was increased to 1 min.

Today the pressure is recorded by a VAISALA PTB 202-A pressure transmitter with timing similar to the intensity logging. The data is processed on a daily basis and stored in digital files.

The records from the station at Dourbes are published in real time on the web page of the ionospheric section of the institute (http://ionosphere.meteo.be/sun/cosmicRay) (see Fig. 1).

2. Commonly observed errors and noise in the output data

The main source of errors in the recorded data arise during the station's startup and shut down. Additionally, discontinuities in the records – when the scalers and the timers skip logging the timestamps for one or more time intervals – are also observed. Apart from this, a sequence of high value counts given by one or more tubes for a considerable period of time (Fig. 2) is often observed. The pulse height (count numbers) of these spikes vary as well.

3. Automatic data correction and reduction

Despite all efforts to reduce noise during measurements and data recording, sometimes it is difficult and even impossible to completely avoid it. This holds especially when the measurements are carried out over prolonged periods. In most cases the raw measurement data has to be processed and monitored for noise and spurious peaks (Rosner, 1975; Rosner, 1977; Yanke et al., 2011; Paschalis et al., 2012). In Dourbes we have developed an automatic data correction program which will test the recorded information by several criteria. Errors and noise in the data will be corrected before the final total counts and pressure corrected data are reduced, stored and displayed. The philosophy behind the data correction algorithm is to perform a correction at a given instant if and only if it is based on data values obtained at the same time and not on values from the neighbouring moments – in the past or in the future. For example, if one (or more) tube(s) of the station gives an erroneous result at time t, their value will be corrected using the measurements from the other tubes at the same time as explained hereafter.

3.1. Atmospheric pressure measurements

Malfunctions of the pressure sensor are not common, but when present, they produce wrong pressure corrected values over a prolonged interval (hours and sometimes, days). Properly handling such malfunctions is important for real-time services. The first check on the pressure data is made to verify that the recorded pressure value is within the range of the pressure sensor, 590–1150 hPa. If the measurement falls outside this range, it is extrapolated from the previous 5 valid values. If the data file is corrected on a daily basis, that is not in real-time (which is the case with old files), the erroneous pressure values are interpolated from the neighbouring pressure values. This is carried out only if the length of the interval of missing values does not exceed the 60 minutes. If the gap length exceeds this threshold, a message is generated that an external pressure data has to be provided. However, such occurrences are rare; in the case of Dourbes station, over several decades of operation, there was only one case when the pressure had to be taken from external sources.

3.2. Tube intensities data

To preserve original data and avoid introduction of spurious peaks or artificial values, the identification of an erroneous tube signal has to be carried out according to the already mentioned two principles: the faulty tube has to be correctly identified as such, and the correction should not be based on neighbouring (previous or subsequent measurements). The latter condition is important in order to prevent introduction of spurious peaks, e.g. when using median filters. It also allows the data processing to take place in real time (that is the algorithm does not require

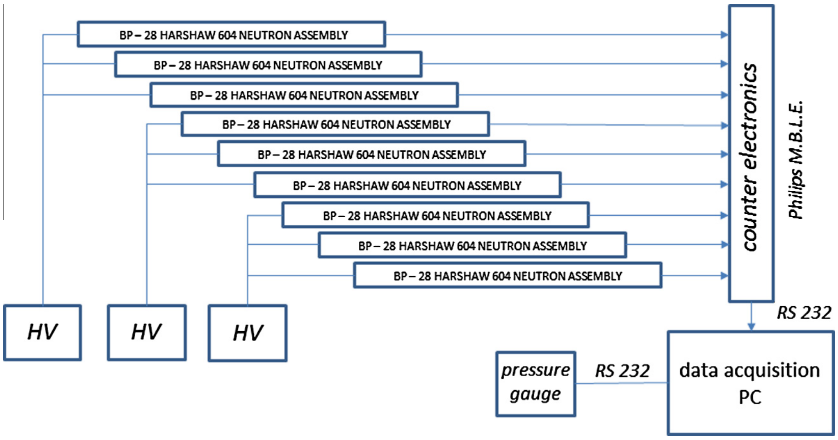

Fig. 1. Principle schematics of Dourbes neutron monitor station.

measurements at neighbouring time instants for the correction of the current value).

The detailed data processing algorithm for identification and the correction of the intensity measurements at the Dourbes cosmic ray station is shown in Fig. 3 and consists of three important steps: continuity test and low pass initial filtering, ratios, test and the derivative test.

3.2.1. Continuity check and low pass filter

The continuity check verifies the consistency of the data – that is, whether recorded value exists for every logging interval. The high-pass filter will remove wrong values occuring during the station's startups and shutdowns, when the acquisition time is shorter. This shorter acquisition time results in lower counts per time interval and introduce significant fluctuations in the recorded values. The threshold of the filter is verified manually by an operator.

3.2.2. Tube ratios

So far, data from tubes that have been identified by the first continuity test as faulty, were marked for correction. The remaining tube values of the record at this instance are subjected to the ratios test, which consists in comparing the ratios between the tubes, for example the ratio between tube i and tube j, R_{ij}, with a reference value, $R_{ij_{ref}}$, which is calculated during a normal operation of the station. If we have n detector tubes which have passed the differential discriminator of Section 3.2.1, then a single tube, say tube i is identified as faulty if the following condition is satisfied (Ruffolo, 2012):

$$\bigwedge_{j=1, j\neq i}^{n} |R_{ij} - R_{ij_{ref}}| > 3\sigma_{ij} \tag{1}$$

where the logical multiplication is over all valid intensity measurements ($n = 9$ for the station in Dourbes). The

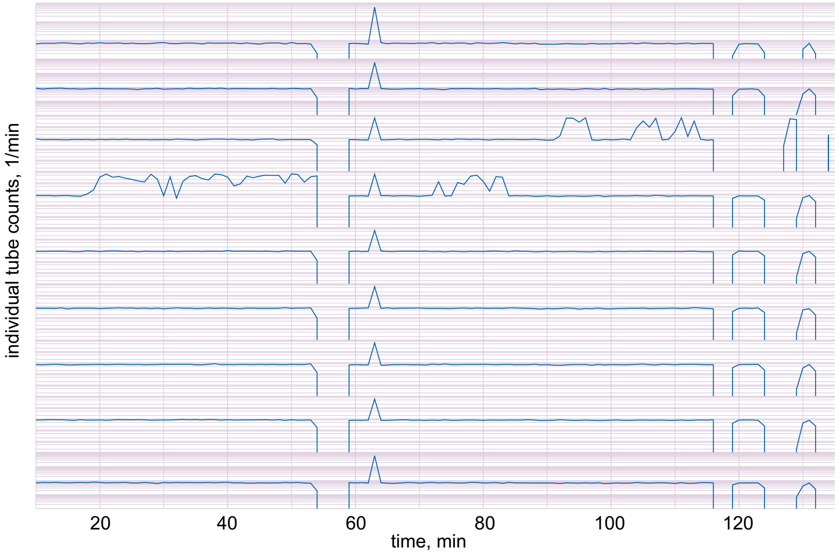

Fig. 2. Some commonly observed faults in the station output that need to be corrected (normal detector operation).

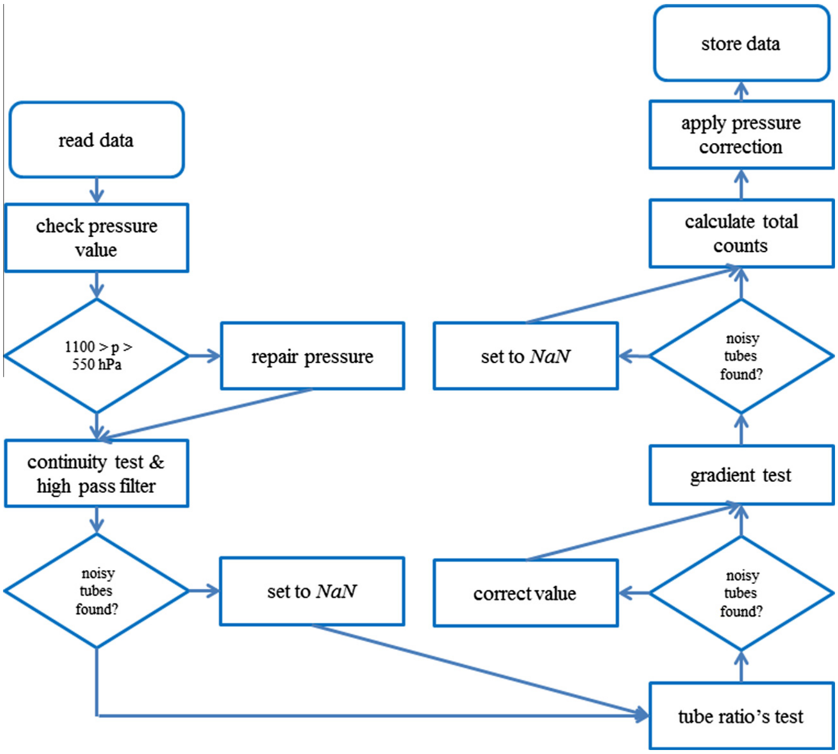

Fig. 3. Data correction algorithm flowchart.

automatic correction based on this condition is carried out only if the sum of the logical test in Eq. (1) equals $n-1$. Differences smaller than 3σ are not significant and are not addressed.

This was the simplest scenario when a single tube is the sole source of noise. In the case when there are more tubes at fault, the underlying logic in this algorithm can become very complicated, depending on the number of tubes with "noisy" measurements and the magnitude of this noise. An important case is when only one tube is giving a correct value. Then the ratios algorithm will give a result that this tube is "noisy". Fortunately, in the majority of the cases the extrema test would have filtered the $n-1$ noisy tubes. If however they have passed the discriminator, the derivative test (see Section 3.3.1) will probably identify the resulting intensity spike.

3.3. Correction of a noisy tube measurement

After a noisy tube is identified, its value is corrected from measurements of the remaining tubes using the relative contribution (tube fraction) of the noisy tube. The count rate of the noisy tube i, N_{fi}, is subtracted from the total uncorrected count rate, N_f, to obtain the reduced count rate, N':

$$N' = N_f - \sum_{i=1}^{n} N_{fi}$$

The corrected total count rate, N, of the station will be obtained by multiplying N' by a factor, k, that accounts for the noisy tube(s) contribution to the total count rate:

$$N = kN' = \frac{1}{1 - n + \sum^{n} \frac{1}{k_i}} N'$$

where $k_i = \frac{N}{N - N_i}$ and is determined by an operator from manually verified data during normal operation of the station (when $n = 1,\ k = k_i$). For the station of Dourbes, the relative contribution coefficients of each detector tube k_i are plotted in Fig. 4. The coefficients k_i as well as the reference tube ratios $R_{ij_{ref}}$ are determined at regular time intervals to account for slow changes in the station parameters like aging etc.

3.3.1. Single peaks filtering

An important noise contribution to the signal of a neutron monitor station, especially at a minute time resolution, are singlets or doublets (and very rarely, triplets) with large absolute amplitudes. They can be caused by high voltage discharges (positive spikes) or may occur just before or after the station shutdowns when the acquisition interval is shorter (the low amplitude peaks). While the continuity check and the low-pass filter may sieve out most of the high amplitude peaks, the low amplitude peaks may remain unnoticed. The peaks need to be identified and in most of the cases removed as they do not represent a real intensity measurement. In the same time, an increase due to a ground level enhancement event should be left intact. This renders the use of a simple band filter unsuitable. The average relative increase between two subsequent values, the gradient, is calculated for 24 hours prior to the current moment. The condition for a valid peak is that the gradient of the peak should be less than the **median** of the gradients

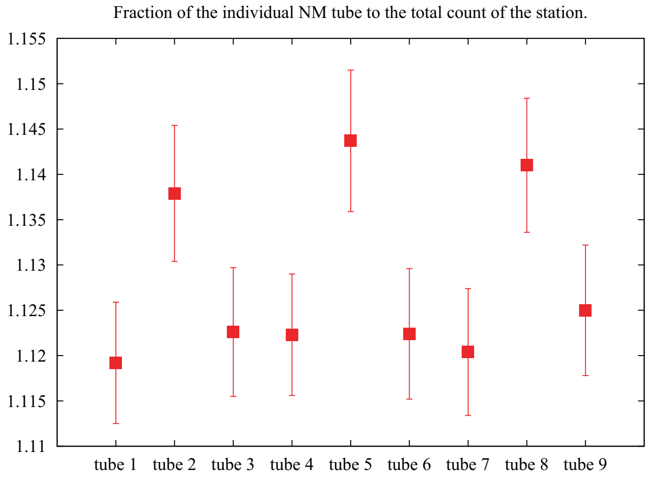

Fig. 4. Relative contribution of each detector tube of the cosmic ray station in Dourbes. Knowing these numbers allows to recalculate the value of a tube which has produced a noisy or has not produced any value.

Table 1
Dourbes neutron monitor station parameters.

Position	50.097 N, 4.590 E
Geomagnetic cutt-off	3.18, GV
Elevation	225 m
Average count rate, Solar max.	5840, counts/s
Average count rate, Solar min.	6480, counts/s

for the day within a confidence interval 5× MAD (median absolute deviation). The same procedure is used for filtering of erroneous values near detector stop and start when the acquisition time was shorter and the resulting count rate lower.

An important case is when the spike results from large number of detector tubes – that is when the tube ratios test cannot be used to identify the noisy tube(s). If at least one tube had given a correct measurement, then the values of the remaining tubes could be theoretically restored. If the assumption that the tube is correct is false (despite that it has passed the criteria) its value will be still used for correction of the remaining tubes. This will result in value which is greater or smaller by more than 3 times the standard deviation. In this case the time derivative condition would remove the resulting total value. In reality, the tube ratios test will only be triggered if not more than half of the tubes have been identified as noisy.

4. Calculation of the station parameters

The entire data has been subjected to the above explained criteria which were implemented by a computer code. The results were used to calculate several station parameters which, together with the position and the elevation, are presented in Table 1.

The reported geomagnetic cut-off rigidity was calculated by using the TJI95 computer code utilizing the trajectory-tracing technique (Smart and Shea, 2001). From the corrected data during the solar extrema, the average maximum and minimum counts/min for the station have been calculated. The average count rate at solar minimum is used for the calculation of the relative increase in the measurement which is calculated by

$$I_{\%} = \frac{I - I_R}{I_R} 100\%$$

where $I_R = 6480 \pm 50$ counts/min (see Table 1). Relative increase has been used in order to provide a better ground for comparison between different stations. The absolute values of a neutron monitor are difficult to standardise and experience has shown that relative intensity remains more accurate than the inter-calibration (Ruffolo, 2012). The pressure corrected intensity is carried out by the formula:

$$N = N_0 \exp \beta (p - p_0)$$

where p_0 is the reference atmospheric pressure 740 mmHg, and β is the barometric coefficient 0.00985 1/mmHg. The station uses the clock of the data acquisition PC which is synchronised via the network time protocol (NTP) with the NTP server of the Royal Observatory of Belgium.

4.1. Results

The work of the automatic data correction is illustrated in Fig. 5, where noisy peaks, resulting from single or multiple detector tubes, and other low and high magnitude erroneous data explained so far have been filtered. The resulting CR intensities are plotted in the same figure and ready for the final pressure correction. The complete and corrected data of the GCR intensity measurements from the station at Dourbes is plotted in Fig. 6. The sunspot number evolution throughout the years is given to illustrate the correlation between the CR intensity and the solar activity.

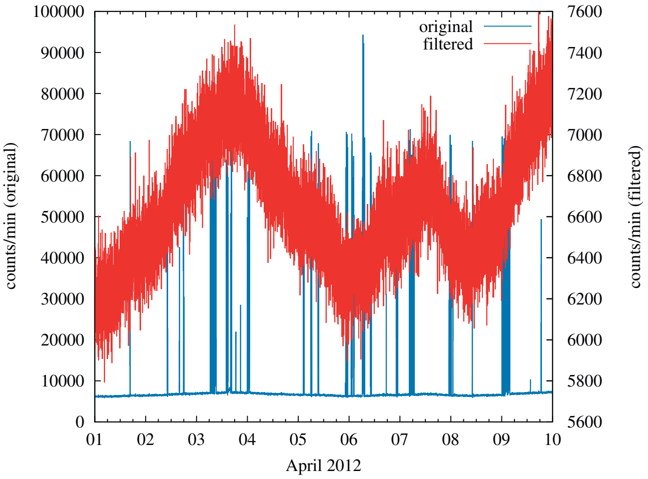

Fig. 5. The working of the ADCA – minute data from the first 10 days of April 2012 before and after processing (not corrected for pressure).

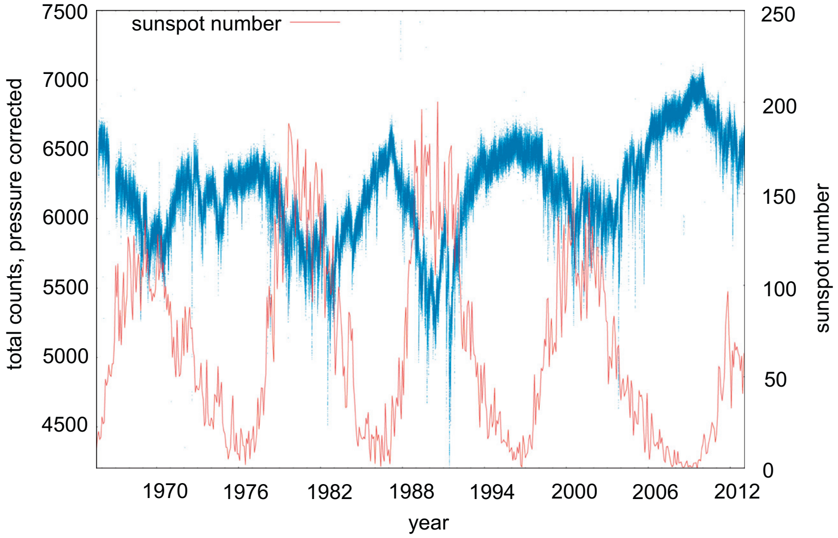

Fig. 6. GCR intensity data, in counts per minute plotted against the sunspot number from ROB (2013) from October 1966 until January 2013.

5. Conclusions

An automatic data quality control (ADQC) procedure for correction of data from the Dourbes Neutron Monitor was developed and implemented in real-time. The ADQC is capable to filter noisy, erroneous signals from individual sections as well as spikes from the entire station. Based on the three criteria explained here, the majority of the occurring spikes can be removed and where possible – data recovered. The recovery is carried out only if more than half of the tube sections produce the correct value. In case of pressure sensor malfunction, the ADQC will extrapolate the available data or require external data for the pressure correction. Thus the corrected data can be already used for analysis, modelling and development of algorithms for finding precursors of solar events. If successful, the algorithms will be included in the real-time data processing.

Acknowledgements

The authors wish to acknowledge the use of the sunspot number data from the Royal Observatory of Belgium and the international neutron monitor database for use of neutron monitors data. This work is funded by the Royal Meteorological Institute (RMI) via the Belgian Solar-Terrestrial Centre of Excellence (STCE).

References

Belov, A.V., Bieber, J.W., Eroshenko, E.A., et al., 2001. Pitch-angle features in cosmic rays in advance of severe magnetic storms: neutron monitor observations. In: Proceedings of the 27th International Cosmic Ray Conference, 07–15 August, 2001. Hamburg, pp. 3507–3510.

Grieder, P., 2001. Cosmic Rays at Earth. Elsevier.

Hatton, C.J., Carimichael, H., 1964. Experimental investigation of the NM-64 neutron monitor. Canadian Journal of Physics 42 (12), 2443–2472.

Kudela, K., Storini, M., 2006. Possible tools for space weather issues from cosmic ray continuous records. Advances in Space Research 37 (8), 1443–1449.

Leerungnavarat, K., Ruffolo, D., Bieber, J.W., 2003. Loss cone precursors to Forbush decreases and advance warning of space weather effects. The Astrophysical Journal 593, 587–596.

Munakata, K., Bieber, J.W., Yasue, Shin-ichi, et al., 2000. Precursors of geomagnetic storms observed by the muon detector network. Journal of Geophysical Research: Space Physics 105 (A12), 27457–27468.

Nagashima, K., Fujimoto, K., Morishita, I., 1991. Galactic cosmic-ray anisotropy and its heliospheric modulation, inferred from the sidereal semidiurnal variations observed in the rigidity range 300–600 GV with multidirectional muon telescope at Sakashita underground station. Planetary and Space Science 39 (12), 1637–1655.

NAS, 2008. National Academy of Sciences, Severe Space Weather Events–Understanding Societal and Economic Impact. <http://www.nap.edu/catalog.php?recordid=12507>.

Paschalis, P., Sarlanis, C., Mavromichalaki, H., 2012. Primary data processing algorithms for neutron monitors. In: Proceedings of the 23rd European Cosmic Ray Symposium.

ROB, 2013. Royal Observatory of Belgium, Solar Influence Data Analysis Center. <http://sidc.oma.be/sunspot-index-graphics/sidcgraphics.php>.

Rosner, B., 1975. On the detection of many outliers. Technometrics 17 (2), 221–227.

Rosner, B., 1977. Percentage points for the RST many outlier procedure. Technometrics 19 (3), 307–312.

Prof. D. Ruffolo (2012). Personal communication.

Smart, D.F., Shea, M.A., 2001. Geomagnetic cutoff rigidity computer program. Tech. Rep. NAG5-8009, NASA.

Yanke, V., Belov, A., Eroshenko, E., et al., 2011. Primary processing of multichannel cosmic ray detectors. In: Proceedings of the 32nd International Cosmic ray Conference.