STUDY OF COSMIC-RAY VARIABILITY USING GROUND-BASED AND SPACE-BORNE DATA

SERGEY A. KOLDOBSKIY
Study of cosmic-ray variability using ground-based and space-borne data

Sergey A. Koldobskiy

University of Oulu Graduate School
University of Oulu
Faculty of Science
Space Physics and Astronomy Research Unit

Academic dissertation to be presented, with the assent of the Graduate School of the University of Oulu, for public discussion in the Auditorium L2, Linnanmaa, on November 27, 2020, at 10 o’clock noon.
Opponent
Dr. Igor V. Moskalenko, Stanford University, California, USA

Custos
Prof. Ilya G. Usoskin, University of Oulu, Finland

Pre-examiners
Dr. Konstantin Herbst, Christian-Albrecht University of Kiel, Germany
Dr. Athanasios Papaioannou, National Observatory of Athens, Greece

Supervisors
Prof. Ilya G. Usoskin, University of Oulu, Finland
Prof. Kalevi Mursula, University of Oulu, Finland

ISSN 1239-4327

Punamusta Oy 2020
Koldobskiy, Sergey A,

Study of cosmic-ray variability using ground-based and space-borne data
Space Physics and Astronomy Research Unit, University of Oulu, Finland
Report No. 134 (2020)

Abstract

Cosmic rays (CRs) are charged particles with energies above 1 MeV, accelerated in sources outside of the Earth’s magnetosphere. The variability of their fluxes, caused by solar modulation and solar eruptive events, is an important field of astroparticle physics, and represents the main focus of this study.

The main instruments to study the cosmic-ray variability are neutron monitors (NMs), located at different locations around the globe, and space-borne experiments. A NM is an integral detector so that its response is integrally related to CR fluxes via NM response function. Space-borne experiments have particle detectors that allow to directly register different CR particles.

In this work both data from NMs and space-borne experiments are combined to study the variability of CRs. In particular, the reconstruction of the solar modulation potential $\phi$ applying the simplified force-field (FF) model of the solar modulation of cosmic rays was performed using in-situ cosmic-ray fluxes measured by both the PAMELA and AMS-02 experiments together with NM data for the period from 2006 to 2017. Validation of the FF model for periods of different solar activity levels was further performed, and it was found that such an approximation performs better during solar minima, but disagrees with the observations of up to $\approx 10\%$ during solar maximum. This makes the FF model approach not well suited for detailed studies of the solar modulation processes. At the same time, this precision is adequate to quantify the condition of the heliospheric modulation and to study its long-term variability.

To study solar energetic particle (SEP) fluxes using NM data, a new method of “effective rigidity” was proposed, allowing to reconstruct high-energy SEP integral fluences recorded during ground level enhancement (GLE) events. A significant advantage of this novel method is that it is a non parametric one and thus the spectral shape of the SEP fluence can be deduced directly from the reconstructed data. Reconstructions of the SEP fluences for two recent GLEs, #69 and #71, using this newly developed method, yield a very good agreement with the laborious method of the full fluence reconstruction using NM data, and with PAMELA measurements (for GLE #71), but disagree with earlier simplified estimations based on NM data.

The NM data analysis was performed using the NM yield function by Mishev et al. [2013], which was validated using the AMS-02 data for protons and helium for a time period between 2011 to 2017 and showed the best performance among other modern yield functions.

Improved knowledge of the CR variability is crucially important for e.g. CR-induced atmospheric effects, including the production of cosmogenic isotopes.

Keywords: cosmic rays, ground level enhancements, solar energetic particles, solar modulation, space climate
To my Mom, Dad and Katya Crab!
Preface

First I want to gratefully acknowledge my primary supervisor Prof. Ilya Usoskin and my colleague Dr. Gennady Kovaltsov for their wish to transfer their priceless experience in physics and in data analysis to me. I also want to thank Prof. Ilya Usoskin and Prof. Kalevi Mursula for the opportunity to work in an excellent group of researchers that create new knowledge in space physics.

I express high gratitude to my colleagues and friends Alexander Mishev, Stepan Poliuanov, Pauli Väisänen and Maria Poliuanova for interesting discussions, warm company and help in living in Oulu.

I truly thank my pre-examiners Dr. Konstantin Herbst and Dr. Athanasios Papaioannou for their valuable comments and advices. I am also grateful to Dr. Igor V. Moskalenko for coming to Oulu as my opponent.

Finally, I greatly appreciate efforts of PAMELA and AMS-02 collaborations, PI of neutron monitors, and data managers that maintain NMDB (nmdb.eu), GLE (gle.oulu.fi) and SSDC (https://tools.ssdce.asi.it/CosmicRays/) databases, without their work my study was impossible.

Oulu, January 31, 2020

Sergey Koldobskiy
Original publications

The thesis is based on the following publications, which are referred in the text by their Roman numerals I–IV and are given in the chronological order:


The author’s contribution to these publications was crucial and can be specified as follows.

In Paper I the author created a code for comparison of the data obtained from the PAMELA satellite-based experiment and from neutron monitors using the yield-function formalism, made the comparison in terms of the solar modulation potential, and wrote the first draft of the article.

In Paper II the author created a code for calculations of the effective rigidity for ground-based NMs, tested it with the ideal single power law form of the solar energetic particle fluence and a more realistic representation with the Band function, and wrote the first draft version of the article.
In Paper III the author created a code for reconstruction of the NM scale factor $\kappa$ and tested it with the data from several sea-level NMs and with four NM yield functions, made a comparison of the solar modulation potential $\phi$ reconstructed with data from the PAMELA, AMS-02 experiments and from the NM network, and wrote the first draft of the article.

In Paper IV the author created a program for reconstruction of the solar energetic particle fluence using the effective rigidity method, tested it with four different NM yield functions, and wrote the first draft version of article.
## Contents

Abstract  
Preface  
Original publications  
Contents  
1. Introduction ................................................. 7  
   1.1. Brief history of cosmic-ray studies .................. 7  
   1.2. Composition and spectra ............................. 9  
   1.3. Measurements of cosmic rays ......................... 13  
      1.3.1. PAMELA and AMS-02 experiments ................. 14  
      1.3.2. Neutron monitors ............................... 16  
2. Heliosphere and modulation of galactic cosmic rays ....... 21  
   2.1. Solar wind ........................................ 23  
   2.2. Heliosphere ....................................... 24  
   2.3. Transport equation ................................ 27  
   2.4. Local interstellar spectrum ......................... 28  
   2.5. Force-field parametrization ......................... 28  
3. Solar eruptive events .................................... 31  
   3.1. Solar flares ...................................... 31  
   3.2. Coronal mass ejections ............................ 33  
   3.3. Solar energetic particles ......................... 35  
   3.4. Ground level enhancements ......................... 38  
4. Cosmic rays and Earth .................................... 42  
   4.1. The terrestrial magnetic field and geomagnetic cutoff .... 42  
   4.2. Atmospheric cascade ................................ 45  
5. Summary .................................................. 48  
References .................................................. 50  
6. Original publications .................................... 62
1. Introduction

1.1. Brief history of cosmic-ray studies

Cosmic rays (CRs) are charged particles with energies above 1 MeV, accelerated in sources outside of the Earth’s magnetosphere. They were discovered in the beginning of the nuclear-physics era in 1912 by Victor Franz Hess and independently by Domenico Pacini. Hess [1912] used an ionization chamber and discovered that the ionization rate is increasing as a function of the altitude (see right panel of Fig. 1.1). These results have been correctly interpreted by him as an existence of radiation with non-terrestrial nature. Pacini [1912] used an ionization chamber and disputed the dogmatic paradigm that the background radiation is induced from the soil, by measuring the count rate above the water surface and below it. He found that the ionization rate does not change for both soil and water surfaces, while it decreases under water. He interpreted these results as a cosmic nature of the background radiation.

In subsequent years cosmic-ray physics was one of the pioneer fields in nuclear and particle physics. Before the accelerator era, it was the only way to discover new particles. In 1932 Carl David Anderson discovered the existence of the first antiparticle, positron, using cosmic-ray data. Shortly after that, in 1936, he and V. Hess received the Nobel Prize in physics for their advances in nuclear physics. In 1930–1950’s, pions, muons, kaons and first detected hyperons were discovered in CRs using cloud chambers and nuclear emulsions.

In 1930, extensive air showers were predicted and a few years later discovered by Pierre Victor Auger, signaling the beginning of the era of high-energy cosmic ray studies.

At the end of the 1950’s – beginning of the 1960’s, the space era started with advances in rocket science, the launch of the first satellite named 'Sputnik' and the first human in space. This led to the discovery of the radiation belts outside the Earth. Soon thereafter many scientific satellites have been launched into space, expanding our understanding of the solar system, our Galaxy, and the Universe. Several dozens of these missions were devoted to the study of cosmic-ray physics [Mewaldt, 2013].
Fig. 1.1: (1.1a) Victor Hess is ready for the balloon flight; (1.1b) The ionization rate as a function of the height measured by V. Hess. Pictures taken from CERN website.

In 1957–1958, the International Geophysical Year [Stoker, 2009] was organized by the geophysical community and a network of so-called neutron monitors (NMs) was installed on the surface of Earth [Simpson, 2000]. These instruments detect the nucleonic component of particle showers induced by cosmic-ray interactions in the atmosphere. These devices allowed to record long-term cosmic-ray variability due to the Sun’s activity. Short-term cosmic ray variability caused by transient changes in the heliospheric magnetic field [Forbush, 1937] and by solar energetic particles (SEPs) was discovered before the neutron monitor era with ionization chambers [Forbush, 1946]. Together with satellite observations, the NM data allowed to get a better understanding of the cosmic-ray variability processes.

Since 1960’s, it became possible to construct cosmic-ray detectors with particle identification capabilities and to expand the studied energy range with the help of modern satellite and balloon-borne experiments together with specialized ground-based installations.

Nowadays, cosmic-ray measurements allow us to understand the principles of acceleration and transport of energetic particles in the Galaxy and in the solar system. Cosmic-ray physics together with modern astronomical observations in different wavelengths provide us with information, allowing to evaluate magnetic fields and the chemical composition of interplanetary and interstellar medium. CRs are accelerated at astrophysical sources, such as supernova remnants [Büssing et al., 2005; Berezhko, 2014], and propagate through the interstellar medium, where additional acceleration processes may take place, which are not reproducible on Earth with modern technology, so that cosmic-ray physics together with modern
astronomy provide a unique opportunity to study our Galaxy, the Universe and the extreme physical conditions within.

Cosmic-ray physics is closely related to astronomy, nuclear physics and partly geophysics. Many dedicated experiments are currently or have been in operation [see, e.g., Mewaldt, 2013, for review], while several others (e.g., HERD, GAPS, AMS-100) are planned. Results of CR measurements are used for planning space missions and predicting fluxes of harmful space radiation, which is especially important for missions outside the Earth magnetosphere, where the effect of SEPs can be deathly harmful for humans and electronics [Schrijver and Siscoe, 2010].

1.2. Composition and spectra

Before the entry into the Earth’s atmosphere cosmic rays are “primary”, and after the entry they interact with nuclei of atmospheric gases, resulting in development of nuclear and electromagnetic cascades.

The main sources of CRs near the Earth are the Galactic sources, such as supernova remnants and pulsars (CRs accelerated in these sources are called galactic cosmic rays, GCRs), and the Sun (solar cosmic rays, SCRs, or SEPs). In addition, so-called anomalous cosmic rays (ACRs) are known, which are produced as a result of penetration of neutral atoms of the local interstellar medium into the heliosphere, their subsequent ionization by solar ultraviolet radiation or due to charge exchange with solar wind ions, when they become pickup ions (PUIs), which are subsequently picked up by the solar wind, convected out and accelerated at the termination shock, and finally re-enter into the heliosphere [Fisk et al., 1974].

Figure 1.2 shows that the elemental composition of cosmic rays does not exactly reflect the composition in the sources. It is because some elements, such as Li, Be and B as well as Sc, Ti, V, Cr and Mr, which are much less abundant in the solar system compared to their abundance in the GCRs (due to the high probability of nuclear interaction and fragmentation to other elements during stars evolution), are produced in interactions of GCRs with the interstellar medium. Several nuclei, such as nitrogen, have a mixed production both in astrophysical sources and by collisions of heavier nuclei with the interstellar medium [Aguilar et al., 2018b]. It was expected (e.g., on the base of IMP-8 and ACE/CRIS measurements) that different GCR species should have different spectral indices, that was recently confirmed with high precision by the AMS-02 experiment [Aguilar et al., 2018d].

The energy range of GCRs assumed to be from several MeV to almost ZeV. Some number of ultra-high energy cosmic rays exceeding this energy have been detected using ground-based experiments. Such high-energy CRs cannot be produced in our Galaxy from the theoretical point of view, so these particles supposed to have extra-galactic cosmic rays [Berezinsky, 2014]. Processes of their acceleration are still unclear.

The flux of GCR with energies smaller than several tens of GeV is significantly changed by solar modulation [see a review by Potgieter, 2013]. First measurements
Fig. 1.2: Normalized composition of different elements in the Solar system and in GCR. Courtesy of NASA.

of GCR outside the heliosphere were made recently on board of Voyager-1 and Voyager-2 spacecraft [Stone et al., 2013, 2019], but only for low-energy range up to 0.5 GeV.

Contrary to GCRs, SCRs and ACRs are less energetic and different in composition (in comparison to GCRs and between each other). The former consist of protons, helium and light ions, which typically do not exceed energies of about 100 MeV/nucleon, but during solar flares and coronal mass ejections (CMEs) fluxes of solar energetic particles with energies ranging from several MeV to ∼10 GeV can be produced and accelerated, that can be extremely dangerous for space missions. Figure 1.3 shows fluxes of protons during the SEP event on 13 December 2006 together with the averaged GCR spectrum measured by the PAMELA experiment for period from July 2006 to March 2008. Contrary to SCRs, in ACRs there are more helium than protons, and much more oxygen than carbon, the reason of this is a different production mechanism of these particles, related to ionization and acceleration of neutral interstellar atoms in the heliosphere [Mewaldt et al., 1998].

Figure 1.4 shows spectra of CRs, gamma-rays, and neutrinos obtained in different experiments. For the energy range below several TeV it is possible to distinguish different cosmic-ray species due to direct cosmic-ray observations with space- and balloon-borne missions, and for energies above it the all-particle spectra, reconstructed from measurements of the CR secondaries with ground-based detectors, are measured. In a wide energy range the spectrum can be roughly approximated by a power-law function with the index ∼ −2.7. Of particular interest are the energy regions of $10^{15} – 10^{16}$ eV, the so-called “knee”, and $10^{18} – 10^{19}$ eV (“ankle”), where changes of the spectrum indices are observed, indicating different mechanisms of acceleration or transport. Greisen [1966]; Zatsepin and Kuz'min [1966] proposed that the cosmic-ray flux is theoretically energy-limited due to interaction of ultra-high-energy ($>5 \times 10^{19}$ eV) CRs with the cosmic microwave background.
This spectral feature is known as the Greizen–Zatsepin–Kuz’min (GZK) limit.

The average intensity of CRs is considered to be roughly constant on the millions-year timescale [Wieler et al., 2013].

The working range of the Large Hadron Collider (LHC) is also shown on Figure 1.4, it becomes apparent that some processes of interactions of the very-high energy cosmic rays can not be reconstructed using the collider, and this fact highlights a particular aspect of cosmic-ray physics. In addition to astrophysical studies, cosmic-ray physics still can be used for nuclear physics studies as a century ago, because highest cosmic-ray energies are not achieved by the collider experiments (and are not expected to be achieved in the near future).

Together with the data of cosmic rays, which are charged particles, spectra of gamma-rays and neutrinos are shown, which also carry important information about processes in the astrophysical sources and in the interstellar medium. Moreover, since these particles do not have charge, they are not deflected in the magnetic fields, so that they can be used to point to the source locations.

Traditionally, particles observed in CRs are divided into several groups according to their charge number $Z$ [Grieder, 2001]: $p$ ($Z = 1$), $\alpha$ ($Z = 2$), L ($Z = 3 \ldots 5$), M ($Z = 6 \ldots 9$), H ($Z \geq 10$), VH ($Z \geq 20$) (respectively, protons, alpha-particles, light, medium, heavy and very heavy). Almost half (43%) of the energy of cosmic rays is contained in protons, another 23% – in the energy of helium nuclei ($\alpha$-particles) and 34% of the energy are originated in another CR particles.
Fig. 1.4: Set of cosmic-ray, gamma-ray and neutrino measurements in a wide energy range, made by different modern experiments (see legend). For energies below several TeV separate spectra for different CR species are given, above it the all-particle spectrum is shown. Two CR spectral features (“knee” and “ankle”) are also noted. The upper limit of the collision energy in the Large Hadron Collider is denoted with red line. Ultra-high energy cosmic rays are suppressed because of the GZK limit. Picture adopted from Evoli [2018].
By the number of particles, cosmic rays are by 92% composed of protons, 6% are helium nuclei, about 1% are heavier elements, and about 1% are electrons [see, e.g., Adriani et al., 2014; Tanabashi et al., 2018]. When studying sources of cosmic rays outside the Solar System, the proton-nuclear component is mainly detected by the orbital gamma-ray telescopes through its emission produced in the interstellar medium. These gamma-rays are generated through the decay of $\pi^0$, produced in nuclear interaction of CRs with atoms of interstellar gas. Electrons (and positrons) are detected through accompanied inverse Compton scattering and synchrotron emission. Typically, this radiation is seen in the meter waves, but in case of strong magnetic fields in the region of the cosmic ray source, also the higher frequency ranges of radiation can be detected. Therefore, the electronic component can be detected by ground-based astronomical instruments.

### 1.3. Measurements of cosmic rays

Depending on the energy of cosmic rays, different detector systems are used. To measure cosmic rays with energies from several MeV to several TeV, space-borne, balloon-borne and ground-based experiments are used. Only several space- and balloon-borne detectors are/were able to measure the spectrum of more energetic particles (up to several tens of TeV), such as ATIC, CREAM, NUCLEON, CALET and DAMPE [see, e.g., Atkin et al., 2017; Adriani et al., 2018, and references herein]. The upper energy limit for the space- and balloon-borne particle registration is caused by two factors. The first one is related to a very small flux of high-energy particles, so that the probability of registration of high-energy particle by space- or balloon-borne experiment with small aperture is low, and the second is that experiments are limited in mass so that it is hardly possible to install super-heavy calorimeters (which would be able to measure the energy of the high-energy CRs) and to build experiments with very big aperture, that would be able to detect CRs with energies more then tens of TeV. Both these problems will be possibly solved in the future, with construction of new super-heavy rockets. As a payload for these new rockets, the new AMS-100 project was proposed recently [Schael et al., 2019].

Henceforth, for observations of more energetic particles, ground-based experiments are used, such as Pierre Auger Observatory [Góra, 2018]. These detectors register secondary-particle showers generated as a result of interaction of high-energy primary cosmic-ray particle with the Earth’s atmosphere (see details in Section 4.2). The reported highest reconstructed energy for cosmic rays is $(3.2 \pm 0.9) \cdot 10^{20}$ eV [Bird et al., 1995], which is slightly higher than the GZK limit.

Many cosmic-ray experiments are currently in operation and were carried out in the past. Since the topic of this work is the variation of CR due to the solar activity, two recent important experiments, PAMELA and AMS-02, working in the energy range of the solar modulation, are discussed here.
1.3.1. PAMELA and AMS-02 experiments

Several recent space-borne cosmic-ray experiments can be mentioned with respect to study the solar modulation of cosmic rays and solar energetic particles. The most sophisticated ones are PAMELA and AMS-02 missions.

The PAMELA (Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics) experiment [Adriani et al., 2014] is an international project aimed at studies of cosmic rays in a wide energy range from tens of MeV to about 1 TeV. It was installed onboard the Russian satellite “Resurs DK-1”, launched in June 2006 with continuous data taking process beginning in July 2006. The mission was in operation until January 2016. The experiment consists of several detector systems with different registration techniques (see Figure 1.5a): the time-of-flight system allows to measure the velocity of a particle, the magnetic spectrometer can reconstruct its rigidity, the calorimeter allows to separate particles depending on patterns of the particle shower development in the calorimeter material, and the neutron detector can separate interactions of high-energy hadrons and leptons. Moreover, scintillation and semiconductor detectors of time-of-flight system and magnetic spectrometer, correspondingly, are able to measure particle’s energy losses. Out-of-aperture particles are discarded from analysis with the help of an anticoincidence shield made from scintillation detectors.

In the field of CR variations, the PAMELA mission gave us unique results, such as variations of the GCR proton spectrum for the period between 2006 and 2014 with the Carrington rotation\(^1\) time resolution [Adriani et al., 2013; Martucci et al., 2018], variation of GCR electrons for the period from 2006 to 2009 [Adriani et al., 2015a], and variation of the positron-to-electron ratio for the period of 2006–2016 [Adriani et al., 2016b]. Moreover, PAMELA was able to study short-term variations occurring during the mission: in Adriani et al. [2011a] the temporal evolution of fluxes during strong SEP events in December 2006 was shown and in Bruno et al. [2018] the integral fluxes for all SEP events observed by PAMELA were shown and fitted with the Ellison-Ramaty functional form [Ellison and Ramaty, 1985], in Adriani et al. [2015b] some SEP features (anisotropy, temporal evolution) during the 17 May 2012 SEP event were studied, in Adriani et al. [2016a] measurements of geomagnetic cutoff variations during the 14 December 2006 geomagnetic storm were presented, and in Munini et al. [2018] the amplitude and recovery times for several GCR components during the Forbush decrease in December 2006 were reconstructed for the first time.

The AMS-02 (Alpha Magnetic Spectrometer) experiment [Aguilar et al., 2015] was launched on the final Space Shuttle mission in 2011 and was installed onboard the International Space Station (ISS). Soon after deployment, the data taking was started and is continuing until now. AMS-02 detector consists of several detector systems (see Figure 1.5b), including the magnetic spectrometer, time-of-flight system and calorimeter (with the same principle of particle identification as in PAMELA experiment). Moreover, AMS-02 is equipped with a Ring Image Cherenkov detector (which is effective for precise velocity measurements) and the

\(^1\)One Carrington rotation corresponds to the period of 27.2753 days, starting from 9 November 1853.
transition radiation detector, which is especially effective for hadron/lepton separation in the high-energy region. AMS-02 has a twenty times bigger aperture than PAMELA, allowing to collect more data leading to more statistically significant results.

In the field of cosmic ray variability, until now, the AMS-02 collaboration has published three studies: in Aguilar et al. [2018a] variations of GCR proton and helium fluxes were studied for the period from May 2011 to May 2017 with Bartels rotation\(^2\) (BR) time resolution, in Aguilar et al. [2018c] a similar study was performed for electrons and positrons. In Aguilar et al. [2019] a study of the solar modulation of helium isotopes covering the period 2011–2018 was presented for the first time.

Among other applications, data from both experiments provide us with unique opportunity to better understand the processes in the heliosphere and to validate the data of neutron monitors using direct CR measurement for the first time.

\(^2\)One Bartels rotation corresponds to the period of 27 days, starting from 8 February 1832
1.3.2. Neutron monitors

NMs are detectors that register hadron secondaries of the CR-induced atmospheric showers. The first design of a NM (IGY-type) was proposed by John Simpson in 1948, and in the early 1960’s Carmichael proposed an updated NM64 design with improved characteristics [Hatton and Carimichael, 1964, and references herein]. Neutron monitors are crucial for studying CR physics, in particular, several pioneering results were obtained using these devices, such as proving the existence of the solar modulation of GCRs and the discovery of high-energy solar neutrons [Simpson, 2000, and references herein]. Today NM data are used to study the solar modulation of GCRs and acceleration and propagation processes of SEPs.

The principal scheme of a standard NM is shown in Figure 1.6. It consists of four main components:

1. **The Reflector** is an outer shell made of proton-rich material, such as paraffin or polyethylene. Low-energy neutrons undergo elastic scattering with this material, so that they cannot penetrate into this material. Therefore background neutrons, not directly associated with the particle shower, are kept out of the NM and low-energy neutrons generated in the producer are kept in. At the same time this material is almost transparent to the cosmic-ray induced cascade neutrons, since they are more energetic.

2. **The Producer** is made of lead, and by weight it is the main component of a neutron monitor. Fast neutrons undergo nuclear reactions with lead to produce, on average, about ten lower-energy neutrons, amplifying the cosmic-ray signal.

3. **The Moderator** is also made of a proton-rich material. Its main function is to slow down the neutrons, generated within the reflector, to lower energies via
scattering, which makes them more likely to be detected in the proportional counter.

4. *The Proportional Counter*. In the proportional counter, neutrons with energies close to thermal encounter a nuclei of the filling gas and undergo neutron capture reaction. There are two main types of counters: based on $^{10}$B and $^3$He. In the standard NMs, $^{10}$B is used in form of BF$_3$ gas, which produces a signal via the reaction $n + ^{10}$B $\rightarrow \alpha + ^7$Li. Another type of proportional counters uses the reaction $n + ^3$He $\rightarrow p + ^3$H. However, because of very high leakage ability of helium, $^3$He-based counters are less stable on the long run. These nuclear reactions produce energetic charged particles ($p$ and $\alpha$) that ionize gas in the counter, producing an electrical signal.

After 1957, when the International Geophysical Year was carried out, about twenty standard NMs were deployed and launched all over the world, and this number increased over the time. A historical overview of ground-based observations in the energy range sensitive for the solar modulation can be found in Stoker [2009]. About 40 neutron monitors are in operation today, a map of these NMs is given in Figure 1.7. It is very important to have many NMs for several reasons. First, a network of neutron monitors located at different latitudes/longitudes (and having different cutoff rigidities, see Section 4.1) and altitudes allow us to study fluxes of cosmic ray particles and their anisotropy. Second, it allows us to make a cross-calibration of neutron monitors to check and find possible trends in data.

Over the world, several so-called “bare” NMs [Nuntiyakul et al., 2018] are also
installed in different locations. They do not have a lead producer, and this fact allows them to have enhanced sensitivity to low-energy nuclei, but simultaneously these NMs have reduced total efficiency [Vashenyuk et al., 2007]. The response of a “bare” NM can be used for a rough estimation of the CR spectrum (mostly, for SEPs), since the number of registered neutrons in ms timescales is proportional to the energy of the incoming particles. This so-called multiplicity study approach is perspective but requires precise simulation of CR propagation through the atmosphere and the detector material.

The common way to analyse the neutron monitor count rate is to use the yield function (YF), which relates the NM count rate to the primary CR flux. In this approach, a theoretical response of NM to CR in units [counts per second per counter], \( N \), can be calculated as follows:

\[
N(P_c, h, t) = \frac{1}{\kappa} \sum_{i=1}^{n} \int_{P_c}^{\infty} Y_i(R, h) J_i(R, t) dR,
\]

where \( P_c \) is the cutoff rigidity (see Section 4.1) for the given NM, \( Y_i \) are neutron monitor yield functions for different cosmic ray species in units of \([\text{cm}^2 \text{ster/counter}]\), \( h \) is the depth of the atmosphere for the given NM in units of \([\text{g/cm}^2]\). \( J_i \) are intensities of corresponding CR in the units of \([\text{nuc}/\text{m}^2 \text{ster s GV}]\) and summation is over different CR species. \( \kappa \) is a scaling factor (typically in the range 0.8–1.4) accounting for the “nonideality” (local surrounding, exact electronic setup, efficiency of counters, etc.) of each NM. For example, \( \kappa \) of any real NM in a real world depends on the thickness and material of the ceiling, walls and roof around it, the setup of the discrimination level and dead-time of the measuring electronics, etc. Such scaling factors need to be defined experimentally as it was done by Usoskin et al. [2017].

Clem and Dorman [2000] were the first who calculated the NM yield function (hereafter CD00 YF) using the modern Monte-Carlo (MC) simulation of cosmic ray transport through the atmosphere and the neutron monitor detector.

After that pioneering work several other NM YFs have been developed, in particular:

- CM12 [Caballero-Lopez and Moraal, 2012] YF was empirically constructed on the basis of NM latitudinal surveys, when the NM is mounted on a mobile platform, such as a vessel, and this platform has an expedition over different locations, corresponding to different cutoff rigidities [for example, Nuntiyakul et al., 2014]; thus the YF was defined only for the rigidities below 15 GV and without separation between different CR species, extension to higher energies/ rigidities and to \( \alpha \)-particles was made theoretically;
- Mi13 [Mishev et al., 2013] YF was computed using the PLANETOCOSMICS Geant4 MC simulation tool [Desorgher et al., 2005, 2009] considering, for the first time, the finite lateral size of the atmospheric cascade and the NM’s electronic dead time;
- Ma16 [Mangeard et al., 2016] YF was computed using the FLUKA MC package [release 2011, see Böhlen et al., 2014];
Modern yield functions are usually considered for two main components of cosmic rays: for protons and for α-particles, since other nucleonic CR components can be roughly scaled to helium. Figure 1.8 shows all the discussed yield functions for protons and α-particles.

Previously, it was noted that YFs depend on the atmospheric depth (or altitude), so it was suggested to use a simple barometric correction for high-altitude NM count rates. However, air shower development features on different altitudes are not considered in this approach. In order to quantify this effect, Flückiger et al. [2008] evaluated the NM YF altitude dependence using the MC simulation and showed that a simple barometric correction disagrees with MC simulations.

The recent AMS-02 proton and helium data [Aguilar et al., 2018a] with BR time resolution, together with measurements of heavy nuclei [Aguilar et al., 2017, 2018b, d], allowed us to check several assumptions in the neutron monitor yield function formalism and to validate the NM yield functions, that was done in Paper III. In particular, the nucleonic ratio of the heavier species to protons (including all species with Z >1) was previously taken fixed as 0.3 [Usoskin et al., 2017], since there were no helium data with a fine time resolution. In Paper III direct AMS-02 proton and helium measurements were used to calibrate neutron monitor responses and validate several NM yield functions. Data from seven sea-level NMs were used: Inuvik (geomagnetic cutoff rigidity $P_c=0.3$ GV), Apatity (0.65 GV), Oulu (0.8 GV), Newark (2.4 GV), Moscow (2.43 GV), Hermanus (4.58 GV), and Athens (8.53 GV). Real count rates from these NMs were compared with the corresponding calculated count rates (using Eq. 1.1 and all four YFs described above), and the scaling factors $\kappa$ and their variation in time were further constructed. The performed analysis showed that the best performance is achieved
with the use of the Mi13 yield function. CD00 and Ma16 YF tend to overestimate the response from low-energy CR, and therefore can not be recommended for an analysis of GLE. CM12 YF showed no significant trends due to the solar cycle, but it is defined for rigidities below 15 GV and for protons only.
2. Heliosphere and modulation of galactic cosmic rays

The Sun is our home star, and scientists have studied it for several centuries. Observations of the Sun with telescopes began in the 17th century and soon after that sunspots were discovered, i.e. areas of lower luminosity, which are visible as dark spots on the surface of the Sun.

Between 1761 and 1776, Christian Horrebow made regular observations of sunspots at the Rundetrán observatory in Copenhagen and proposed that the variation in number and size of sunspots is periodic [Jørgensen et al., 2019]. Samuel Heinrich Schwabe [1844] after 17 years of continuous observations of the Sun proved a periodic variation in the average sunspot number (SSN) and roughly calculated the variation period (≈ 11 years), so that the solar activity cycles were found and later called Schwabe cycle [see a review by Hathaway, 2015].

George Ellery Hale [1908] showed that sunspots are strongly magnetized (it was the first detection of magnetic fields beyond Earth). In 1919 he and his colleagues showed that the magnetic polarity of sunspot pairs is constant throughout a cycle. It is opposite across the equator throughout a cycle and reverses itself from one cycle to the next. These observations revealed that the complete magnetic cycle spans two solar cycles, or 22 years, before returning to its original state.

It was observed that the solar surface is magnetized also outside sunspots, that the surface magnetic field is weak and to the first-order approximation a dipole, and that this dipole undergoes polarity reversals with the same period as the sunspot cycle. Parker [1955], and soon thereafter Babcock [1961] considered the solar cycle as a spatiotemporal magnetic process and developed the first qualitative models of the solar dynamo. Leighton [1969] upgraded Babcock’s model and created a semi-qualitative approach based on both observations and solution of magnetohydrodynamics (MHD) equations. This model, which is known as the Babcock–Leighton model of the solar dynamo, is widely used nowadays.

Due to direct observations of solar activity with space-borne and ground-based experiments and accurate work with archive sunspot number data, today we know that the solar activity has several different periods of variations: the most pronounced are ∼27-day variation (the synodic period of the Sun’s rotation), 11- and
22-year variations. There are also known Rieger variation (≈137 days) and quasi-biennial oscillations (QBO), whose nature is still unknown. For shorter periods, there are transient sporadic events, such as solar flares and coronal mass ejections. Figure 2.1 shows the monthly averaged SSN [see Clette et al., 2014, for details] from 1750’s to nowadays, where 11-year cycles are clearly seen.

At larger timescales, periods of Grand minima (Maunder, Spörer) and Grand maxima of solar activity are clearly seen in cosmogenic proxies, such as $^{14}$C and $^{10}$B [Eddy, 1977; Stuiver and Braziunas, 1989; McCracken and Beer, 2007; Usoskin et al., 2007; Inceoglu et al., 2015]. In addition, detailed studies of the cosmogenic proxies and SSN data exhibit variations with typical time scales of ≈88 yr (Gleissberg cycle), ≈205 yr (de Vries/Suess cycle) and ≈2400 yr (Hallstatt cycle) [Damon and Sonett, 1991; Vasiliev and Deryachev, 2002; McCracken et al., 2013; Usoskin et al., 2016].

All these variations modulate the cosmic-ray particle flux near Earth, and therefore solar activity can be reconstructed from records of the CR variability.

Soon after deployment of NMs it was found that fluxes of secondary cosmic ray particles vary with ∼11-year cycle and are in anti-correlation with the SSN, so that for periods of high solar activity the flux of cosmic ray particles is lower and vice versa. Variability of monthly averaged SSN and the Oulu NM count rates is given in Figure 2.2, where a clear anti-correlation between the two records is visible. Detailed studies show that the correlation coefficient between the SSN and the cosmic ray flux in the energy region below 20 GV (or the NM count rate) is very high [about 0.8, Tomassetti et al., 2017], moreover there is a time lag between the SSN number and cosmic ray intensity varying from 2–3 months to 1.5 years.

Fig. 2.1: Monthly averaged international sunspot number v. 2.0. The plot was obtained using the SILSO data, Royal Observatory of Belgium, Brussels (http://www.sidc.be/silso/home).
Fig. 2.2: Monthly averaged international sunspot number v.2.0 (dashed line) and Oulu NM count rate. The plot was made using from the SILSO data, Royal Observatory of Belgium, Brussels (http://www.sidc.be/silso/home) and Oulu NM database (http://cosmicrays.oulu.fi/).

[Usoskin et al., 2001; Tomassetti et al., 2017]. In order to describe these results it was required to consider the processes in the heliosphere, which affects the particle transport, what was done for a first time by Eugene Parker [1965].

To better understand the principles of the solar modulation of cosmic rays we need to know what the main properties of the solar wind and the heliosphere are.

2.1. Solar wind

The solar wind (SW) is a stream of ionized particles flowing out from the solar corona at a speed of 300–1200 km/s into the surrounding space. It is one of the main components of the interplanetary (IP) medium.

An assumption of the existence of a stream of particles emitted from the Sun was first expressed by the British astronomer Richard Carrington. In 1859, Carrington and Richard Hodgson independently observed what was later called solar flare (see Chapter 3). The next day, a geomagnetic storm occurred, and Carrington suggested a connection between these phenomena.

Observations of comet tails by Ludwig Biermann, who suggested the existence of an ionized gas flowing outward from the Sun at velocities of $\sim 500–1500$ km s$^{-1}$, helped to inspire the solar wind theory of Parker [1958], who proposed a supersonic, radial expansion of the solar corona.

Due to the high conductivity of solar wind plasma, the solar magnetic field is
“frozen” into the outgoing wind flows and is observed in the IP medium in the form of the IP (or heliospheric) magnetic field (IMF/HMF) [Owens and Forsyth, 2013]. Due to the rotation of the Sun with the sidereal period of ~25.38 days, the magnetic field lines twist into Archimedean spirals.

Long-term observations of the solar wind showed that its characteristics depend on the solar activity and the solar latitude [McComas et al., 2008]. The solar wind is structured and has two distinct components: (a) the slow solar wind (about 300–500 km/s near the Earth’s orbit) is emitted from the “usual” corona of the Sun and (b) the fast solar wind (500–800 km/s near the Earth’s orbit) originates in “coronal holes”. Areas of coronal holes increase systematically as the solar activity declines, reaching the maximum value at the minimum phase and vice versa [Tokumaru et al., 2010]. The heliospheric current sheet (HCS), a shear within the heliosphere, across which the polarity of the magnetic field of the Sun changes [Smith, 2001], is sometimes considered as a region of the stationary wind, which is close to the slow SW in its properties. Interactions between slow and fast solar wind may form so-called stream interaction regions (SIRs) with special features [Richardson, 2018].

Due to the solar wind, the Sun loses about 1.5 million tons of matter every second. The solar wind consists mainly of protons, α-particles and electrons, while nuclei of other elements form very small quantities in the SW. While the solar wind comes from the outer layer of the Sun, it does not reflect the composition of the elements in this layer.

The solar wind pressure significantly decreases on the boundary of the heliosphere (see details in next Section), where SW becomes subsonic and deflects from outward propagation. These processes form a boundary of the heliosphere, which prevents penetration of interstellar plasma into the solar system. IMF and solar wind significantly modulate the flux of low energy (<20 GeV) GCRs coming from the boundary of the heliosphere inwards.

Local changes in the SW/IMF due to eruptive solar events lead to short-term changes in the CR flux as observed on Earth, such as fluxes of solar energetic particles, which will be discussed in details in Chapter 3, and Forbush decreases [Belov, 2008].

Interaction between solar wind/IMF and planetary magnetospheres leads to such phenomena as auroras, geomagnetic storms and energetic particles precipitations on solar-system planets with a magnetic field.

2.2. Heliosphere

The heliosphere is the region of the near-Sun space (within \( \approx 120 \text{AU} \)) hydromagnetically controlled by the supersonic solar wind and the interplanetary magnetic field. From the outside, the heliosphere is bounded by the termination shock arising in the solar wind due to its interaction with the interstellar plasma and the

\(^{1}\)The standard distance unit in the heliosphere is the astronomical unit (AU), which is equal to the mean distance between the Sun and the Earth \( (\approx 150 \times 10^6 \text{ km}) \).
interstellar magnetic field.

An artist’s representation of the heliospheric structure is shown in Figure 2.3. Within first 100 AU from the Sun the solar wind speed is on average 500 km/s, it is supersonic and expands outward from Sun. On the distance of about 100 AU its density becomes low enough, SW slows down, becomes subsonic, and disturbances in plasma start significantly influencing it, leading to the formation of a special heliospheric region, where a shear between the solar wind and interstellar plasma appears. The border at which the solar wind slows down is called the termination shock; the border along which the pressure of the solar wind and the interstellar medium is balanced is called the heliopause. Detailed properties of the boundary region are still an object of intensive studies: it was believed earlier that the region of a “bow shock” exists [Baranov et al., 1971], at which the interstellar wind becomes subsonic due to the influence of the HMF, but recent observations made by the IBEX (Interstellar Boundary EXplorer) mission showed an possible existence a bow wave rather than a bow shock [McComas et al., 2012]. Model efforts [Owens and Forsyth, 2013; Scherer et al., 2016; Pogorelov et al., 2017] show that depending on the strength and orientation of the magnetic field within the interstellar medium, its interaction with the heliosphere may or may not involve a standing bow shock or bow wave. Several authors claim that the so-called “hydrogen wall” due to charge exchange reactions between solar wind and interstellar neutrals exists [Baranov and Malama, 1993; Zank et al., 1996], as confirmed by a recent New Horizons mission observations [Gladstone et al., 2018].

The HCS, which separates the regions of opposite polarities of the IMF, serves as a magnetic equator, and solar wind parameters including speed, temperature, and
density vary with distance from the HCS. This surface extends radially from the Sun and reaches the boundaries of the heliosphere. First observations of the HCS by Wilcox and Ness [1965]; Rosenberg and Coleman [1969] showed that the shape of the current sheet is related to the large-scale dynamical flow of the solar wind, resulting in the Parker spiral, which is a kind of Archimedean spiral, resembling a ballerina’s skirt (Figure 2.4), and named after its discoverer Eugene Parker, who predicted it theoretically [Parker, 1958].

The solar cycle affects the HCS (its tilt, width and geometry), these changes are most pronounced at solar minimum and become less significant at solar maximum [Riley et al., 2002]. The thickness of the current sheet at 1 AU is about 10000 km and it increases (approximately linearly) with the heliocentric distance [Smith, 2001]. In the current sheet there is a weak electric current of about $10^{-10}$ A/m$^2$ [Cholis et al., 2016]. The magnetic field produced on the surface of the Sun is approximately $5 \times 10^{-6}$ T [Israelevich et al., 2001]. The magnetic field strength decreases unevenly with increasing distance from the Sun: the radial component of the magnetic field decreases with the distance as $1/r^2$, but azimuthal one decreases as $1/r$. As a result the IMF becomes more azimuthal with the distance and at 1 AU the magnetic field strength is of the order of 4–8 nT. These results agree with the Parker [1958] model predictions [Burlaga et al., 2002].
2.3. Transport equation

Parker [1965] introduced an equation for the heliospheric transport of cosmic rays. Thereafter, Gleeson and Axford [1967], Krymskij [1969], and Toptygin [1985] derived this equation in a strict sense. This differential equation combines the mechanisms of particle transport and adiabatic energy losses in the heliosphere and deals with the CR distribution function $f(r, R, t)$:

$$\frac{\partial f}{\partial t} = -\left( \mathbf{V} \cdot \nabla + \langle \mathbf{v}_d \rangle \cdot \nabla f \right) + \nabla \cdot (\mathbf{K}_s \cdot \nabla f) + \frac{1}{3} (\nabla \cdot \mathbf{V}) \frac{\partial f}{\partial \ln R} \tag{2.1}$$

where $R$ is the particle’s rigidity, $t$ is time, $r$ is the position vector in the heliosphere. Terms on the right hand side of the equation respectively represent convection (part a), with $\mathbf{V}$ being the solar wind velocity; averaged particle drift velocity $\langle \mathbf{v}_d \rangle$ caused by gradients and curvatures in the global HMF as well as in the HCS (part b); diffusion through irregular heliospheric magnetic field (part c), with $\mathbf{K}_s$ being the symmetrical diffusion tensor; and the term describing adiabatic energy changes (part d) due to the divergence of the expanding solar wind. The last one is especially important for GCR modulation in the heliosphere.

Additional source functions may also be added to Eq. 2.1, e.g., to the study solar modulation of Jovian electrons or SEPs [Potgieter, 2013, and references herein].

It is worth to note that particle drifts [Jokipii et al., 1977] are charge-sign depended in formalism of Eq. 2.1 and recent results from the PAMELA [Adriani et al., 2016b] and AMS-02 [Aguilar et al., 2018c] experiments allowed to study the difference in the electron and positron spectra evolution due to this effect in great details. Earlier these effects were studied using proton and electron data.

Special codes are typically used to solve the modulation problem. Historically, Jokipii and Kopriva [1979] were the first who suggested to use finite-difference scheme for the solution of the modulation problem [see, e.g., Kóta and Jokipii, 1991; Potgieter and Moraal, 1988]. This approach is still in use and represented, e.g., by works of the South African CR group [see, e.g., Bisschoff et al., 2019]. Another approach for solving Eq. 2.1 is to use stochastic simulations, and there are two ways to do it. First one is to solve the direct problem [as pioneered by Gervasi et al., 1999], when a pseudo-particle transport from the boundary of the heliosphere to its centre is simulated. Nowadays, the direct problem can be solved with a time-dependent 3D model by Luo et al. [2018]. This direct simulation is straightforward but highly CPU time consuming. Another way to use stochastic simulations is solving the inverse problem [as proposed by Zhang, 1999], where the quasi-particle transport is “backtraced” from the inner heliosphere to its boundary. It is much more effective in the sense of computational power but less straightforward physically. This way of solar modulation problem solving is represented e.g. by the latest revision of 2D HelMod model [Boschini et al., 2018].
2.4. Local interstellar spectrum

Local interstellar spectrum (LIS) is the CR spectrum outside the heliosphere, different CR species have slightly different LIS shapes. The solar modulation of cosmic rays and the absence of experimental CR data outside the heliosphere in the energy region from $\sim 100$ MeV to $\sim 40$ GeV result in the fact that LIS in this energy range is still not precisely known, but it is required in order to solve the modulation problem. In addition, knowledge of this spectrum is crucially important for understanding of processes of CR acceleration and propagation in the Galaxy. Proton/nuclei LIS for the energy region above $\sim 40$ GeV, where the modulation effects become negligible, was evaluated for the first time in early 1970’s [Ryan et al., 1972; Smith et al., 1973]. Today, thanks to results of space- and balloon-borne missions, LIS for several CR species, such as electrons, positrons, and nuclei up to the oxygen, are well known. The LIS in the low-energy region for several CR species such as protons, electrons and heavy nuclei up to Ni, became known in the 2010’s, thanks to the pioneering results of the Voyager-1 mission [Stone et al., 2013; Cummings et al., 2016].

By solving Eq. 2.1, one can “demodulate” the CR spectrum measured near Earth and thus estimate corresponding LIS [for example, Vos and Potgieter, 2015; Aslam et al., 2019; Boschini et al., 2017, 2018]. However, it yields a reference unmodulated GCR spectrum for specific modulation model rather than the true LIS. Another way to estimate the LIS is the use of numerical codes for the calculation of the propagation of relativistic charged particles in the Galaxy, that requires a solution of a system of transport equations (time-dependent partial differential equations) in 3D or 4D (space and momentum) in the whole Galaxy. The most popular code for such kind of calculation is GALPROP [Moskalenko and Strong, 1998; Strong and Moskalenko, 1998; Porter et al., 2019], and also DRAGON [Evoli, 2018] and USINE [Maurin, 2015] are based on similar principles.

Because of solar modulation and the fact that the nature of the heliospheric diffusion coefficients is not fully established, all cosmic ray LIS at kinetic energies $\lesssim 10$ GeV remain very approximate [Potgieter, 2013].

2.5. Force-field parametrization

Gleeson and Axford [1968] proposed a heavily simplified version (spherical symmetry, steady state, and adiabatic cooling) of Eq. 2.1 that reduces all the modulation effects to a single parameter $\phi$ called the modulation potential, which is mathematically interpreted as the averaged rigidity (i.e., the particle’s momentum per unit of charge) loss of a CR particle in the heliosphere. However, it is only a formal index whose physical interpretation is not straightforward, especially on short timescales and during periods of high solar activity [Caballero-Lopez and Moraal, 2004].

Although it has little physical sense because of the simplified assumptions, this model, called force-field (FF), provides a very good and useful parametrization of
the near-Earth GCR spectrum, which can later be used in different applications. One of the most promising applications of the FF model is the evaluation of the solar modulation potential (and therefore solar activity) on the long-time timescale [Usoskin, 2017].

The force-field model links the energy spectrum of GCR particles $J$ of type $i$ (protons, helium, etc.) with the mass and charge numbers $A_i$ and $Z_i$, respectively, near Earth to their reference intensity outside the heliosphere, LIS $J_{LIS}$, so that

$$J_i(T, \phi) = J_{LIS}(T + \Phi_i) \frac{T(T + 2T_i)}{(T + \Phi_i)(T + \Phi_i + 2T_i)}$$

(2.2)

where $T$ is the kinetic energy per nucleon, $T_r = 0.938$ GeV is the proton’s rest mass, and $\Phi_i = \phi \cdot (eZ_i/A_i)$. One can see from Eq. 2.2 that the exact value of the modulation parameter depends on the reference LIS [Usoskin et al., 2005; Herbst et al., 2010, 2017; Asvestari et al., 2017].

Usoskin et al. [2005] performed the first consisted long-term reconstruction of $\phi$, based on the NM network data since 1951 using LIS parametrization by Burger et al. [2000]. That work was updated in Usoskin et al. [2011], and recently in Usoskin et al. [2017], using the PAMELA data and an updated NM yield function by Mishev et al. [2013]. This most recent reconstruction of the solar modulation potential $\phi$ for NM era from 1950’s to nowadays was made using an estimation of the LIS by Vos and Potgieter [2015], the NM YF by Mishev et al. [2013] and PAMELA proton data for 2006–2009 [Adriani et al., 2013]. There were also other reconstructions of solar modulation potential based on different methods [Corti et al., 2016; Gieseler et al., 2017].

New PAMELA proton data for 2010–2014 and AMS-02 proton and helium data for 2011–2017 allowed us to perform the NM response calibration, what was done in Papers I, III.

In Paper I the PAMELA proton data for the period 2010–2014 with the Carrington rotation time resolution [Martucci et al., 2018] was considered by applying the formalism of Eq. 2.2 to find the best-fit values of the solar modulation potential $\phi$ using $\chi^2$-minimization procedure (the LIS from Vos and Potgieter [2015] was used). The obtained $\phi$ values were compared with the values estimated from the NM data [Usoskin et al., 2017], for periods of low and moderate solar activity. An agreement between the satellite data and NM data was found very good, but for high solar activity a systematic discrepancy of $\approx10\%$ was found (Figure 5 in Paper I, cf. Gieseler et al. [2017]). It can be explained by the fact that periods of high solar activity can not be described with the FF model, whose simplifications may not be working well in such conditions. In Usoskin et al. [2017] it was shown that the empirical dependence between the NM count rate and $\phi$ can be deduced. Analysis of PAMELA data for 2010–2014 allowed to find a discrepancy between this empirical model and experimental data, a reason of that being unclear by that time.

In Paper III, the AMS-02 GCR proton and helium data for the period 2011–2017 with BR time resolution [Aguilar et al., 2018a] were considered. Similar to Paper I, the $\phi$ values were estimated using the $\chi^2$-minimization procedure and compared with the NM-based ones. For periods of overlapping observations, AMS-
02 proton spectra and NM-based φ-reconstruction are in good agreement with the PAMELA data, but for periods of low solar activity (2006–2009 for PAMELA and after May 2016 for AMS-02) there is a small (about 20 MV) discrepancy in the estimated φ value (see Figure 9 in Paper III). For high solar activity, the discrepancy between NM-modelled data and satellite-based data can reach up to 50–100 MV (up to 15%). Nevertheless, for practical applications (where the FF model is typically applied) the precision is sufficient, so in this sense FF model passes the validation using AMS-02 and PAMELA data.

In Paper III the standard assumption [e.g., Usoskin et al., 2011] that all heavier-than-proton species can be described by the same model as protons was also checked and it was found out that this assumption does not work for periods of high solar activity. It was shown that a constant scaling coefficient of 0.3 in LIS used earlier to account for the heavier-than-proton species is not correct, since the mean coefficient obtained for the AMS-02 dataset is 0.353 and, importantly, it exhibits a significant solar cycle dependence. The use of this modified coefficient leads to a better agreement (in comparison to Paper I) in an empirical dependence between the NM count rate and φ.
3. Solar eruptive events

Sometimes transient eruptive events occur on the Sun, such as solar flares and coronal mass ejections (CMEs), whose effect can propagate to the heliosphere disturbing it [see reviews by Gopalswamy, 2001; Cane et al., 2002]. During these events, a lot of energy is irradiated in different wavelengths, in particular, in radio wavelengths [so called solar radio bursts, see Dulk, 1985, for details]. Simultaneously, charged particles, such as protons, ions, electrons, can be accelerated near the solar surface, in the solar corona and in the interplanetary medium, these particles are known as solar energetic particles (SEPs), see Section 3.3. There are two main SEP acceleration mechanisms. One is related to magnetic reconnection in flare sites, which can be observed with the hard X-ray, gamma-ray and microwave signatures during the impulsive flare phase. The other is particle acceleration at large-scale shocks often driven by an agent CME, exemplified by interplanetary shocks.

Large solar flares are often accompanied by emergence of fast CMEs into the interplanetary space.

Charged particles, accelerated in solar eruptive events, propagate in the heliosphere along the HMF lines. If the acceleration site on Sun is located on the HMF line connected to Earth, SEPs can reach Earth. From this point of view, geoeffective SEP events usually take place near the western limb. SEP that reach Earth, can produce a ground level enhancement event (GLE, see Section 3.4). High-intensity SEP fluxes play a crucial role in space weather and are very dangerous for both manned and unmanned satellites/space stations, so they posses a serious risk to modern and future technological infrastructures [Gopalswamy, 2018; Schrijver and Siscoe, 2010].

3.1. Solar flares

Solar flares are sudden energy releases on the surface of the Sun. They usually occur in close proximity to active regions and last from few to few tens of minutes. Solar flares are the first transients that became known to us, the first records
of them were made by Carrington [1859] and independently by Hodgson [1859] in white light on 2 September 1859. There were many sunspots in period from August 28 to September 2 and at September 1 afternoon a powerful white light flare occurred, which produced a big coronal mass ejection, that soon reached Earth leading to the most powerful directly observed geomagnetic storm. This event is now known as the Carrington event [Cliver and Dietrich, 2013].

Flares can occur on different scales: from magnetic network to big active regions [Benz, 2017]. Large flares (as in Figure 3.1), however, occur in active regions showing a complex geometry of the 3D magnetic field. Largest flares are associated with arcades of loops spanning a line of zero line-of-sight magnetic field on the surface.

Flares are observed in a wide range of electromagnetic waves such as radio, visible light, X-rays, and gamma rays [Shibata and Magara, 2011]. In extreme case, even the photosphere responds to a big flare, observed as white-light brightenings. Also a flare may accelerate high-energy particles, which travel through the interplanetary space, sometimes having a severe impact on the environment of the Earth.

The key physical processes for producing a flare are: emergence of magnetic field from the solar interior to the solar atmosphere (flux emergence), local enhancement of electric current in the corona (formation of a current sheet), and rapid dissipation of electric current (magnetic reconnection) that causes shock heating, mass ejection, and particle acceleration. Evolution toward the onset of a flare is rather quasi-static when free energy is accumulated in the form of coronal electric current, while dissipation of the coronal current proceeds rapidly, producing various dynamic events that affect lower atmospheres such as the chromosphere and

Fig. 3.1: X28 solar flare at 4 November 2003 as seen by EIT instrument at wavelength 195 Å. Courtesy of SOHO/EIT (ESA & NASA).
Table 3.1: Modern solar flare classification based on soft X-ray peak energy flux, their rate and fraction of events accompanied by CMEs. Adopted from Schrijver and Siscoe [2010].

<table>
<thead>
<tr>
<th>GOES class</th>
<th>1–8 Å peak [W/m²]</th>
<th>CME fraction [%]</th>
<th>Events/year (max/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>&gt;10^{-8}</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>B</td>
<td>&gt;10^{-7}</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>C</td>
<td>&gt;10^{-6}</td>
<td>≈20</td>
<td>&gt;2000/300</td>
</tr>
<tr>
<td>M</td>
<td>&gt;10^{-5}</td>
<td>≈50</td>
<td>300/20</td>
</tr>
<tr>
<td>X</td>
<td>&gt;10^{-4}</td>
<td>≈90</td>
<td>10/one</td>
</tr>
<tr>
<td>X10</td>
<td>&gt;10^{-3}</td>
<td>≈100</td>
<td>few/none</td>
</tr>
</tbody>
</table>

photosphere. In particular, accelerated particles precipitate to the chromosphere, where they heat plasma to high temperatures observed in soft X-rays and hot plasma expands along the loop into the corona, a process termed “evaporation” [Shibata and Magara, 2011]. It is a simplified flare mechanism, that considers only basic processes and detailed understanding of flare mechanisms is still under discussion.

Three stages of a solar flare are typically considered [Fletcher et al., 2011]. First is the precursor stage, during which a release of magnetic energy is triggered, soft X-ray emission is detected in this stage. In the second or impulsive stage, protons and electrons are accelerated to energies exceeding 1 MeV as a result of magnetic reconnection process. During the impulsive stage, radio waves, hard X-rays, and gamma rays are emitted. Gradual build up and decay of soft X-rays can be detected in the third, decay stage. Duration of these stages can be as short as a few seconds or as long as an hour.

The frequency of flares occurrence roughly follows with the Sun’s eleven year cycle. When the solar cycle is at minimum, active regions are small and rare, and only few solar flares are detected [Hathaway, 2015].

The modern view on flare classification defines the flare class depending on the energy flux in soft X-rays as monitored by the GOES and other space-borne experiments [Schrijver and Siscoe, 2010; Fletcher et al., 2011]. In Table 3.1 this classification is given together with accompanying information about CME fraction [Yashiro et al., 2005] in these events and their occurrence frequency.

3.2. Coronal mass ejections

Coronal mass ejections are large-scale eruptive events on the Sun that consist of hot plasma and typically radially expand from the generation region. They were discovered in the early 1970’s with the launch of the first Earth orbiting coronagraph [see, e.g., Howard, 2006, for a historical review].
Fig. 3.2: Evolution of coronal mass ejection on 27 February 2000 as taken by LASCO C2 and C3 coronographs in white light. Courtesy of SOHO/LASCO (ESA & NASA).

According to the original definition, a CME is an observable change in the coronal structure that includes the appearance and outward motion of a new, discrete, bright, white-light feature in the coronagraph field of view [Chen, 2011, and references herein]. Later it was found that CMEs can also be observed in other wavelengths Hudson and Cliver [2001].

From the physical point of view, CMEs are large expulsions of plasma and magnetic field from the Sun’s corona. Figure 3.2 shows an example of a powerful CME as seen by space-borne coronographs. During the CME, coronal material is ejected into the interplanetary space, and accompanying magnetic field is embedded (“frozen-in”) into the material. Typically, the magnetic field within the CME is stronger than the background IMF strength. The heliospheric counterparts of CMEs at the Sun are known as interplanetary coronal mass ejections (ICMEs) and can be identified in situ on the basis of magnetic field, plasma composition and SEP signatures [Zurbuchen and Richardson, 2006]. The interaction between the regular solar wind and ICMEs (which have different velocity and magnetic field) may generate interplanetary shocks and compression/rarefying regions [Kilpua et al., 2017].

CMEs differ with respect to their mass, angular width and morphology, their velocity and energy. In particular, the mass of a CME typically lies in the range from $10^{11}$ to $4 \times 10^{13}$ kg, the angular width of CME projected on the plane of the sky is in the interval from $\sim 2^\circ$ to $360^\circ$. CMEs with angular resolution $<10^\circ$ are considered as “narrow” CMEs and their morphology differs from others [Schwenn, 2006], since they show jet-like motions along open magnetic lines, whereas normal CMEs are
characterized by a closed frontal loop. A typical morphology for normal CMEs is the so-called three-part structure, including a bright frontal loop, a dark cavity and an embedded bright core, that can be seen in Figure 3.2.

The sum of kinetic and potential energies of a typical CME is of the order of $10^{22} - 10^{25}$ J, which is similar to the energy budget of solar flares [Emstke et al., 2004]. The CME speed ranges from tens to thousands km/s with the average value being $\sim$400 km/s [Yashiro et al., 2004].

The occurrence rate of CMEs is basically consistent with the solar cycle (SSN), with a peak delay of 6–12 months [Robbrecht et al., 2009]. Evaluations of the occurrence rate differ from 0.5–2 per day near the solar minimum to $\sim$6–8 per day for the solar maximum [Chen, 2011].

CMEs are usually associated with solar flares but can occur independently [Zhang and Dere, 2006].

### 3.3. Solar energetic particles

Solar energetic particles are high-energy particles accelerated in solar eruptive events, they consist of protons, electrons and ions with energy ranging from a few tens of keV to few GeV.

The first recognized SEP events are attributed to Forbush [1946], who used ground-based ionization chambers to register the secondary cosmic-ray particles associated with a solar flare (so that the events were actually what we now call ground-level enhancement (GLE) events).

The SEP dichotomous classification is generally accepted [Reames, 2013; Kallenrode, 2003], according to which SEP events can be divided into two distinct classes: (a) the gradual and (b) the impulsive events (Table 3.2 and Figure 3.3), which are related to different acceleration mechanisms (Figure 3.4). The gradual events [Desai and Giacalone, 2016], associated with type II radio emission, are believed to be accelerated high in the corona by shocks driven by coronal mass ejections. They are characterized by elemental abundances, charge states, and temperatures typical of the ambient corona, and they produce by far the highest SEP integral intensities near Earth. The impulsive events [Reames, 2017], generally much less intense, are linked to short-duration soft X-ray flare emission from low altitudes and fast-drift type III radio emission reflecting electron escape into the interplanetary medium. They are thought to be accelerated at solar flare sites mostly by processes in association with magnetic reconnection or wave-particle interactions and are characterized by enrichments in $^3$He, electrons and heavy ions such as iron. Thus, gradual SEP events are large, intense, and spatially and temporally extensive. In contrast, impulsive SEP events seen in space are small, weak and compact, but also numerous, where magnetic reconnection includes open magnetic field lines.

Recent studies have shown that SEP events in general originate from a mixture of impulsive and gradual processes, and the event evolution depends on their relative importance and the magnetic connection to Earth, albeit there is still
Fig. 3.3: Example of two classes of SEP events and their intensity-time profiles as registered by ISEE 3 satellite mission. On the left: A large gradual event is produced by a CME-driven shock that populates IMF lines with SEPs over a broad longitudinal extent. On the right: A solar flare produces an impulsive event that populates only well-connected IMF lines. Adapted from McComas et al. [2016].
Fig. 3.4: Sketch of simplified SEP acceleration mechanisms in impulsive (left-hand side) and gradual (right-hand side) event. Picture taken from Cliver [2000].

Table 3.2: SEP events classification

<table>
<thead>
<tr>
<th>Property</th>
<th>Impulsive</th>
<th>Gradual</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEP duration</td>
<td>&lt;1–20 h</td>
<td>&lt;1–3 days</td>
</tr>
<tr>
<td>X-ray duration</td>
<td>~10 min–1 h</td>
<td>≥1 h</td>
</tr>
<tr>
<td>$^3\text{He}/^4\text{He}$</td>
<td>~1</td>
<td>$\sim 4 \times 10^{-4}$</td>
</tr>
<tr>
<td>H/He</td>
<td>~10</td>
<td>~100</td>
</tr>
<tr>
<td>Coronographs</td>
<td>N/A</td>
<td>CME</td>
</tr>
<tr>
<td>Solar wind</td>
<td>N/A</td>
<td>IP shock</td>
</tr>
<tr>
<td>Events/year</td>
<td>~1000</td>
<td>~10</td>
</tr>
</tbody>
</table>
no consensus about the details of the individual mechanisms [Gopalswamy et al., 2012].

In addition to SEPs, there are solar neutrons, which are produced rarely by collisions in the solar photosphere (such as the interactions between the accelerated proton, heavy ion and the surrounding atmosphere). These events are very rare, mostly due to the fact that neutrons should be relativistic to reach Earth, since neutron’s lifetime is 880 s and light travels 1 AU distance during 500 s. On the other hand, neutrons are neutral and do not interact with the HMF and the magnetosphere, so they can propagate to Earth on a straight line and initiate an atmospheric cascade before the arrival of SEPs. Solar neutrons were registered by both space-borne missions and ground-based NMs several times [Yu et al., 2015, and references herein]. Only before two GLE events (see next Section), the neutron precursor signal was recorded.

3.4. Ground level enhancements

Ground level enhancements (in terms of the NM count rates) are produced when ∼GeV SEP-related nuclei initiate a nuclear cascade (see Section 4.2) through the Earth’s atmosphere that can be detected at the ground level as an increase in NM count rate above the GCR background signal.

The formal definition of a GLE is the following: “A GLE event is registered when there are near-time coincident and statistically significant enhancements of the count rates of at least two differently located neutron monitors including at least one neutron monitor near sea level and a corresponding enhancement in the proton flux measured by a space-borne instrument(s)” [Poluianov et al., 2017].

Relatively weak SEP events registered by only high-altitude polar neutron monitors, but with no response from cosmic-ray stations at the sea level, can be classified as sub-GLEs [Mishev et al., 2017]. In the work of Poluianov et al. [2017] the formal definition of a sub-GLE was also introduced: “A sub-GLE event is registered when there are near-time coincident and statistically significant enhancements of the count rates of at least two differently located high-elevation neutron monitors and a corresponding enhancement in the proton flux measured by a space-borne instrument(s), but no statistically significant enhancement in the count rates of neutron monitors near sea level”.

Sometimes count rates of NMs can slightly increase as the result of a transient reduction of the geomagnetic shielding [Belov et al., 2005, but this is only related to detectors located at low or medium latitudes] or crossing of the boundary layer between regions with different heliospheric parameters [Gil et al., 2018]. The latter events are called anisotropic cosmic-ray enhancements (ACRE).

Up to date we know 72 GLE events, associated with high-energy SEPs. The first four events were registered before the NM era and today researchers make some efforts in attempts to find and reconstruct the data for these events. Starting from GLE #5, all GLE data from NM stations are stored in the International GLE database [Tuohino et al., 2018], hosted by Oulu university (gle.oulu.fi).
The database provides NM information in a useful form with visualizations and for several events there are estimates of the effective doses of radiation during the event [Tuohino et al., 2018].

Figure 3.5 shows temporal evolution of GLE #5 event, which was the most powerful registered one, as recorded by Leeds NM. Several characteristics of the GLE signal are typically considered: the maximal increase in data (in percents above to GCR background), the GLE integral increase (in units of [%×hour] in respect to GCR background) and temporal evolution of the GLE signal. Some or all these characteristics can be used for the reconstruction of the SEP fluxes. In particular, the GLE temporal evolution is required for careful reconstruction of the SEP time-depended differential fluxes, accounted for the anisotropy of fluxes and transient geomagnetic effects [Mishev and Usoskin, 2016], this method is known as “full reconstruction”. This method allows to reconstruct the temporal evolution and anisotropy of SEP fluxes near Earth and to study mechanisms of SEP acceleration and transport in details, as in Kocharov et al. [2017]. But this method requires a high accuracy in the assessment of the GLE effect and is very laborious.

Tylka and Dietrich [2009] proposed to use the GLE integral increase as recorded by NMs for assessing the integral fluence of SEPs. They created an original method of GLE analysis using the NM yield function formalism (on the basis of the NM YF by Clem and Dorman [2000]) and a prescribed SEP fluence functional shape in high energy as a single power law, without consideration of anisotropic effects, so that the method is much more simple in calculation and it is easier to get a stable solution in comparison to the method of full reconstruction. These authors also considered lower (in comparison to NM data) energies, where the data from
GOES/SAMPEX/IMP-8 satellites or similar experiments or with ionospheric detection techniques for early GLE’s were used. To describe both datasets with one analytical function they used the Band function [Band et al., 1993], originally created for the description of gamma-ray burst spectra in astronomy. This function has the following form:

\[
F(\geq R) = \begin{cases} 
  J_0 \left( \frac{R}{R_1} \right)^{-\gamma_1} \exp \left( -\frac{R}{R_0} \right), & \text{if } R < R_1 \\
  J_0 \left( \frac{R}{R_1} \right)^{-\gamma_1} \exp \left( \frac{-R}{R_1} \right) \left( \frac{R}{R_0} \right)^{-\gamma_2}, & \text{if } R \geq R_1,
\end{cases}
\]  

(3.1)

where \(F(\geq R)\) is the omnidirectional event-integrated fluence of SEP in units of \([\text{cm}^{-2}]\), \(J_0\) is an overall fluence normalization coefficient, \(\gamma_1\) is the low-rigidity power law index, \(\gamma_2\) the high-rigidity power law index and \((\gamma_2 - \gamma_1)R_0 \equiv R_1\) is the breakpoint rigidity. The Band function is constructed in such a way that both the function and its first derivative are continuous. Results for 58 GLEs were presented and tabulated in a recent work of Raukunen et al. [2018]. For the first four GLEs there is no data, for eight the signal was too weak for reliable reconstruction, and finally, GLE #72 took place after the submission of the paper.

Despite of the evident high significance of this work, there are two issues which should be addressed. First, it is the choice of the prescribed function. Several authors claimed that a single power-law energy spectrum is unphysical and that some kind of cutoff or rollover must exist at high energies of SEPs [for example, Ellison and Ramaty, 1985], that was recently confirmed by direct measurements with the PAMELA experiment [Bruno et al., 2018]. Bruno et al. [2019] supposed also to use the Band function with an additional exponential cutoff for the description of event-integrated SEP intensity in a wide energy range. Second, it is the use of the YF by Clem and Dorman [2000], since the new AMS-02 measurements have made it possible to validate NM YFs (see Subsection 1.3.2 and Paper III) and it was shown that the CD00 YF overestimates the NM response for low rigidities (<10 GV), where the GLE effect takes place. It is worth to mention that the AMS-02 data for the NM YFs validation became available after the publication of Raukunen et al. [2018], so that the authors were not aware of it.

These features were addressed in Papers II and IV, where the new “effective rigidity” method for the SEP fluence reconstruction was presented and verified on two recent GLEs: #69 and 71. GLE #69 occurred on 20 January 2005 and was one of the most powerful GLEs ever registered and was characterized by strong anisotropy of SEP fluxes. GLE #71, occurred on 17 May 2012 was much weaker than GLE #69, but also anisotropic. Importantly, this latter event was observed by the PAMELA mission, and recently the integral fluence reconstructed from the PAMELA data was published by Bruno et al. [2018], where it was fitted with the Ellison-Ramaty function. For the studied GLE event, this result is in agreement with results of a combined analysis of the PAMELA, GOES, and ACE experiments [Bruno et al., 2019]. Moreover, the results of the “effective rigidity” reconstruction were compared with the full reconstruction of the time-evolving SEP spectrum, considering realistic anisotropy [Mishev et al., 2018]. The obtained fluence values using the Mi13 YF are in good agreement with the full reconstruction and PAMELA data for GLE #71 but disagree with the previous evaluations of
Raukunen et al. [2018], which are based on the CD00 YF. Moreover, for GLE #69, the new method allowed to observe an expected rollover of the fluence energy dependence in the NM region. A revised reconstruction of integral fluences for all the observed GLE events using the “effective rigidity” method will be published later.
4. Cosmic rays and Earth

4.1. The terrestrial magnetic field and geomagnetic cutoff

The Earth has a magnetic field which partially shields our planet from harmful cosmic ray radiation. The structure of the magnetic field is complicated, but in the first approximation it can be considered as an eccentric tilted dipole. At several Earth radii, the structure of the magnetic field becomes significantly asymmetrical due to the influence of the solar wind and the IMF [Pulkkinen, 2007]: it is compressed on the dayside and has an extended tail on the nightside (see Figure 4.1).

Due to the Earth’s magnetic field, charged particles cannot penetrate into the Earth’s atmosphere if they do not have sufficient energy (with exception for magnetic pole regions). Because of this fact, proper consideration of the geomagnetic field is essential for the data analysis of cosmic-ray experiments, both ground-based and space-borne. One of key concepts within this consideration is so-called geomagnetic cutoff rigidity, whose physical meaning is simple: only particle with the rigidity greater than that can penetrate the Earth atmosphere. The cutoff rigidity depends on the geomagnetic field strength and inclination so that it is high (up to about 17 GV) in equatorial regions and zero at magnetic poles, moreover, it depends on the particle arrival direction. It is standardly defined using back-tracing charged particles in the model geomagnetic field, so that the particle starts at a given location at the height of 20 km and is traced backwards along its computed trajectory until it either exits the magnetosphere, hits the Earth’s surface or becomes trapped (spends too much time bouncing inside the magnetosphere). Trajectories of particles that can exit the magnetosphere are called “allowed”, and, on the contrary, “forbidden” trajectories correspond to particles, that cannot exit the Earth’s magnetosphere. Due to a complex structure of the geomagnetic field, cutoff rigidity for a given location has a complicated structure, since there is an interval, called penumbra, between the so-called lower cut-off rigidity $R_L$, below which all charged particle trajectories in the magnetosphere are forbidden (umbra region) and, upper cut-off rigidity $R_U$, above which all trajectories are allowed [Cooke et al., 1991]. Fig. 4.2 shows the trajectory-derived cosmic-ray cutoff rigidity and penumbra structure in the vertical direction for location of Newark NM,
Fig. 4.1: Basic structure of the Earth magnetosphere. Courtesy of ESA.
Fig. 4.2: Illustration of trajectory-derived cosmic-ray cutoff rigidity (X-axis, in GV) and the cosmic-ray penumbra structure in the vertical direction for Newark NM for IGRF epoch 1980. White indicates allowed rigidities, black indicates forbidden rigidities. The lower $R_L$ and upper $R_U$ cutoff rigidities are also indicated. Picture adopted from Cooke et al. [1991].

where white indicates allowed rigidities, black indicates forbidden rigidities. For this NM, the lower cut-off rigidity equals to 1.9 GV and upper one to 2.3 GV.

While the cutoff rigidity depends on the particle arrival direction, often the so-called effective vertical cutoff rigidity is used, which is equal to the minimum rigidity that a charged particle must possess to reach the middle atmosphere (20 km altitude) in the vertical directions [Cooke et al., 1991]. The cutoff rigidities are not constant in time, and they change because of (a) the slow change of geomagnetic field [see, e.g., Gvozdevskii et al., 2016] and (b) transient changes in the solar wind and the IMF, that affects outer Earth’s magnetosphere [Chu and Qin, 2016]. To describe these features, special models are used. There is an international standard model of the geomagnetic field, The International Geomagnetic Reference Field (IGRF), which represent mathematical description of the large-scale structure of the Earth’s main magnetic field and its secular variations, and is performed using Gauss spherical harmonic decomposition of the surface magnetic field. It is updated at 5-year intervals, reflecting the most accurate measurements available at that time. The current 12th edition of the IGRF model [Thébault et al., 2015] allows one to calculate the Earth magnetic field for period from 1900 until 2020. Special codes for external, disturbed by HMF magnetosphere also exist, a number of them were created by Nikolay Tsyganenko [TS models, see Tsyganenko and Andreeva, 2015, and references herein], which use observable near Earth solar wind and HMF parameters to produce a model of the external magnetosphere. Transient disturbances of HMF also might affect the magnetosphere and cause short, but significant changes in cutoff rigidity. This effect was observed, e.g., by Adriani et al. [2015b] in PAMELA experiment during the 17 May 2012 SEP event.

Considering all the facts mentioned above, particle motion in the magnetosphere has a complicated path and can be described by a set of linear differential equations and typically solved with the fourth order Runge-Kutta integration [Smart et al., 2000], which allows to backtrace the particle trajectories. Special software package PLANETOCOSMICS [Desorgher et al., 2005, 2009] allows to compute particle trajectories using different models of the planets magnetosphere and external magnetic fields. As an example of these computations, the map of vertical
4.2. Atmospheric cascade

Earth’s atmosphere is thick, with the total amount of matter in the atmosphere being $\sim 10^{33}$ g/cm$^2$ on average for sea-level, corresponding to $\sim 12$ nuclear interaction lengths [Grieder, 2001]. Therefore, energetic hadrons, after penetration into the Earth atmosphere, interact with nuclei of atmospheric gases, producing secondary particles. This phenomenon is known as the atmospheric cascade or the air shower. If the energy of the incoming particle is more then $10^{14}$ eV the resulting air shower is very wide and called an extensive air shower.

A simplified scheme of the air shower is shown in Figure 4.4 (without neutrino component). After the interaction of primary CR (mostly protons and $\alpha$-particles) with a nucleus in the atmosphere, daughter products of the interaction propagate farther and may again interact with atmospheric atoms. On the ground level, a shower of secondary particles can be observed. The vertical (directed on the right angle to Earth) fluxes of cosmic rays in the atmosphere with $E > 1$ GeV as a function of the atmosphere depth are shown on Figure 4.5. One can see that the
Fig. 4.4: Scheme of an air shower generated by primary cosmic-ray particle in the Earth atmosphere. Picture adopted from https://en.wikipedia.org/wiki/Air_shower_(physics).

The main components of particle showers on the Earth surface are neutrinos, muons, protons and neutrons. Air showers have three main components: hadronic, electromagnetic \((e^+, e^-, \gamma)\) and muon \((\mu^+, \mu^-)\). Hadronic component mostly consists of protons, neutrons and \(\pi\)-mesons. Pions itself are hadrons, but they decay fast, typically into the \(\mu\nu\) or \(\gamma\gamma\)-pairs, depending on the pion charge, and produce EM shower component. In addition, neutrino component is often considered as a separate topic of study since the detection of neutrino forms a very hard task.

Knowledge of the cross-section of nuclear reactions and features of air shower developments crucial for the ground-based cosmic-ray experiments data analysis, in particular, for NM response analysis and interpretation. Modern way of use the NM response is data analysis using yield functions (Subsection 1.3.2), and modern way to calculate NM YFs is usage of Monte-Carlo simulations of the charged particle transport through the atmosphere and the detector material.
Fig. 4.5: Vertical fluxes of particles of CR-induced cascades in the atmosphere with $E > 1$ GeV. Picture adopted from Tanabashi et al. [2018]
5. Summary

The thesis includes results of a study of both periodic and transient variations in the cosmic-ray flux as registered by the space-borne experiments PAMELA and AMS-02 and by the ground-based network of neutron monitors.

In Paper I, a comparison of the CR proton spectrum directly measured by the PAMELA experiment during 2006–2014 with data of polar neutron monitors for the same time interval is presented. It is shown that the measured spectra are well described by the force-field model for the modulation potential range 350–750 MV. The obtained modulation potential agrees with that calculated from the data of the world neutron monitor network for low solar activity between 2006 and 2012 but diverges during the solar cycle maximum. An empirical relation between the modulation potential and the inverted neutron monitor count rate appears somewhat steeper than the modelled one, as confirmed also by data from fragmentary balloon-borne measurements. A reason for the discrepancy was unclear on the moment of publication.

In Paper II the “effective” rigidity of a neutron monitor for a GLE event is defined for the first time so that the event-integrated fluence of solar energetic protons with rigidity above it is directly proportional to the integral intensity of the GLE as recorded by a polar neutron monitor, within a wide range of solar energetic-proton spectra. This approach provides a direct way to assess the integral fluence of a GLE event based solely on neutron-monitor data. The effective rigidity/energy for polar sea-level NM was found to be 1.13–1.42 GV (550–800 MeV). A small model-dependent, systematic uncertainty in the value of the effective rigidity is caused by uncertainties in the low-energy range of the neutron-monitor yield function.

In Paper III the spectra of protons and helium directly measured in space by the AMS-02 experiment for the period 2011–2017 were used to calibrate ground-based NMs. Calibration of several stable sea level NMs (Inuvik, Apatity, Oulu, Newark, Moscow, Hermanus, and Athens) was performed using these spectra. Four modern NM yield functions were verified: Mi13 [Mishev et al., 2013], Ma16 [Mangeard et al., 2016], CM12 [Caballero-Lopez and Moraal, 2012] and CD00 [Clem and Dorman, 2000]. The Mi13 yield function was found to realistically represent the NM response to galactic cosmic rays. The use of the CM12 yield function leads to a small skew in the solar cycle dependence of the scaling factor, which represents
the ratio between calculated and expected NM count rates and is due to local features of the given NM. On the other hand, Ma16 and CD00 yield functions tend to overestimate the NM sensitivity to low-rigidity (<10 GV) cosmic rays, that may be important for an analysis of ground level enhancements, leading to a potential underestimate of fluxes of solar energetic particles as based on NM data. The Mi13 yield function is recommended for quantitative analyses of NM data, especially for ground level enhancements. The validity of the force-field approximation was also studied, and it was found that it fits well the directly measured proton spectra, within a few percent for periods of low to moderate activity and up to ≈ 10% for active periods. The results of this work strengthen and validate the method of the cosmic ray variability analysis based on the NM data and yield function formalism and improve its accuracy. Moreover, the difference in the empirical relation between the modulation potential and the inverted neutron monitor count rate, which was observed in Paper I, was explained by the incorrect consideration of heavy nuclei, and a new semi-empirical model was presented to correct this effect.

In Paper IV a new method to reconstruct the high-rigidity part (≥ 1 GV) of the spectral fluence of SEP for GLE events, based on the world-wide NM network data, is presented. The method is based on the effective rigidity $R_{\text{eff}}$ and scaling factor $K_{\text{eff}}$ definitions, which were introduced in Paper II. In contrast to many other parametric methods, based on derivation of the best fit parameters of a prescribed spectral shape, it provides a true non-parametric estimate of the fluence. Reconstruction of the SEP fluences for two recent GLE events, #69 (20 Jan. 2005) and #71 (17 May 2012), using four NM yield functions: Mi13, Ma16, CM12, CD00, was presented. The results were compared with the full reconstructions and direct measurements by the PAMELA instrument. While reconstructions based on Mi13 and CM12 yield functions are consistent with the measurements, those based on CD00 and Ma16 ones underestimate the fluence by a factor of 2–3. It is also shown that the often used power-law approximation of the high-energy tail of SEP spectrum does not properly describe the GLE spectrum in the NM-energy range. Therefore, the earlier estimates of GLE integral fluences need to be revised.
References

Adriani, O., et al., Observations of the 2006 December 13 and 14 solar particle events in the 80MeV-1-3 GeV n-1 range from space with the PAMELA detector, Astrophysical Journal, 742, 102, 2011a.


Raukunen, O., et al., Two solar proton fluence models based on ground level enhancement observations, Journal of Space Weather and Space Climate, 8, A04, 2018.


Richardson, I. G., Solar wind stream interaction regions throughout the heliosphere, 2018.


Wieler, R., J. Beer, and I. Leya, The galactic cosmic ray intensity over the past 10 6 -10 9 years as recorded by cosmogenic nuclides in meteorites and terrestrial samples, Space Science Reviews, 176, 351–363, 2013.


6. Original publications