Cosmic rays: bringing messages from the sky to the Earth’s surface
Raios cósomicos: trazendo mensagens do céu para a superfície da Terra

Alessandra Abe Pacini
University of Vale do Paraíba, São José dos Campos, SP, Brazil


All the evolution of the scientific knowledge about Cosmic Rays (CRs) in the 20th century leads to the following statement: Cosmic Rays constitutes a time and spatial variable flux of fully ionized energetic atoms and sub-atomic particles with specific abundances and energies that continuously permeates the solar system, reaching the Earth and producing ionization and nuclear reactions in our atmosphere. Since the primary CR flux is composed of charged particles, it is expected that any magnetic irregularity and/or turbulence affects their large scale transport and consequently its energy spectrum. Moreover, since the interaction of the CR particles with the terrestrial atmosphere produces not only atmospheric ionization, but also nuclear reactions (creating a cascade of secondary particles), it is easy to understand that near ground CR measurements (direct or indirect) may allow us to study the physical processes involved in the journey of CR particles (including the birth conditions of the lower energetic ones and the propagation of the higher energetic ones through the heliosphere) before reaching the terrestrial environment, and also along the transport of the secondary particles into the Earth’s atmosphere. In this sense, considering a cosmic ray particle as a “celestial messenger” is reasonable. Therefore, several works have been using the CR footprints found in the terrestrial surface (through the ground-based measurement of cosmogenic tracers) as a tool for solar and atmospheric studies. It is present here a discussion concerning the time evolution of the CR scientific knowledge and a brief review of the physical concepts of those celestial messengers.

Keywords: Cosmic Rays, ground-based measurements, celestial messengers.

1. Historical overview

In the beginning of 20th century, the scientific community was intrigued by the discovery of the perma-
nent atmospheric ionization. The main question was related to the source of such ionization and the most acceptable explanation at that time linked its origin with the natural radioactivity (mostly manifested as gamma-rays photons) presented in certain waters, soils and salt deposits. Still in the first decade of the last century, the improvement of the instrumental techniques permitted new kinds of experiments, and the effect of this atmospheric ionization was measured by more sensitive electrosopes in high altitudes, on board of balloons or under the sea level. All different experiments suggested that the air ionization has an altitude dependence and its causing radiation has a penetrating component. Besides those evidences, the dominant opinion about this radiation origin didn’t change so easily, and for a decade it was attributed to radioactive materials in the Earth’s crust [1].

Only in the year of 1912, Domenico Pacini (in February) first considered that “a sizeable cause of ionization exists in the atmosphere, originating from penetrating radiation, independent of the direct action of radioactive substances in the soil”, and Victor Hess (in August) demonstrated by balloon measurements that there was a new source of such radiation from above. After Hess conclusion (that the origin of the measured ionization was extra-terrestrial), the term ”cosmic radiation” became common.

With the end of the World War I, the number of investigations concerning the cosmic radiation grew up in the United States, and Robert Millikan start using the term “Millikan rays” or “cosmic rays” (CRs) to designate the source of the air ionization. The Fig. 1, extract from an article published in the American journal “Popular Science Monthly” in 1928 [2] shows the electromagnetic spectrum as it was understood, still considering the cosmic rays as radiation. Only in 1929, Walter Bothe and Werner Kolhörster, in Germany, developed a coincidence experiment that indicate the corpuscular nature of the cosmic rays [3].

Although the discovery of CRs was the result of different experiments, Hess was awarded by the 1936 Nobel Prize in Physics for been the “first to ascribe with confidence this increased ionization to radiation coming from outside the Earth”, as Arthur Compton said on the occasion, and is known as the father of CR research [4]. Many other scientists of 20th century had focused their researches to better understand the nature of CRs and due to the results of their experiments and studies, incredible new areas of elementary particle physics, space physics, geophysics and astrophysics were opened.

Among the considered pioneers, it is worth to mention: Werner Kolhörster (who first proposed a possible latitude dependence of CR, in 1919); Jacob Clay (who measured such geomagnetic dependence of the CR flux in 1927-28); Arthur Compton (who measured in 1931-33 the intensity of the atmospheric ionization around the world, characterizing its geomagnetic dependence, and performed an experiment that measured the East-West effect); Luis Alvarez, Thomas Johnson and Bruno Rossi (who also found the East-West effect in three independent experiments developed between 1931 and 1934, showing that most primary cosmic rays are positively charged.

Figure 1: The electromagnetic spectrum as it was known in the beginning of last century, including the “Millikan rays” or “cosmic rays”. It was characterized as a conceptual mistake years later, when the nature of CR as particles, and not “rays” (radiation) was understood (Extract from [2]).

Revista Brasileira de Ensino de Física, vol. 39, nº 1, e1306, 2017
DOI: http://dx.doi.org/10.1590/1806-9126-RBEF-2016-0168
particles - protons); Carl Anderson (who discovered positrons in 1933, and with the collaboration of Seth Neddermeyer, also discovered the muons in 1937); Pierre Auger (discovering the EAS, or atmospheric cascade, in 1938); Cesar Lattes, Giuseppe Occhialini and Cecil Powel (who discovered the pions in 1947, whose existence was predicted by Hideki Yukawa in 1935); Enrico Fermi (who proposed a theory explaining the mechanisms of CR acceleration, in 1949); John Linsley (who first observed a CR at highest energies, in 1962), and many others.

The Fig. 2 shows the portraits of two of the pioneers in this scientific area (Domenico Pacini, in the left and Victor Hess, in the right) representing all the scientists who dedicated their lives, with or without public recognition, for the exciting investigation concerning the nature and effects of cosmic rays.

After the cosmic rays discovery, there were more than 100 years of great advances in the field, due to the improvement of ground- and space-based instruments and numerical and theoretical models. Moreover, the association of a modern modelling of the Extensive Air Showers, EAS, with new measurement techniques have been providing new insights about the nature of the most energetic CRs detected on Earth [5]. There are still some big questions to be solved (like the astrophysical origin of the highest energy component and the turnovers in its energy spectrum) and besides the development of particle accelerators, the role of CR studies in nuclear physics is still important.

2. The nature of cosmic rays

All the evolution of the scientific knowledge about CRs in the last century leads to the following statement: Cosmic Rays constitutes a time and spatial variable flux of fully ionized energetic atoms and sub-atomic particles with specific abundances and energies that continuously permeates the solar system, reaching the Earth and producing ionization and nuclear reactions in the terrestrial atmosphere. The flux of CR particles near Earth is strongly dependent on their energies, which varies from hundreds of MeV up to $10^{21}$eV. There are many possible mechanisms that can contribute to the energy of those particles (Fermi shocks acceleration, magnetic reconnection, stochastic heating among others. For a good review, see [6]). Those natural particle acceleration processes can take place far in the Galaxy (at supernovae remnants or active galactic nuclei), at the boundaries of the region dominated by the solar magnetic field - heliosphere - (particularly at its termination shock and at the heliopause [7]) or even closer to the Earth, in the solar atmosphere and in the interplanetary medium.

When the acceleration process is associated with solar flares and/or Coronal Mass Ejections (CMEs), the CR particles released to the interplanetary medium are called Solar Cosmic Rays (SCRs) or Solar Energetic Particles (SEPs). The flux of SEPs is strongly variable (as well as its composition) but its energies are generally below 1 GeV (rarely reaching 10 GeV). The CRs with energies above 10 GeV measured near Earth are originated from outside of the Solar Sys-

Figure 2: Pictures of two pioneers of the Cosmic Rays Sciences: in the left, Domenico Pacini (1910) and, in the right, Victor Hess (1936), who made the scientific community to raise their eyes to the sky.
tem, either in our Galaxy or outside it and are called Galactic Cosmic Rays (GCRs). Both GCR and SEP compositions have a major contribution of protons and $\alpha$-particles with a smaller, but important, contribution of heavier nuclei. The goal of this section is present an overview concerning the GCR features observed in the Earth’s environment in order to help the reader to understand our analysis, thus from here on, the SEPs will be neglected.

Fig. 3 presents the mean relative abundances measured in the GCR flux by CRIS instrument on board of ACE satellite, compared with the Solar System abundances.

One can see that the hydrogen and helium nuclei are a few orders of magnitude more abundant than any other element in the GCR flux. Moreover, there is a significant excess of some nuclei (specially the light ones Li, Be and B) in the GCR flux compared with the solar system abundances. This excess is due to the collisions suffered by the GCR particles with those of interstellar medium, which produces thus isotopes of such nuclei. Accordingly, these elements are not primary (accelerated at the source, e.g., at supernovae) but rather secondary (produced along GCR transport in galaxy). Fig. 4 shows the compilation of several measurements of the GCR energy spectrum (updated from the famous Swordy plot first presented in [8]): in the left panel, in a broad energy range (from $10^8$ up to $10^{21}$ eV), where the main features of the spectrum are clear and, in the right panel, a zoom-in on the ultra high energy range (above $10^{17}$ eV), evidencing the more recent features observed.

One can see in the left panel that between ~ tens of GeV and ~ $10^{15}$ eV, the GCR energy spectrum can be approximated by a power law with a negative spectral index. Around the energy of about $10^{15}$ eV the spectrum becomes steeper (softer), defining the "knee" in the spectrum. It is still not clear if this change is related to the galactic diffusion or to the decrease of efficiency of the supernova acceleration mechanism for higher energies. It is generally assumed that at energies above ~ $10^{17} - 19$ eV there is a domination of extragalactic CR particles over the galactic ones, and it seems to justify the hardening of the GCR spectrum which define the "ankle" (see e.g. [9]). For energies above the ankle, a significant suppression of the GCR flux is expected due to its interaction with the Cosmic Microwave Background (knowing as Greisen - Zatsepin - Kuz’min effect (GZK effect), see e.g. [10]). Recent experiments on the higher energy GCR range have been finding strong evidences supporting the existence of the GZK effect and are providing new insights of the nature of the ultra high energetic CR particles [5]. Below energies about 30 GeV, the GCR spectrum clearly not follow a power law. The reason is the solar modulation affecting such GCR particles during their passage through the heliosphere, which affects the propagation of low energies GCR. The next section of this paper briefly describes the modulation suffered by the CR flux in the heliosphere and terrestrial environment.

3. The journey of cosmic ray

3.1. In the heliosphere

The heliospheric (or solar) modulation is related to the Interplanetary Magnetic Field (IMF) and solar wind conditions and varies according to the solar magnetic cycles. The standard theory of charged particles propagation in the Heliosphere was developed by [11] and its famous basic transport equation describes the temporal variations of the GCR flux as a combination of four main modulation processes: convection (due to the solar wind), drift (due to the curvature and gradient of IMF), diffusion (due to fluctuations of the IMF strength) and the adiabatic energy losses. A good review concerning the physical basis of all the four solar modulation processes and
the state of art of the current theoretical models can be found in [12] and references therein.

The most important interplanetary structure in the drift-dominated scenario (that usually take place during solar minima periods) is the Heliospheric Current Sheet - HCS (also known as heliospheric neutral sheet), which is the thin surface that separates the two hemispheres where the IMF points toward or away from the Sun. The generally wave form of the HCS, resulting of the dipolar solar magnetic field lines by the solar wind into the heliosphere, is known as “ballerina skirt”. When this Archimedean spiral field (Parker spiral) is outward (or inward) in the northern hemisphere, the heliospheric polarity scenario $A > 0$ (or $A < 0$) is defined, being switched after $\sim 11$ years. Thus, for solar minimum periods with $A < 0$ conditions, the GCR protons drift path so long that diffusion becomes the more effective transport mechanism. For $A > 0$ cycles, when the positively charged particles are drifting inward in the polar region, the large tilt of the HCS in solar maximum years induces a high energy losses. Moreover, the presence of transient and propagating disturbances in the heliosphere (specially the Global Merged Interaction Regions - GMIRs) creates a more diffusive environment for the GCR propagation and reduces the flux of GCR reaching the Earth [13]. For periods with moderate solar activity, both drift and diffusion mechanisms are important for the GCR modulation.

A good parameter to represent the GCR heliospheric modulation is the modulation potential, $\Phi$ in [MV] [14]. $\Phi$, also known as the “force-field parameter”, is the mean rigidity (momentum per charge) loss of a GCR particle inside the heliosphere without specifying the mechanism and provides a very useful and accurate formal parametrization of the GCR spectrum near Earth, described by a single parameter [14]. Fig. 5 shows the temporal variation of the solar activity (represented by its most traditional index: the number of sunspots in the solar photosphere) along with the modulation po-

Figure 4: Cosmic ray spectrum obtained by a compilation of data from different instruments, available at [www.physics.utah.edu/~whanlon/spectrum.html]
Figure 5: Temporal evolution of the sunspot number, the heliospheric modulation potential ($\Phi$) and the GCR flux reaching the Earth in a polar region (represented by the ground-based measurements of a neutron monitor located at Oulu, Finland) during the solar cycles 21-24.

The 11-yr cycle observed in the GCR measurements (anti-correlated to the solar cycle) is due to the more complicated IMF configuration and more turbulent solar wind, which causes a large scattering of the GCR particles during solar maxima periods. Some other solar features can be noted in a higher time resolution data as short-term modulations, and they are mainly related to the solar rotation, flares or CMEs. The 22-yr Hale cycle (seen as the alternation of sharp and at peaks in the GCR data during solar minima periods in Fig. 5) is a result of the 22-yr polarity reversal of the solar magnetic field (periods of $A < 0$ are indicated by the red box) [12].

The heliospheric modulation is much more important for low energetic GCR particles, while it produces only a few percent changes in the flux of higher energetic GCRs (above 10 GeV). For periods of solar minimum with positive polarities, the GCR spectrum is usually higher than the ones obtained for equivalent conditions during $A < 0$ periods.

The role of drift as the dominant solar modulation in the inner heliosphere (< 10 AU) during solar minima periods (and the diffusion dominating the solar maxima modulation) had been verified by the observations (ground- and spacebased ones) for the last solar cycles and became an accepted conceptual paradigm, until the beginning of the 24th solar cycle, when the scenario changed dramatically [15]. In fact, the solar modulation scenario looks different in the last solar cycle, and it is a challenge to understand the relative importance of the four mechanisms associated to GCR modulation during a very weak solar cycle. It is, e.g., suggested that during periods of very weak solar activity (like the Maunder Minimum), the relation between solar activity and cosmic rays may be inverted [16]. The new features found at the heliospheric modulation of GCR measured by ground-based detectors are discussed in [17] and references therein.
3.2. In the earth’s environment

After been modulated by the heliosphere, the CR flux reaches the Earth almost isotropically. However, before eventually reaching the ground, the CR particles will still find two main obstacles to overcome: the **geomagnetic field** and the **terrestrial atmosphere**.

The **geomagnetic obstacle** can be roughly described by a dipole magnetic field, with the magnetic poles in the opposite orientation to the geographic ones and inclined about 22 degrees from the geographical axis. This configuration exists nowadays, but it may change quite essentially with time. The CR charged particles are affected by Lorentz forces due to the geomagnetic field when they reach the vicinity of the Earth. Although the geomagnetic field is not strong (it is less than 1 \( \mu \)T on the surface, decreasing with the cube of the distance in outer space), it is spatially huge, changing the trajectories of high-energy charged particles considerably. The flux reaching the equatorial regions is affected by the perpendicular geomagnetic lines, creating a sign-dependent drift: positively charged particles will drift from the West to the East, and the opposite for the negative charged particles.

The descriptions of the energetic particles trajectories in a dipolar geomagnetic field become possible after the famous theoretical studies of auroral particles motion from Carl Størmer in the beginning of 20th century (before the discovery of the CRs). In his approach to investigate the transport of auroral particles in the geomagnetic field, Størmer realized that it was easier to trace the trajectory of electrons from the Earth towards the Sun, and not on the other way round. In the times when computers were not available, the calculation of each trajectory (by hand) was an impressive work (which involved, of course, many students and collaborators), summarized at his book [18]. One of the important conclusions raised from Størmer’s calculations is that there are some allowed trajectories to be taken for a charged particle in the geomagnetic field, and others forbidden. He concluded that if a positive charged particle is entering the geomagnetic field at any geographical position, there is a cone of forbidden trajectories with the axis pointing toward the east, predicting the eastwest effect [1].

The modern description of a GCR particle trajectory is, of course, much more complex (since GCR flux is arriving to the Earth from all directions and have a much higher energy) and considers a much more realistic model of the geomagnetic field (considering the internal and external variabilities). Fig. 6 shows (in the left side) some Størmer’s calculated trajectories for electrons represented with copper wires in an incredible picture (seen at [19], pag 81) and, (in the right side) some back-trajectories of positively charged particles (CRs) with different rigidities leaving the Earth from a location 20 km altitude above Switzerland [20] computed by a modern backtracing codes (e.g. CORSIKA, GEANT-4), which combines a full description of the internal geomagnetic field and external magnetospheric field.

One can see that that, for the specific location, the lower is the CR magnetic rigidity, the more complicated is the particle’s movement in the geomagnetic

![Figure 6: Some back-trajectories of solar electrons reaching the Earth in different locations calculated by Størmer and collaborators in the beginning of 20th century and represented by copper wires (left side) and some back-trajectories of CR reaching the Earth above Switzerland with different rigidities computed by modern codes (left side). (Adapted from [19] and [20]).](image-url)
field. Moreover, there is a critical rigidity (in this case, 4.52 GV) below which a CR can not reach the location coming from the outer space. The minimum rigidity that a CR particle must have in order to overpass the geomagnetic blocking and interact with the terrestrial atmosphere at an specific latitude/longitude/altitude position defines the cutoff rigidity (Pc) value [21]. Pc values can be easily calculated for a centered or eccentric dipole geomagnetic model as shown in the Appendix of [22]. Due to the geomagnetic pole migration, the cutoff rigidity from low latitudes have been changing notably during the last decades. [21] computed the globally distribution of the Pc values and the absolute changes occurred during one century Fig. 7.

It is also important to calculate the allowed cones and the asymptotic directions of the primary GCR for different ground-based stations [23]. An asymptotic direction of a CR particle represents the direction of its motion before entering into the magnetosphere.

Once the GCR charged particle has overcome the geomagnetic shielding, it suffers another obstacle before possibly reaching the ground level: the terrestrial atmosphere. The atmospheric constituents play a role of target for this flux of energetic particles, and nuclear collisions between the GCRs and the atmospheric gas particles are very probable. The altitude of the first interaction roughly corresponds to the lower stratosphere. Below this region, the atmospheric column density increases exponentially, increasing the probability of nuclear collision and interaction as well.

There is an "atmospheric cutoff" (∼ 500 MeV/nucleon) corresponding to the lowest energy that a CR particle must have in order to produce nuclear reactions during its interaction with air constituents [20]. After the first interaction, CR particles suffer a series of successive interactions creating a shower of secondary particles and nuclides. The cosmogenic products of these interactions define the cosmic ray cascade (or extensive air showers - EAS), and corresponds to the main source of air ionization below 40 km.

The first interaction of the primary GCR particle with the nucleus of an atmospheric atom is usually a nuclear spallation, sputtering off nucleons from the target nucleus and leaving behind a lighter nucleus. The sputtered nucleons (protons, neutrons and mesons, mainly pions: π) start a nuclear cascade in the atmosphere. As it is basically shown in the scheme of Fig. 8, a neutral pion (π⁰) decays almost instantaneously (in about 10⁻¹⁶ s) creating photons, electrons and positrons. The charged pions (π±) have longer lifetime (∼ 10⁻⁸ s) relative to the neutral ones and decay into a muon (µ±) and neutrino.

Therefore, the CR cascade possesses three main components, which are briefly described below:

- Electromagnetic component (also called "soft" component) is basically formed by electromagnetic gamma-ray, electron and positrons resulting from the decay of π⁰ and µ±.
- Mesonic component (also called "hard" component) is formed by muons from the decay of the µ±. Muons are formed in the stratosphere and have a half-life of 2.2 µs, which is smaller than their time of flight in the atmosphere (40 µs) but they can be measured on the ground-level (even under-ground) due to the effects of time dilatation caused by its relativistic speed. During its way towards the ground, the muons are weakly interacting with the atmo-
spheric matter (producing mainly ionization) and losing about 2 GeV of its energy [25].

- Nucleonic component, in the ground level, is \( \sim 98\% \) composed by secondary neutrons from the cascade, whose number density is larger than that of the protons, because they do not suffer ionization during their transport in the atmosphere.

The sub-atomic particles from the cascade can also induces nuclear reactions (mainly by spallation, but also by thermal neutron or by negative muon captures) depending on their energy. The new nuclei formed by the interaction of primary or secondary cosmic ray particles with the atmosphere or terrestrial targets are called Cosmogenic Nuclides. These cosmogenic nuclides can be both stable (e.g. helium-3) and radioactive (e.g. the radioactive isotopes of the elements aluminium, chlorine, calcium, iodine and beryllium). After their production in the atmosphere, the cosmogenic nuclides can be attached to aerosols or react with atmospheric components, following the correspondent cycle before reach the Earth’s surface. Finally, they are stored in different natural matrices (e.g. icecores, speleotems, sediments and tree rings. More details in [24]).

Radioactive environmental techniques for tracing large-scale air-mass transport have been applied in studies of atmospheric dynamics for decades (see, for example, [26]), and they are becoming more and more precise due to the improvement of the instrumental sensitivity and associated modeling. Measurements of cosmogenic isotopes in the near-ground air can provide information on physical processes involved in their production and transport (e.g., for studies of air dynamics using beryllium isotopes, see [27]; [28]; [29]; [30] and [31]).

**The measurements of the flux of secondary GCR particles from the cascade or the cosmogenic isotope concentration in the near-ground air provides the access to the information concerning the different modulation**

---

DOI: http://dx.doi.org/10.1590/1806-9126-RBEF-2016-0168
processes suffered by the primary GCR during their transport in the heliosphere, magnetosphere and atmosphere, and also the journey of the secondary particle. To be able to extract all these information from the cosmogenic tracers measured in the terrestrial surface, it is necessary to understand its modulations sources. The use of CR footprints found in the terrestrial surface (through the ground-based measurement of cosmogenic tracers) as a tool for solar, climate and atmospheric studies is not completely established, leaving plenty of room for new studies in this exciting area of the Space Physics and Atmospheric Sciences.

Acknowledgment

The author thanks CNPq Brazilian agency for financial support (proc. 47843/2013-7).

References