The Space Radiation Environment

Sébastien Bourdarie, and Michael Xapsos, Senior Member, IEEE

Abstract—The effects of the space radiation environment on spacecraft systems and instruments are significant design considerations for space missions. Astronaut exposure is a serious concern for manned missions. In order to meet these challenges and have reliable, cost-effective designs, the radiation environment must be understood and accurately modeled. The nature of the environment varies greatly between low earth orbits, higher earth orbits and interplanetary space. There are both short-term and long-term variations with the phase of the solar cycle. In this paper we concentrate mainly on charged particle radiations. Descriptions of the radiation belts and particles of solar and cosmic origin are reviewed. An overview of the traditional models is presented accompanied by their application areas and limitations. This is followed by discussion of some recent model developments.

Index Terms—galactic cosmic rays, solar particle event, trapped protons, trapped electrons

I. INTRODUCTION

Before the space era the only manifestations of the presence of radiations in space were the deformation of the ionized tail of comets caused by the solar wind, the aurora borealis whose origin was not well understood and the ionization of air, secondary cosmic-ray showers and the isotopes (carbon-14 for example) produced by cosmic radiation. It was not until shortly after the beginning of the satellite era in 1958 that the presence of high energy charged particles around the earth (Van Allen belts) was discovered. Since then, it has become evident that the space environment is a highly disruptive medium for space missions. Beyond the natural protection provided by the Earth's atmosphere, various types of radiation can be encountered. Their characteristics (energy and nature), origins and distributions in space are extremely variable. This environment degrades electronic systems and on-board equipment in particular and creates radiobiological hazards during manned space flights.

Based on several tens of years of this space era, a detailed analysis of the problems on satellites shows that the part due to the radiation environment is significant. It appears that the malfunctions are due to problems linked to the space environment (9 to 21%), electronic problems (6 to 16%), design problems (11 to 25%), quality problems (1 to 8%), other problems (11 to 33%) and problems that are still unexplained (19 to 53%) [1]. It is clear that the unexplained problems are either problems linked to the space environment, to the electronics, to the design, or otherwise but the information collected on the ground is generally/often not sufficient to define the origin of the problem. The space environment is largely responsible for about 20% of the anomalies occurring on satellites and a better knowledge of that environment could only increase the average lifetime of space vehicles.

This naturally leads to a detailed study of the space environment and of the effects that it induces on space vehicles and astronauts. The nature of this environment varies greatly between low orbits and the higher altitudes such as the geostationary orbit and beyond. Among its components, we only examine the ionizing charged particles here, i.e. the particles trapped in the radiation belts, the solar particle event and galactic cosmic radiation particles. From the point of view of the effects, the degradations will differ according to the energy of the particles, to their nature and to the satellite orbit. The degradations and disturbances induced by space radiation in the materials and the electronic components are phenomena that have been studied for many years. Two categories of effects should be noted:

- the cumulative effects such as the degradation of thermal control coatings, optics and electronics and the erosion of materials;
- the sporadic effects such as noises in the detectors and optics, single event effects in highly integrated electronic circuits and electrostatic discharges.

The increasingly frequent presence of man in space and the projects such as the distant and long-duration missions (Lunar base, flight to Mars, etc.) pose a problem in induced biological effects, essentially in the long term, by high energy radiation. Radiation can have two possible types of biological effects:

- immediate, permanent or delayed non stochastic effects (destruction or modification of cells), the speed with which the symptoms appear and their seriousness increase in proportion to the exposure to the radiation;
- stochastic effects, associated with the modifications to the cells, whose probability of appearing in the long term increases in proportion to the irradiation (cancers, leukemia, genetic effects).
The sun is both a source and a modulator of space radiations. Understanding its cyclical activity is an important aspect of modeling the space radiation environment. The solar activity cycle is approximately 11 years long [2]. During this period there are typically 7 years during solar maximum when activity levels are high and 4 years during solar minimum when activity levels are low. In reality the transition between solar maximum and solar minimum is a continuous one but it is often considered to be abrupt for convenience. At the end of each 11-year cycle the magnetic polarity of the sun reverses and another 11-year cycle follows. Thus, strictly speaking the total activity cycle is approximately 22 years long. Of the space radiations considered here the magnetic polarity apparently only affects the galactic cosmic ray fluxes and not the trapped particle or solar particle event fluxes. Thus, things are often viewed on an approximately 11-year cyclical basis.

Two common indicators of this approximately 11-year periodic solar activity are sunspot numbers and solar 10.7 cm radio flux (F10.7). The most extensive record is that of observed sunspot numbers, which dates back to the 1600s. This record is shown in Fig. 1. The numbering of sunspot cycles began in 1749 and it is currently near the end of solar cycle 23. The record of F10.7 began part way through solar cycle 18 in the year 1947 and is shown in Fig. 2.

Although sunspot numbers and F10.7 are commonly accepted indicators of solar activity, quantitative relations to measured radiation events and fluxes are not necessarily straightforward. Large solar particle events are known to occur with greater frequency during the declining phase of solar maximum [3]. Trapped electron fluxes also tend to be higher during the declining phase [4]. Trapped proton fluxes in low earth orbit (LEO) reach their maximum during solar minimum but exactly when this peak is reached depends on the particular location [5]. Galactic cosmic ray fluxes are also at a maximum during solar minimum but in addition depend on the magnetic polarity of the sun [6].

There has been considerable effort put into forecasting long-term solar cycle activity. A review of a number of the methods is presented by Hathaway [7]. These include regression methods, which involve fitting a function to the data as the cycle develops. Also discussed are precursor methods, which estimate the amplitude of the next cycle based on some type of correlation with prior information. These methods can also be combined. In addition, physically based methods are being developed based on the structure of the magnetic field within the sun and heliosphere [8,9].

However, accurate methods for predicting future solar cycle activity levels prior to the start of the cycle have thus far been elusive. A potential breakthrough, however, has recently been reported that uses a combination of computer simulation and observations of the solar interior from instrumentation onboard the Solar and Heliospheric Observatory (SOHO) [9]. Given the current state of this modeling, probabilistic models of solar activity are useful. Such a model of F10.7 is shown in Fig. 3 [10]. This also illustrates the general behavior of the observed cyclical properties, at least over recent cycles. The greater the peak activity of a cycle, the faster the rise-time to the peak level. Furthermore the cyclical activity is asymmetric such that the descending phase of the cycle is longer than the ascending phase.

III. GALACTIC COSMIC RAYS

A. General Characteristics

Galactic cosmic rays (GCR) are high-energy charged particles that originate outside of our solar system. Although
their origin is not completely understood it is believed that supernova explosions are a significant source. Some general characteristics of GCR are listed in Table 1. They are composed mainly of hadrons, the abundances of which are listed in the Table. A more detailed look at the relative abundances is shown in Fig. 4. All naturally occurring elements in the Periodic Table (up through uranium) are present in GCR, although there is a steep drop-off for atomic numbers higher than iron (Z=26). Energies can be as high as \(10^{11}\) GeV, although the acceleration mechanisms to reach such high energies are not understood. Fluxes are generally a few \(\text{cm}^{-2}\text{s}^{-1}\), and vary with the solar cycle. Typical GCR energy spectra for a few of the major elements during solar maximum and solar minimum are shown in Fig. 5. It is seen the spectra tend to peak around 1 GeV per nucleon. The flux of the ions with energies less than about 10 GeV per nucleon is modulated by the magnetic field in the sun and solar wind. During the high activity solar maximum period there is significantly more attenuation of the flux, resulting in the spectral shapes shown in Fig. 5.

Single Event Effects (SEE) are the main radiation effects caused by GCR in microelectronics and photonics. The metric traditionally used to describe heavy ion induced SEE is linear energy transfer (LET). LET is the energy lost by the ionizing particle per unit path length in the sensitive volume. For SEE analyses energy spectra such as those shown in Fig. 5 can be converted to LET spectra. Such integral LET spectra for solar maximum and solar minimum conditions are shown in Fig. 6. These spectra include all elements from protons up through uranium. The ordinate gives the flux of particles that have an LET greater than the corresponding value shown on the abscissa. Given the dimensions of the sensitive volume this allows the flux of particles that deposit a given amount of charge or greater to be calculated in a simple approximation. In some modern devices, the LET metric may not be useful because of the highly scaled nature of devices, the complexity of the geometry, or the increased use of metal over-layers. In addition, if nuclear reactions play a significant role in producing SEE, the LET metric is not valid for this situation. Exposure of astronauts to GCR is a serious consideration for manned missions because GCR are difficult to shield against given the typical weight constraints of missions.

### Table 1. Characteristics of Galactic Cosmic Rays.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>87% protons</td>
<td>Up to (10^{11}) GeV</td>
<td>1 to 10 (\text{cm}^{-2}\text{s}^{-1})</td>
<td>SEE</td>
<td>LET</td>
</tr>
<tr>
<td>12% alphas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% heavier ions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5. GCR energy spectra for protons, helium, oxygen and iron during solar maximum and solar minimum conditions [6].

Fig. 4. Abundances of GCR up through Z = 28.
The LET spectra shown in Fig. 6 are applicable to geosynchronous and interplanetary missions where there is no geomagnetic attenuation. The earth’s magnetic field, however, provides significant protection. Due to the basic interaction of charged particles with a magnetic field, the charged particles tend to follow the geomagnetic field lines. Near the equator the field lines tend to be parallel to the earth’s surface. Thus all but the most energetic ions are deflected away. In the polar regions the field lines tend to point toward the earth’s surface, which allows much deeper penetration of the incident ions. The effect of the geomagnetic field on the incident GCR LET spectrum during solar minimum is discussed for various orbits in [12].

B. Galactic Cosmic Ray Models

The original Cosmic Ray Effects in MicroElectronics (CREME) suite of programs of Adams [13] was developed specifically for microelectronics applications. It turned out to be a very useful and popular tool and has been updated since then. CREME96 is the current version [14] and uses the GCR model of Moscow State University (MSU) [15]. Recent work has begun on a subsequent update [16].

In principle the MSU model is similar in approach to a GCR model that was originated independently at NASA by Badhwar and O’Neill [6]. Both models are based on the diffusion-convection theory of solar modulation [17]. This is used to describe the penetration of cosmic rays into the heliosphere from outside and their transport to near earth at 1 Astronomical Unit (AU). The solar modulation is used as a basis to describe the variation of GCR energy spectra over the solar cycle, as shown in Fig. 5. However, the implementation of the solar modulation theory for the two models is different. The Badhwar and O’Neill model estimates the modulation level from GCR measurements at 1 AU. Correlations to ground-based neutron monitor counting rates are then made to establish long-term predictive capability. The MSU model is not as direct but uses multi-parameter fits to ultimately relate solar cycle variations in GCR intensity to observed sunspot numbers.

Comparisons of the GCR proton and alpha particle spectra of the two models above plus that used in the QinetiQ Atmospheric Radiation Model (QARM) show discrepancies among all three models for narrow time ranges [18]. Examples of this are shown in Fig. 7 for protons. This is not surprising considering the details of the solar modulation implementation are different. However, similar predictions are seen for the total fluence over the course of a solar cycle.
A recent development led by the California Institute of Technology is to use a transport model of GCR through the galaxy preceding the penetration and subsequent transport in the heliosphere. [19]. During the initial propagation of GCR through the galaxy use is made of knowledge of astrophysical processes that determine the composition and energy spectra of GCR. Comparisons of the fitted model spectra to the ACE satellite measurements are shown in Fig. 9. The model spectra do not reflect solar modulation so the model is still a work in progress. The fitting is done to demonstrate that the model has the potential to closely reproduce the ACE measurements. The elements C and Fe are GCR primaries while B, Sc, Ti and V are GCR secondaries produced by fragmentation of primaries on interstellar H and He.

IV. SOLAR PARTICLE EVENTS

A. General Characteristics

It is believed that there are 2 categories of solar particle events and that each one accelerates particles in a distinct manner. Solar flares result when the localized energy storage in the coronal magnetic field becomes too great and causes a burst of energy to be released. They tend to be electron rich, last for hours, and have an unusually high $^3\text{He}$ content relative to $^4\text{He}$. A Coronal Mass Ejection (CME), on the other hand, is a large eruption of plasma (a gas of free ions and electrons) that drives a shock wave outward and accelerates particles. CMEs tend to be proton rich, last for days, and have a small $^3\text{He}$ content relative to $^4\text{He}$. A review article by Reames gives a detailed account of the many observed differences between solar flares and CMEs [21].

CMEs are the type of solar particle events that are responsible for the major disturbances in interplanetary space and the major geomagnetic disturbances at earth when they impact the magnetosphere. The total mass of ejected plasma in a CME is generally around $10^{15}$ to $10^{17}$ grams. Its speeds can vary from about 50 to 1200 km/s with an average speed of around 400 km/s. It can take anywhere from about 12 hours to a few days to reach the earth. Table 2 lists some further general characteristics of CMEs.
TABLE 2. CHARACTERISTICS OF CMEs.

<table>
<thead>
<tr>
<th>Hadron Composition</th>
<th>Energies</th>
<th>Integral Fluence (&gt;10MeV/nucleon)</th>
<th>Peak Flux (&gt;10MeV/nucleon)</th>
<th>Radii</th>
<th>TID</th>
<th>DD</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>96.4% protons</td>
<td></td>
<td>Up to ~10^16 cm^2</td>
<td>Up to ~10^5 cm^2 s^(-1)</td>
<td>TI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5% alphas</td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>~0.1% heavier ions</td>
<td></td>
<td></td>
<td></td>
<td>SE</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All naturally occurring chemical elements ranging from protons to uranium are present in solar particle events. They can cause permanent damage such as Total Ionizing Dose (TID) and Displacement Damage (DD) that is due mainly to the proton and possibly the alpha particle component. Just because the heavy ion content is a small percentage does not mean it can be ignored. Heavy ions, as well as protons and alpha particles in solar particle events, can cause both transient and permanent SEE.

Figs. 10 and 11 illustrate the periodic yet statistical nature of solar particle events. They are plots of the daily solar proton fluences measured by the Interplanetary Monitoring Platform-8 (IMP-8) and Geostationary Operational Environment Satellites (GOES) over an approximately 28 year period. Fig. 10 shows > 0.88 MeV fluences while Fig. 11 shows > 92.5 MeV fluences. The solar maximum and solar minimum time periods are shown in the Figs. to illustrate the dependence on solar cycle.

The available solar particle data that cover the largest period of time are for protons. Since the available solar heavy ion data are not nearly as extensive, solar proton models and solar heavy ion models will be discussed separately.

B. Solar Proton Models

The probabilistic methods applied to solar proton event data are described here. Section 1 discusses the probabilistic nature of events. Section 2 then describes the distribution of event magnitudes. Sections 3 and 4 describe modeling cumulative fluences over the course of a mission, and section 5 discusses worst-case events during a mission.

1) Self-Organized Criticality and the Probabilistic Nature of the Energy Release Process

Substantial efforts have been put into studies of the occurrence of solar particle events. One of the main goals is to find a reliable predictor of events. Despite this significant international effort, solar particle events can occur suddenly and without obvious warning. In addition to potential problems with electronic systems and instrumentation, this is an especially serious concern for new space initiatives that plan to send manned spacecraft to the moon, Mars or interplanetary space. Thus, there is strong motivation to develop predictive methods for solar particle events. It is hoped that the apparent stochastic character can be overcome and predictability achieved if precursor phenomena such as x-ray flares or magnetic topology signatures can be properly interpreted or if the underlying mechanisms are identified. Whether the nature of the energy release process for solar particle events is deterministic or stochastic is a very basic question. More specifically, the question is whether it is possible to predict the time of occurrence and magnitude of solar particle events or if probabilistic methods are necessary.

The self-organized criticality (SOC) model is a phenomenological model originated by Bak, Tang and Wisenfeld [22] that can give insight into the basic nature of a system. It postulates that a slow continuous build-up of energy in a large interactive system causes the system to evolve to a
critical state. A minor, localized disturbance can then start an energy-releasing chain reaction. Chain reactions and therefore energy releasing events of all sizes are an integral part of the dynamics, leading to a "scale invariant" property for event sizes. This scale invariance results in power function distributions for the density functions of event magnitudes and waiting times between events. As a result of this basic nature it is generally assumed in the literature that accurate predictions of the magnitude and time of occurrence of such events are not possible. A system in a SOC state is therefore generally assumed to be probabilistic in nature.

Applications for the theory of SOC have been found in natural phenomena such as earthquakes, avalanches and rainfall. It has recently been shown that the energy release due to solar particle events is consistent with the dynamics of a SOC system [23]. This was based on three analyses of 28 years of solar proton data taken by the IMP-8 and GOES series of satellites. The first was rescaled range (R/S) analysis, which was used to determine that events show long-term correlation. The second was a demonstration of fractal properties of event sizes, which suggests scale invariant behavior. The third was an analysis of the number density distribution of fluence magnitudes, which was shown to be a power function. These are hallmark features of systems that exhibit self-organized criticality.

The third of these analyses is a necessary characteristic of SOC phenomenon [24-26]. The number density distribution of monthly solar proton fluences for a 28-year period is shown in Fig. 12. The ordinate represents the number of occurrences when the monthly fluence exceeds that shown on the abscissa. It is seen that this distribution is a straight line on a semi-logarithmic plot that spans about 4 orders of magnitude. The number density function is

$$ \frac{dN}{d\Phi} = \frac{-29.4}{\Phi} $$

where $\Phi$ is the monthly fluence.

In this case the density function turns out to be exactly proportional to the reciprocal of the fluence. Thus, the solar event data can be represented by a power function of a type commonly referred to as $1/f$ [22]. It can therefore be viewed as $1/f$ noise, also known as flicker noise. It is well known that this type of noise results when the dynamics of a system is strongly influenced by past events. Thus, an especially compelling argument can be made that solar particle events are a SOC phenomenon.

The general behavior of a SOC system is that of a non-equilibrium system driven by a slow continuous energy input that is released in sudden bursts with no typical size as indicated by the power function distribution shown in equation (1). Although research involving SOC is still a developing field and there is much yet to be learned about the sun's dynamics [27-29], these results strongly suggest that it is not possible to predict that a solar particle event of a given magnitude will occur at a given time. It is therefore a reasonable approach to model solar particle events as a probabilistic phenomenon.

2) The Maximum Entropy Principle and the Distribution of Solar Proton Event Magnitudes

Given that the occurrence of solar particle events is a stochastic phenomenon, it is important to accurately model the distribution of event magnitudes. However, in general it can be rather difficult to select a probability distribution for the situation where the data are limited. There have been a number of empirical assumptions that the event magnitudes can be represented by certain distributions. For example, lognormal distributions [30,31] and power function distributions [32,33] have been used. The lognormal distribution describes the large events well but underestimates the probability of smaller events. On the other hand power functions describe the smaller events well but overestimate the probability of larger events.

The Maximum Entropy Principle was developed by E.T. Jaynes [34] using the concept of entropy originated by Shannon [35]. Jaynes showed in his studies of statistical mechanics that the distribution that should be selected is the one that maximizes the mathematical expression for entropy subject to the constraints imposed by available information. Choosing the distribution with the greatest entropy avoids the arbitrary introduction of information. Jaynes established the principle as a procedure for making an optimal selection of a probability distribution using limited data.

The probability distribution's entropy, $S$, is defined [34,36]

\[ S = - \int p(M) \ln(p(M)) \, dM \]  \hspace{1cm} (2)

where \( p(M) \) is the probability density of the random variable \( M \). For the case of solar particle event fluences, \( M \) is conveniently taken as the base 10 logarithm of the event fluence. A series of mathematical constraints are imposed upon the distribution, drawing from known information. In this case the constraints are [37]:

a) The distribution can be normalized.
b) The distribution has a well-defined mean.
c) The distribution has a known lower limit in the event fluence. This may correspond to a detection threshold, for example.
d) The distribution is bounded and consequently infinitely large events are not possible.

The resulting system of equations are used along with equation (2) to find the solution \( p(M) \) that maximizes \( S \). This has been worked out for many situations [36] and can also be solved using the LaGrange multiplier technique [38]. The following result for solar proton event fluences is obtained for the solar maximum time period:

\[ N = N_{\text{tot}} \left[ \frac{\phi^{-b} - \phi_{\text{max}}^{-b}}{\phi_{\text{min}}^{-b} - \phi_{\text{max}}^{-b}} \right] \]  \hspace{1cm} (3)

where \( N \) is the number of events per solar maximum year having a fluence greater than or equal to \( \phi \), \( N_{\text{tot}} \) is the total number of events per solar maximum year having a fluence greater than or equal to \( \phi_{\text{min}} \), \( -b \) is the index of the power function, and \( \phi_{\text{max}} \) is the maximum event fluence. Equation (3) is a truncated power function in the event fluence. It behaves like a power function with an index of \( -b \) for \( \phi << \phi_{\text{max}} \) and goes smoothly to zero at the upper limit \( \phi_{\text{max}} \).

Fig. 13 shows > 30 MeV solar proton event data compared to the best fit to equation (3). The data are from the 21 solar maximum years during solar cycles 20 – 22. It is seen that the probability distribution derived from the maximum entropy principle describes the data quite well over its entire range. This strong agreement indicates that this probability distribution captures the essential features of a solar proton event magnitude distribution. It is a power function for small event sizes and falls off rapidly for very large events. The interpretation of the maximum fluence parameter \( \phi_{\text{max}} \) is interesting in itself and will be discussed further in section 5.

3) Cumulative Fluence During Solar Maximum

During a space mission the solar particle event fluence that accumulates during the solar maximum time period is often the dominant contribution to the total fluence. Thus, much prior work focuses on this period of the solar cycle. A solar cycle typically lasts about 11 years. A commonly used definition of the solar maximum period is the 7-year period that spans a starting point 2.5 years before and an ending point 4.5 years after a time defined by the maximum sunspot number in the cycle [39]. The remainder of the cycle is considered solar minimum.

Once the initial or underlying distribution of event sizes during solar maximum such as that shown in Fig. 13 is known, it can be used to determine the accumulated fluence for a period of time during solar maximum. Due to the stochastic nature of the events, confidence level approaches are often used so that risk-cost-performance tradeoffs can be evaluated by the designer. The first such model was based on King’s analysis of >10 to >100 MeV protons during solar cycle 20 [30,40]. One “anomalously large” event, the well-known August 1972 event, dominated the fluence of this cycle so the model predicts the number of such events expected for a given mission length at a specified confidence level. Using additional data, a model from JPL emerged in which Feynman et al. showed that the magnitude distribution of solar proton events during solar maximum is actually a continuous distribution between small events and the extremely large August 1972 event [31]. Under the assumptions that this underlying distribution can be approximated by a lognormal distribution and that the occurrence of events is a Poisson process, the JPL Model uses Monte Carlo simulations to calculate the cumulative fluence during a mission at a given confidence level [31,39]. An example of this is shown in Fig. 14 for > 30 MeV protons. Thus, according to this model, there is approximately a 10% probability of exceeding a proton fluence of \( 10^{10} \, \text{cm}^{-2} \) for a 3-year period during solar maximum. This corresponds to a 90% confidence level that this fluence will not be exceeded.
An underlying assumption of the JPL Model is that the year to year fluences during solar maximum have such a large variation that each year during solar maximum can be treated the same statistically. Other Monte Carlo based models have been developed that parameterize the number of events that are predicted to occur as a function of time from the beginning of the solar cycle. One such model, developed at ONERA, uses measured event numbers for this [41]. Another such model assumes the event number is directly proportional to the sunspot number and thus relies on knowledge of sunspot numbers during the mission time period [33].

It has also been demonstrated that the cumulative fluence distribution during solar maximum is consistent with a lognormal distribution for periods of time up to at least 7 years [42]. This was shown using the Maximum Entropy Principle, Bootstrap-like methods [43] and by Monte Carlo simulations using the initial distribution shown in Fig. 13. Thus the cumulative fluence distribution is known once the parameters of the lognormal distribution are determined. These parameters depend on the proton energy range and the mission duration. They have been determined from the available satellite data and well-known relations for Poisson processes. Fig. 15 shows examples of the annual proton fluences for >1, >10 and >100 MeV protons plotted on lognormal probability paper. This Fig. is constructed so that if a distribution is lognormal, it will appear as a straight line. The fitted data can also be used to determine the lognormal parameters for different periods of time and is used in the ESP Model [44].

Fig. 16 shows a representative comparison of the models discussed above. In addition it shows an update of the ESP Model, called PSYCHIC [45], in which the data were extended to cover the time period from 1966 to 2001 and the proton energy range extended to over 300 MeV. Results shown are for the 90% confidence level and for a mission length of two solar maximum years. In all cases the energy range shown corresponds to the data range on which the statistical models are based, i.e. no extrapolations are used. Thus, the model differences seen are an indicator of model uncertainties. The spectral shape for the King Model is based on the August 1972 event and is therefore somewhat different than the other model results. The JPL91, ESP, and PSYCHIC models all agree reasonably well for their common 1 to 60 MeV energy range.
4) Cumulative Fluence During Solar Minimum

It has often been assumed that the solar particle event fluence during the solar minimum time period can be neglected. However, for missions that are planned mostly or entirely during solar minimum it is useful to have guidelines for solar particle event exposures, especially considering the current frequent use of commercial-off-the-shelf (COTS) microelectronics, which can exhibit rather low TID failure levels.

Due to the relative lack of events during solar minimum, models are more difficult to construct for this period. However, Monte Carlo based models that parameterize the number of events that are predicted as a function of time throughout the solar cycle are useful for predicting cumulative fluences during solar minimum. Two such models are those developed at ONERA [41] and by Nymmik [33]. In addition, recent solar minimum time periods have been analyzed to obtain 3 average solar proton flux levels that allow varying degrees of conservatism to be used [45].

5) Extreme Value Theory and Worst Case Events

An important consideration for spacecraft designers is the worst-case solar particle event that occurs during a mission. One approach is to design to a well-known large event such as that which occurred in October 1989 [14], or a hypothetical one such as a composite of the February 1956 and August 1972 events [46]. Energy spectra of some of the most severe solar proton events during solar cycles 19-22 are shown in Fig. 17. In addition, there are event classification schemes in which the magnitudes range from “small” to “extremely large” that can be helpful for design purposes [48,49].

![Image](http://rahome.gsfc.nasa.gov/image.png)

Fig. 17. Some of the most severe solar event energy spectra in solar cycles 19-22 [47].

Additional information can be provided to the designer if a confidence level associated with the worst case event is known for a given mission length. The designer can then more systematically balance risk-cost-performance tradeoffs for the mission in a manner similar to what is done for cumulative fluences. Once the initial probability distribution such as that shown in Fig. 13 is determined it becomes possible to construct such a statistical model using extreme value theory. In the usual central value statistics, the distribution for a random variable is characterized by its mean value and a dispersion indicator such as the standard deviation. Extreme value statistics, pioneered by Gumbel [50], focuses on the largest or smallest values taken on by the distribution. Thus, the “tails” of the distribution are the most significant. For the present applications the concern is with the largest values.

Examples of extreme value modeling of environmental phenomena such as floods, wave heights, earthquakes and wind speeds can be found in a number of places [50-52]. This modeling was first applied to radiation effects problems by Vail, Burke and Raymond in a study of high density memories [53]. It has turned out to be a very useful tool for studying the response of large device arrays to radiation. Other radiation effects applications have been found for arrays of gate oxides [54,55], sensor arrays [56,57] and EPROMs [58].

For the application to solar particle events the interest is in the worst-case event that will occur over a period of $T$ solar maximum years. Since the number of events that can occur over this period is variable, the expression for the extreme value distribution must take this into account. Assuming that event occurrence is a Poisson process [39], it can be shown that the cumulative, worst case distribution for $T$ solar maximum years is [59]

$$F_{\text{max}}(M; T) = \exp\left\{ - N_{\text{tot}} T \left[ 1 - P(M) \right] \right\} \quad (4)$$

where $P(M)$ is the initial cumulative distribution, which is closely related to equation (3) [37].

Fig. 18 shows results for worst-case event fluences for mission lengths of 1, 3, 5 and 10 solar maximum years. The ordinate represents the probability that the worst-case event encountered during a mission will exceed the $>30$ MeV proton fluence shown on the abscissa. Also shown in the Fig. by the vertical line denoted by “Design Limit” is the maximum event fluence parameter, $\phi_{\text{max}}$. As will be discussed next, this parameter can be used as an upper limit guideline. Results analogous to these have also been obtained for peak solar proton fluxes during events [60], which are very relevant for SEE. The event fluence magnitudes are discussed here because of the interesting comparison that can be made with historical data to help validate the model.
A unique feature of this model is the upper limit parameter for a solar proton event fluence, \( \phi_{\text{max}} \). For the case of > 30 MeV protons this turns out to be \( 1.3 \times 10^{10} \text{ cm}^{-2} \). However, this is a fitted parameter that was determined from limited data. There must be some amount of uncertainty associated with the parameter. Thus, it should not be interpreted as an absolute upper limit. One method of estimating its uncertainty is the parametric "bootstrap" technique [43]. This method attempts to assess the uncertainty of the parameter due to the limited nature of the data. The idea is to randomly select event fluences according to the distribution given by equation (2) until the number of events in the distribution is simulated. The equation is then fitted to the simulated data, and the parameters extracted. The procedure is repeated, and each time the parameters have different values. After a number of simulations, the standard deviation of the parameter of interest can be determined. This technique showed the upper limit parameter plus one standard deviation equaled \( 3.0 \times 10^{10} \text{ cm}^{-2} \) [37].

A reasonable interpretation for the upper limit fluence parameter is that it is the best value that can be determined for the largest possible event fluence, given limited data. It is not an absolute upper limit but is a practical and objectively determined guideline for use in limiting design costs.

Constraints on the upper limit of solar proton event sizes can be put on models as a result of studies of historical-type evidence. Relatively small fluctuations of \(^{14}\text{C} \) observed in tree rings over a long period of time [61] and measured radioactivity in lunar rocks brought back during the Apollo missions [62] are consistent with the upper limit parameter but are not especially restrictive. The strictest constraint to date comes from analysis of approximately 400 years of the nitrate record in polar ice cores [63]. The largest event reported was estimated to be \( 1.9 \times 10^{10} \text{ cm}^{-2} \) for > 30 MeV protons. This was the Carrington event that occurred in September 1859. Fig. 19 shows a bar graph of the upper limit parameter, \( \phi_{\text{max}} \), for > 30 MeV protons including the one standard deviation uncertainty that was estimated from the parametric bootstrap method. This is compared with the reported value for the Carrington event. It is seen that these quantities are well within the uncertainties. Also shown for reference is the value for the October 1989 solar particle event that is commonly used as a worst-case event.

C. Solar Heavy Ion Models

Solar heavy ion models are generally not as advanced as solar proton models due to the large number of heavy ion species, which complicates measurements of individual species. For microelectronics applications, solar heavy ion models are needed primarily to assess SEE. Astronaut exposure is also a serious concern for manned missions.

1) Cumulative Fluences

One quantity of interest is the average SEE rate during a mission. This means that models for cumulative solar heavy ion fluence must be developed to be used along with GCR Models. Tylka et al. used a Monte Carlo procedure similar to the JPL91 solar proton model [39] to predict cumulative fluences for certain elements during a mission at a specified confidence level [64]. This was done for 2 broad energy bins each for alpha particles, for the CNO group, and for Fe. It is based on the University of Chicago Cosmic Ray Telescope (CRT) data taken between 1973 and 1996.

The most complete model for cumulative solar heavy ion fluences is the PSYCHIC Model [65]. Here measured alpha particle energy spectra are scaled to proton energy spectra based on measurements from the IMP-8 and GOES instrumentation during the time period 1973 to 2001. The energy spectra of remaining major heavy elements – C, N, O, Ne, Mg, Si, S and Fe – are scaled to the alpha particle energy spectra using measurements of the Solar Isotope Spectrometer (SIS) onboard the ACE spacecraft over the most recent 7 year solar maximum period. An abundance model is used for the
remaining minor heavy elements. It is based on measurements from the International Sun-Earth Explorer-3 (ISEE-3) spacecraft and current knowledge of solar photospheric abundances and processes. Results for differential fluence-energy spectra are shown in Fig. 20 for some of the major elements and a summed spectrum for atomic number $Z > 28$. Also shown by the points in the Fig. are cumulative fluence results for alpha particles and iron for the same conditions based on the modeling effort of Tylka [64].

2) Worst Case Events

In an attempt to model worst-case events, the original CREME model [13] and subsequently the CHIME model [66] scaled heavy ion abundances to protons for individual events. However, this assumption that individual events with the highest proton fluxes should also be heavy ion rich turned out to be inconsistent with subsequent data [21] and led to worst-case event models that were too conservative [67]. Modifications of the original CREME code were made in the MACREE model [68] to define a less conservative worst-case solar particle event. MACREE gives the option of using a model based on the measured proton and alpha particle spectra for the well-known October 1989 event and an abundance model that is 0.25 times the CREME abundances for atomic numbers, $Z > 2$. A model that originated at JPL [69] characterizes the distribution of 1 to 30 MeV per nucleon alpha particle event fluences using a lognormal distribution in order to assign confidence levels to the event magnitudes. The alpha particle data are based on measurements from the IMP-8 satellite for solar maximum years between 1973 and 1991. For ions heavier than $Z = 2$ an abundance model is used and the fluxes are scaled to the alpha particle flux for a given confidence level [67]. The current version of the widely used CREME code, CREME96, uses the October 1989 event as a worst-case scenario. It provides 3 levels of solar particle intensity [14]. These are the “worst week”, “worst day” and “peak flux” models, which are based on proton measurements from the GOES-6 and -7 satellites and heavy ion measurements from the University of Chicago CRT on the IMP-8 satellite. The most extensive heavy ion measurements in the model are for C, O and Fe ions [70]. It is noteworthy that the energy spectra of these 3 elements extend out to roughly 1 GeV per nucleon.

Comparisons to the CREME96 worst case models have been made with data taken by the Cosmic Radiation Environment Dosimetry (CREDO) Experiment onboard the Microelectronics and Photonics Test Bed (MPTB) between 2000 and 2002 [71]. The data show that 3 major events during this time period approximately equaled the “worst day” model. An example of this is shown in Fig. 21 for an event that occurred in November 2001.

![Fig. 20. Differential fluence-energy spectra for protons, alpha particles, oxygen, magnesium, iron and summed spectra for $Z > 28$ elements for a 2-year mission during solar maximum at the 90% confidence level. Lines are spectra calculated with the PSYCHIC model [65] and points are obtained from [64].](image)

![Fig. 21. Comparison of a solar heavy ion event that occurred in November 2001 with the CREME96 “worst day” model. The progression of daily intensities is indicated with the peak intensity occurring on day 2929 of the mission [71].](image)
V. EARTH RADIATION BELTS

A. Overview and background

The Earth's magnetosphere can be seen as a natural cavity in the interplanetary medium in which the Earth is relatively well protected against external influences. It is compressed on the solar side and highly extended on the anti-solar side. In this structure, at the level of the poles, two horns, flaring out towards space offer the particles from the interplanetary medium a possibility of penetrating into the upper atmosphere. Close to Earth, the charged particles present in the magnetosphere can be trapped by the magnetic field and form the radiation belts.

In the magnetosphere, the radiation belts only occupy a relatively restricted internal region (Fig. 23). The region closest to Earth is well known and constant over time: this is the upper atmosphere. The external limit, however, is poorly defined and depends on the conditions in the solar wind through the magnetic field as we will see later on. They therefore extend from the upper atmosphere (some hundreds of km) up to geostationary orbit and beyond.

In the Earth's magnetosphere, the magnetic field is the sum of two terms, one of internal (main component) and the other of external origin. The internal magnetic field is probably due to the convection motion in the core of the planet; in addition to this main term, there is the permanent residual field of the Earth's crust. At the zero order the field can be considered to be dipolar. However the single dipolar approximation is not rigorous. It is then more appropriate to take an off-center and tilted dipolar magnetic field as approximation. This gives a dipole whose center is not at the center of the Earth and whose axis is not parallel to the Earth's rotation axis (Fig. 24). The result of this geometry of the magnetic field is an anomaly, a zone in which the field is weaker. This region is situated at the level of Brazil, and is known as the South Atlantic Anomaly. More realistic models of the internal field exist, the old models such as Jensen and Cain 1962 [72], GSFC 12/66, and the International Geophysical Reference Field model (IGRF) [73]. These models consider the terms of a multipolar higher order. It is important to note that the Earth's field is subject to long-term changes (secular drifts), in particular the South Atlantic Anomaly is drifting south-eastwards. At the present time, we note:

- a decrease in the intensity of 27 nT/year (0.05 % a year),
- a drift of the axis, resulting in a westward rotation of the southern end of the dipole (0.014° a year) and an increase in the shift towards the West Pacific close to 3 km a year.
Up to a distance of some Earth radii, 5 to 6, the magnetic field is close to a dipolar field and the magnetosphere is more or less in revolution. Beyond that distance, the external fields become less and less negligible and contribute to the deformation of the dipolar internal field. These external fields are the sum of several components: the compression on the day side and the blast in the anti-solar direction, forming the tail of the magnetosphere. Several models of external field exist, we can mention for example the models of Tsyganenko [74, 75], Tsyganenko and Stern [76], Olson and Pfitzer [77, 78], and Alexeev et al. [79]. These latter – the sum of the fields transported by the solar wind and induced by the currents in the magnetosphere – are subject to rapid variations. The variations in the interplanetary environment have an impact on the magnetosphere. The variations in the speed of the solar wind (400 to 1000 km/s) and therefore of the energy transported are, depending on the orientation of the interplanetary magnetic field, more or less well transferred to the magnetosphere, increasing the instabilities of the external magnetic field. For example the compression of the sub-solar zone may be sufficient to place a geostationary satellite temporarily beyond the magnetopause; likewise, during these geomagnetic storms, injections of high-energy particles are observed in the radiation belts.

In order to understand and reproduce the dynamics of the charged particles present in the magnetosphere, it is common to define magnetic coordinates (Fig. 25). r is the distance from the center of the dipole to the point under consideration, \( \lambda \) its latitude (and \( \theta \) its colatitude: \( \theta = \pi/2 - \lambda \)), and \( \phi \) its magnetic longitude. A field line (or force line) is defined by the McIlwain parameter, \( L \) [80] roughly equal (only true with a dipolar field) to the distance (expressed in planet radii) from the center of the planet to the intersection point of that force line with the magnetic equatorial plane. A point on a force line is defined by the \( B \) parameter, modulus of the magnetic field at the point under consideration. \( B \) and \( L \) then represent a coordinates system linked to the model of the magnetic field under consideration.

All charged particles immersed in an electromagnetic field will be subject to the Lorentz force: \( \vec{F} = q(\vec{v} \times \vec{B}) + q \vec{E} \) where \( q \) is the particle's charge, \( \vec{v} \) its speed, \( \vec{B} \) the magnetic field and \( \vec{E} \) the electric field. If the magnetic field is very strong and the energy of the particles is great (and therefore their speed too) then the effect of the electric field can be ignored and the Lorentz force is reduced to \( \vec{F} = q(\vec{v} \times \vec{B}) \). Under these conditions, the movement of the high-energy particles can be generally broken down into three basic periodic movements.

**Gyration:** a charged particle immersed in a magnetic field will have a rotation movement around the field line. This movement is called gyration. It is then possible to define some magnitudes relative to this movement:

- the Larmor radius, \( r_L = \frac{mv_{\perp}}{qB} \) where \( m \) is the relativistic mass of the particle, \( v_{\perp} \) the component perpendicular to the magnetic field of the particle speed, \( q \) its charge and \( B \) the modulus of the magnetic field;
- the relativistic magnetic moment, \( \mu = \frac{mv_{\perp}^2}{2B} = \text{constant} \).

**Bounce:** If a particle has one component of its speed parallel to the magnetic field then it will move along the field line. When making any movement the particles keep their relativistic magnetic moment, \( \mu \) constant. Since the magnetic moment has to remain constant, the particle which moves from the equator (point where the magnetic field is weakest along the field line) towards the higher latitudes will see an increasingly strong magnetic field. It is necessary that the perpendicular component of the speed should increase in order for \( \mu \) to remain constant. This will be possible until the perpendicular speed is equal to the particle's total speed, parallel speed then being null. At this particular point the particle stops, it is at its mirror point. A weak force due to the gradient of the magnetic field enables this particle to go backwards to its other mirror point situated in the other hemisphere. The particle therefore has a back and forth movement between its two mirror points, this is the bounce movement.

It is possible to define the angle that the particle's speed vector must have with respect to the magnetic field when it crosses the equator such that its mirror point is in the upper
atmosphere. The particle is then lost and will not be able to come back. This allows us to define a loss cone, that is to say if the speed vector is within the cone then the particle cannot bounce and will be lost.

Drift: in order to simplify the problem, we place ourselves on the magnetic equator. Since the magnetic field of the planets has a radial gradient, the gyration cannot take place in a constant Larmor radius. Indeed, the magnetic field along a gyration becomes stronger if the particle approaches the planet, the Larmor radius is then smaller and therefore the radius of the trajectory's curve is also smaller. The particle will thus be able to move away from the planet, the magnetic field will be weaker and therefore the Larmor radius and the radius of the trajectory's curve will be greater. The particle therefore does not go through a simple circle but along a more complex trajectory. This movement breaks down into a simple gyration (circular) and a rotation movement around the planet: this is the drift movement.

A charged particle submitted to these three basic and periodic movements then moves through torus shaped surfaces around the Earth, which are commonly called drift shells (Fig. 26). The periods associated with each of these basic movements for a 3 MeV electron at L=3 are respectively 2.14 × 10^4 s, 0.19 s and 504 s. The disparity between the periods is very great, a factor of the order of 1000 should be noted between each of them going from the gyration movement to the drift movement.

A view of the radiation belts is given in the following Fig.s. A single maximum is observed for the proton belt (Fig. 27) for a value of L that depends on the energy (L=1.7 for 10 MeV protons); the flux is very stable there and the maximum energies can reach between some MeV and some hundreds of MeV depending on the position.

The electron belt is more complex (Fig. 28) and has two maximums respectively corresponding to the internal and external zones:

- the first one centered on L = 1.4 extends up to L = 2.8; the electron populations are relatively stable there and can reach maximum energy levels of the order of 10 or even 30 MeV;
- the second one, centered on L = 5, extends from L = 2.8 to L = 10; the electron flows there are much more variable and the energy levels can be as high as 7 MeV.

<p>| Table 3: Characteristics of the Earth's radiation belts. |
|-----------------|----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Particle</th>
<th>Energy</th>
<th>Extension (Earth radii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>e-</td>
<td>1 keV-7 MeV</td>
<td>1-10</td>
</tr>
<tr>
<td>p+</td>
<td>1 keV-300 MeV</td>
<td>1-7</td>
</tr>
</tbody>
</table>

A single maximum is observed for the proton belt (Fig. 27) for a value of L that depends on the energy (L=1.7 for 10 MeV protons); the flux is very stable there and the maximum energies can reach between some MeV and some hundreds of MeV depending on the position.

The electron belt is more complex (Fig. 28) and has two maximums respectively corresponding to the internal and external zones:

- the first one centered on L = 1.4 extends up to L = 2.8; the electron populations are relatively stable there and can reach maximum energy levels of the order of 10 or even 30 MeV;
- the second one, centered on L = 5, extends from L = 2.8 to L = 10; the electron flows there are much more variable and the energy levels can be as high as 7 MeV.

B. Description of radiation belts

The magnetic field in the vicinity of the Earth becomes such that all relativistic charged particles are trapped and their movement is then quasi-periodic. These special conditions are thus favorable to the accumulation of high-energy charged particles in certain regions of space which creates the radiation belts. Given the trajectories of the particles the radiation belts have a toroidal shape which surrounds the Earth. The Earth's atmosphere is the lower limit of the radiation belts since it causes the loss of all the trapped particles. The upper limit, however, is less clear and is defined by the minimum intensity in the presence of disturbances of the magnetic field such that the particles are always trapped.

Discovered during the first space missions by J. Van Allen, the particles trapped in the radiation belts (or Van Allen belts) are essentially protons and electrons. The energy ranges commonly encountered go from some keV up to some tens or even hundreds of MeV. Table 3 summarizes the properties of the Earth's radiation belts.

![Fig. 26. Composition of a charged particle's three periodic movements: gyration, bounce and drift. The particle then follows a torus surface called a drift shell.](image-url)

![Fig. 27. Proton radiation belt.](image-url)
At zero order the radiation belts can be considered to be symmetrical in longitude in a region going from the Earth's surface up to the geostationary orbit, i.e. as long as the magnetic field is not too different from a dipole. However, at high altitudes, the field differs from a dipole and the belts are no longer axisymmetric. At the level of the geostationary orbit, the fluxes of high-energy particles (electrons between 100 keV and some MeV and protons between 100 keV and 1 MeV) then have a maximum on the day side and a minimum on the night side. We then speak of day-night asymmetry due to the topology of the magnetic field (external).

Since the Earth's dipole is tilted and off-centered by 500 km towards the West Pacific, the radiation belt (protons and electrons) goes down to a low altitude over the South Atlantic, the populations of charged particles being attached to the magnetic field. A satellite in low orbit (LEO : Low Earth Orbit) will thus only be exposed to radiation on certain fractions of the orbit as far as the trapped particles are concerned when passing through (Fig. 29):

- the polar horns (electrons below 1000 km, electrons and protons above that altitude),
- the South Atlantic Anomaly (protons and electrons at all altitudes).

As can be seen in Fig. 30 the position of Kourou, close to the SAA, means that the launcher's trajectory passes through a zone with a great flux of energetic trapped protons when being injected into geostationary transfer orbit. This must be taken into account when designing the on-board electronics which may be sensitive to the singular effects induced by protons.

C. Dynamics of the radiation belts

Given the measurements of the trapped particles in the Earth's environment, it is now certain that a static view of the radiation belts is obsolete. The American CRRES satellite in the early 1990s clearly evidenced the extreme dynamics of the trapped electrons and protons. As stated earlier, the radiation belts are linked to the existence of the Earth's magnetic field and the populations of particles are the result of an equilibrium between:

- the sources, injections from the tail of the magnetosphere and creations by nuclear reactions between atoms in the upper atmosphere and energetic ions (solar or cosmic),
- the losses by precipitation in the upper atmosphere or by charge exchange with the atoms and molecules from the exosphere (extended atmosphere).

These various terms can vary over time and are highly dependent on the magnetic field; any transient disturbances and time drifts of the magnetic field result in rapid fluctuations (magnetic storms) and long-term variations of the fluxes in the belt.

It is therefore judicious to look at the dynamics of belts at different time scales as a function of the effect (of the degradation) to be studied. If you are only interested in the...
cumulative effects such as the dose effects then variations on the scale of the minute (sub-storm) or of the week (storm) serve no purpose. However the variations on the solar cycle scale will be fundamental. On the contrary, in the case of studies on charging environments, the time scales go from some hours to some days, and the dynamics of the belts must be described on the scale of the sub-storm (surface charge) or of the magnetic storm (surface and internal charge).

1) Dynamics on the scale of the solar cycle—Protons

The proton radiation belt (high-energy component > 10 MeV) varies slowly as a function of the solar cycle [4]. The flux levels are roughly at their highest when the solar cycle is at its lowest and vice versa. This is the result of two physical processes that condition the dynamics of the protons, the absorption of the protons by the upper atmosphere on the one hand and the modulation of the CRAND source (Cosmic Ray Albedo Neutron Decay) on the other hand. When the solar cycle is at its maximum the upper atmosphere is heated up and the densities at constant altitude increase. It can then be understood that the losses of trapped protons induced by the charge exchange increase. However, when the solar cycle is at its maximum, the fluxes of cosmic radiation fall due to the intense solar activity and the source is reduced. The balance is shown in Fig. 31.

Another important characteristic of low altitude proton fluxes is an overall decrease in the fluxes from one cycle to the next. The comparison of the fluxes from one solar minimum to the next shown in Fig. 31 very clearly illustrates this slow variation. It is due to the secular drift of the Earth’s magnetic field. At the present time this poses problems for forecasting the low altitude proton fluxes from one cycle to the next.

2) Dynamics on the scale of the solar cycle—Electrons

The variations in the electron belt are above all known in the external zone, the geostationary orbit being particularly well documented [82, 83]. In geostationary orbit (Fig. 32), the fluxes of electrons are at their lowest when the solar cycle is at its highest and are at their highest three or four years after the top of the cycle (just before the solar cycle is at its lowest). This modulation as a function of the radio-solar flow (F10.7) at 10.7 cm shows that the amplitude is all the greater if we examine the high energies (MeV and above). However at low energy levels (some hundreds of keV) the modulation is virtually inexistent. The strong fluxes of electrons are linked to the presence of coronal holes on the surface of the sun which in turn involve intense and long-duration magnetic storms at the level of the Earth’s magnetosphere (see next paragraph).
Even if the charge phenomena on satellites are induced by instantaneous fluxes of electrons, it is nevertheless possible to use this curve to define the unfavorable periods that lead to surface or internal charging. Surface charging can appear at any moment during the solar cycle, since the low-energy electrons involved in this process are not modulated by the solar activity. As for the internal charge, it will preferably appear some years after the maximum of the solar cycle when the coronal holes have an influence on the Earth's environment.

3) Dynamics on the scale of the magnetic storm-Proton

The low-energy protons (some tens to some hundreds of keV) are very sensitive to magnetic storms. The fluxes of particles therefore follow the Earth's magnetic activity in a region going from L=2 to L=6 with time scales going from a minute to several hours. A view from the CRRES satellite (MEB instrument) makes it possible to view 14 months' dynamics of the belt of 62 keV protons (Fig. 33) in various different regions (the satellite crosses the magnetic field lines close to the equator) [84, 85].

At higher energy levels (several tens of MeV) the belt is generally very stable but major events can dramatically change the flux levels in intermediate regions of the radiation belts [86]. In fact, if there is a solar flare in progress when a very intense magnetic storm is building up then the solar particles can be trapped and thus significantly increase the fluxes of trapped particles in a region between L=2 and 2.5 [87]. The example of the event in March 1991 is striking (Fig. 34).

4) Dynamics on the scale of the magnetic storm-Electrons

To better understand the environmental conditions that lead to charging phenomena on satellites the dynamics of the electrons will have to be controlled during major magnetic storms. According to the observations made in geostationary orbit the low-energy electrons (which induce surface charges) appear right from the first instants of the disturbance whereas the higher-energy electrons (which induce the internal charge) are detected some days after the beginning of the event (Fig. 35).
A distinction can also be made between various different classes of events [88] (Fig. 36): (1) storms where the magnetic activity index Kp is high (up to 8 or 9) but which do not last long (less than 1 day) and which do not produce any or only a few high-energy electrons in geostationary orbit (Fig. 36 on the left) and (2) the storms where the magnetic activity index Kp is moderate (up to 6 or 7) but which last longer (several days) and which produce large quantities of high-energy electrons in geostationary orbit (Fig. 36 in the middle). The most surprising thing is that in the second case the storms where the magnetic activity index oscillates between 2 and 4 for several days produce nearly as many high-energy electrons in geostationary orbit as a more violent storm with a Kp of 6 (Fig. 36 on the right).

In order to understand the electron acceleration phenomena we must examine in detail the physical processes that affect the dynamics of these particles. The low-energy electrons are first of all transported from the tail of the magnetosphere towards the inside of the radiation belts by an increase in the radial diffusion at the beginning of the storm. As they come closer to Earth, they see an increasingly strong magnetic field and, under the effect of the Lorentz force, they drift around the Earth firstly in the night-morning sector. In the vicinity of the plasmapause (\( L=4 \)), these electrons are not only submitted to radial diffusion but also to the wave-particle interaction. The combination of the two leads to a slow but continuous acceleration of the electrons. A large proportion of these electrons (now at high energy levels) will diffuse radially towards the internal zone of the radiation belts. There, they will be globally lost by precipitation in the loss cone due to the wave - particle interaction which is then preponderant. However a small proportion of the high-energy electrons will be able to diffuse from the plasmapause towards the exterior of the belt and finally reach the geostationary orbit. This acceleration mechanism makes it possible to explain, in particular, the delay between the low-energy electrons and the high-energy electrons in geostationary orbit.

Events such as these are very frequent, the CRRES period makes it possible to view the dynamics of the electrons for 14 months still with the major storm of March 1991 (Fig. 37).

5) Extreme events in the Earth electron belts

It is clear that there are many magnetic storm affecting the Earth radiation belts and some of them can be very extreme. Of course they are likely to happen rarely but can affect any space system dramatically. One full solar cycle is represented in Fig. 38 to appreciate the occurrence of such strong storms where extreme electron events recorded along two different orbits are highlighted. The sunspot number is plotted on the top panel to identify the phase of the solar cycle. The middle panel shows 5.35 MeV electron fluxes measured at LEO onboard NPOES-15 spacecraft (800 Km-98°) and the bottom one shows 5.5-7.1 MeV electron fluxes measured along Polar orbit (HEO). Over this long time period extreme events are only recorded during the declining phase of the solar cycle at LEO, this was during the large July-August 2004 and November 2004 storms. During that time very energetic electron were produced. Usually environment at LEO is assumed to be mainly a proton environment and any anomaly, like single event transient (SET) is assumed to be induced by definition by a single particle. In such a case what is called a
SET could just be a transient anomaly induced by an internal discharge. On the other hand along Polar orbit the situation is different. Along this time period the extreme events (regarding 5.5 – 7.1 MeV electrons) are seen in August-September 1998 during the rise of the solar cycle. Also the location of maximum flux values are at larger L shells compared to the August 2004 events. As a conclusion, conditions under which an extreme electron environment built up at LEO can be very different to the ones at HEO orbits. In other words, when an "extreme event" is seen along a given orbit does not mean it is extreme for all orbits. Non linearity in particle dynamics makes it difficult to extrapolate any single, local measurements to all locations in radiation belts.

For proton radiation belts, extreme event results from a combination of a solar flare (which is from the radiation belt point of view a source term) and a large magnetic storm (in this case usually due to a coronal mass ejection). An example is given in Fig. 39. The top panel shows 9-15 MeV solar protons measured by GOES-08 spacecraft at geosynchronous, middle panel shows in a L versus time map 9.65-11.35 MeV protons measured at LEO onboard SAC-C and the bottom shows the magnetic activity index Kp. On the middle panel proton flares are clearly observed at L values greater than 4 (Note that L of 4 highlighted by the horizontal dashed line show the average magnetospheric shielding) whereas trapped particles in the proton radiation belts are encountered at L below 3. On march the 31st, a large coronal mass ejection impacted the Earth magnetosphere and induced a large magnetic storm. The net effect on protons was to lower down the magnetospheric shielding leading access to low L shells to solar protons. Then a large amount of these solar particles have been trapped while the magnetospheric shielding was low. The trapped proton enhancement is clearly seen in the L=2 range.

Once new protons are trapped in the belts they can remain there for months. Fluxes can decrease slowly because particles are lost by friction processes and/or charge exchange or can be lost suddenly because a new large magnetic storm occurs with no source term from solar flares (see Fig. 40).

In other words, some extreme events combined with a solar flare can lead to large flux enhancements whereas others not being combined with a flare can lead to large flux decreases. It makes such events difficult to predict because the net results before and after the extreme event depends on how the storm and the flares are synchronize at the Earth.

7) Static models-NASA's AP8 and AE8 models
Numerous measurements performed between 1958 and 1978 have made it possible to familiarize ourselves with the fluxes and energies of the electrons and protons trapped in the Earth's radiation belts. Empirical models have been derived
from these measurements, giving a general but static view of the belts (Fig. 41). The most recent ones, developed by the NASA in the 1970s, AP8 ("Aerospace Corporation Proton version 8") and AE8 ("Aerospace Corporation Electron version 8"), give proton and electron spectrums at the solar minimum (AE8 MIN and AP8 MIN) and maximum (AE8 MAX and AP8 MAX) at all geomagnetic coordinate points (B, L) in the magnetosphere. Their validity range extends respectively from L=1.15 to 6.6 for protons with an energy level comprised between 100 keV and 400 MeV and from L=1.2 to 11 for electrons with an energy level comprised between 40 keV and 7 MeV [89, 90, 91].

These models, however, are now obsolete even if they remain a reference for all industrial companies working in the space sector [91].

First, the magnetic field has drifted and the South Atlantic Anomaly is now further to the East and South than it was in the 1970s. This problem, however, is not necessarily a limitation. When the fluxes of particles along a satellite orbit are assessed with a view to calculating the total dose for a mission, an average must be calculated for a large number of orbits. The exact position of the Anomaly is then no longer of interest. However, if for a certain mission, an Ariane launch from Kourou for example, the exact position of the South Atlantic Anomaly is required, it would be judicious to use the model of the magnetic field from that time and then make the latitude and longitude transformations to take into account the Anomaly's drift. It has also been demonstrated that the fluxes of protons are underestimated (by a factor of 1.6 to 2) by the AP8 models for altitudes comprised between 300 and 500 km, i.e. in the vicinity of the cutoff induced by the atmosphere. This is partly due to the interpolation technique. Finally, the East-West asymmetry of the particles measured at those altitudes is not reproduced by the AP8 models.

The AE8 model also has its shortfalls. At low altitude, at the level of the internal belt, the maximum energy given by the model is 5 MeV whereas the CRRES satellite has measured electrons with an energy level of 30 MeV. In the outer belt, however, the flows are overestimated by a factor of at least 3 at high energies. Likewise in geostationary orbit, the fluxes predicted by the AE8 MIN and MAX models are identical, the variations due to the solar cycle are not correctly reproduced here.

8) Static models-The NOAAPRO model

The NOAAPRO (NOAA proton) model developed by S. Huston for the NASA is the first model that takes into account the variations of the fluxes of protons trapped in the radiation belts at low altitude as a function of the solar activity. It is based on the MEPED measurements performed by the NOAAPROTIROS satellites. It determines the mean omnidirectional integrated fluxes of protons with energy levels higher than 16, 30 and 80 MeV at an altitude of 800 km as a function of the date and of the radio-solar flow at 10.7 cm, F10.7 [4]. This model is of course still very limited from the energy range and altitude viewpoint, however for a polar orbit at 800 km it represents a great step forward with respect to AP8 and is therefore positioned as a good base for a future low-altitude proton model.

9) Static models-The IGE2006 model

The IGE2006 model ("International Geosynchronous Electron" previously called POLE "Particle Onera Lan Environment") developed by ONERA-DESP in cooperation with the Los Alamos National Laboratory (LANL) and JAXA is the first model that takes into account variations of the fluxes of electrons trapped in the radiation belts in geostationary orbit and of a function of the solar activity [3]. It is based on the measurements made by the LANL's and JAXA’s geostationary satellites. It determines the mean omnidirectional differential fluxes of electrons with an energy level comprised between 1 keV and 5.5 MeV as a function of the solar activity. Once again, this model is limited to a single orbit but, here too, it represents a good step forward with respect to AE8 and a first step to develop a mean model of the electron environment in the external belt.

10) Dynamic models-The CRRESPRO and CRRESELE models

Subsequent to the CRRES mission in 1990 and 1991 (14 months in all) empirical and dynamic models of the radiation belts saw the light of day. The CRRESPRO (CRRES proton) model provides fluxes of trapped protons with an energy level comprised between 1 and 100 MeV for values of L between 1.15 and 5.5 [93]. It is based on the measurements provided by the PROTEL telescope. It calculates the fluxes of protons averaged over an orbit for two states, one so-called active and another quiet state (CRRESPRO ACTIVE and CRRESPRO QUIET). This model reflects, in particular, the creation of a second belt of protons subsequent to a solar flare synchronized with a major geomagnetic storm. As for the CRRESELE (CRRES electron) model it provides the fluxes of electrons with an energy level comprised between 700 keV and 5 MeV for L values of 2.5 to 6.5 and six levels of magnetic activity, plus a worst case and a mean state [94, 95]. It is based on the measurements given by the HEEF detector. The inputs are only Ap (geomagnetic activity index) averaged over 15 days. This model makes it possible to take into account the creation of a third belt and the effects of recurrent
storms at 27 days. The major problem of these models is their representativeness given the short period of measurement acquisition (14 months during a solar maximum).

11) Dynamic models-The ESA SEE1 model

This model is also deduced from the measurements provided by the CRRES satellite, and more precisely by the MEA detector. This model provides fluxes of electrons with energy levels higher than 100 keV as a function of the magnetic activity index Kp [96]. It is associated with a neuronal network which predicts the fluxes on the basis of the magnetic activity index Kp throughout the solar cycle. This model still poses the problem of the representativeness of the CRRES measurements (14 months as opposed to the 11 years of the solar cycle). Furthermore, the fluxes of electrons with energy levels higher than 1.5 MeV are extrapolated and are consequently overestimated.

12) Dynamic models-The Salammbô models

Since the 1990s, ONERA/DESP has been developing physical models of the proton and electron radiation belts, called the Salammbô codes. At the present time, these codes represent a family of models (Salammbô 4D, 3D and 2D) which provide a more or less well-refined description of the belts as a function of what one wants to reproduce and of the desired resolution of the result [84, 86, 97, 98]. These models make it possible to describe the dynamics of the proton and electron belts with energy levels of 10 keV-300 MeV and 10 keV-10 MeV respectively in the region going from L=1 to 7 with a time resolution of between one minute and several hours. The inputs to these models are the magnetic activity indices Kp and Dst and a boundary condition deduced from geostationary measurements. At present these models make it possible to understand the dynamics of the charged particles trapped subsequent to magnetic storms of variable intensities. Notably, the creation of the second proton belt seen by CRRES has been reproduced, as have the effects of long and short magnetic storms on the external belts of electrons. In the future, this model should make it possible to define the conditions required to obtain a worst case for radiation belts. It also offers the possibility of validating, or not, the representativeness of the measurements and even of extrapolating measurements over time. Finally another application of the model is to interpolate and extrapolate the measurements in order to reconstruct a complete and dynamic cartography of the radiation belts.

13) Discussion relative to the various models

All the models that have been developed to date are based on various different data bases compiled at different times. It is clear that the representativeness of the measurements is a major problem as far as the development of models is concerned. It can be noted, in particular, that AP8 and AE8 are based on non-continuous measurements acquired during a small solar cycle, the NOAAPRO and POLE models result from data collected during solar cycles of the same level, and the CRRESPRO, CRRESELE and ESA SEE1 models are based on a data base that is extremely limited over time and only during a solar maximum (Fig. 42).

Another problem is the L, B and energy coverage of each of these models. The most complete at the present time are still AP8 and AE8. All the others only partially cover this three-dimensional space. A comparison is given in Fig. 43.

From the engineering model point of view, the AP8 and AE8 models are the ones that cover the largest domain. Updating works are in progress, they have given rise to the creation of new models which are still limited but which offer new bases for developing a complete model of the radiation belts. Another shortfall at the present time is the definition of the worst case environment for assessing the transient effects that can appear on satellites.

D. D. Sensitivity of orbits to the radiation

It is clear that, given the distribution of the high-energy charged particles in the radiation belts and the magnetospheric shielding to protect against solar or cosmic particles, the environment close to satellites is highly dependent on the orbit. Here we propose to give a rapid overview of this environment based on the AP8, AE8 and JPL91 models for low orbits, 800 km 98° and 30°, 1400 km circular, geostationary orbit and 20000 km 55°.

A projection of each of these orbits in a meridian plane is shown in Fig. 44 in order to appreciate the belt regions passed
through by each of the satellites. It can immediately be seen that the high-altitude orbits are not subjected to the fluxes of high-energy trapped protons. However, these orbits are constraining from the relativistic electron viewpoint.

![Image of various orbits](image)

Fig. 44. Projection of the various orbits in a meridian plane with, in the background, omnidirectional integrated fluxes of protons (AP8 MIN, energy E>10 MeV) on the left and of electrons (AE8 MIN, E>500 keV) on the right. The red arrows indicate the solar or cosmic protons with the field line where they are stopped.

If we look in greater detail at the differences from one orbit to another (Fig. 45) it can be seen, from the point of view of the protons, that the high orbits (35000 and 20000 km) are not very sensitive to the trapped protons but, on the other hand, they are not protected by the magnetospheric shielding, hence the presence of solar protons. In low orbit the situation is quite different, the trapped protons are not negligible whatever their level of energy. However, if a low orbit is greatly tilted then the presence of solar protons will be great, but there are virtually none for slightly tilted orbits, which are naturally protected by the magnetospheric shielding. As for the electrons, as it has been stated earlier, the high orbits are greatly exposed.

![Image of omnidirectional flux spectra](image)

Fig. 45. Omnidirectional differential flux spectrums for trapped protons (AP8 min) on the left, for trapped electrons (AE8 max) in the center and for solar protons (Feynman 80 %) on the right.

VI. FUTURE CHALLENGES

A number of future challenges exist for space radiation environment models if they are to continue to help produce reliable, cost-effective spacecraft designs and have utility in implementing new space technologies. There should be a goal to produce more dynamical and more physical models of the environment. Such results should allow more accurate projections for future missions. Galactic cosmic ray models are closely tied to solar activity levels, which modulate the fluxes of these energetic ions. Challenges for these models are to incorporate an improved description of the solar modulation potential and to develop cosmic ray transport models that account for relevant atmospheric processes. Solar particle events demonstrate a strongly statistical character. A major challenge for these models is to develop a description of the energy storage and release processes in the solar structure. This could lead to a more detailed probabilistic model of the cyclical dependence of event frequencies and magnitudes.

Developing and implementing a strategy to deal with the radiation environment for manned and robotic space missions is critical for new interplanetary exploration initiatives. Getting astronauts safely to Mars and back will involve unprecedented strategies. For example, the lack of predictability of solar particle events indicates a potential strategy of establishing a measurement system in the inner heliosphere for the early detection and warning of events. Once an event is detected, accurate predictions must be made of the transport process to Earth, Mars and possibly beyond so that properties such as time of arrival, duration, intensity and energy spectra can be transmitted well ahead of the arrival time. The current GCR models depend on knowing solar activity levels in order to predict GCR fluxes. Thus, the lack of an established method for predicting future solar cycle activity is a concern in the planning for new exploration initiatives. Higher than expected GCR fluxes are a serious problem for long-term manned missions because they are difficult to shield against.

For trapped particle radiations, efforts are being made to develop particle maps for various climatological conditions that occur throughout the solar cycle for the full range of particle energies and geomagnetic coordinates covered by the AP-8 and AE-8 models. Ultimately, the goal is to develop an accurate description of the source and loss mechanisms of trapped particles, including the influence that magnetic storms have on the particle populations.

Thus, although there has been recent progress in modeling the space radiation environment, there are many future challenges that remain.

REFERENCES


