A Brief History of Space Climatology: From the Big Bang to the Present

Michael Xapsos, NASA Goddard Space Flight Center
NSREC 2018 Short Course

Outline

I. Background .................................................................................................................. 2
II. Introduction ................................................................................................................ 6
III. The Early Universe ..................................................................................................... 6
   A. Big Bang Nucleosynthesis ......................................................................................... 7
   B. Stellar Nucleosynthesis ............................................................................................ 8
   C. Extreme Event Nucleosynthesis ............................................................................... 9
   D. Abundances and Radiation Effects of the Elements .............................................. 10
IV. Transition to Modern Times ..................................................................................... 11
V. Modern Times – Space Climatology ....................................................................... 14
   A. Definition of Space Climatology and Space Weather ........................................... 15
   B. Galactic Cosmic Rays .............................................................................................. 16
      1. Properties ............................................................................................................... 16
      2. Models ................................................................................................................... 20
   C. Solar Particle Events ................................................................................................ 23
      1. Properties ............................................................................................................... 23
      2. Models ................................................................................................................... 25
         a) Cumulative Fluence Models ............................................................................. 26
         b) Worst Case Event Models ................................................................................ 30
      3. Current Issue: Use of Statistical Models vs. Worst Case Observations ............... 36
   D. The Van Allen Belts .................................................................................................. 37
      1. Trapped Particle Motion in the Magnetosphere ..................................................... 37
      2. Trapped Protons ....................................................................................................... 40
         a) Properties ........................................................................................................... 40
         b) Models ................................................................................................................. 45
      3. Trapped Electrons ................................................................................................... 47
         a) Properties ........................................................................................................... 47
         b) Models ................................................................................................................. 49
         c) Current Issue: The Case of the Missing Electrons ............................................ 52
   E. Example Environments Including Shielding ............................................................ 54
      1. Total Ionizing Dose ................................................................................................. 54
      2. Single Event Upset .................................................................................................. 57
VI. Summary ................................................................................................................... 59
VII. Acknowledgments ..................................................................................................... 60
VIII. References ............................................................................................................... 60
I. Background

This manuscript is focused on space climatology - the radiation environment observed over an extended period of time at a given location, corresponding to a space mission duration and orbit. Electronic devices and integrated circuits must be designed for this climatology in order to operate reliably. The main concerns covered are three categories of high-energy particle radiations. The first is galactic cosmic rays, a comparatively low-level flux of ions originating outside of our solar system. The second is solar particle events, bursts of radiation emitted by the sun characterized by high fluxes of protons and heavy ions. The third is the Van Allen Belts, primarily protons and electrons trapped by the Earth’s magnetic field. The plasma environment and spacecraft charging effects will not be discussed here. However, an excellent Short Course presentation of this topic was given in 2015 [Ma15].

In order to have successful, cost-effective designs and implement new space technologies, the climatology must be understood and accurately modeled. Underestimating radiation levels leads to excessive risk and can result in degraded system performance and loss of mission lifetime. Overestimating radiation levels can lead to excessive shielding, reduced payloads, over-design and increased cost. The pervasive use of commercial-off-the-shelf (COTS) microelectronics in spacecraft to achieve increased system performance must be balanced by the need to accurately predict their complex responses in space, which begins by understanding the environment in which they function.

This section gives a top level description of the basic radiation effects and introduces the metrics that describe them. There are three categories of effects that are considered – single event effects (SEE), total ionizing dose (TID) and total non-ionizing dose (TNID) or displacement damage. The following manuscripts in this Short Course provide much greater detail of these effects on electronic devices and systems.

SEE can be defined as any measureable effect in a circuit caused by a single incident particle. It can be either non-destructive such as single event upset (SEU) and single event transients (SETs) or destructive such as single event latchup (SEL) and single event gate rupture (SEGR).

SEE can be caused by direct ionization as the incident particle passes through a sensitive region of a device, as shown in figure 1 [Ba13]. This is usually the case for incident heavy ions. The metric used to characterize the incident ion is Linear Energy Transfer (LET). LET is the ionizing energy lost by the ion per unit path length in the material of the sensitive region. One form LET units take is MeV/cm. However, it turns out that the general expression for LET is directly proportional to the material density. That dependence can be removed by dividing by the density in units of mg/cm³, giving LET units of MeV·cm²/mg. This is the form of LET most often used by the radiation effects community. This can also be viewed as the energy lost per unit path length, where path length is an areal density. The metric used to calculate single event rates in space is the charge collected at the sensitive volume junction, $Q$.

$$Q = C \cdot LET \cdot s$$

where $s$ is the ion’s path length through the sensitive volume and $C$ is a unit conversion.
Figure 1. SEE caused by direct ionization [Ba13].

SEE can also be caused by nuclear reaction products from the interaction of the incident particle with a nucleus in the device material. This is usually the case for incident protons and always the case for incident neutrons. For this situation the relevant metric for single event rate calculation remains the charge collected in the sensitive volume but the calculation is more complex because there can be several emerging reaction products that deposit charge.

Total dose effects in electronic and photonic parts are cumulative, long-term degradation due to ionizing or non-ionizing radiation - mainly primary protons and electrons but secondary particles arising from interactions between these primary particles and spacecraft materials can also contribute. For the case of TID, the concern is mainly its effects in insulating regions of metal-oxide semiconductors (MOS) and bipolar devices, most commonly composed of SiO$_2$. In MOS devices, ionizing radiation causes threshold voltage shifts due to exposure of gate oxides, shown in figure 2 [Sc02]. The effect first appears as parametric degradation and can eventually result in functional failure. Ionizing radiation can also cause leakage currents in MOS devices due to exposure of field oxide regions used for isolation and result in increased power consumption. Eventually the effect becomes so pronounced the transistor cannot be switched to the off state. For bipolar devices loss in performance is caused by gain degradation and increased leakage currents.

The TID metric is the ionizing energy deposited per unit mass of material in the sensitive volume. It is given by
where $\phi$ is the incident particle fluence in cm$^{-2}$ and $C$ is a unit conversion. The MKS units for TID are the Gray (Gy), where 1 Gy = 1 J/kg. The cgs units are the rad, where 1 rad = 100 erg/g.

\[
TID = C \cdot LET \cdot \phi 
\tag{2}
\]

The effects of TNID or displacement damage dose are also cumulative effects caused when the incident radiation displaces atoms in a semiconductor lattice or optical material. TNID is the energy going into displaced atoms per unit mass of material in the sensitive volume. This is shown in figure 3, which illustrates defect production by a recoiling 50 keV silicon atom set in motion by an incident proton [Sr13]. Resulting defects can cause material property changes such as carrier lifetime shortening, mobility decreases and degradation of optical transmission. Displacement damage effects are commonly observed in solar cells, focal planes and optocouplers, which are often a component in power devices. There are two commonly used metrics for TNID. The first is completely analogous to the TID metric shown by equation (2).

\[
TNID = C \cdot NIEL \cdot \phi 
\tag{3}
\]

Figure 2. Illustration of the TID effect in the gate region of an MOS device [Sc02].
where NIEL, the Non-Ionizing Energy Loss [Su93], is the displacement damage analog of LET. It is the non-ionizing energy lost by the ion per unit path length in the material of the sensitive region. The second metric is to use an equivalent fluence of particles that causes an effect or is equivalent to the TNID value of interest. Equivalent fluences of 50 MeV protons are often used for space applications and in the case of solar cells 10 MeV protons and 1 MeV electrons are used.

Figure 3. Illustration of the displacement damage produced by a 50-keV Si recoil [Sr13].
II. Introduction

With the previous background information on basic radiation effects terminology, this section starts discussion of the history of space climatology. This will be developed by following a timeline starting with the Big Bang and ending at the present. A description of the early universe from a radiation effects perspective will be presented, featuring the origin and abundances of particles relevant for radiation effects – electrons, protons, neutrons and heavy ions. An interesting feature here is a recent development that is changing the view of the origin of ultra-heavy elements in the Periodic Table. It will be seen that the origin and abundances of these radiations are generally related to the radiation effects they cause in electronic devices and even some of the design requirements that are levied. A transitional period leading to modern times will then be discussed that involves the discovery of sunspots, the solar cycle and the sun’s pervasive influence on space climatology. This leads to the main discussion about modern space climatology on galactic cosmic rays, solar particle events and trapped particles. Radiation properties such as elemental composition, fluxes, energies, and dependence on solar cycle phase and spacecraft orbit will be discussed, with emphasis on variability of these properties. Radiation models used for space system design will be presented along with a current issue associated with each of the three types of radiation. This will bring the reader up to date and complete the journey along the space climatology timeline.

III. The Early Universe

It is now well established that the size of the universe is expanding with time. Therefore looking backward in time would reveal a universe that encompasses smaller and smaller volumes the farther back we go. Remarkably, scientists have been able to explain many phenomena by continuing to trace this contraction back to a time about 13.8 billion years ago, considered to be the age of the universe. At this point it is assumed to be a singularity of infinitesimal size and infinitely dense mass. This generally accepted Big Bang Theory of the birth and evolution of the universe is described in a number of interesting publications for a general audience [Ha88], [Ha05], [Li13], [Ma08]. Figure 4 shows an overall timeline beginning with the Big Bang and continuing through different eras to the present [after Li05].
The following discussion of the early universe is limited to the origin and abundances of radiations that are significant for radiation effects in electronic devices and circuits – electrons, protons, neutrons and heavy ions. It involves three types of nucleosynthesis processes – Big Bang, stellar and extreme event nucleosyntheses.

A. Big Bang Nucleosynthesis

A tiny fraction of a second after the Big Bang it is theorized that elementary particles called quarks existed. There are 6 types of quarks – up, down, top, bottom, strange and charm. The most stable of these are the up and down quarks, which are the building blocks of nucleons. At times on the order of microseconds after the Big Bang the early universe had expanded and cooled enough to allow quarks to come together and form stable nucleons. Two up quarks and one down quark form a proton while two down and one up quark form a neutron. Electrons, which are known to be particles with no internal structure, also existed a tiny fraction of a second after the Big Bang. Continued expansion and cooling allowed protons and neutrons to coalesce into simple nuclei. At an age of about 380,000 years the universe had cooled enough to allow electrons to orbit nuclei and form simple atoms, mainly hydrogen and helium. This portion of the timeline is shown in figure 5.

Figure 4. Timeline from the Big Bang to the present [after Li05].
B. Stellar Nucleosynthesis

The formation of the elements in the Periodic Table is a complex subject and there can be more than one pathway to the synthesis of an element. The purpose of the next two sections is not to exhaustively describe this for each element but to simply give a general description of elemental origins so they can ultimately be connected to the radiation effects they cause.

Over a long period of time on the order of hundreds of millions of years, the elements created after the Big Bang, primarily hydrogen, began accumulating into gaseous structures such as the iconic image shown in figure 6 taken by the Hubble Space Telescope and known as the “Pillars of Creation”. These features of the Eagle Nebula are about 4 to 5 light years in their largest dimension. A star will be born when branches or portions of these structures separate and the hydrogen atoms are close enough to start fusing. It is believed this is how the first stars formed.
Figure 6. Timeline for the formation of the first stars.

At this point in time stars would have consisted almost entirely of hydrogen and helium. The gravitational attraction of the star’s enormous mass is balanced by the energy release of fusion reactions to form helium, and keeps the star from collapsing in on itself. When the hydrogen is mostly used up, the star begins to contract. This raises the temperature of the core and if the star is large enough (much larger than our sun) helium begins to fuse and additional energy is released to balance the gravitational force. Thus, during the lifetime of large stars a chain of nuclear fusion reactions starting with hydrogen and helium produce elements from carbon to iron in the star’s core. Iron is the element with the highest binding energy in the Periodic Table and is therefore the most stable. When the star’s core is entirely iron, fusion is no longer possible because the reaction requires energy to be provided rather than resulting in its release. The star’s life is then over. It implodes and becomes a supernova as described in the next section. This production of the elements from C to Fe was first proposed by Hoyle [Ho46], [Li13].

C. Extreme Event Nucleosynthesis

There are two basic conditions that are required for the production of ultra-heavy elements, i.e., those heavier than iron. The first is that there must be enormous energy available in order to overcome the unfavorable energetics of forming these ultra-heavy elements from lighter elements. The second is that there must be an abundance of neutrons available, which is seen by examining the excess of neutrons relative to protons in the nuclei of the ultra-heavy elements in the Periodic Table. There are few known processes in the universe where this could occur. The two most likely happen after the active lifetimes of large stars. One is due to a supernova explosion, which is initiated when a star’s fuel is used up and the core consists entirely of iron. With no remaining energy to support itself against gravity, the star collapses. Protons and electrons are crushed together to form neutrons and there is a tremendous release of energy from the collapse making the production of the ultra-heavy elements possible. A second process is the collision/merger of two neutron stars, observed for the first time August 17, 2017.
A neutron star is the remnant of a large star after a supernova explosion that has collapsed to the density of nuclear material and consists mainly of neutrons. Visible light was detected from this event and gave evidence that ultra-heavy elements such as platinum and gold were formed in significant amounts. This led some scientists to postulate it could be the dominant process for formation of ultra-heavy elements.

D. Abundances and Radiation Effects of the Elements

With that general background on the origin of elements, their abundances are now examined. Figure 7 presents the solar abundances of elements in the Periodic Table as a function of mass number. This generally represents the elemental abundances of the solar system [An89]. Protons and alpha particles existed shortly after the Big Bang so it is not surprising that the elements H and He are the most abundant. The elements ranging from C to Fe are synthesized in stars larger than the sun in nuclear chain reactions. They are therefore less abundant than the lighter elements H and He. Since the sun ejects these heavy elements during solar particle events but cannot synthesize them, this has the interesting consequence that these heavy elements originated in previous generation stars. Finally, note the rapid decline of the elemental abundances beyond Fe. These ultra-heavy elements are likely only produced in the rare explosive processes discussed in section C.

Figure 7. Solar abundances of the elements [An89].
A Periodic Table of Radiation Effects can now be constructed that shows the different effects these radiations produce. This is shown in figure 8 in which the effects are color coded. The blue color indicates that the radiation generally produces total dose effects, including both TID and TNID. A green color indicates SEE and the lavender color indicates charging effects. The table is geared toward radiation effects so electrons and neutrons are included alongside protons. The most abundant radiations, electrons and protons, are largely responsible for cumulative total dose effects that require large numbers of particle strikes in devices. The less abundant alpha particles can contribute to total dose effects to a limited extent as can neutrons. In space neutrons are produced primarily by interactions of protons with spacecraft materials, planetary atmospheres and planetary soils. Due to their large numbers, electrons are mainly responsible for charging, another cumulative effect. The heavy elements C through Fe are not abundant enough to contribute significantly to these cumulative effects but they are important for SEE. Beyond the Fe, Co, Ni group the elemental abundances and therefore the particle radiation fluxes in space are very low. This is shown in the figure by shading only a small portion of the elemental box green. It can, however, be important to consider their effects for high confidence level applications such as destructive or critical SEE. The three remaining elements that have not yet been discussed, Li, Be and B are relatively rare and produced mainly by fragmentation of heavier galactic cosmic ray ions. This will be shown later in figure 13.

![Periodic Table of Radiation Effects](image)

**Figure 8. Periodic Table of Radiation Effects.**

**IV. Transition to Modern Times**

Now that the origin of radiations in the early universe have been discussed along with their abundances and effects on electronic devices, let’s move on to the transition period to modern times when the era of space climatology emerged. A timeline of this era is shown in figure 9.
Figure 9. Timeline for the emergence of space climatology.

The telescope was invented in 1608 by the Dutch lens maker Hans Lippershey. Shortly thereafter Galileo Galilei improved its magnification and was the first to use a telescope to study space. These studies could be regarded as the start of modern experimental astronomy. He was one of the first to observe sunspots through a telescope and hypothesized they were part of the solar surface as opposed to objects orbiting the sun.

Today sunspots are regarded as a proxy to solar activity. They are active regions having twisted magnetic fields that inhibit local convection. The region is therefore cooler than its surrounding and appears darker when viewed in visible light. The connection of sunspots to solar activity can be seen in figure 10, which compares two images taken at the same time, one in visible light and the other in ultraviolet (uv) light. The bright areas in the uv image indicate high activity and correspond almost exactly to the areas of sunspots, as seen in visible light.
Later in the century, in 1687, the first edition of Isaac Newton’s monumental *Principia Mathematica* was published. This historical book mathematically described the laws of motion and the universal law of gravitation. Significantly, it showed that the law of gravitation could be used to derive Kepler’s empirical laws of planetary motion. This could be viewed as the beginning of modern theoretical astronomy.

However, there was something troubling about the orbit of the planet Mercury that could not be entirely explained by Newton’s law of gravitation. In particular the observed orbital precession did not exactly match the calculations. It was suspected that there may be an unknown planet inside of Mercury’s orbit that was perturbing it and would be difficult to detect due to its proximity to the sun. In 1826 Heinrich Schwabe began a study in an attempt to understand this. It turned out the puzzle of Mercury’s orbit would not be solved until Einstein applied his model of general relativity to it. However, Schwabe became interested in studying sunspots, and 17 years of meticulous studies later he published a paper describing the sunspot cycle. The era of modern space climatology began to take form in 1843 with this discovery.

Today it is recognized that understanding the sun’s cyclical activity is an important aspect of modeling the space radiation environment. The record of sunspots dates back to the early 1600s, while numbering of sunspot cycles begins in 1749 with cycle number 1. Currently sunspot cycle 24 is nearly over. The sunspot cycle in recent times has been approximately 11±1 years long and can have significantly different levels of activity from one cycle to the next. This 11-year period is often considered to consist of 7 years of solar maximum when activity levels are high and 4 years of solar minimum when activity levels are low. In reality the transition between solar maximum and solar minimum is a continuous one but it is sometimes considered to be abrupt for convenience. The last 6 solar cycles of sunspot numbers are shown in figure 11 [WDC].

Figure 10. Images taken of the sun at the same time on February 3, 2002. The left image is in visible light and the right image is in ultraviolet light. Credit: ESA and NASA (SOHO)
Another common indicator of the approximately 11-year periodic solar activity is the solar 10.7 cm radio flux ($F_{10.7}$). This closely tracks the sunspot cycle. The record of $F_{10.7}$ began part way through solar cycle 18 in the year 1947. The sun’s influence on space climatology and space weather is pervasive. It is a source of solar protons and heavy ions, as well as trapped protons and electrons. Furthermore, it modulates these trapped particle fluxes as well as galactic cosmic ray fluxes entering our solar system. Galactic cosmic ray fluxes interact with the atmosphere and are the main source of atmospheric neutrons. These neutrons decay to protons and electrons and supply additional flux to the trapped particle population. The sun is either a source or a modulator of all energetic particle radiations in the near-Earth region. These radiations are discussed next in section V.

V. Modern Times – Space Climatology

The prior section brings us to the beginning of the era of space climatology. This modern era is shown by the timeline in figure 12. It is marked by the discovery of the space radiations and their impact on electronics that are used in spacecraft.

Galactic cosmic rays (GCR) were discovered in 1912 by Victor Hess using electrosopes in a balloon experiment [Ro64]. The penetrating power of this radiation was clear to Hess from these initial observations. It would turn out to be many orders of magnitude more energetic than particles emitted from radioactive materials, which were known at the time. Solar energetic particles were subsequently discovered by Scott Forbush in 1942 [Fo46]. It had been known for nearly 100 years prior that bursts of electromagnetic radiation could be emitted by the sun and have an effect on Earth communications but this was the first indication that energetic particles could also be a problem. Shortly after that the transistor was invented at Bell Telephone
Laboratories in William Shockley’s group [Ga13]. The launch of the first satellites, Sputnik I and II, by the Soviet Union in 1957 was followed by the launch of Explorer I and III by the United States in 1958. The Explorer satellites led to the discovery of the Van Allen Belts by James Van Allen [Ro64]. Research began to analyze the effects of radiation on bipolar transistors, primarily for Department of Defense applications. With the beginning of this work the first Nuclear and Space Radiation Effects Conference (NSREC) was held at the University of Washington in 1964 [Ga13], [Pe13]. By 1975 SEU was reported to occur in spacecraft [Bi75], although it was apparently observed three years prior to this by the same group [Pe13a]. The NSREC was continuing to expand and held its first Short Course in 1980 [Pe13]. The Radiation and its Effects on Components and Systems (RADECS) Conference began in 1989. By 1991 the NSREC had recognized the importance of space environment research and began to include an environment session in the conference. Twenty-seven more years along the timeline brings us to the current conference in 2018.

Figure 12. Time line for modern space climatology from the year 1900 to the present and its relation to radiation effects conferences.

With this perspective, the following discusses modern space climatology with emphasis on the items in figure 12 highlighted with blue arrows. Section A begins with a definition of space climatology and space weather. Sections B, C and D discuss properties, models and a current issue for each topic of galactic cosmic rays, solar particle events and the Van Allen Belts, respectively. Section E then applies the models discussed and shows examples of TID and SEU environments, including the effect of shielding.

A. Definition of Space Climatology and Space Weather

It is not difficult to find long and complex definitions of space climatology and space weather, especially the latter. These terms are generally defined here as the condition of the upper atmosphere and beyond, more specifically the conditions of the space radiation environment for a given location or orbit. For space weather the time period of interest is the short term, e.g., daily conditions, whereas for space climatology the time period is an extended
one such as a mission duration. This has implications for model use in the design and operation of spacecraft. Climatological models are used during the mission concept, planning and design phases of spacecraft in order to minimize mission risk. These are generally statistical models that allow risk projection well into the future over the mission duration. Space weather models are used during the launch and operation phases in order to manage residual risk. They are generally nowcast or short-term forecast models of the radiation environment. The following discussion deals mainly with the climatological aspects of the radiation environment.

B. Galactic Cosmic Rays

1. Properties

Galactic cosmic rays (GCR) are high-energy charged particles that originate outside of our solar system. Some general characteristics are listed in Table 1. They are composed mainly of hadrons, the abundances of which are listed in the Table [Xa13]. A more detailed look at the relative abundances compared to solar abundances is shown in figure 13. The two abundance distributions are generally similar. The main differences result from fragmentation of GCR ions that tend to smooth out the GCR distribution relative to the solar abundances. This is particularly noticeable for the elements Li, Be and B (Z=3 to 5), which are produced mainly from fragmentation of heavier GCR ions such as C and O in occasional collisions with interstellar hydrogen or helium. 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).

Table 1. Characteristics of Galactic Cosmic Rays.

<table>
<thead>
<tr>
<th>Hadron Composition</th>
<th>Energies</th>
<th>Flux</th>
<th>Radiation Effects</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>90% protons 9% alphas 1% heavier ions</td>
<td>Up to $\sim 10^{20}$ eV</td>
<td>1 to 10 cm$^{-2}$s$^{-1}$</td>
<td>SEE</td>
<td>LET</td>
</tr>
</tbody>
</table>
The amazing variation in energy range of GCRs is shown in figure 14 based on data compiled by Swordy [Sw01]. Energies can be up to the order of $10^{20}$ eV, although the acceleration mechanisms to reach such extreme energies are not understood. GCR with energies less than about $10^{15}$ eV are generally attributed to supernova explosions within the Milky Way galaxy and more recently neutron star collisions. These fluxes, on the order of a few ions cm$^{-2}$s$^{-1}$, are significant for SEE. On the other hand the origins of GCR with energies greater than about $10^{15}$ eV are largely unknown. It is often stated that the origin of GCR with energies beyond $10^{18}$ eV is extragalactic [Mr12]. A theoretical limit, the Greisen-Zatsepin-Kuzmin (GZK) limit [Gr66] shown in figure 14, is an upper limit in energy that a GCR proton cannot exceed if it travels a long distance as would occur if it originated in another galaxy. The reasoning is that the proton would interact with the omnipresent Cosmic Microwave Background (CMB) and lose energy to it. The CMB is residual electromagnetic radiation left from the Big Bang [Ma08]. However, this limit appears to have been exceeded many times and is a source of controversy. This illustrates how little is known about these ultra-high energy particles. Fortunately particle fluxes at these extreme energies are so low that they are not significant for SEE.
Figure 14. Differential flux vs. energy for GCRs [Sw01].

Typical GCR energy spectra for a few of the major elements during solar maximum and solar minimum are shown in figure 15 [Ba96]. The flux of the ions with energies less than about 20 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 the figure.
Figure 15. GCR energy spectra during solar maximum and solar minimum for some of the most abundant ions [Ba96].

For SEE analyses energy spectra such as those shown in figure 15 are often converted to LET spectra. Integral LET spectra for solar maximum and solar minimum conditions are shown in figure 16. 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 device sensitive volume this allows the flux of particles that deposit a given amount of charge or greater, and therefore an SEE rate, to be calculated in a simple approximation [Pe02]. The final manuscript in this Short Course by Brian Sierawski discusses the applicability and shortcomings of the LET parameter for calculating SEE rates in space.
Figure 16. GCR LET spectra for solar maximum and solar minimum conditions. From CREME96: [https://creme.isde.vanderbilt.edu](https://creme.isde.vanderbilt.edu).

The LET spectra shown in Figure 16 are applicable to geosynchronous missions where there is no significant geomagnetic attenuation. The Earth’s magnetic field, however, needs to be accounted for at altitudes lower than geosynchronous. Due to the basic interaction of charged particles with a magnetic field, the 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 or away from the Earth’s surface, which allows much deeper penetration of the incident ions. The effect of the geomagnetic field on incident GCR LET spectra is discussed for various orbits in the final manuscript of this Short Course and elsewhere [Ba97].

2. Models

There has been a long-time interest in developing models of GCR fluxes to aid in design of electronic systems, which began with James H. Adams’ development of the GCR model in the Cosmic Ray Effects in Microelectronics 1986 (CREME86) code [Ad81], [Ad87]. This section focuses on two popular models used for calculating SEE rates in space, although there are other interesting models that are available [Da01], [Ku17], [Le06], [Ma13].
One model is that developed by R. Nymmik of Moscow State University (MSU) [Ny96]. It is currently used in CREME96 [Ty97], the updated version of the 1986 suite of codes hosted on the Vanderbilt University website, https://creme.isde.vanderbilt.edu. The other is the Badhwar-O’Neill model developed at the NASA Johnson Space Center [ON10], [ON15]. The two models are based on the idea that the energy spectra of GCR ions outside of the heliosphere is given by Local Interstellar Spectra (LIS). A diffusion-convection theory of solar modulation is used to describe the GCR penetration into the heliosphere and transport to near Earth at 1 Astronomical Unit (AU). This solar modulation is used as a basis to describe the variation of GCR energy spectra over the solar cycle, as shown in figure 17 for iron ions [ON15]. Both models currently use sunspot numbers as input for solar activity leading to solar modulation. The implementation, however, is different. The MSU model uses multi-parameter, semi-empirical fits to relate the sunspot numbers to GCR intensity. The Badhwar and O’Neill model solves the Fokker-Planck differential equation for the solar modulation parameter as a function of sunspot number. This implementation and various sources of GCR data are described by Pat O’Neill in [Xa13]. Figure 18 shows a comparison of the two models with data. Although both of these models are successfully used for SEE applications, the Badhwar-O’Neill model incorporates a broader and more recent data base and is used extensively by the medical community.

![figure 17](image-url)

Figure 17. Illustration of solar modulation for GCR iron ions [ON15].

In previous sections the solar modulation of GCR flux has been described. Lower solar activity levels result in higher GCR fluxes. As shown in figure 11, the most recent complete solar minimum period between cycles 23 and 24, approximately centered at the year 2009, was quite “deep” and prolonged. In fact it was the deepest solar minimum of the space era and resulted in the highest GCR fluxes observed in this era. This has raised concerns about solar cycles trending toward this behavior and how elevated the GCR fluxes could get in the future [Ma13].

One of the advantages of basing the solar modulation on sunspot numbers is that there is a continuous detailed record of sunspots dating back to 1749. This allows the GCR fluxes to be estimated over this period of time that covers 24 solar cycles. An example is shown in figure 19 for 80 MeV/amu oxygen [ON15]. It is seen that over this extended period of time the peak flux values for each solar minimum have not varied by more than about 30%. The recent deepest minimum of the space era in 2009 can be compared to the deepest since 1750, which occurred in 1810. It can also be compared to the 1977 solar minimum that is used in CREME96 as a typical minimum. The GCR models should be adequate for design of electronic systems as long as appropriate consideration is given to the recent trend in GCR fluxes.
C. Solar Particle Events

1. Properties

The terms solar flare, coronal mass ejection (CME) and solar particle event are often used colloquially and interchangeably. These terms actually have different meanings so this section begins by describing them. Figure 20 is a schematic showing solar energetic particle production. These particles are likely energized by magnetic reconnection, a process that converts stored magnetic energy to kinetic energy, thermal energy and particle acceleration. One emission process of the sun is electromagnetic in nature. Irradiance is a comparatively low intensity emission that varies with the solar cycle. By contrast a solar flare is a burst of electromagnetic radiation characterized by a sudden brightening as shown on the right-hand side of figure 20. It turns out that solar flares are often, but not always, accompanied by solar energetic particles. The second general type of emission process of the sun is mass emission. The solar wind is a steady stream of plasma (a gas of free ions and electrons) consisting of protons, alpha particles and electrons in the eV to keV energy range and has an embedded magnetic field. A CME is a large eruption of plasma that carries an embedded magnetic field stronger than that of the solar wind. A CME image is shown on the left-hand side of figure 20. A CME that has a high enough speed will drive a shock wave that further accelerates particles. This is analogous to an airplane creating a shock wave if it exceeds the speed of sound. If the CME driven shock reaches Earth it can cause geomagnetic disturbances. CMEs are the second source of solar energetic particles, as shown in the figure. Further
properties of solar flares and CMEs are discussed in a review article by Reames giving a detailed account of the many observed differences [Re99].

![Diagram of the Sun and solar processes](image)

**Figure 20. Solar energetic particle production. Image credits: NASA and ESA**

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. Therefore the focus here is mainly on CMEs. The mass of magnetized plasma ejected in an extreme CMEs can be on the order of $10^{17}$ grams. CME speeds can vary from about 50 to 2500 km/s with an average speed of around 450 km/s. It can take anywhere from hours to a few days to reach the Earth. Table 2 lists some further general characteristics of CMEs.

<table>
<thead>
<tr>
<th>Hadron Composition</th>
<th>Energies</th>
<th>Integral Fluence ($&gt;10$ MeV/amu)</th>
<th>Peak Flux ($&gt;10$ MeV/amu)</th>
<th>Radiation Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>96.4% protons</td>
<td>Up to $\sim$GeV/amu</td>
<td>Up to $\sim10^{10}$ cm$^{-2}$</td>
<td>Up to $\sim10^6$ cm$^{-2}$s$^{-1}$</td>
<td>TID TNID SEE</td>
</tr>
<tr>
<td>3.5% alphas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sim0.1%$ heavier ions</td>
<td></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 TID and TNID that is due mainly to protons with a small contribution from alpha particles. Heavy ions are not abundant enough to significantly contribute to these cumulative effects. An extreme CME can deposit a few krad(Si) of dose behind 100 mils (2.5 mm) of aluminum shielding. Even though the heavy ion content is a small percentage of the total it cannot be ignored. Heavy ions, as well as protons and alpha particles in solar particle events, can cause both transient and permanent SEE.

The solar cycle dependence of both solar particle event and GCR fluxes is shown in figure 21 in which the differential flux of all carbon, nitrogen and oxygen ions in the 25 to 250
MeV/nucleon range is shown during the time period 1974 to 1996 [Ba97]. Superimposed in blue are the sunspot numbers during that time period illustrating the activity of solar cycles 21 and 22. The solar particle event fluxes are seen as the sharp spikes in the figure, which indicate the statistical and periodic nature of these events. Note that the events occur with greater frequency during the solar maximum time periods. They are superimposed on the low level background flux of GCR approximately on the order of $10^{-4} \text{ (cm}^2\text{-s-sr-MeV/n)}^{-1}$ that slowly varies with the solar cycle as discussed in section V.B. The GCR fluxes are approximately anti-correlated with the solar cycle.

![Differential flux of 25 – 250 MeV/nucleon C, N and O measured with IMP-8 spacecraft instrumentation between 1974 and 1996. Superimposed in blue are the sunspot numbers from solar cycles 21 and 22 [Ba97].](image)

**Figure 21.** Differential flux of 25 – 250 MeV/nucleon C, N and O measured with IMP-8 spacecraft instrumentation between 1974 and 1996. Superimposed in blue are the sunspot numbers from solar cycles 21 and 22 [Ba97].

2. **Models**

There have been a number of climatological models for solar particle events developed over the years for spacecraft design. Due to the stochastic nature of events confidence level based approaches have often been used to allow the spacecraft designer to evaluate risk-cost-performance trades for electronic parts [Xa06]. The first such model was King’s analysis of solar cycle 20 data [Ki74]. One “anomalously large” event, the well-known August 1972 event
dominated the fluence of this cycle so the model was often used to predict the number of such events expected for a given mission length at a specified confidence level [St74]. Using additional data a model from JPL emerged in which Feynman et al. showed the distribution of solar proton event magnitudes is continuous between small events and extremely large events such as that of August 1972 [Fe90]. The JPL model is a Monte Carlo based approach [Fe93]. Other probabilistic models followed based on more recent and extensive data. A model from Moscow State University introduced the full solar cycle dependence by assuming the event numbers are directly proportional to sunspot numbers [Ny99]. The NASA Emission of Solar Protons (ESP) and Prediction of Solar Particle Yields for Characterization of Integrated Circuits (PSYCHIC) models are based on Maximum Entropy Theory and Extreme Value Statistics [Xa99], [Xa00]. A summary of a number of statistical models is given in [Xa06a]. The European Space Agency (ESA) Solar Accumulated and Peak Proton and Heavy Ion Radiation Environment (SAPPHIRE) model using the Virtual Timelines method invokes a Levy waiting time distribution [Ji12] and continues to evolve [Ji18]. A new model is also under development that updated the data base of the ESP model [Ro18] and incorporates a new approach to solar cycle dependence of event numbers [Ad11].

a) Cumulative Fluence Models

Models for cumulative solar proton fluence are useful for evaluating damage due to TID and TNID. They can also be used to determine long-term SEE rates for devices vulnerable to protons. This can be helpful for estimating the probability of a destructive SEE over the course of a mission.

The most straightforward cumulative solar proton fluence model is ESP/PSYCHIC. It is based on measured annual proton fluences during solar maximum. An advantage of this approach is that it is not necessary to know specific details about the time series of events such as the waiting time distribution, for which there are different approaches [Ju07], [Ji12]. It is implicit in the data. This is shown in figure 22 where total fluences from 21 solar maximum years are shown as points for 3 different energies [Xa00]. This graph is shown on lognormal probability paper on which a lognormal distribution appears as a straight line. The fitted distributions can then be used to obtain the lognormal parameters for N-year distributions. An example result is shown in figure 23 for 10 years in geostationary Earth orbit (GEO). Solar protons are not attenuated significantly by the geomagnetic field in GEO. As is the case for all the climatological models discussed above the output spectra are obtained at a user specified level of confidence for the mission duration. The confidence level represents the probability that the calculated spectrum will not be exceeded during the mission.
Figure 22. Cumulative annual solar proton event fluences during solar maximum periods for 3 solar cycles plotted on lognormal probability paper. The straight lines are fits to the data [Xa00].
Figure 23. ESP/PSYCHIC model results for cumulative fluence over a 10 year period including 7 years during solar maximum in GEO. Energy spectra are shown for confidence levels ranging from 50 to 99%.

Comparison of the JPL, ESP/PSYCHIC and SAPPHIRE models is shown in figure 24 for a 2-year solar maximum period at the 95% confidence level [Ji18]. The JPL and SAPPHIRE models are both Monte Carlo based approaches. It is seen that the largest differences between models occurs at high proton energies. A new statistical model, the Ground Level Enhancement (GLE) model, is also shown [Ra17]. It is based on randomly sampling parameters from fitted proton spectra based on neutron monitor data analyzed by Tylka [Ty09]. This model is expected to have its best accuracy at high proton energies, where it is seen to fall in the middle of predicted fluences by the other models.

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. 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 [Fe93]. The remainder of the cycle is considered solar minimum. Fluences that accumulate during solar minimum can be found in a number of publications [Ny99], [Xa04], [Ji12].
Solar heavy ion models are not as advanced as solar proton models primarily because the data are much more limited. For microelectronics applications they are needed to assess SEE. The ESP/PSYCHIC cumulative fluence model for solar heavy ions is described in [Xa07]. Due to the limited data available the probabilistic model is limited to long-term (approximately 1 year or more) cumulative fluences and not worst case events. The approach taken was to normalize the alpha particle fluxes relative to the proton fluxes based on measurements of the Interplanetary Monitoring Platform-8 (IMP-8) and Geostationary Operational Environmental Satellites (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 normalized relative to the alpha particle energy spectra using measurements of the Solar Isotope Spectrometer (SIS) onboard the Advanced Composition Explorer (ACE) spacecraft for the 7 year solar maximum period of solar cycle 23. Remaining naturally occurring minor heavy elements in the Periodic Table are determined from measurements made by the International Sun-Earth Explorer-3 (ISEE-3) spacecraft or an abundance model. Example results for 2 years during solar maximum at the 50% (median) confidence level behind 100 mils of aluminum shielding are shown in figure 25.

LET spectra used for SEE analysis have a somewhat unusual shape. Figure 25 demonstrates that this shape is due to the elemental contributions. Interestingly, this can be
related back to the nucleosynthesis of elements in the Periodic Table described previously. The maximum LET that an ion can have in a material is called the Bragg Peak. Therefore on an LET plot such as figure 25, the fluence an ion contributes to the total LET spectrum drops sharply to zero at the Bragg Peak. For example, in silicon this occurs for protons at an LET less than 1. It is seen that protons and alphas produced in Big Bang nucleosynthesis contribute LET values to the total LET spectrum up to about 1 MeV-cm$^2$/mg. Elements formed in stellar nucleosynthesis contribute up to an LET of about 29 MeV-cm$^2$/mg, while elements formed from extreme event nucleosynthesis contribute over the full range of LET values.

![LET spectra for cumulative fluences of solar protons and heavy ions for 2 solar maximum years at the 50% confidence level behind 100 mils of aluminum shielding. The total fluence is multiplied by a factor of 1.5 for clarity. Also shown are the contributions to the total LET spectrum due to protons, alphas, Z = 3 (Li) to 26 (Fe), and Z = 27 to 92 (trans Fe) [Xa07].](image)

Figure 25. LET spectra for cumulative fluences of solar protons and heavy ions for 2 solar maximum years at the 50% confidence level behind 100 mils of aluminum shielding. The total fluence is multiplied by a factor of 1.5 for clarity. Also shown are the contributions to the total LET spectrum due to protons, alphas, Z = 3 (Li) to 26 (Fe), and Z = 27 to 92 (trans Fe) [Xa07].

b) Worst Case Event Models

Another consideration for spacecraft design is the worst case solar particle event that occurs during a mission. It is important to know how high the SEE rate can get during such an event. The most straightforward approach is to design to a well-known large event. The
radiation effects community most often uses the October 1989 event while the medical community uses the August 1972 event. Hypothetical events such as a composite of the February 1956 and August 1972 events have been proposed [An94]. There are also event classification schemes in which the magnitudes range from “small” to “extremely large” that may be useful [Ny96a], [St96]. At one time the so-called Carrington Event of 1859 was widely quoted as being a worst case event over the last 400 years based on the nitrate record in polar ice cores [Mc01]. However, the glaciology and atmospheric communities disagreed with this interpretation, as the Carrington Event was not observed in most ice cores [Wo12]. Although this event resulted in a severe geomagnetic storm it is now recognized that the solar proton fluences for this event are not reliably known.

The commonly used October 1989 event is provided for use as a worst case scenario in the CREME96 suite of codes at three levels of solar particle event intensity [Ty97]. They are the “worst week”, “worst day” and “peak flux” models based on proton measurements from the GOES-6 and -7 satellites and heavy ion measurements from the University of Chicago Cosmic Ray Telescope on the IMP-8 satellite. The peak flux model covers the highest 5-minute intensity during the event. Comparisons of these models have been made with data taken by the Cosmic Radiation Environment Dosimetry (CREDO) Experiment onboard the Microelectronics and Photonics Test Bed (MPTB) during a very active period of solar cycle 23 [Dy02]. The data show that 3 major events during this time period approximately equaled the “worst day” model. An example of this is shown by the LET spectra in figure 26.

Figure 26. Comparison of a major 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 [Dy02].
Another approach to worst case event models is to use statistical methods. The idea is analogous to cumulative fluence models where a worst case event would be calculated for a given confidence level and mission duration. There have been several methods proposed for this including extreme value statistics [Xa98], [Xa99], semi-empirical approaches [Ny99], and Monte Carlo calculations [Ji12], [Ji18].

The field of extreme value statistics is one with both an extensive theoretical and applied history. It has frequently been used to describe extreme environmental phenomena such as floods, earthquakes, high wind gusts, and very appropriate for NSREC 2018, volcanoes [Gu58], [An85], [Ca88]! It has turned out to be a useful radiation effects tool when applied to large device arrays such as high density memories [Mc00], gate oxides [Va84], [Xa96], and sensors [Bu88], [Ma89].

Extreme value statistics focuses on the largest or smallest values taken on by a distribution. Thus, the “tails” of the distribution are the most significant. Here the focus is obtaining the extreme value distribution of a random process when information is known about the initial distribution.

Suppose that a random variable, \( x \), is described by a probability density \( p(x) \) and corresponding cumulative distribution \( P(x) \). These are referred to as the initial distributions. Figure 27 shows an initial probability density for a Gaussian distribution [Bu88]. If a number of observations, \( n \), are made of this random variable there will be a largest value within the \( n \) observations. The largest value is also a random variable and therefore has its own probability distribution. This is called the extreme value distribution of largest or maximum values. Examples of these distributions are shown in the figure for \( n \)-values of 10 and 100. Note that as the number of observations increases the extreme value distribution shifts to larger values and becomes more sharply defined. The extreme value distributions can be calculated exactly for any initial distribution. The probability density for maximum values is

\[
\begin{align*}
  f_{max}(x;n) &= n[P(x)]^{n-1} p(x) \\
  F_{max}(x;n) &= [P(x)]^n
\end{align*}
\]

(1)

The corresponding cumulative distribution of maximum values is

\[
F_{max}(x;n) = [P(x)]^n
\]

(2)
Figure 27. Extreme value distributions for $n$-values of 10 and 100 compared to the initial Gaussian distribution [Bu88].

As $n$ becomes large, the exact distribution of extremes may approach a limiting form called the asymptotic extreme value distribution. If the form of the initial distribution is not known but sufficient experimental data are available, the data can be used to derive the asymptotic extreme value distribution. There are 3 types of asymptotic extreme value distributions of maximum values – the type I or Gumbel, type II and type III distributions [Gu58], [An85], [Ca88].

Next, an example will be shown that is directly related to the previous discussion of figure 27 before returning to the radiation environment. This is a problem about device processing distributions first identified as a radiation effects issue by Tim Oldham [Ol93]. It is also related to Marc Gaillardin’s presentation on process variations and radiation effects in devices in this Short Course.

Suppose a threshold voltage adjust implant is used on an array of $10^6$ NMOS transistors to tune the threshold voltage to the desired level. Post-implant measurements on a limited number of devices show threshold voltages with a Gaussian distribution having a mean of 700 mV and standard deviation of 5.1 mV. This is shown in figure 28. The problem is how to
determine the expected maximum and minimum threshold voltages in the array without making 10^6 measurements. This has significance for TID applications because the threshold voltage of an irradiated device can determine its radiation tolerance. The dose at which the array of transistors first begins to fail is more important than the dose at which the average transistor fails. Applying the methodology described above results in the extreme value distributions of maximum and minimum values for n = 10^6 transistors, shown in the figure [Xa96]. Notice how far from the mean the extreme values appear for this large number of devices. The expected minimum and maximum threshold voltages in the array are 676 and 724 mV, respectively, close to 5 standard deviations from the mean value.

![Figure 28. Probability densities for the initial Gaussian distribution and extreme value distributions of maximum and minimum values of threshold voltage for 10^6 transistors after a threshold voltage adjust implant [Xa96].](image)

With this background the problem of worst case event models for solar particle events is now considered. In order to determine a worst case event probabilistically, either by extreme value theory or by Monte Carlo simulation, information about the initial distribution must be known. The first description of the complete initial distribution was determined using Maximum Entropy Theory [Xa99]. This is a mathematical procedure for making an optimal selection of a probability distribution when the data are incomplete by avoiding the arbitrary introduction or assumption of information that is not available. It can therefore be argued that this is the best
choice that can be made using the available data [Ja57], [Ka89]. The result is a truncated power law in the distribution of event magnitudes, shown in figure 29 for the case of > 30 MeV proton event fluences. This describes the essential features of the distribution. The smaller event sizes follow a power law and there is a rapid falloff for very large magnitude events. Note that the figure also shows the October 1989 event used as a worst case event in CREME96. A variant of this distribution has subsequently been proposed [Ny07] but there is no significant improvement in the overall fit to data [Ji18], resulting in the use of both functional forms.

![Graph showing the comparison of the truncated power law distribution to 3 solar cycles of data during solar maximum [Xa99].](image)

Figure 29. Comparison of the truncated power law distribution to 3 solar cycles of data during solar maximum [Xa99].

Given the initial distribution of event magnitudes such as the one shown in figure 29, the extreme value method can be applied to obtain a worst case event over the course of a mission. However, this situation is a little more complex. The number of events that occur during a mission is variable, so this must be taken into account. If it is assumed the event occurrence is a Poisson process [Fe93] the worst case distribution can be calculated according to [Xa98a], [Bo08]. Example results are shown in figure 30 for > 30 MeV proton event fluences [Xa99]. The probability of exceeding the fluence shown on the y-axis equals one minus the confidence level.

An interesting feature of this model is the “design limit” shown in the figure. A reasonable interpretation 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 an objectively determined engineering guideline for use in limiting design costs.

Figure 30. Probability for worst-case event proton fluences expected during the indicated time periods during solar maximum [Xa99].

Other worst case statistical models have been developed for both solar proton event fluences and peak fluxes [Xa98], [Xa98a], [Ny99], [Ji12], [Ji18]. There are worst case event statistical models for heavy ions but these are limited due to the lack of data [Ny99], [Ji18a]. There is also a probabilistic model for solar electrons that is part of an interplanetary electron model [Ta11].

3. Current Issue: Use of Statistical Models vs. Worst Case Observations

As seen in the last section there are two types of approaches for evaluating worst case solar particle events. One is to use a statistical model to calculate the worst case event that will occur during the mission at a specified level of confidence. The other is to use a worst case observation such as the event that occurred in October 1989, as in CREME96. Figure 29
illustrates this difference. This section compares the approaches and discusses their advantages and disadvantages.

A statistical model uses the entire data base of events. There is much to consider as events can have very different characteristics in terms of magnitudes (fluence or peak flux), time duration, energy spectra and heavy ion content. On the other hand the worst case observation approach is straight forward and based on a single well characterized event. As a result the proton and heavy ion characterization of the event are self-consistent for a worst case observation. This is not necessarily true for the worst case statistical model in which the proton and heavy ion fluxes are analyzed independently. For example, proton and heavy ion fluxes can peak at different times, leaving open different approaches to what characterizes the worst case.

Another interesting comparison is that the statistical model allows the designer to make risk, cost, performance trades when selecting electronic parts. For example, a higher risk environment can be designed to in return for a higher performance or less expensive part. When using a worst case observation such as the October 1989 event there is little flexibility in the design environment. This can make requirements difficult to meet for high risk missions such as CubeSats. Thus, considering the type of mission can be important for deciding on an approach.

Lastly, it is worth noting the current state of development of these models. The worst case observation approach has a long history of successful use. Worst case statistical models for solar protons are also successfully used. However, heavy ion models are a developing area of research.

D. The Van Allen Belts

1. Trapped Particle Motion in the Magnetosphere

The Earth’s magnetosphere consists of both an external field due to the solar wind and an internal magnetic field. The internal or geomagnetic field originates primarily from within the Earth and is approximately a dipole field. The solar wind and its embedded magnetic field tends to compress the geomagnetic field. During moderate solar wind conditions, the magnetosphere terminates at roughly 10 Earth radii on the sunward side. During turbulent magnetic storm conditions it can be compressed to about 6 Earth radii. The solar wind generally flows around the geomagnetic field and consequently the magnetosphere stretches out to a distance of possibly 1000 Earth radii in the direction away from the sun.

Figure 31 shows the geomagnetic field, which is approximately dipolar for altitudes of up to about 4 or 5 Earth radii. It turns out that the trapped particle populations are conveniently mapped in terms of the dipole coordinates approximating the geomagnetic field. This dipole coordinate system is not aligned with the Earth’s geographic coordinate system. The axis of the magnetic dipole field is tilted about 11.5 degrees with respect to the geographic North-South axis and its origin is displaced by a distance of more than 500 km from the Earth’s geocenter. The standard method is to use McIlwain’s \((B,L)\) coordinates \([Mc61]\). Within this dipole coordinate system, \(L\) represents the distance from the origin in the direction of the magnetic equator, expressed in Earth radii. One Earth radius is 6371 km. \(B\) is simply the magnetic field strength. It describes how far away from the magnetic equator a point is along a magnetic field line. \(B\)-values are a minimum at the magnetic equator and increase as the magnetic poles are approached. Further background information on the magnetosphere and \((B,L)\) coordinates can be found in \([Wa94]\, [Ba97]\, [Ma02]\, [Bo08]\.
Next the basic motion of a trapped charged particle in this approximately dipole field will be discussed. Charged particles become trapped because the magnetic field can constrain their motion. As shown in Figure 32 the motion a charged particle makes in this field is to spiral around and move along the magnetic field line. As the particle approaches the polar region the magnetic field strength increases and causes the spiral to tighten. Eventually the field strength is sufficient to force the particle to reverse direction. Thus, the particle is reflected between so called “mirror points” and “conjugate mirror points”. Additionally there is a slower longitudinal drift of the path around the Earth that is westward for protons and eastward for electrons. This is caused by the radial gradient in the magnetic field. Once a complete azimuthal rotation is made around the Earth, the resulting toroidal surface that has been traced out is called a drift shell or $L$-shell. A schematic of $L$-shells is shown in Figure 33. The $L$-shell parameter indicates magnetic equatorial distance from Earth’s center in number of Earth radii and represents the entire drift shell. It is important to understand because it provides a convenient global parameterization for a complex population of particles.
Figure 32. Motion of a charged trapped particle in the Earth’s magnetic field. After E.G. Stassinopoulos [Ba97].

Figure 33. Illustration of the geometry of L-shells.
2. Trapped Protons

a) Properties

Some of the characteristics of trapped protons and their radiation effects are summarized in Table 3 and shown in figure 34. The $L$-shell range is from about $L = 1.14$ at the inner edge of the trapped environment out beyond geosynchronous orbits to an $L$-value of around 10. Trapped proton energies extend up to the GeV range. The energetic trapped proton population with energies $> 10$ MeV is confined to altitudes below 20,000 km, while protons with energies of a few MeV or less are observed at geosynchronous altitudes and beyond. The maximum flux of $> 10$ MeV protons occurs at an $L$-value around 1.7 and exceeds $10^5 \text{ cm}^{-2} \text{s}^{-1}$. The atmosphere limits the belt to altitudes above about 200 km. Trapped protons can cause TID, TNID and SEE.

<table>
<thead>
<tr>
<th>$L$-Shell Values</th>
<th>Energies</th>
<th>Fluxes* ($&gt;10$ MeV)</th>
<th>Radiation Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.14 – 10</td>
<td>Up to $\sim$GeV</td>
<td>Up to $\sim 10^5 \text{cm}^{-2} \text{s}^{-1}$</td>
<td>TID, TNID, SEE</td>
</tr>
</tbody>
</table>

*long-term average
Figure 34 is a contour plot of the trapped proton population with energies > 10 MeV shown in a dipole coordinate system. The x-axis is the radial distance along the geomagnetic equator in units of Earth-radii while the y-axis is the distance along the geodipole axis, also in units of Earth-radii. Thus, a y-value of zero represents the geomagnetic equator. A semi-circle with a radius of one centered at the point (0,0) represents the Earth’s surface. It is seen that it is a particularly convenient way to reduce a large quantity of information and get an overview of the particle population on a single plot.

Trapped proton fluxes in Low Earth Orbit (LEO) are approximately anti-correlated with solar cycle activity. This is most pronounced near the belt’s inner edge as shown in figure 35 [Hu98]. Here $F_{10.7}$, the solar 10.7 cm radio flux, is used as a proxy for solar activity. As solar activity increases the atmosphere expands and causes greater losses of protons to the atmosphere during solar maximum. In addition there is a decreased production of protons in the atmosphere during solar maximum coming from the Cosmic Ray Albedo Neutron Decay (CRAND) process. The CRAND process is the production of atmospheric neutrons from GCR that subsequently
decay to protons (and electrons) and can become trapped. As discussed previously, GCR fluxes are lower during solar maximum.

Figure 35. Approximate anti-correlation of low altitude trapped proton flux (points) with $F_{10.7}$ as an indicator of solar activity (gold curve) [Hu98].

For spacecraft that have an orbit lower than about 1000 km the so-called “South Atlantic Anomaly” (SAA) dominates the radiation environment. This anomaly is due to the fact that the Earth’s geomagnetic and rotational axes are tilted and shifted relative to each other as discussed before. Thus, part of the proton belt’s inner edge is at lower altitudes as shown in figure 36. This occurs in the geographic region around South America. It is shown in figure 37 as a contour plot on geographic coordinates for > 35 MeV proton fluxes at an altitude of about 840 km [Jo15].
Figure 36. Illustration of the South Atlantic Anomaly (Da96).
Figure 37. Contour plot of proton fluxes > 35 MeV in the SAA at an altitude of about 840 km measured by the Polar Orbiting Earth Satellite (POES) from July 1998 to December 2011 [Jo15].

Higher energy protons are generally fairly stable in the proton belt. However, during the 1990-1991 Combined Release and Radiation Effects Satellite (CRRES) mission the Air Force Research Laboratory (AFRL) discovered the formation of a transient proton belt in the $L$-shell 2 to 3 region [Gu96]. It is now known that CMEs can cause geomagnetic storms that suddenly reconfigure the belt. Figure 38 shows that enhanced fluxes can occur in the $L$-shell 2 to 3 region if a CME is immediately preceded by another event [Bo08]. Note that although the enhanced flux begins to decay immediately it can remain measurable for well over a year. The figure also shows that a CME can cause reduction of an enhanced flux. The details of these belt reconfigurations are not fully understood.
Figure 38. Sudden changes in 9.65 to 11.35 MeV trapped proton fluxes caused by solar particle events measured on the Satellite for Scientific Applications (SAC-C) [Bo08].

b) Models

The general approach to a trapped particle model calculation is to first use an orbit generator to obtain the geographical coordinates of the spacecraft – latitude, longitude and altitude. Next the geographical coordinates are transformed to a dipole coordinate system in which the particle population is mapped. The trapped particle environment is then determined external to the spacecraft. The Space Environment Information System (SPENVIS) suite of programs has implemented a number of trapped particle models for unrestricted use at http://www.spenvis.oma.be/.

The well-known Aerospace Proton-8 (AP-8) trapped proton model is the eighth version of a model development effort led by James Vette. Over the years these empirical models have been indispensable for spacecraft designers and for the radiation effects community in general. The trapped particle models are static maps of the particle population during solar maximum and
solar minimum based on data from the 1960s and 1970s. Because these models provide the mean flux values of the environment, a Radiation Design Margin (RDM) is used for design specifications. Details of the AP-8 model and its predecessors can be found in [Sa76], [Ve91].

The shortcomings of AP-8 and the need for updates have been discussed [Da96]. Consequently there have been a number of notable efforts to develop new trapped proton models [Gu96], [Hu98], [He99], [Hu02], [Bo14]. Comparisons of these models with AP-8 and each other for different orbits are given by Lauenstein and Barth [La05].

Recently more comprehensive models have been developed. One such model was initially called AP-9 and is now undergoing a name change to the International Radiation Environment Near Earth (IRENE) model [Gi13], [Jo15]. AP9/IRENE allows 3 methods of calculation. There is a statistical model for the mean or percentile environment. There is a perturbed model that adds measurement uncertainty and data gap filling errors. Thirdly, there is a Monte Carlo capability that includes space weather variations. AP9/IRENE is based on data taken between 1976 and 2016. It does not include solar cycle variation, i.e., output is averaged over the solar cycle. As a result of its probabilistic approach and use of percentiles, confidence levels can be used for design specifications. The other recent comprehensive model is the Global Radiation Earth Environment model (GREEN) [Si18]. GREEN is an integration of AP-8 with other models that have been developed in order to expand the overall energy and orbital capabilities. Results for the GREEN model were not available at the time of this writing.

Figure 39 is a comparison of AP-8 and AP-9/IRENE for a polar LEO. The orbital parameters used were those of the Landsat-8 satellite. This provides a reasonable overall comparison as the spacecraft flies through varying portions of the proton belt multiple times each day. Although there are large differences between the models at energies less than 1 MeV, these energies are not significant for most applications. Over most of the remaining energy range the AP8 model shows higher fluxes during solar minimum compared to solar maximum, as expected, while AP9/IRENE generally results in the highest fluxes. AP9/IRENE also extends to higher energies, which is due to the incorporation of the NASA Van Allen Probes data.
3. Trapped Electrons

a) Properties

Some of the characteristics of trapped electrons are summarized in Table 4 and shown in figure 40. There is both an inner and an outer zone or belt of trapped electrons. These two zones are very different so the characteristics are listed separately. As is also the case for trapped protons the boundaries of the zones are not sharp and they are to some extent dependent on particle energy. For the purposes of this discussion the inner zone is assumed to be between $L$-values of 1 and 2. It was originally thought that electron energies range up to approximately 5 MeV but that has not been observed recently. This electron population tends to remain relatively stable but a long-term average is difficult to ascertain as will be seen in section c). The outer zone has $L$-values ranging between about 3 and 10 with electron energies generally less than about 10 MeV. Here fluxes peak between $L$-values of 4.0 and 4.5 and the long-term average value for $> 1$ MeV electrons is roughly $3 \times 10^6$ cm$^{-2}$ s$^{-1}$. This zone is very dynamic and the fluxes can vary by orders of magnitude from day to day. An interesting feature of the outer belt is that it extends down to low altitudes at high latitudes. Trapped electrons contribute to both TID and TNID effects.
Figure 40. Trapped electron fluxes $> 1$ MeV according to the AE-8 model during solar maximum [Bo08].

Table 4. Trapped Electron Characteristics.

<table>
<thead>
<tr>
<th>Zone</th>
<th>L-Shell Values</th>
<th>Energies</th>
<th>Fluxes* ($&gt; 1$ MeV)</th>
<th>Radiation Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inner Zone</td>
<td>1 - 2</td>
<td>Up to 5 MeV?</td>
<td>uncertain</td>
<td>TID</td>
</tr>
<tr>
<td>Outer Zone</td>
<td>3 - 10</td>
<td>Up to 10 MeV</td>
<td>Up to $3 \times 10^6$ cm$^{-2}$ s$^{-1}$</td>
<td>TNID</td>
</tr>
</tbody>
</table>

* long-term average
The distribution of trapped particles is a continuous one throughout the inner and outer zones. However, between the two zones is a region where the fluxes are at a local minimum during quiet periods. However, the fluxes are quite variable. This is known as the slot region. The location of the slot region is assumed to be between $L$-values of 2 and 3 for this discussion. This is an attractive one for certain types of missions due to the increased spatial coverage compared to missions in LEO.

**b) Models**

The long-time standard model for trapped electrons has been the Aerospace Electron-8 (AE-8) model [Ve91], [Ve91a]. It consists of two static flux maps of trapped electrons – one for solar maximum and one for solar minimum conditions (see figure 40). Due to the variability of the outer zone electron population, the AE-8 model is valid only for long periods of time. A conservative rule of thumb is that it should not be applied to a period shorter than 6 months.

A feature of the outer zone is its high degree of volatility and dynamic behavior. This results from geomagnetic storms and substorms, which cause major perturbations of the geomagnetic field. Measurements from the Upper Atmosphere Research Satellite (UARS) illustrate the high degree of variability of electron flux levels prior to and after such storms. Figure 41 shows the electron energy spectra for $3.25 < L \leq 3.5$ after long-term decay from a prior storm (day 235) and two days after a large storm (day 244) compared to the average flux level over a 1000 day period [Pe01]. It is seen for example, at 1 MeV, that the difference in the one-day averaged differential fluxes over a 9-day period is about 3 orders of magnitude. This illustrates the difference between the long-term average space climate and the short-term space weather in the outer zone.
Figure 41. Total electron flux before and after a geomagnetic storm compared to a long-term average as measured onboard the UARS [Pe01].

Due to the volatile nature of the outer zone, it seems natural to resort to probabilistic methods. This is the case for the new AE-9/IRENE trapped electron model [Gi13], [Jo15], which uses the same methodology as described before in the discussion on trapped protons. Other statistical analyses have also been used for both the outer zone and slot region [Wr00], [Ko01], [Pe01], [Br04]. Another approach used to describe outer zone fluxes has been to relate them to the level of disturbance of the geomagnetic field by using geomagnetic activity indices such as $A_p$ [Br92] and $K_p$ [Va96].

An important orbit in the outer zone that is widely used for telecommunications satellites is GEO. Figure 42 shows a comparison between the AE8 and AE9/IRENE mean values. AE8 has no solar cycle dependence in GEO so there is no distinction between solar maximum and solar minimum, as was the case in figure 39. It is seen that AE8 gives more conservative fluxes over most of the energy range. The group at ONERA, the French National Aerospace Research Center, has also done considerable work on trapped electron models for GEO. Their most recent model is IGE-2006 [Si06], which gives the option of a maximum (worst case), mean or minimum (best case) flux output. When calculation of the mean flux is done in SPENVIS and compared to figure 42, results show lower fluxes than both AE8 and AE9/IRENE except at
energies approximately less than 0.1 MeV. However, the IGE-2006 model has been incorporated into the group’s new comprehensive GREEN model for trapped electrons so more detailed comparisons are deferred until GREEN becomes available for use.

Figure 42. Comparison of the AE8 and AE9/IRENE (version 1.5) models for GEO.

Figure 43 gives a good overall view of the dynamic behavior of trapped electrons for about a 3.5 year period as measured by Van Allen Probes instrumentation [Cl17]. Fluxes of 0.75 MeV electrons are mapped out according to $L$-shell values as a function of time. Color coding of electron intensities are shown along the top of the graph. The 2 boxed areas indicate the most severe storm periods. The figure shows the volatile nature of the outer zone ($L > 3$). During storm periods electrons can be injected into the slot region ($2 < L < 3$). Here they are fairly short-lived as the decay period is about 10 days. During severe storms electrons can also be injected into the inner zone ($1 < L < 2$). Note the stability of the inner zone as the injected electrons decay away very slowly and persist strongly more than a year after the storm.
Figure 43. Fluxes of 0.75 MeV electrons mapped according to \( L \)-shell as a function of time for approximately 3.5 years. Fluxes are background corrected [C117].

c) **Current Issue: The Case of the Missing Electrons**

Figure 43 is a good indicator of the behavior of the electron belts in recent times for energies up to about 0.75 MeV. The inner zone is fairly stable for long periods of time, as evidenced in the figure. When high energy (> 1.5 MeV) electron data are similarly examined as shown in the top portion of figure 44 [C117], nothing looks out of the ordinary. The outer belt looks volatile and the inner belt appears stable. While inner zone fluxes predicted by models in current use such as AE8 and AE9/IRENE are not large for energies between 1.5 MeV and a maximum of about 5 MeV, they are ordinarily accounted for in radiation effects analysis. However, the top portion of the figure has not been corrected for background contamination, which is mainly due to high energy protons. The Van Allen Probes instrumentation has improved capability in this regard and when background contamination is removed the result is shown in the bottom portion of figure 44. The high energy electrons of the inner zone are almost completely gone! In fact there is no evidence of > 1.5 MeV electrons in the inner zone since the Van Allen Probes were launched in 2012. This is the case of the missing electrons.
Figure 44. Fluxes of 1.58 MeV electrons mapped according to $L$-shell as a function of time for approximately 3.5 years. The top graph is uncorrected for background contamination and the bottom graph is corrected. Note the difference in the inner zone ($1 < L < 2$) [Cl17].

The question of what happened to this portion of the inner zone remains. Instrumentation prior to the Van Allen Probes has not had the same capability for analyzing background contamination. It therefore seems fairly certain that some of the older data reported as trapped electrons were actually due to high energy proton contamination. In addition the situation may also reflect a difference in time periods. The injection of > 1.5 MeV electrons into the inner zone may require extreme magnetic storms while the storms during the Van Allen Probes era have been fairly mild.

Finally in this section the question of how this affects TID requirements is addressed. As an example the LEO corresponding to the Hubble Space Telescope is examined and presented in figure 45. Electron fluence-energy spectra are shown calculated with 2 models. The first is the AE8 model, which consists of older data from the 1960s and 1970s. The other is AE9/IRENE, which is based on Van Allen Probes data and CRRES data for the inner zone. The CRRES data includes the severe storm of March 1991. It is seen that the models agree well out to about 1 MeV. Above energies of 1 MeV it is not surprising from the above discussion that the AE8 model shows higher fluxes. Analysis of TID behind 2.5 mm of aluminum shielding for the Hubble orbit shows that if AP8/AE8 is used electrons contribute less than 20% of the TID. If AP9/AE9/IRENE is used electrons contribute less than 2% of the TID. Thus, although inner belt electrons present an interesting scientific challenge they are unlikely to drive radiation effects problems except possibly surface effects.
Figure 45. Comparison of the AE8 and AE9/IRENE (version 1.5) models for the LEO of the Hubble Space Telescope.

E. Example Environments Including Shielding

1. Total Ionizing Dose

In this example a total ionizing dose vs. shielding depth curve will be obtained. A highly elliptical orbit is chosen because it is exposed to all particle populations that contribute significantly to TID – trapped protons, trapped electrons and solar protons. The orbital parameters used were those of the first portion of the NASA Magnetic MultiScale (MMS) mission. Shielding calculations were done for a solid aluminum sphere geometry with dose in silicon calculated at the center of the sphere. In a later presentation of this Short Course by Renaud Mangeret, discussion of more complex shielding at the box level will be presented.

There are now 2 basic options for methods of design specification. The first is the traditional margin-based approach. In this case the dose-depth curve is calculated and a margin is applied to that. Although different margins are used for different applications, government organizations use a margin of times 2 for many situations. This is shown by the red curve in figure 46 where the AP8, AE8 and ESP/PSYCHIC models were used. The newer method is a confidence level based approach, where the dose-depth curve is calculated for a given level of confidence. This necessitates the use of the AP9/IRENE and AE9/IRENE models because they are the only trapped particle models with this capability. The blue curve in figure 46 shows results using the newer trapped particle models along with the ESP/PSYCHIC model to calculate the dose-depth curve at the 95% confidence level. No margin is applied to this. It is seen that the results agree well out to about 6 mm of aluminum shielding. Beyond this the difference is
primarily due to higher proton flux levels predicted by AP9/IRENE and secondarily the fact that the newer models extend to higher proton and electron energies. For those readers interested in transitioning to the confidence level based approach, dose-depth curves at the 95% confidence level are fairly consistent with using a mean environment with times 2 margin for various orbits.

Figure 46. Dose-depth curves for a highly elliptical orbit using 2 specification methods. The orbit of 1.2 x 12 Earth radii (1274 km perigee x 70.080 km apogee) at a 28.5 degree inclination includes trapped protons, trapped electrons and solar protons contributing to TID.

Figure 47 shows the different radiation contributions to the dose-depth curve in figure 46. The results show the 95% confidence level but the situation is similar for the case of times 2 margin for the same orbit. For lightly shielded electronics the trapped electron component dominates TID. However, electrons can be effectively shielded against and at higher levels of shielding the trapped proton contribution is the major contributor. Solar protons and bremsstrahlung radiation contribute at a low TID level for the range of shielding shown. The latter is photon emission due to the slowing down of electrons in material. It should be noted that these results are particular to the spacecraft orbit considered.
The confidence level based TID approach has several advantages over the traditional margin based approach. When convolved with laboratory test data it allows the device TID failure probability to be calculated for a mission at the desired level of shielding [Xa17]. An example of this is shown in figure 48 for bipolar transistors used for high speed, low power applications for several orbits. This is a better characterization of a device radiation performance in space. It allows more systematic trades during the design process. It is also amenable to reliability analyses, which is not possible if only a TID margin is known.
2. Single Event Upset

Next SEE environments are considered. The examples presented here are restricted to single event upset (SEU) data and calculations. Figure 49 shows SEU data from the Air Force Research Laboratory SeaStar spacecraft detected on a solid state recorder for more than 4 years [Po02]. The spacecraft was in a polar LEO. The SEU count per day is shown on the y-axis. There is a slowly varying background of upsets due to trapped protons and GCR. In this case it is believed most of these SEU were due to trapped protons. Superimposed on this background are sharp increases in the upset rate due to radiation bursts from solar particle events. The largest event spikes were due to the July 14-15, 2000 and November 9, 2000 events. In addition rate spikes due to subsequent smaller events are also seen. Although the environment here is different than what is observed in figure 21, note the general similarity in that the effects are due to background radiation that varies slowly with solar cycle superimposed with solar particle events.
FIGURE 49. SEU count per day for a solid state recorder on AFRL’s SeaStar spacecraft in a polar LEO. Data were obtained for more than 4 years beginning in January 1999. Upsets due to the solar particle events of July and November 2000 are identified [Po02].

Finally, calculated SEU rates are shown for the same highly elliptical orbit considered previously for figures 46 and 47. In this orbit the calculations must account for GCR and solar heavy ions. Additionally if the device is sensitive to proton-induced upset, solar protons and trapped protons must also be considered. SEU rates are shown in figure 50 that were calculated for a 4 Gbit NAND flash memory [Pe10]. The sensitive volume was obtained from process reverse engineering and publicly available data. It is seen that increased shielding reduces SEU rates for the worst case solar particle event, the October 1989 event, used in the CREME96 suite of programs. Proton-induced upsets, both those caused by solar protons and trapped protons, can also be reduced with increased shielding. However, the GCR rates are fairly constant with increased shielding. Thus, the GCR rate provides a lower limit for the SEU rate that is not practical to reduce significantly. The rates provided here for heavy ions do not include fragmentation processes in shielding. For discussion of this the reader is referred to [ON15a].
Figure 50. SEU rates calculated for a 4 Gbit NAND flash memory for a worst case solar particle event, solar protons, trapped protons and GCR during solar minimum and solar maximum as a function of shielding [Pe10].

VI. Summary

This work presented a space climatology timeline ranging from the Big Bang to NSREC 2018. It began with a description of the early universe including the origin and abundances of particles significant for radiation effects. It continued to a transition period to modern times when the era of modern space climatology began to emerge due to discoveries of sunspots and the solar activity cycle, along with development of early astronomical methods. The timeline concluded in the modern era with a description of the radiation environments of galactic cosmic rays, solar particle events and the Van Allen Belts.

A general theme is that the space radiation environment is highly variable and must be understood to produce reliable, cost-effective designs for successful space missions. This includes both long-term variations of space climate and short-term variations of space weather.
VII. Acknowledgments
The author thanks Pat O’Neill of Johnson Space Center for providing the calculations and data for figures 17 – 19. He further thanks Craig Stauffer of AS&D, Inc. for assistance with calculations and Martha O’Bryan of AS&D, Inc. for assistance with graphics.

VIII. References


