Solar energetic particles: Is there time to hide?

Don Reames

NASA/Goddard Space Flight Center, Greenbelt, MD 20771, USA

Received 13 March 1998; received in revised form 3 December 1998

Abstract

In the large solar energetic particle (SEP) events that constitute a serious radiation hazard, particles are accelerated at shock waves driven out from the Sun by coronal mass ejections (CMEs). A self-regulating mechanism of wave formation by the streaming particles limits SEP intensities early in the event. Hazardous intensities do not occur until the arrival of the shock itself. This provides an opportunity to warn astronauts to take shelter after the onset of the event at the Sun and before arrival of the shock, a time of ~12 h or more. The actual time history of particle intensities depends strongly on the longitude of the event at the Sun, on the width the CME, and especially on the speed of the shock. Fortunately, hazardous events are relatively rare. Unfortunately, this gives us few events to study, so we are forced to extrapolate knowledge gained at lower energies in the frequent smaller events. It is essential that the spacecraft with our best instrumentation be positioned outside the Earth’s magnetosphere where they can observe these rare large events when they do occur. © 1999 Published by Elsevier Science Ltd. All rights reserved.

1. Introduction

It is well known that large solar energetic-particle (SEP) events can be an extreme threat to astronauts on deep space missions such as those to the Moon or Mars, and even in high-latitude regions of the orbits of space stations. In fact, it is usually assumed that shielding will be provided in some small region of the spacecraft where astronauts can go to seek protection in times of severe radiation hazard. However, we are unlikely to be able to predict the onset of SEP events in the near future, so there is a critical question of providing a reliable and timely warning for the astronauts to seek shelter before they receive a high radiation dose. This problem is aggravated when members of the crew are involved in extra-vehicular activities (EVAs) or exploration of lunar or planetary surfaces.

We have learned a lot about SEP events in the last few years, in fact, there has been a complete change in the paradigm of particle acceleration in these events (Reames, 1993, 1995, 1997). It was once believed that the particles were accelerated in solar flares and somehow diffused across solar magnetic fields and out into interplanetary space. This monolithic “solar flare myth” (Gosling, 1993) has been replaced because we now have measurements that distinguish the flare particles from those accelerated at shock waves driven into interplanetary space by huge coronal mass ejections (CMEs).

Particles from solar flares have unusual abundances of elements and isotopes produced by wave–particle interactions and they are highly ionized by the hot flare plasma where they originate (Reames, Meyer and von Rosenvinge, 1994; Reames, 1990, 1995). Most of the accelerated particles are confined to magnetic loops in the flare region and they eventually plunge into the solar atmosphere at the footpoints of the loops to produce X-rays, γ-rays and heat, which contributes to the brilliant flash phase of the flare. Those few particles
accelerated on open magnetic field lines stream out to Earth over a limited cone of solar longitude. Because of their limited intensities and duration, impulsive-flare events do not constitute a radiation hazard. We will not consider them further in this paper.

The events of greatest interest here are the large “gradual” events where an expanding CME drives a shock wave that can easily cross magnetic field lines and accelerate particles as it goes, first in the solar corona and then outward into interplanetary space. To reliably accelerate particles above ~10 MeV, the shock speed must exceed ~750 km/s (Kahler et al., 1987; Reames et al., 1997a), thus, it is only the largest and fastest ~1% of CMEs that produce SEP events. In the largest of these, shock speeds reach ~2500 km/s and, as we will see, particle intensities can increase as the 4th or 5th power of the shock speed. The ionization states of Fe in large gradual events indicate that Fe from unheated material at ambient coronal temperatures (Boberg et al., 1996) is being accelerated to energies of at least 200–600 MeV/amu (10–30 GeV) (Tylka et al., 1995). Proton acceleration up to 21 GeV does not reach peak intensity until the shock is 5–10 solar radii from the Sun (Kahler, 1994). At lower energies, acceleration can continue into the heliosphere far beyond the orbit of Earth.

Fortunately, the enormous SEP events, which can be fatal to astronauts, occur only rarely. However, this rarity makes them especially difficult to study or to predict from statistics alone. Our greatest success has come from studying particles at lower (1–10 MeV) energies that are produced in a large number of events. If we can fully understand the physics of that process and the structure and evolution of small events, then we can begin to extend that knowledge to ~100 MeV particles in the largest events, since those particles obey the same physical laws. This paper takes a first step in that direction.

2. Shock acceleration

In this section we discuss the physics of particle acceleration at CME-driven shocks. Initially, we will assume that an observer remains on a single magnetic flux tube as the shock wave begins accelerating particles at the solar end of the tube and continues to do so as it propagates out to the Earth at 1 AU. A shock with a speed of ~1600 km/s will make the trip in about one day. By comparison, protons of ~10 MeV have a speed of about 1 AU/h so they stream rapidly out to the observer while the shock is still near the Sun. Scattering of these particles by irregularities in the magnetic field has some effect on particle transport to 1 AU. However, scattering has a large effect on transport and acceleration near the shock where, as we shall see, both magnetic turbulence and particle intensities become large. We will defer to the next section the discussion of the spatial structure of these events that causes the particle intensities to change as the observer’s connection point is swept across a shock surface of varying strength. Resulting differences in the intensity–time profiles of events from eastern and western solar longitudes will be discussed at that time.

Some particles from the super-thermal tail of the distribution function in the solar corona will be scattered back and forth across a newly formed shock by the ambient magnetic turbulence. These particles will gain an increment in velocity on each traversal of the shock. Some of these particles will begin to stream away from the shock. This streaming distribution is unstable to the production and amplification of Alfvén waves in the plasma, a process that is well-described in standard texts on plasma physics (e.g., Stix, 1962; Melrose, 1980). These new waves begin to scatter more particles of the similar energy that come behind and thus reduce the streaming. This increased scattering...
efficiently traps particles near the shock where they are further accelerated. As particles of higher and higher energy stream away they produce additional self-generated waves with frequencies in resonance with their magnetic gyrofrequency. Lee (1983) examined this process of acceleration at low energies and found that an equilibrium is established in which the intensities of particles and waves decreased with distance from the shock. Sufficiently far from the shock, there are not enough particles to produce enough waves to stop the streaming and the particles stream freely away.

The particle intensity–time profile seen by an observer of the process we have described is shown in Fig. 1a. Early in the event, particles are seen streaming out from the shock that is still near the Sun. If the energetic particles and the source or “seed population” have the same radial divergence as they come out, the profile will remain flat. Eventually the shock itself will pass, accompanied by the particles that are trapped near it by self-generated waves. Historically this peak near the shock has been called an “energetic storm particle (ESP) event” if the shock is propagating quasi-parallel to the magnetic field and called a “shock spike” if the shock is propagating quasi-perpendicular to (across) the field. We will see that both can exist in different spatial regions (solar longitudes) of a single event.

It is these high particle intensities at the time of shock passage that pose the greatest radiation hazard to astronauts. Fluences of > 30 MeV protons that occurred at shock passage in the 1972 August 4 event approaching $10^7 \text{ cm}^{-2}$ would have been a serious hazard without shielding. Doses of 3765 rem (skin) and 91 rem (depth) were calculated from the observed protons (McKinnon, 1972), although instrument saturation could be a factor in these calculations. The role of the shock and the late arrival of the particles were not understood at that time and intensities predicted from flare observations (“the flare myth”) fell far short of observations. The officially-predicted probability of a proton event by the NOAA/Boulder forecast center at 2200 UT on August 3, less than 24 h before peak intensity in the August 4 event, was 20% (McKinnon, 1972).

If the shock strength or efficiency decreases with time and distance from the Sun, profiles like that in Fig. 1b are produced. In small or weak events, accelerated particle intensities may decrease and the trapping structure may dissipate before the shock arrives at the observer; in these cases the intensity bump at the shock will not be seen. This will occur first at high energies, then at low, because there are always fewer high-energy particles to produce resonant waves. Generally, the presence of a shock spike or ESP event means that the shock is still actively accelerating particles of that particular energy as it passes.

Fig. 2 shows a superposition of the intensity–time profiles of 3–6 MeV protons observed by Helios 1 are shown for 6 large events. Streaming-limited flat profiles are seen early in the events. Shock peaks are seen to arrive later in several of the events with intensities increasing by factors up to ~100 at the shock (after Reames, 1990).

Fig. 2. Superposed intensity–time profiles of 3–6 MeV protons observed by Helios 1 are shown for 6 large events. Streaming-limited flat profiles are seen early in the events. Shock peaks are seen to arrive later in several of the events with intensities increasing by factors up to ~100 at the shock (after Reames, 1990).
particles out from the Sun. These numerical calculations were not specific to shock acceleration but simply considered a source of arbitrary intensity near the Sun and followed the evolution of the energetic particles and waves as a function of time and distance outward. The maximum intensity of ~1 MeV particles at several distances from the Sun are shown in Table 2 of Ng and Reames (1994) as a function of source intensity.

Fig. 3 shows this maximum plateau intensity near Earth as a function of source intensity. As the source intensity increases, the observed plateau intensity initially increases linearly, then rolls over to a broad maximum, and finally even declines slightly. If there were no wave growth, we would find the linear relationship shown in Fig. 3. As the source intensity increases, wave growth increases, the scattering mean-free-path declines, and the broad plateau of the streaming limit is produced. The calculations of Ng and Reames (1994) do not show the late peak at shock passage since their source is assumed to remain near the Sun.

Fig. 4 shows intensity–time profiles at 3 different energies for 6 large SEP events during the last solar cycle measured on the NOAA/GOES spacecraft. First, note that the profiles for the 1989 October 19–22 event are quite similar in appearance at all energies, even above 100 MeV, to the low-energy profiles in Fig. 2 and the profile in Fig. 1a. As the shock speed increases, high-energy particles begin to behave like the low-energy particles seen in smaller events. Specifically, high-energy particles are accelerated out to 1 AU and beyond, and their intensity peaks near the time of shock passage. Second, note that the 8.7–14.5 MeV and 39–82 MeV protons, each have nearly identical plateau intensities in all of the events. These are the streaming limits at these energies and are indicated by the dashed horizontal lines running across the entire figure. The 110–500 MeV channel appears to reach the streaming limit in the first 4 events, but not in the later 2. As the shock strength and speed increases, higher and higher energies are driven to the streaming limit (Reames and Ng, 1998).

It is important to realize that the streaming limit does not apply to the intensities at the shock. The shock is the particle source. The streaming limit only applies to particles that must propagate from the source to the observer along interplanetary magnetic field lines. However, the streaming limit does provide a strategy for warning astronauts of an impending radiation hazard. Streaming-limited fluences of >30 MeV protons (<10^7 cm^-2) are not a serious hazard for astronauts. This provides a time of ~12 h (depending on shock speed), after the onset of the event, for the astronauts to seek shelter before arrival of the shock. We do not need to predict the onset of the event at the Sun; after the onset, we do need to predict the strength and arrival time of the shock and the intensity of particles near it.

3. SEP spatial distributions

The spatial distribution of the energetic particles is controlled by the spatial configuration of the CME and the shock wave ahead of it. Kahler et al. (1984, 1987) found that the SEP intensities are correlated with the speed and width of the CME. Cane et al. (1988) examined the time profiles of 235 proton events as a function of the solar longitude of the source, in effect, mapping the spatial structure of a typical SEP event.

The particle view of a typical CME-driven shock structure is shown in Fig. 5 for observers at 3 solar longitudes. Ahead of the shock, magnetic field lines are drawn into a spiral pattern by the expanding solar wind from the rotating Sun. The pattern is modified behind the shock by the faster plasma of the CME. As this structure expands radially outward, a fixed observer encounters field lines that connect him to a point on the shock that moves further and further to the east with time. Thus, in Fig. 5, when the observer on the east flank of the shock first views an event at solar...
longitude W51°, he is extremely well connected to the nose of the shock as it leaves the Sun. However, when the shock arrives at Earth, this observer crosses it 51° to the east of the nose where the shock is much weaker and intensities are much lower. Thus, he sees intensities that rise rapidly early in the event and then decline as he connects to a weaker and weaker part of the shock. In contrast, an observer on the west flank of the shock views an event at E45° solar longitude in the figure. This observer begins to see particles when the shock comes around the corona and crosses his field line. However, the intensity he sees does not peak until after he has passed through the local shock, 45° west of the nose, and encounters the cluster of compressed field lines behind shock that connect him to the strong acceleration region near the nose of the shock from behind. If the CME is very wide, an observer near central meridian will be connected to a shock that varies little in strength with time and he will see the flat profiles we described in the previous section.

Studying the longitude variation of particle profiles, even for statistically large sample of events, does not tell us how the absolute intensities vary with longitude in a single event or how this pattern varies with CME width and speed. For this we must view a single event with multiple spacecraft spaced around it. Obviously, the number of opportunities to do this are limited, especially because of poor tracking coverage of interplanetary spacecraft. However, the late 1970s provided a unique opportunity for observations with the IMP 8 and Helios 1 and 2 spacecraft (Reames et al., 1996, 1997a).

Fig. 6 shows observations of a small CME event in 1979 March by all 3 spacecraft. The inset in the figure shows the location of the 3 spacecraft relative to the CME which is projected downward from the Sun, just as it was shown in the drawing in Fig. 5. In the upper panel of Fig. 6, the intensity profile at Helios 1 shows the slow rise to a streaming-limited plateau, followed by a peak at shock passage, just as we would expect.
for an event close to central meridian. Helios 2 and IMP 8, further to the west, see eastern sources and show a dramatically slower and slower rise to maximum even at modestly different longitudes. Note however, that all 3 of the intensity profiles merge late in the event. This late spatial and spectral invariance is shown the lower panels of the figure where spectra taken early and late in the event are contrasted.

Fig. 7 shows profiles similar to those in Fig. 6, but for the much larger and more powerful event of 1978 September 23. Unfortunately, there are also more data gaps. Here the shock itself is seen at spacecraft separated by an angle of 160° and the intensity profiles show only modest differences despite the huge angular span. Spatially invariant spectra are again seen late in this large event.

The phenomenon of spatially invariant spectra that decline slowly in time is related to particles that are essentially trapped in an expanding magnetic bottle in the vicinity of the quiescent eastern flank of the shock (Reames et al., 1997a). In this region, data from the WIND spacecraft show spectral shapes that remain invariant over orders of magnitude in energy for periods of several days (Reames et al., 1997b).

Another characteristic feature of the spatial distribution of energetic particles around a shock is the bidirectional streaming distributions seen on the loops behind the CME (e.g., Richardson and Reames, 1993).

Fig. 8 shows the location of the energetic particle characteristics we have discussed, superposed on a map of the CME and shock. The precise location of the different particle populations and their relative importance depends on the speed and width of the CME and shock. Despite their different names, shock spikes and ESP events form a continuous distribution across the face of the shock. The energy at which these peaks become large depends upon shock speed.

4. SEP intensity and shock speed

Kahler et al. (1984, 1987) showed that the SEP intensity was correlated with both the angular width and the speed of a CME. Recently (Reames et al.,
1997a) reexamined a list of CMEs and associated shocks that were obtained as the SOLWIND coronagraph near Earth observed the CMEs and Helios 1, located off the solar limbs, measured the shock (Sheeley et al., 1985). This configuration minimized the projection effect of the CME speed. The original list was produced with no reference to SEPs, so it is unbiased in that respect.

Fig. 9 shows the correlation of peak 3–6 MeV proton intensity with CME speed and with the in situ shock speed. Correlation coefficients for the two cases are 0.54 and 0.82, respectively. Despite the favorable location of the Helios 1 spacecraft off the solar limbs, it is clear that the SEP intensities are better correlated with shock speed than with the measured CME speed. Factors like identifying the appropriate portion of the CME for speed measurement and the projection of the features against the plane of the sky still seem to affect the correlation. In fact, if we remove the 1979 May 27 event where the CME speed is 270 km/s and the shock speed 605 km/s, the correlation coefficient for the intensities vs CME speed rises from 0.54 to 0.70. Spatial (longitude) variations and evolution of the shock between the Sun and observer have also not been taken into account.

Nevertheless, it is clear that the peak particle intensities are a strong function of the shock speed and that shock speed is a good predictor of these intensities. We cannot easily extend these studies to higher energy protons because the local shock we observe has usually
weakened to the point where it no longer accelerates high-energy protons in these relatively small events.

The importance of shock speed cannot be overemphasized. A fast shock is required to accelerate the particles in an SEP event. Late on 1997 Jan 6 a CME was launched from the Sun and its dense plasma and huge magnetic cloud arrived at Earth on Jan 10, producing large geomagnetic storms and spectacular aurorae. The event was widely reported in newspapers and the popular press and 50 scientific papers on the event were presented at the spring AGU meeting. However, no SEP event whatsoever was observed, even with the extremely sensitive detectors on the WIND spacecraft. The shock, with an average transit speed of <490 km/s, was inadequate to accelerate particles.

The physics of geomagnetic storms and SEP events are completely different. Only the fastest ~1% of CMEs produce SEP events. CMEs and magnetic plasma clouds will not harm astronauts in route to Mars unless the shock they produce can accelerate high-energy particles.

5. Radial gradients

The study of Ng and Reames (1994) also determined the radial gradient of the streaming-limited plateau, a dependence that was very close to $R^{-3}$. To examine this experimentally Reames and Ng (1998) measured the plateau intensity of protons, at 3–6 MeV on Helios.
Fig. 8. The location of energetic particle features is shown on the map of a CME and shock.

Fig. 9. Peak intensity of 3–6 MeV protons at Helios 1 are shown as a function of the associated CME speed (left panel) and shock speed (right panel). Correlation coefficients are shown in each panel (data from Reames et al., 1997a).
1 and 2 and at 4.2–8.7 MeV on GOES, of all events in which the intensity exceeded $10^{-1}$ $(\text{cm}^2 \text{ sr} \text{ s MeV})^{-1}$. A plot of the observed radial distribution of the plateau intensity is shown in Fig. 10, along with the calculated values from Table 2 of Ng and Reames (1994). While the statistical sample is not large enough to completely determine the radial dependence close to the Sun, the observations are clearly consistent with theory.

Unfortunately, at higher energies, too few events reach plateau intensity to verify their radial gradient experimentally. Nevertheless, it is clearly safer to go to Mars than to Venus from the standpoint of SEP radiation. The $R^{-3}$ dependence is likely to apply to the shock peaks and the streaming plateau, but not to the invariant spectral region behind the shock where the radial gradients are small.

6. Future prospects

Our ability to provide a clearer strategy for astronaut protection depends upon a clearer understanding of the physics of SEP events, and of the underlying parameters involved. Two paths to this understanding are improved observations and better theories and models. Since the time scale between observations of large events is long, theory may provide results more rapidly.

First, we should explore the streaming limit and its energy dependence in greater detail, with emphasis on determining the underlying parameters that control variations in the plateau intensity. Next, we must try to understand the connection between properties of the shock and SEP intensities along the shock surface so
that remote measurements of shock, such as radio type II bursts, can be used to predict peak particle intensities. Modeling means different things to different people, but detailed models of the plasma evolution driven by the CME and of SEP acceleration in the same environment could provide a framework for a joint understanding of both plasma and SEP observations.

Since large SEP events are rare, it seems essential that we should not miss one if it occurs. Poor coverage of spacecraft and the practice of redefining their missions and orbits have had a disastrous effect on the continuity of CME and SEP observations. SEP experiments generally have lower priority than planetary, imaging or solar wind experiments. In 1983, the ISEE 3 spacecraft, designed for local interplanetary observations, was hijacked for a mission to a comet. The low tracking coverage of the spacecraft (1–2 h/day) resulted in the effective loss of an entire solar cycle of SEP data. In 1985, the Air Force actually shot down the only operating coronagraph (SOLWIND) in a weapons test. In 1997 the magnetospheric community made a strong effort to hijack the WIND spacecraft, 3 years after its launch in 1994, and move it into the inner geomagnetic tail. This occurred before WIND had an opportunity to observe any large SEP events of the new solar cycle. Fortunately, disaster was temporarily averted and WIND was placed in orbits where SEP measurements are obscured only part of the time. The complement of spacecraft near Earth can provide excellent new measurements of SEP events during the next solar maximum, if we can resist the schizophrenic compulsion to redefine missions at every whim.

Sometimes it is naively assumed that all energetic-particle instruments are equivalent. However, a high-speed instrument with onboard identification of particles by energy and species at rates of over 10^4 particles/sec is different from an isotope experiment that must telemeter each particle measurement to the ground for careful analysis. Angular distributions, that play an important role in understanding SEP transport, are easily measured when the spacecraft spin axis is normal to the ecliptic. Spacecraft that do not spin or those with spin axis in the ecliptic usually do not permit measurement along the direction of the magnetic field. IMP 8 instruments provide clean background-free measurements of much better quality than those on GOES; however, IMP 8 measurements saturate in large events while those at GOES do not. One can only admire NOAA’s relentless commitment to provide continuous coverage.

There is presently no counterpart to the fleet of spacecraft in the inner heliosphere that existed in the 1970. SEP events cannot be imaged remotely; the only way to study spatial distributions of SEPs is with in situ measurements at multiple locations. Such instruments should be capable of measuring protons and the dominant elements up through Fe above 1 MeV/amu (or lower if possible). They should provide onboard analysis at extremely high rates without saturation. They should be carried on spinning spacecraft with spin axes normal to the ecliptic so they can measure angular distributions.

A different kind of issue that is of long-range concern is the small number of research scientists actively studying SEP events. There are probably less than a dozen scientists, representing 4 or 5 laboratories, who have actively published papers on these events in the refereed literature over the last 5–10 years. This number continues to decline as more effort placed elsewhere. By comparison, there are probably hundreds of scientists studying the magnetosphere and the effects of geomagnetic storms. Often, SEP event studies are used as a justification in proposals for new missions but receive little support thereafter. One must ask whether any knowledgeable scientists will still be available in a decade or so when deep space missions are about to begin.

Under the circumstances, the manned program would be wise to seriously consider its own program of small reliable satellites placed at different solar longitudes in the inner heliosphere. The emphasis would be on providing continuous coverage of SEP events and interplanetary plasma (shocks) during the early phases and on providing a real-time SEP warning system during manned missions. The SOHO coronagraphs and the Radio Mapping Experiment on WIND, or their successors, should also be components of an SEP warning system.

7. Conclusions

The answer to the question we posed in the title seems to be yes. Protons streaming outward early in SEP events generate waves that scatter the particles and impede their flow. Proton intensities on the streaming-limited plateau are a minimal radiation hazard to astronauts. Hazardous intensities can occur when the CME-driven shock wave arrives at the spacecraft and the astronauts must be shielded at that time, nominally ~12 h or more after the event onset at the Sun. This means that it is not necessary to attempt to predict the onset of an event before it occurs, a feat that is presently beyond our capabilities. It is only necessary to measure or predict the intensities at the shock before it arrives at the spacecraft, a nontrivial but more tractable problem. Additional theoretical effort is required to confirm this result and establish the degree of constancy of the plateau intensities.

The best parameter for determining the peak particle intensities is the shock speed, but variations in the
magnetic connection to the shock and the shock history on a given flux tube blur the correlation with speed. This is especially true at higher energies where the strongest acceleration occurs close to the Sun in small events. In general, we have made our greatest progress by studying low energy particles in smaller, but more numerous, events. There is no other practical option. When sufficiently large events do occur, the higher-energy particles seem to confirm the behavior we expect. High-quality theory and models can provide powerful assistance and allow us to extrapolate to higher energies and larger events with greater confidence.

The spacecraft presently available are beginning to provide significant new measurements over an expanded energy range during the new solar cycle. This can continue only if they are allowed to remain outside the magnetosphere, in interplanetary space. These measurements are beginning to probe new areas of the physics of acceleration and transport. Improved understanding of the spatial structures and topology must come from study of the historical multi-spacecraft database because new spacecraft far from Earth in the inner heliosphere will not be available during the next solar maximum.

Acknowledgements

I would like to express my thanks to C. K. Ng for many helpful discussions and to NOAA for providing the CD-ROM containing GOES data.

References


