Preface

Editors' notes

This volume is being developed over the course of several years of the Heliophysics Summer School, starting with the first chapter in 2012. Chapters are being added as they become available from the authors/lecturers, after which this volume will be completed as the 5th in the Heliophysics series. This volume will be available as a freely accessible online volume to complement the four printed Heliophysics volumes published by Cambridge University Press. We recommend that the reader occiasionally check the School's website (see below) for updates. Until the volume is complete, the numbering of chapters, figures, and tables is subject to change.

Additional resources

The texts were developed during a summer school series for heliophysics, taught at the facilities of the University Corporation for Atmospheric Research, in Boulder, Colorado, funded by the NASA Living With a Star program. Additional information, including text updates, lecture materials, (color) figures and movies, and teaching materials developed for the school can be found at http://www.vsp.ucar.edu/Heliophysics. Definitions of many solar-terrestrial terms can be found via the index of each of the first four volumes; a comprehensive list can be found at http://www.swpc.noaa.gov/info/glossary.htm.

Heliophysics

helio-, pref., on the Sun and environs, from the Greek helios.
physics, n., the science of matter and energy and their interactions.

Heliophysics is the

- *comprehensive new term for the science of the Sun - Solar System Connection.*
- *exploration, discovery, and understanding of our space environment.*
- *system science that unites all of the linked phenomena in the region of the cosmos influenced by a star like our Sun.*

Heliophysics concentrates on the Sun and its effects on Earth, the other planets of the solar system, and the changing conditions in space. Heliophysics studies the magnetosphere, ionosphere, thermosphere, mesosphere, and upper atmosphere of the Earth and other planets. Heliophysics combines the science of the Sun, corona, heliosphere and geospace. Heliophysics encompasses cosmic rays and particle acceleration, space weather and radiation, dust and magnetic reconnection, solar activity and stellar cycles, aeronomy and space plasmas, magnetic fields and global change, and the interactions of the solar system with our galaxy.

From NASA's "Heliophysics. The New Science of the Sun - Solar System Connection: Recommended Roadmap for Science and Technology 2005 - 2035."

Space weather

Space weather refers to the variable state of the coupled space environment related to changing conditions on the Sun and in the terrestrial atmosphere, specifically those conditions that can influence the performance and reliability of space-borne and ground-based technological systems, and that can directly or indirectly endanger human well-being.

1

Introduction

Carolus J. Schrijver et al.

[Text excerpted from Schrijver et al. (2015), with minimal modifications.]

Space weather is a real and permanent hazard to society that needs to be, and can be, addressed by combining scientific research with engineering ingenuity: protecting society from space weather requires that we adequately understand the physical processes of space weather, that we characterize the conditions to which technological infrastructures need to be designed, that we learn to effectively forecast space weather, and that the consequences of acting on such forecasts are accepted as necessary for the protection of societal infrastructure.

Societal use of, and dependence on, ground-based electrical systems and space-based assets has grown tremendously over the past decades, by far outpacing population growth as society continues to grow its electrical/electronic and space-based technologies. Global electricity use has increased by a factor of about 1.6 over the 15-year period between 1997 and 2012 (International Energy Agency, 2013). The global satellite industry revenue has multiplied by a factor of about 4.2 over that period (to US$190 billion per year for 2012, part of a total value of US$304 billion for the overall space industry, with over 1,000 operating satellites from over 50 countries; Satellite Industry Association, 2013). In contrast, the global population grew by approximately 20% over that period (Population reference bureau, 2013), demonstrating our increasing use of electrical power and satellite-based information per capita.

With that growth in electrical/electronic and space-based technologies comes increasing vulnerability to space weather: where a century ago the main risk was associated with the telegraph systems we now see impacts in the electric power grid, in satellite functionality, in the accuracy of navigation and timing information, and in long-range high-frequency (HF; cf. Table 5.1, and also Fig.

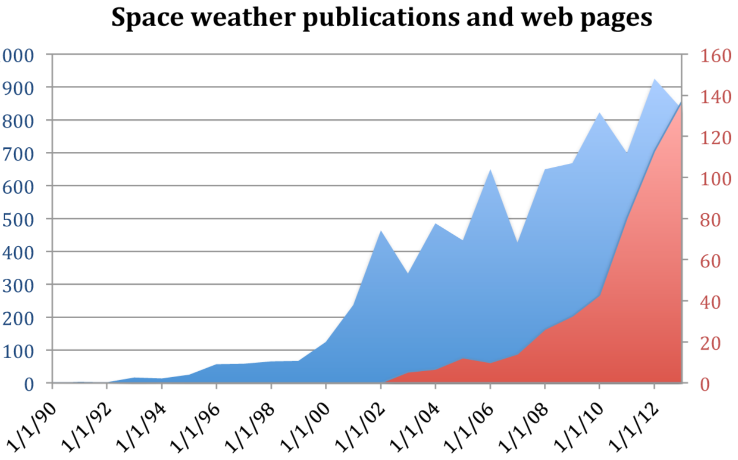

Fig. 1.1. Number of publications per year with "space weather" in the abstract in NASA/ADS (blue; left axis), and the number of web sites returned by a Google search for "space weather" within calendar years (since 2003) in thousands (red; right axis).

4.1 in Heliophysics Volume II) radio communication. We see an increasing interest in understanding space weather impacts and the threats these pose are spread over a variety of civilian sectors [...]. Selected reports on these impacts (that themselves provide information on more literature on the subject) are compiled in an on-line resource list† that accompanies this report; that resource list also includes a glossary of solar-terrestrial terms‡, and links to a National Geographic introduction to space weather accessible via YouTube§, and lectures related to space weather, its impacts, and its science in the NASA Heliophysics Summer School¶.

The reality of the threat to society posed by space weather is increasingly acknowledged - reflected, for example, in the exponential growth of the number of web pages on space weather (totaling over 130,000 new entries in 2013; see Figure 1.1) and in the number of customers subscribing to alert and forecast services (exceeding, for example, 40,000 for the US Space Weather Prediction Center). A core difficulty facing any study that attempts a cost-benefit analysis for space weather is inadequate knowledge of the technological and economic impacts of ongoing space weather and of the risk posed by extreme space storms. This hampers the quantitative identification of the most significant impacts of space weather and consequently the prioritization of the

---

† http://www.lmsal.com/∼schryver/COSPARrm/SWlibrary.html
‡ http://www.swpc.noaa.gov/info/glossary.html
§ http://www.lmsal.com/∼schryver/COSPARrm/SWlibrary.html#youtube
¶ http://www.vsp.ucar.edu/Heliophysics/

research areas and the deployment of infrastructure to protect against space weather. [A]chieving a quantification of the SWx impact on societal technologies is important for the process of allocating the required resources for research and forecasting, and of determining what sectors of society should be involved in appropriating which resources. For example, although the threat posed by geomagnetic storms is broadly recognized as real (e.g., Krausmann, 2011; Langhoff and Straume, 2012), establishing the vulnerability and consequences of such an event has proven to be very difficult (Space Studies Board, 2008; DHS Office of Risk Management and Analysis, 2011; JASON, 2011), hampering a cost-benefit assessment of investments that could make impacted systems less vulnerable by suitable engineering or by improved forecasting (DHS Office of Risk Management and Analysis, 2011).

Like terrestrial weather, space weather manifests itself as a variety of distinct phenomena, and like terrestrial weather, it ranges from benign to extremely severe. Most frequently, space weather is very weak in intensity with apparently little impact on technology. Strong, severe, or extreme geomagnetic conditions (as measured by the Kp index, on NOAA's G scale† occur only 5% of days through a solar magnetic cycle. Even though there are no reports of catastrophic failures in the US high-voltage power grid, there is an increase by 40%±20% in insurance claims for industrial electrical and electronic equipment on the 5% most geomagnetically active days (as measured by the rate of change in the geomagnetic field strength) relative to quiet days, and there is an increase of 30%±10% in the occurrence frequency of substantial disturbances in the US high-voltage power grid (Schrijver et al., 2014). Overall, approximately 4% of the disturbances in the US high-voltage power grid reported to the US Department of Energy are attributable to strong but not extreme geomagnetic activity and its associated geomagnetically induced currents (GICs; Schrijver and Mitchell, 2013).

Other aspects of space weather can adversely affect satellites. Severe solar energetic-particle (SEP) storms, for example, impact satellites directly, while expansion of the terrestrial upper atmosphere by magnetospheric variability (through Joule heating) may affect low-orbiting satellites by modifying their orbits through increased drag which is an issue for on-orbit operations, collision avoidance, and could eventually lead to early re-entry (e.g., Rodgers et al., 1998). A series of such storms during the 2003 October-November time frame, for example, saw considerable impacts on satellites through electronic single-event upsets (SEUs), solar-array degradation, modified orbit dynamics for spacecraft in low-Earth orbits, and noise on both housekeeping data and instrument data. Another manifestation of space weather is the enhance-

---

† http://www.swpc.noaa.gov/NOAAscales/

ment of radiation belt (RB) particles and of magnetospheric plasma that cause charging/discharging phenomena or state upsets in satellite electronics. For 34 Earth and space science missions from NASA's Science Mission Directorate, for example, 59% of the spacecraft experienced such effects (Barbieri and Mahmot, 2004). A graphic laboratory demonstration of discharging inside dielectric materials (a critical space weather impact for satellites in geosynchronous and middle Earth orbit) is available on YouTube‡.

A third impact for space weather occurs via severe modification of trans-ionospheric signals by highly variable plasma density in space and time, thus affecting customers of GNSS services. The economic impact of this type of space weather has yet to be investigated, being complicated by the fact that it will mostly occur well downstream of the immediate service providers and also by the fact that GNSS technology is rapidly evolving even as the total numbers of users and uses increases, increasingly in layered applications that may hide just how GNSS-dependent a system is.

The threat posed to society by the most severe space storms that occur a few times per century is largely unknown and the magnitude of such a threat is consequently highly uncertain: the technological landscape evolves so rapidly that our modern-day highly-interconnected societal infrastructure has not been subjected to the worst space storms that can occur. Some reports put the threat by the most severe space storms among the significant threats faced by our technology-dependent society. Geomagnetic disturbances (GMDs) on electrical systems have been known to impact technology for over 150 years, starting with the telegraph systems, and showing a clear correlation with the sunspot cycle (Boteler et al., 1998). Among the insurance-industry reports that review the space weather risk landscape from the industry perspective (e.g., Hapgood, 2011), one (Lloyd's, 2013) concludes for the US in particular that "the total population at risk of extended power outage from a Carrington-level storm [i.e., unusually strong but likely to occur approximately once per century] is between 20-40 million, with durations of 16 days to 1-2 years", while recognizing that even for weaker storms "the potential damage to densely populated regions along the Atlantic coast is significant." The World Economic Forum (2013) includes vulnerability to geomagnetic storms explicitly in its listing of top environmental risks deemed to be able to significantly impact the global economy. The US National Intelligence Council (2013) noted that "[u]ntil cures' are implemented, solar super-storms will pose a large-scale threat to the world's social and economic fabric". A report on a risk analysis for space weather impacts in the UK (Royal Academy of Engineering, 2013) stated that the reasonable worst case scenario would have a significant impact

---

‡ http://youtu.be/-EKdxzZ52zU

on the national electricity grid." Impacts of geomagnetic storms are not limited to grids at high latitudes: although geomagnetic latitude is an important factor in GMD strengths, "geological conditions tend to override the effect of latitude" (NERC, 1989). In addition, lower-latitude regions can experience GICs arising from fluctuations in the magnetospheric ring current as demonstrated by the serious impact of the Halloween 2003 storms on the power grid in South Africa (e.g., Gaunt, 2013).

Even if we disregard the uncertain impact of extreme space weather, we know that impacts of moderate to extreme space storms that occur a few hundred times per 11-year solar cycle have consequences that merit substantial attention and investment. The overall cost to the US economy alone, for example, of non-catastrophic disturbances in the US power grid attributable to geomagnetically induced currents appears to be of the order of some US$5-10 billion/year (Schrijver et al., 2014). Economic impacts of space weather through other technological infrastructures have yet to be established, but threat assessments suggest "that space weather is the largest contributor to single-frequency GPS errors and a significant factor for differential GPS" (American Meteorological Society, 2011), for an industry that is worth of order US$100 billion/year worldwide (American Meteorological Society, 2011; Pham, 2011). A recent study (Schulte in den Baumen et al., 2014) made a first attempt to couple a GIC impact model with an economic model of global trade showing how GIC impact in three different regions (China, Europe and North America) would drive impacts across the world economy. It reinforces the message that space weather is a global problem - that a physical impact in one region can damage economies far from the impact site.

Given the persistent presence of the threat, society's increasing exposure to space weather, and the likely low-frequency but high-impact extreme-storm scenarios, it is not surprising that calls for the preparation for, forecasting of, mitigation against, and vulnerability assessment for space weather impacts by the international community echo in various studies over the past decade, including reports from academia (Hapgood, 2011), from the US National Research Council (2008), the UN Committee on the Peaceful Use of Outer Space (2013), and from the Organisation for Economic Co-operation and Development (OECD, 2011).

The user base interested in forecasts of space weather is growing rapidly with the increased awareness of space-weather threats and impacts. The official US space-weather forecast center (the Space Weather Prediction Center of the US National Oceanographic and Atmospheric Administration), for example, sees a continuing rapid growth in subscribers to its "Product Subscription

Service"† that was initiated in 2005 and that exceeded 40,000 individual subscribers early in 2014. A survey of the subscribers to the SWPC service in 2013 enabled an assessment of the interests from the user side (Schrijver and Rabanal, 2013), which concluded that "[s]pace weather information is most commonly obtained for reasons of [indirect impacts through interruptions of power or communications on] human safety and continuity or reliability of operations. The information is primarily used for situational awareness, as aid to understand anomalies, to avoid impacts on current and near-future operations by implementing mitigating strategies, and to prepare for potential near-future impacts that might occur in conjunction with contingencies that include electric power outages or GPS perturbations. Interest in, anticipated impacts from, and responses to the three main categories of space weather [- geomagnetic, radiation, and ionospheric storms -] are quite uniform across societal sectors. Approximately 40% of the respondents expect serious to very serious impacts from space weather events if no action were taken to mitigate or in the absence of adequate space weather information. The impacts of space weather are deemed to be substantially reduced because of the availability of, and the response to, space weather forecasts and alerts." It appears that many users of space weather forecasts apply the forecast information to avoid impacts on their systems and operations, either by increased monitoring given situational awareness or by taking preventive mitigating actions. As with terrestrial weather, this means that the economic value of space weather forecasts likely significantly exceeds the total costs of detrimental impacts, and that this value would increase as forecast accuracy and specificity would increase. Other valuable uses of space weather information as indicated by the subscribers to space-weather information lie in anomaly analysis and system design specification (Schrijver and Rabanal, 2013).

The study of space weather is important because of its societal relevance and much headway has been made in recent years [...]. Space weather also teaches us about the physical processes of the local cosmos that is our home within the Galaxy. More commonly known as the field of Sun-Earth connections, or as heliophysics particularly within the US, this is the science of the astrophysical processes that occur in the deep solar interior, in the vast reaches of the solar atmosphere that extend beyond the furthest planet out to the interstellar medium, and that includes the variety of coupling processes to the planets and natural satellites of our own solar system. Understanding all of these processes and interactions between diverse environments is our stepping stone to understanding what happens elsewhere in the universe in similar

---

† http://www.swpc.noaa.gov/pss/

environments, as much as to understanding the distant past and future of our own planetary system and its host star.

For a collection of reading materials on space weather and its societal impacts, see http://www.lmsal.com/∼schryver/SWlibrary.html.

2

Space weather: impacts, mitigation, forecasting

Sten Odenwald

2.1 Introduction

Normal, terrestrial weather is a localized phenomenon that plays out within a volume of 4 billion cubic kilometers over scales from meters to thousands of kilometers, and times as diverse as seconds to days. Whether you use the most humble technology found in remote villages in Bangladesh, or the most sophisticated computer technology deployed in Downtown Manhattan, terrestrial weather can and does have dramatic impacts all across the human spectrum. During 2011 alone, annual severe weather events cost humanity 2000 lives and inflicted damages upwards of $37 billion dollars (Berkowitz, 2011). The public reaction to terrestrial weather is intense, and visceral, with armies of meteorologists reporting daily disturbances around the globe, and weather forecasting models that have decades of development behind them and that have improved in reliability over the years.

In contrast to terrestrial weather and to our methods of mitigating its impact, we have the arena of space weather, which occurs within a volume spanned by our entire solar system, over time scales from seconds to weeks and spatial scales from meters to billions of kilometers. Unlike the impacts caused by terrestrial weather, space weather events on the human scale are often much more subtle, and change with the particular technology being used. There are, for example, no known space weather events in the public literature that have directly led to the loss of human life. The public reaction to space weather events when announced, seldom if ever reaches the level of urgency of even an approaching, severe thunderstorm. Despite the fact that, since the 1990s, we have become more sophisticated about communicating to the public about the potential impacts of severe space weather, these alerts are still only consumed and taken seriously by a very narrow segment of the population with technology at risk; satellite owners, power grid operators, airline pilots and the like. The historical record shows that in virtually all instances, space weather events have only led to nuisance impacts; disrupted radio communication; occasional short-term blackouts; and occasional satellite losses that were quickly replaced. Yet, when translated into the 21st Century, these same impacts would have a significantly larger impact in terms of the numbers of people affected. For instance, the Galaxy 4 satellite outage in 1998 deactivated 40 million pagers in North America for several hours. Pagers at that time were heavily used by physicians and patients for emergency surgeries, to name only one type of direct impact. Numerically, and in terms of function, we are substantially less tolerant of "outages" today than at any time in the history of space weather impacts.

In this chapter, I review the various technologies and systems that have historically-proven susceptibilities to space weather, why they are susceptible, methods being used to mitigate these risks, and how one might estimate their social impacts. I hope to demonstrate that, although we have a firm understanding of why technologies are at risk from basic physics considerations, we are still a long ways from making the case that extraordinary means need to be exerted to improve the reliability of present-day forecasts. One of the reasons for this is that we have been living through a relatively moderate period of solar activity spanning the majority of the Space Age. Without a major "Hurricane Katrina" equivalent in space weather, perhaps akin to the 1859 Carrington-Hodgson Superstorm, there is not much public outcry, commercial foresight, or political will, to significantly improve the current preparedness situation. Moreover, the progress of technology has been so rapid since the beginning of the Space Age in the late 1950s, that many of the technologies that were most susceptible to space weather, such as telegraphy, compass navigation, and short-wave communication, have largely vanished in the 21st Century, to be replaced by substantially more secure, albeit more inter-dependent, consumer technologies.

2.1.1 Open-air radio communication

Although telegraphic communication was the dominant victim of solar geomagnetic activity during the 1800s, by the mid-20th Century, virtually all telegraphic systems had been replaced by land lines carrying telephonic communications, and by the rapid rise of short-wave broadcasting and submarine cables for trans-continental communication (Odenwald, 2010). At its peak around 1989, over 130 million weekly listeners tuned-in to the BBC's World Service. Once the Cold War ended, short-wave broadcasting and listening de-

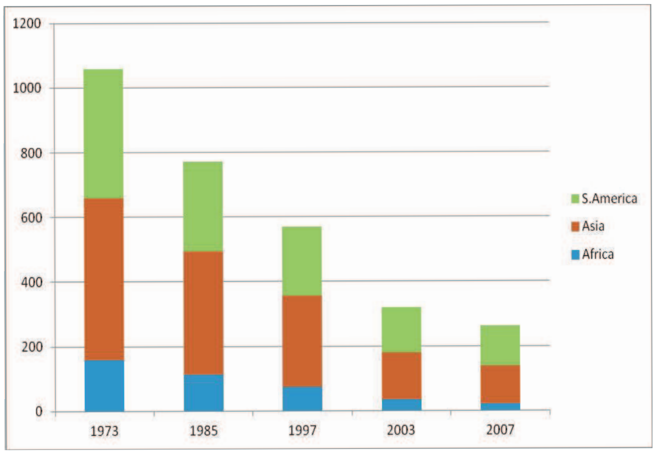

Fig. 2.1. The number of short wave stations (vertical axis) has dropped dramatically since the advent of the World Wide Web and other wireless media, which now provide the main source of news reporting in the 21st century. (Data courtesy Careless, 2010)

clined. As Figure 2.1 shows, less than one third of the stations on the air in 1970s are still operating. Compared to other forms of communication (such as web-based programming) shortwave is very expensive in terms of setting up a radio station and providing operating costs to purchase megawatts of broadcasting power (Careless, 2010, 2011). Nevertheless, by December 2011 an estimated 33% of the human population had access to the Internet and its vast network of formal and informal "news" aggregators, including online counterparts of nearly all of the former shortwave broadcasting stations.

Although shortwave broadcasting is a ghost of its former self, there are still a number of functions that it continues to serve in the 21st Century. It is a backup medium for ship-to-shore radio, delivering state-supported propaganda to remote audiences, time signals (at station WWV, the call sign of the U.S. NIST time signal), encrypted diplomatic messaging, rebel-controlled, clandestine stations, and the mysterious "Numbers Stations". There also continues to be a die-hard population of amateur radio "hams" who continue to thrill at DXing a dwindling number of remote, low-power stations around the world when the ionospheric conditions are optimal. Sometimes, these ham operators serve as the only communication resource for emergency operations. For example, during Hurricane Katrina in 2005, over 700 ham operators formed networks with local emergency services, and were the only medium for rapidly communicating life-saving messages. Despite the lack of public interest or awareness of the modern shortwave band, its disruption could leave many critical emergency services completely blind and unresponsive in a crisis.

Short wave (SW) broadcasting played such a key societal role during the first-half of the 20th century that millions of people were intimately familiar with its quality, program scheduling, and disruptions to this medium. Any disruption was carried as a Front Page story in even the most prestigious newspapers such as the New York Times. Although shortwave stations were routinely jammed by the then Soviet Union or Germany during World War II, these efforts paled in comparison to the havoc wreaked by even a minor solar storm. Known as the Dellinger Effect, a solar flare increases the ionization in the D and F Regions of the ionosphere on the dayside of Earth, spanning the full sun-facing hemisphere. This absorbs shortwave radiation but causes very low frequency (VLF) waves to be reflected. During the four solar cycles that spanned the "Short Wave Era" from 1920 to 1960, there were dozens of flares that delivered radio blackouts, which regularly interfered with trans-Atlantic communication, which was then a major news and espionage flyway for information between Europe and North America. Examples of events reported in the New York Times include:

- July 8, 1941 - Shortwave channels to Europe are affected (p. 10)
- September 19, 1941 - Major baseball game disrupted (p. 25)
- February 21, 1950 - Sun storm disrupts radio cable service (p. 5)
- August 20, 1950 - Radio messages about Korean War interrupted (p. 5)
- April 18, 1957 - World radio signals fade (p. 25)
- February 11, 1958 - Radio blackout cuts US off from rest of world (p. 62)

Although as we noted before, contemporary public contact with shortwave radio is nearly zero, today there are some places where SW is still in limited use, and where the public in those regions would be as conversant with SW fade-outs as the western world was around 1940. For instance, China is expanding its SW broadcasting to remote populations across its territory who do not as yet have access to other forms of communications networks. Even today, short wave outages still make the news:

On August 9, 2011 a major solar flare caused fade-outs in the SW broadcasts of Radio Netherlands World (RNW), but after an hour, broadcasting had returned to its normal clarity. Solar flare disrupts RNW short wave reception (RNP, 2011). This was the first major SW blackout in China since the X7.9-class flare on January 21, 2005, which affected Beijing and surrounding eastern population centers (Xinhuanet, 2005). On February 15, 2011 another large solar flare disrupted southern Chinese SW broadcasting. The China Meteorological Administration reported an X2.2-class flare at that time (Xihuanet, 2011). The January 23, 2012 M9-class solar flare disrupted broadcasts on the 6 20 meters bands across North America, and severely affected the UHF and VHF bands for a period of a few hours (Shortwave America, 2012).

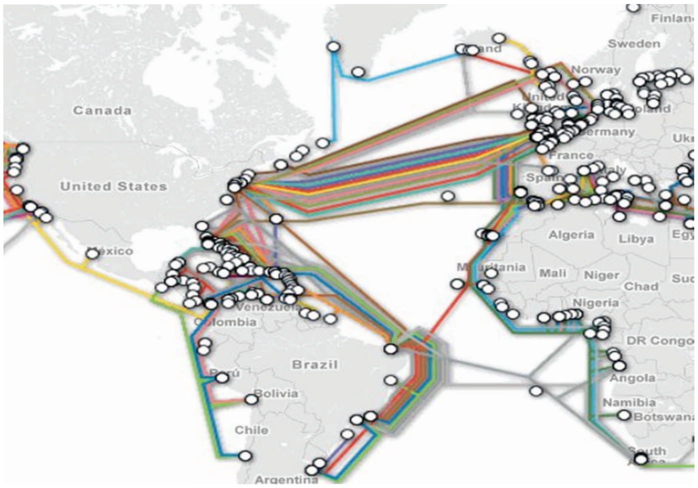

Fig. 2.2. A small portion of a map of the locations of submarine fiber optic cables around 2011. (Courtesy TeleGeography, 2012)

2.1.2 Submarine telecommunications cables

The first copper-insulated, trans-Atlantic cable was deposited on the ocean floor in 1856 between Ireland and New Foundland, but because it was run at voltages that were too high, the insulation broke down and the cable failed within a few weeks. The first successful cable was laid in 1865 between Brest, France and Duxbury, Massachusetts and worked successfully for many years, passing telegraphic signals at a speed of 2 words per minute (≈0.01 bps!). The first copper-insulated, trans-Atlantic telephone cable was laid in 1956. By 1950, over 750,000 miles of copper-based undersea cable had been installed between all of the major continents (International Cable Protection Committee [ICPC], 2009). This was followed by the first fiber optic cable TAT-8 installed between Europe and North America in 1988. By 2009, some 500,000 miles of fiber optic cable had been deployed, and had largely replaced all copper cable traffic due to the much higher bandwidths approaching several terabytes/sec (see Fig. 2.2).

Because signals degrade in strength as they travel through thousands of miles of copper, devices called repeaters are added to the cable every 50 miles or so, and are powered by a current flowing through a separate high-voltage power line that parallels the cable from end to end. Loss of power to a cable can cause immediate loss of signal, so all cables must be continuously powered through connection to the domestic power grid or back-up generators. These voltages can exceed 500 kV, and pose an electrocution hazard to fishing boats that accidentally snag them. Cables are typically broken through fishing accidents, earthquakes and mechanical failure about 150 times a year, causing a loss of communication capacity that may last from days to weeks depending on the depth of the required repair (ICPC, 2009). Because the repair site may only be a meter or so in length, modern repair ships routinely use GPS to reach the proper location of the identified failed repeater, or cable damage. Also, GPS systems are used in deploying fiber optic cables along exact, preplanned routes that minimize cable waste.

There is no formal requirement for communications companies to log cable outage events, especially in a public archive. Consequently, outages only become public knowledge when they impact public telecommunications activities. For example, on February 25, 2012, The East African Marine System (TEAMS) data cable linking East Africa to the Middle East and Europe was severed off the coast of Kenya by a ship that illegally dropped anchor in a restricted area. This cable was already taking the traffic from three other fiber optic cables that had been damaged only 10 days before. It would take three weeks before this cable could be repaired, and data and e-commerce traffic restored to Kenya, Uganda, Rwanda, Burundi, Tanzania and Ethiopia (Parnell, 2012).

Copper-based submarine cables are deployed in a manner similar to the old-style telegraph cables. For this reason they are subject to the same space weather impacts, though for different reasons, and perhaps not the ones you might initially consider. The original telegraphic systems and submarine cables of the 1800s were single conductors through which one-half of the battery was connected. The other half of the battery was grounded to the local Earth to complete the circuit! This works well when the naturally-occurring terrestrial ground current is stable in time, and over large geographic distances comparable to the telegraph network. However, both of these conditions are badly violated during a geomagnetic storm.

During a geomagnetic storm, a strong ionospheric current appears, called the electrojet. This current induces a secondary magnetic field that penetrates the local ground causing ground currents to flow that are called Geomagnetically-Induced Currents or GICs. Any single-wire telegraph system will immediately detect this GIC, which can be much greater than the original battery current, hence the frequent reports about mysterious high voltages and equipment burn out. The older trans-Atlantic cables were not immune from this because they, too, were patterned after the single-wire telegraph system and so GICs were a corresponding problem on these systems. For example, the geomagnetic storm that occurred on 2 August 1972 produced a voltage surge of 60 volts on AT&T's coaxial telephone cables between Chicago and Nebraska. The magnetic disturbance had a peak rate of change of 2200 nT/min., observed at the

Geological Survey of Canada's Meanook Magnetic Observatory, near Edmonton, and a rate of change of the magnetic field at the cable location estimated at 700 nT/min. The induced electric field at the cable was calculated to have been 7.4 V/km, exceeding the 6.5 V/km threshold at which the line would experience a high current shutdown (Space Weather Canada, 2011).

One might think that modern-day fiber optic cables are immune from this GIC effect because they involve a non-conductive optical fiber. High-voltage (HV) power is supplied to the cable at each end, with one end being at V^+ and the other at V^- potential. Just as for telegraph systems, one side of the HV supply is grounded to Earth, which provides a pathway for GICs. Repeaters for boosting the signal are connected in series along the cable axis and supplied by a coaxial power cable. GIC currents can temporarily overload the local power supply, causing repeaters to temporarily fail, and usually require resetting.

Have any incidents involving fiber optic cables ever been reported? We are mindful of the old adage that absence of evidence is not the same as evidence of absence. The fact that there is no impartial way to track outages on modern fiber optic telecommunications cables, and there are no federal regulations that require this reporting, means that reports are voluntary. When we search through public documents and find no cases of space weather-related cable outages, it only means that we cannot choose between two possible situations: Either they do occur and are not reported to save embarrassment, or the public records are unbiased and so lack of examples indicated lack of an impact. There are, however, some notable examples: At the time of the March 1989 storm, a new transatlantic telecommunications fiber-optic cable was in use. It did not experience a disruption, but large induced voltages were observed on the power supply cables (Space Weather Canada, 2011).

2.1.3 Ground-based computer systems

Solar storms can be a rich source of energetic particles via shock-produced Solar Proton Events (SPEs), galactic cosmic ray (GCR) enhancements during sunspot minimum, or events taking place within the magnetosphere during the violent magnetic reconnection events attending a geomagnetic storm. Although high-energy cosmic rays can penetrate to the ground and provide about 10% of our natural radiation background, secondary neutrons can be generated in air showers and penetrate at much higher fluxes to the ground. A number of monitoring stations, such as the Delaware Neutron Monitor, provide day-to-day measurements of the GCR secondary neutron background and detect ground-level enhancements (GLEs). At aviation altitudes, these high-energy neutrons can produce avionics upsets, which are easily corrected by error detection and correction (EDAC) algorithms or multiply-redundant avionics systems. On the ground, and ostensibly shielded by a thick atmosphere, computer systems and chip manufacturing processes have been allegedly affected by solar storm events (Tribble, 2010). Trying to identify even one case where such "computer glitches" were caused by GCR or space weather events remains problematical. Nevertheless, consumers and governments expect their computer systems to function reliably (computer virus attacks excepted), so even manufacturers such as Intel take this issue seriously. US patent 7,309,866, was assigned to Intel for their invention of "Cosmic ray detectors for integrated circuit chips" (Hannah, 2004): "Cosmic particles in the form of neutrons or protons can collide randomly with silicon nuclei in the chip and fragment some of them, producing alpha-particles and other secondary particles, including the recoiling nucleus. [...] Cosmic ray induced computer crashes have occurred and are expected to increase with frequency as devices (for example, transistors) decrease in size in chips. This problem is projected to become a major limiter of computer reliability in the next decade."

Bit-flip errors, in which the contents of a memory cell become switched from a "0" state to a "1" state or vice versa, are a pernicious form of Single Event Upset (SEU) that continues to plague ground based computer systems that use high-density VLSI (very large-scale integration) memory. The mechanism is that a high-energy neutron collides with a substrate or gate nucleus, producing a burst of secondary charged particles. These electrons and ions drift into a memory cell and increase the stored charge until a state threshold is achieved, at which point the cell indicates a high-Q state of "1" rather than a relatively empty, low-Q state of "0"; hence the bit-flip error. Extensive testing and research to identify the origin of these soft-memory errors led to alpha particle emission from naturally occurring radioisotopes in the solder and substrate materials themselves. Extensive re-tooling of the fabrication techniques, however, failed to completely eliminate SEUs. Currently, a system with 1 GB of RAM can expect one soft-memory error every week, and a 1 terabyte system can experience SEUs every few minutes. Error detection and correction (EDAC) algorithms cost power and speed, and do not handle multi-bit errors where the parity does not change (Tezzaron, 2003). According to Paul Dodd, manager for the radiation effects branch at Sandia National Labs: "It could be happening on everyone's PC, but instead everyone curses Microsoft. Software bugs probably cause a lot of those blue-screen problems, but you can trace some of them back to radiation effects" (Santarini, 2005).

Although there are no specific, documented examples of ground-based computer crashes due to specific solar storms, it is legitimate to consider what might be the societal consequences of space weather-induced computer glitches.

If they occur from time to time, it is instructive to consider the impact that other more prosaic glitches have produced:

- March 2, 2012 - Computer glitch hits Brazil's biggest airline. "Brazil's biggest airline says a computer glitch took down its check-in system in several airports across the country, causing long delays" (boston.com, 2012).
- November 5, 2011 - HSBC systems crash affects millions across UK. "HSBC was today hit by a nationwide systems crash thought to have affected millions of customers. The bank's cash machines, branches, debit cards, and internet banking services all stopped working at 2.45pm after a computer glitch" (Paxman, 2011).

2.1.4 Space-based computers

The first documented space weather event on a satellite occurred on Telstar-1 launched in July1963. By November, it had suddenly ceased to operate. By exposing the ground-based duplicate Telstar to various radiation backgrounds, Bell Telephone Laboratory engineers were able to trace the problem to the gate of a single transistor in the satellite's command decoder. Apparently, excess charge had built up on the gate, and by simply turning the satellite off for a few seconds, the problem disappeared. By January, 1963 the satellite was back in commercial operation relaying trans-Atlantic television programs between Europe and North America (Reid, 1963).

During the 1960s, a number of NASA reports carefully documented the scope and nature of space weather-induced satellite and spacecraft malfunctions. There was as yet no significant commercial investment in space, so NASA analyzed glitches to its own satellites and interplanetary spacecraft. Of course, military satellites of ever increasing complexity, cost, and political sensitivity were also deployed, but no unclassified documents were then, or are now, available to compare space weather impacts across many different satellite platforms. This leads to an important issue that is crucial to impact assessment and mitigation. How can we assess risks and prospective economic losses when so much of the required data is protected through national secrecy regulations and commercial confidentiality? Even among the "public domain" NASA satellites, data as to the number and severity of "glitches" is usually buried in the "housekeeping" data and rarely makes it out of the daily briefing room since it is irrelevant to the scientific data-gathering enterprise.

In a perfect world, we would like to have data for all of the 2000+ currently operational satellites that describes the numbers, dates and types of spacecraft anomalies that they experienced. From this we would be able to deduce how to mitigate the remaining radiation effects, identify especially sensitive satellites and quantify their reliability, and to develop accurate models for forecasting when specific satellites will be most vulnerable. In reality, much of what we can learn is by "reading between the lines" in news reports, correlating these biased forms of information against the known space weather events, and hoping that a deterministic pattern emerges. Even this has been a daunting challenge when adjacent satellites in orbit can experience the same space weather conditions, but have very different anomalies, thereby making correlations between space weather conditions and satellite anomalies seem less certain.

2.1.4.1 How does it happen?

Satellite anomalies can be broadly defined to include any event in which some operating mode of a satellite differs from an expected or planned condition. In this context, the term "anomaly" is extremely broad, spanning a continuum of severities from trivial satellite state changes and inconsequential data corruption, to fatal conditions leading to satellite loss. Actual data from satellite-born sensors shows that these events can be quite numerous. For instance, SOHO data from a 2 GB onboard Solid State Recorder typically records over 1000 SEUs/day (Brecca et al, 2004). Only rarely, however, do SEUs actually lead to satellite conditions requiring operator attention a condition commonly termed an anomaly. For SOHO, only ≈60 anomalies during an 8-year period (≈8 anomalies/satellite/year) have required significant operator intervention, despite the literally millions of SEU events recorded during this time.

Anomalies need not be fatal to be economically problematical. On January 20, 1994, the Anik E1 and E2 satellites were severely affected by electrostatic discharges (ESDs). Although the satellites were not fatally damaged, they required up to $70 million in repair costs and lost revenue, and accrued $30 million for additional operating costs over their remaining life spans (Bedingfield et al., 1996). The Anik satellite problems were apparently the result of a single ESD event affecting each satellite (Stassinopoulos et al., 1996), suggesting that large numbers of anomalies are not required to 'take out' a satellite. If anomalies are frequent enough, however, the odds of a satellite failure must also increase, as will the work load to satellite operations. According to FUTRON (2003), satellite operators ordinarily spend up to 40 percent of their time on anomaly-related activities. Ferris (2001) has estimated the cost of dealing with satellite discrepancies (in which some system of the satellite and ground system does not operate as desired) as $4,300/event leading to overall operations impacts approaching $1 million/satellite/year under apparently routine space weather conditions. Anecdotal reports suggest that during major solar storms, far higher operator activity can occur. For example, the GOES-7 satellite experienced 36 anomalies on October 20, 1989, during a single, severe solar storm event (Wilkinson, 1994).

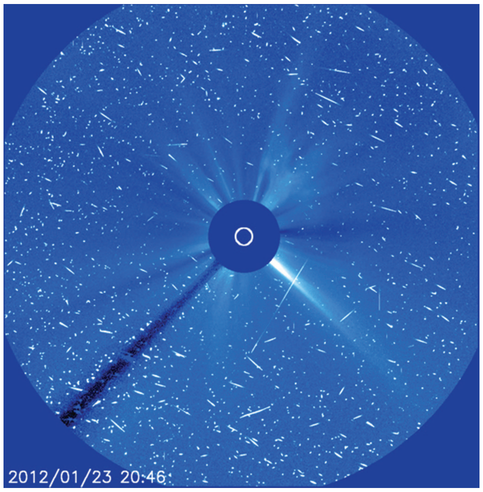

Fig. 2.3. This figure shows the effects of high-energy solar protons on an exposed imager on the Solar and Heliospheric Observatory during the January 23, 2012 solar storm. A greatly reduced flux of particles entering shielded satellite circuitry results in SEUs, many of which are harmless, but a few per year can result in serious operational anomalies.

Table 2.1. *Tabulation of statistics of satellite anomalies (see Sect. 2.1.4 for data and references).*

Rates:	study a: 1-10 /yr/sat			
	study b: 3 /yr/sat for GEO, 2-3× more during enhanced space weather			
		study 1	study 2	study 3
Class:	1: mission failure	8%	6%	Class 1+2 comb.:
	2: interruption	7%	> 1 week 39%	0.019 ± 0.006 /yr/sat
	3: performance decr.	Class 3+4	2h-1 week 35%	
	4: inconsequential	comb.: 84%	<1 h 20%	
Cause:	ESD: 23-49%; SEU: 18-26%; Rad. damage: $\sim 5\%$			

The issue of "how bad can it get?" is an interesting one, especially given our dramatically increased reliance upon GEO satellite systems since ca 1980 that are economically baselined on the assumption of 100% reliability during a 10 to 15 year satellite service life span. The ≈250 GEO satellites now in operation produce an annual revenue of $80 billion (Ferster, 2005) so any space weather impact is potentially costly, and can involve more than one satellite at a time. Satellite designers use sophisticated tools to assess radiation hazards under “worst case” conditions (e.g. the August 1972 and March 1991 events) however, recent studies of extreme space weather conditions suggest that the period since ca 1960 has not been typical of the historical record of severe storms during the last 500 years (McCracken et al., 2001; Townsend, 2003). Moreover, there is a large discrepancy between models that predict, for example, SEU events and actual satellite observations of them (e.g. Hoyos, Evans and Daly, 2004). Some recent studies have attempted to estimate the economic consequences to commercial GEO satellites for severe solar storm episodes (e.g. Odenwald and Green, 2007), but the studies were hampered by the lack of detailed knowledge of how the frequencies of satellite anomalies vary in severity with storm intensity. Consequently, the loss of a satellite during a severe space weather event could not be modeled realistically, nor its economic impact properly assessed. Most reported anomalies, broadly defined, are nuisances involving recoverable data corruption, easily-corrected phantom commands, or 'bit flips' often caught by onboard EDAC algorithms. These are not the kinds of anomalies that lead to significant economic consequences for a commercial satellite. Other less frequent anomalies cause sub-system failures, out-of-spec satellite operations, attitude and telemetry-lock errors and even outright satellite failures. These are most certainly the kinds of anomalies that have economic consequences. Some authors have also classified anomalies by satellite orbital location (e.g. LEO, MEO, GEO), recognizing that each environment has its own physical drivers for anomaly generation, but more often than not, these classes are aggregated together. Here is one possible scheme:

- Class 1 - Mission-Failure: The satellite ceases operation as a consequence of an unrecoverable system malfunction (e.g. Telstar-401).
- Class 2 - Mission interruption: Involves a recoverable damage to sub-systems. Only built-in redundancies, if available, are capable of mitigating some of these problems, where the satellite's safe mode may be enabled, or a back-up subsystem has to be activated (e.g. Anik-E1). These may take hours of effort to remedy, at a cost to satellite revenue and operator overhead charges.
- Class 3 - Performance decrease: Can include spacecraft pointing errors, attitude control system error, or a brief loss of data or telemetry usually corrected by a manual or automatic system reset.
- Class 4 - Inconsequential: Memory bit-flips and switching errors easily corrected using on-board EDAC software, or simple operator action (e.g. TDRSS-1 telemetry; cosmic ray corruption of Hubble Space Telescope data).

One of the earliest, and most detailed, publically available studies of satellite anomalies and reliability is the work by Hecht and Hecht (1986: the Hecht

Report). The study was based on 2,593 anomaly reports for 300 satellites launched between ca 1960 and 1984. There were ≈350 satellites in operation by 1984, so the Hecht Report is relatively complete. This ground-breaking study analyzed the detailed reports provided by 96 satellite programs. A 'failure' was defined as "...the loss of operation of any function, part, component or subsystem, whether or not redundancy allowed the recovery of operation". Their study identified 213 Class 1 and 192 Class 2 anomalies out of a total collection of 2593 anomalies for a mission failure rate defined by our Class 1 of about 405/2593 or 1 in 6. No attempt was made to correlate the anomalies with space weather conditions.

One of the most widely used, recent starting points for anomaly studies is the archive assembled by Wilkinson and Allen (1997; National Geophysical Data Center, hereafter NGDC) which identifies most of the 259 satellites by name, or code, along with orbital location and/or altitude information. The date and type of anomaly is provided for many of the 5,033 events spanning the time period from 1970 to 1997, so that a proper assessment can be attempted of the various category-specific anomaly rates as a function of date and satellite type. There are 3,640 events that have been tagged according to type and system impact, including 647 SEU events and 848 ESD events. The NGDC archive contains 43 commercial GEO satellites included in the archive, accounting for a total of 480 anomalies spanning 20 years, and also appears to contain about 40% of all operating satellites during the sample time span, and is relatively complete for our purposes. The average annual anomaly rate of the GEO satellites was found to be about 3 anomalies/satellite/year, but can rise to twice or three times this rate during enhanced space weather conditions.

Robertson and Stoneking (2005: Goddard) examined 128 severe (Class 1 and 2) anomalies among 764 satellites. The data were culled from web-based satellite anomaly lists including the 'Airclaims Space Track' as well as NASA documents and the Aerospace Corporation 'Space Systems Engineering Database', and only included satellites from the US, Europe, Japan or Canada. The total number of satellites (military + commercial) operating during this interval is 827, so the sample contains about 92% of all possible operational systems during the 1990-2001 time period. A total of 35 anomalies were Class 1, which led to what was considered the total loss of the satellites. Four each anomaly in Class 1 there are three in Class 2. Their calculated anomaly rate was based on the number of anomalies recorded, divided by the number of satellites launched during a given year. Re-normalizing their mishap rates to, instead, reflect the annual operating satellites, the average mishap rate for Classes 1+2 is about 0.019 ± 0.006 anomalies/sat/year. The inverse of this rate is 166 which is sometimes called the mean time to failure (MTF). Clearly for commercial satellites expected to last 10 to 15 years before replacement, a MTF of 166 years is good news! The correlation between these anomalies and space weather events was not studied.

The extensive studies by Belov et al. (2004) and Dorman et al. (2004) included satellite anomaly reports based on 300 satellites and ≈6,000 anomalies spanning the time period from 1971 to 1994. The data was drawn from NASA archives, the NGDC archive and unpublished reports from 49 Kosmos satellites (1971-1997). The term 'anomaly' was never precisely defined, but since the survey included the NGDC archive without distinction, we can assume that all Class 1-4 events were grouped together. The sample included 136 satellites in GEO orbits. They deduced that there were typically 1 to 10 anomalies/satellite/year. Specifically, the LEO Kosmos satellites experienced 1-7 anomalies/satellite/year, however some Kosmos satellites (Kosmos 1992 and 2056) reported ≈30 anomalies/satellite/year. Their statistical analysis indicated that anomalies occur during days when specific space weather parameters (electron/proton fluxes, Dst, Ap, etc) are disturbed. The largest increases coincide with times when electron and proton fluences are large, and can cause enhancements up to a factor of 50 in anomalies over quiet-time conditions. There appears to be a threshold of 1,000 pfu ($E > 10$ Mev) for proton fluxes, below which there are few anomalies reported. The anomalies continue to remain high for two days after the SPE event.

Koons et al. (1999) published "The Impact of the Space Environment on Space Systems", which investigated a sample of 326 anomaly 'records' collected from a diverse assortment of satellites culled from the NGDC 'Satellite Anomaly Manager', Orbital Data Acquisition Program (ODAP: Aerospace Corp.), NASA's Anomaly reports (Bedingfield et al. 1996, and Leach and Alexander, 1997), and the USAF Anomaly Database maintained by the 55th Space Weather Squadron. The specific number of satellites involved was not stated, however, the ODAP archive contains information from 15 USAF and 91 non-Air Force 'programs' no doubt drawn from LEO, MEO and GEO satellite populations. Although no information was provided as to the time period spanned by the study, the individual archives extend from 1970 to 1997. The definition of a record in terms of anomalies can vary enormously. Each record contained information for one class of anomalies for one 'vehicle'. Anomalies of a similar class were of the same functional type. Approximately 299 records out of 326 (92%) have causes diagnosed as 'space environment' but this does not necessarily correlate with a count based on anomaly frequencies. An example cited is that one record for the MARECS-A satellite included 617 anomalies. About 51 of 326 records were from commercial satellite systems and programs. In terms of the distribution of the records with anomaly diagnosis, 162 (= 49%) were associated with Electrostatic Discharges, 85 (= 26%) with SEUs, and 16 (=5%) with 'total radiation damage'. Based on 173 reports of how quickly the anomalies were rectified, the Koons et al. (1999) study indicates that the number of mission failures represents 9/173 reports for a frequency rate of 1 in 19. The rates for the other classes are: Class 2 (More than 1 week) = 39%; Class 3 (1 hr to 1 week) = 35% and Class 4 (Less than 1 hour) = 20%.

Ferris (2001) analyzed 9,200 satellite operations discrepancy reports from 11 satellites between 1992-2001. A 'discrepancy' was defined as "the perception by a satellite operator that some portion of the space system had failed to operate as desired." The satellites were selected on the basis of which operators and owners were willing to divulge detailed anomaly logs for this study, which is a strong bias probably in favor of systems that had low absolute rates and few critical failures. Only three of the satellites were communications satellites; none were for civilian commercial use. This, of itself, is a problem since we cannot know to what extent these satellites are typical, or whether they are pathological. This is often the case when working with studies in which the satellite identities are not publically revealed. Of the discrepancies catalogued, only 13% involved the satellites themselves. The vast majority, 48%, involved issues with the ground segment, and specifically, most were discrepancies generated by software issues (≈61% of total discrepancies). Typical discrepancy rates involving 1,200 events imply ≈13 discrepancies/satellite/year. There were, however, higher rates recorded in 1996 involving 160 events for 4 satellites for a rate of 40 discrepancies/satellite/year or about one every 9 days. The study was the first one published in the open literature that also provided an assessment of the cost of rectifying these anomalies. Routine problems that require no more than 10 minutes to resolve by a team of 8 people cost $800 per event. More significant problems requiring 3-8 hours and more people cost $4,300 per event. The estimate only included labor hours and an average of the resolution times for the logged events, and not the cost of equipment or materials. In the latter case an 'event' may include the replacement of part of the ground station, processors or other mechanical items.

Cho and Nozaki (2005) investigated the frequency of ESDs on the solar panels of five LANL satellites between 1993-2003. During this period, LANL 1989-046 experienced 6038 ESDs/year while LANL-92A recorded 290 ESDs each year. Although the cumulative lifetime ESD rates on solar panels can exceed 6,000 events/kW over 15 years, the chances of a catastrophic satellite failure involving substantial loss of satellite power, remains small, though not negligible. For example, in 1973, the DSCS-9431 satellite failed as a result of an ESD event. More recently, the Tempo-2 (1998) and ADEOS-2 (2003) satellites were also similarly lost. Koons et al. (1991, 2000) and Dorman et al. (2005) have shown that ESDs appear to be ultimately responsible for half of all mission failures (e.g. Class 1 anomalies) and correlated with space weather events.

Wahlund et al. (1999) have studied 291 ESD events on the Freja satellite (MEO orbit) and have found that the number of ESDs increases with increasing Kp. A similar relationship between increasing Kp and anomaly frequency was found by Fennell et al (2000) for the SCATHA satellite (near-GEO orbit). These results are consistent with earlier GOES-4 and 5 satellite studies by Farthing et al. (1982) and by Mullen et al. (1986). In addition to Kp, Fennell et al. (2000) and Wrenn, Rogers and Ryden (2002) identified a correlation between 300 keV electron fluxes and the probability of internal ESDs from the SCATHA satellite. The probability increases dramatically for electron fluxes in excess of 100,000ṗfu. A similar result was found a number of years earlier by Vampola (1987). At daily total fluences of $\approx 10^{12}$ electrons/cm^2 the probability of an ESD occurring on a satellite exponentially reaches 100% (e.g. Baker, 2000).

2.1.4.2 That was then – this is now

During the 23rd Sunspot Cycle (1996-2008) there were dozens of satellite malfunctions and failures noted soon after a major solar storm event, beginning with Telstar-401 (1996) and ending with the Japanese research satellite ASCA on October 29, 2003. The 24th Cycle had its own satellite outages and malfunctions of note.

On August 25, 2011, South Africa's $13 million LEO satellite SumbandilaSat failed. The explicit cause was stated publically to be 'damage from a recent solar storm', which caused the satellite's onboard computer to stop responding to commands from the ground station. This was not, however, the first time this satellite was damaged by radiation. Shortly after its launch in September 2009, radiation caused a power distribution failure that rendered the Z-axis and Y-axis wheel permanently inoperable, meaning that the craft tumbles as it orbits and has lost the ability to capture imagery from the green, blue and xantrophyll spectral bands. The reason given for the lack of proper radiation hardening was that there was not enough money to do this properly, and the satellite was built from commercial off-the-shelf (COTS) equipment. Moreover, SumbandilaSat was intended only as a technology demonstrator (Martin, 2012).

The case of the Anik F2 'technical anomaly' on October 6, 2011 is a replay of similar stories during the 23rd Sunspot Cycle. The satellite entered a Safe Mode that caused it to stop functioning and turn away from Earth. The Boeing satellite was launched in 2004 and was expected to function for 15 years. The owner of the satellite, Telsat, indicated in public news articles that they did not believe the problem had to do with the arrival of a CME that reached Earth early the same morning, but was caused by some other unspecified internal issue with the satellite itself. It is the first serious anomaly of its kind since the satellite was launched in 2004. What the news reports failed to mention was that the Sun has been relatively quiet for the majority of this 7 year period (Mack, 2011).

The temporary outage of Anik F2 caused a number of problems that impacted millions of people covered by this satellite service. WildBlue satellite ISP in the United States uses Anik F2 to provide broadband services to about a third of its customers. A total of more than 420,000 subscribing households mostly in parts of rural America lost service for several days, along with ATM service. Canadian Broadcasting Corporation indicated that 39 rural communities, and 7,800 people lost long-distance phone service. The satellite is also used for air traffic control, causing the grounding of 48 First Air flights, and 1000 passengers, in northern Canada. Communities in the North West Territories were instructed to activate their emergency response committees, and start using their Iridium phones (Mack, 2012; CBS News, 2012; Marowits, 2011).

On April 5, 2010, Galaxy-15 experienced an electrostatic discharge that caused a severe malfunction, rendering the satellite capable of re-transmitting any received signal at full-power, but not able to receive new commanding de Shelding, 2011). Reports cited a space weather event on April 5 as the probable cause of the electrostatic discharge that was the likely triggering event, however although Intelsat acknowledged the ESD origin, they categorically refuted the space weather cause in the April 5 solar event, preferring to declare that the orgin of the ESD was unknown. A consequence of this type of satellite failure is that Galaxy-15 was potentially able to interfere with other GEO satellites as it came within 0.5 degrees of their orbital slots. Thanks to careful, and complex, maneuvering of the satellites to maximize their distance from this satellite as it entered their orbital slots, AMC-11, Galaxy-13, Galaxy-18, Galaxy-23 and SatMex-6 and Anik F3 were able to reduce or eliminate interference, and no impacts to broadcasting were reported or acknowledged. "The fact that you haven't heard about channels lost or interference is the proof that we have been able to avoid issues operationally," said Nick Mitsis, an Intelsat spokesperson. "I don't want to underplay that" (Clark, 2010). In January 2011 commanding of the satellite resumed and its “zombisat” moniker has been changed to “phoenix”.

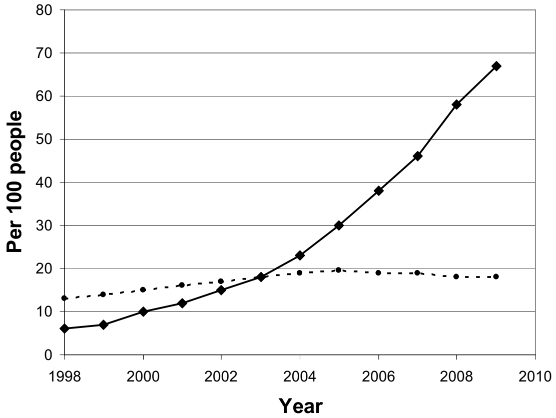

Fig. 2.4. A report from the International Telecommunications Union finds that at the end of 2009, 67 percent of all people on Earth were cell phone subscribers (solid line). The number of land line subscribers is now in decline (dotted line) having reached a maximum of 19% of world inhabitants in 2005 (Duncan, 2010).

2.1.5 Cellular and satellite telephones

Although telephone calls by land lines are among the safest communication technology, and the most resistant to space weather effects, they have also been in rapid decline thanks to the wide spread adoption of cellular and mobile phones, especially among the under-30 population. According to an article in The Economist (2009) customers are discontinuing landline subscriptions at a rate of 700,000 per month, and that by 2025 this technology will have gone the way of telegraphy. Between 2005 and 2009, the number of households with cell phone-only subscriptions rose from 7% to 20%. In terms of space weather vulnerability, there is one important caveat. Without an electrical power grid, conventional land-lines fail, and cell phones may not be recharged even though the cell towers may have emergency back up power capability. An example of this vulnerability occurs whenever natural disasters strike and cell towers are unavailable, or the crushing load of cell traffic renders the local tower network unusable. Moreover, one does not have to wait for power grid failure to have an impact on cell phone access during episodes of solar activity.

A seminal paper by Lanzerotti et al. (2005) demonstrates that solar radio bursts, which occur rather often in an active photosphere, can cause enhanced noise at the frequencies used by cellphones (900 MHz to 1900 MHz), when the observer's angle between the cell tower and the Sun is small. This interference effect shows up in the Dropped-Call statistic for east-facing receivers at sunrise or west-facing receivers at sunset. For a given cell phone and cell tower in the optimal line-of-sight geometry with respect to the Sun on the horizon, dropped calls occur about once every 3 days during solar maximum, and every 18 days during solar minimum. The article notes that the detailed, direct, evidence for solar-burst influence on cell phones remains a proprietary issue not openly available for investigation. The authors note that ”solar bursts exceeding about 1000 sfu (solar flux units, 1 sfu = $10^{-22}\,\mathrm{W\,m^2\,Hz^{-1}}$) can potentially cause significant interference when the Sun is within the base-station antenna beam, which can happen for east- or west-facing antennas during sunrise and sunset at certain times of the year.” Because base stations are only vulnerable for about two hours each day during sunrise and sunset, a typical station might be affected about one day out of 42 for solar maximum, and one day in 222 during solar minimum.

2.1.6 GPS-based systems

Navigation by satellite is not a new technology. It was first introduced by the US Navy in 1960 with the orbiting of five Transit satellites. This system was replaced by the NAVSTAR-GPS system in the 1970s. The first commercial use of satellite-based global positioning systems came less than 1 year after the next generation, 24-satellite 'Block I-GPS' constellation had been deployed in 1994, when Oldsmobile offered the GuideStar navigation system for its high-end automobiles. The GPS satellites provided an L1 channel at 1575 MHz capable of 10-meter-scale precision, that in 1990 was 'selectively degraded' to 100-meter precision. In 1999, President Clinton ordered that selective availability be turned off, and on May 1, 2000 the modern era of non-military GPS was ushered-in. Since 2000, the commercial applications of GPS have enormously expanded to include, not only car navigation aids, but oil extraction, fiber optic cable deployment, civilian aviation, emergency services, and even expanding public cellphone services, called apps, to locate nearby stores, restaurants and even parking spaces in downtown Manhattan! A report by Berg Insight (2011) indicates that GPS-enabled mobile phones reached 300 million units in 2011, and is expected to reach nearly 1 billion units by 2015.

Although the GPS constellation is stationed in polar orbits that frequently pass through the van Allen radiation belts in MEO, they are well-shielded and are upgraded every 5-10 years through replacement satellites such as the Block-II and Block-IIII systems. Although the details of the frequency of satellite anomalies is highly classified, it can be surmised that a legacy of 40 years of space operations has left the GPS system with a broad assortment of mitigation strategies for essentially eliminating outages. Nevertheless, there is one aspect of GPS system operation that cannot be so easily eliminated.

GPS signals must be delivered to ground stations by passage through the ionosphere. Because radio propagation through an ionized medium causes signal delays, and accurate timing signals are important in locating a receiver in 3-dimensional space, any changes in ionospheric electron content along the propagation path will cause position errors in the final solution (see Ch. 5). Space weather events, especially X-ray flares, cause increased ionization and introduce time-dependent propagation delays that can last for many hours until the excess ionospheric charge is dissipated through recombination. This also causes amplitude and phase variations called scintillation, which causes GPS receivers to loose lock on a satellite. Since a minimum of 4 satellites are required to determine a position, excess scintillation can result not just in a bad position solution, but can cause a loss-of-lock so that not enough satellites are available for various locations at various times during the event.

When civilian, single-frequency GPS systems using the L1 frequency are used, the anomalous propagation problem has to be mitigated by reference to a 'GPS Ionospheric Broadcast Model' and making the appropriate corrections. The resulting accuracy is about 5 meters. But this correction can only work for a limited period of time and so the path-delay problem is only partially solved. The result is that most civilian GPS systems can be easily disturbed by solar activity. Dual-frequency GPS systems that operate at L1 (1575 MHz) and L2 (1228 MHz) can measure the differential propagation of the satellite signal in real-time, and by relating this to the plasma dispersion equation, calculate the instantaneous total electron content (TEC) along a path, and then use this to make the requisite on-the-spot timing correction. In fact, this method can be turned around by using networks of GPS receivers to actually map out the changing ionospheric structure over many geographic locations. Figure 2.5 shows one such 'TEC' calculation for April 20, 2012 for 19:00 UT developed by JPL. The black spots are the GPS receivers in the network. Green indicates a TEC of about 50×10^{16} electrons/m^2 while red indicates 80×10^{16} electrons/m^2. Generally, a TEC of 6×10^{16} electrons/m^2 corresponds to an uncorrected position error of about 1 meter. The figure displays potential position errors as high as 13 meters over Chile.

Although the L1 carrier signal can be received without special instrumentation, the L2 timing information is coded and not accessible to non-military receivers. However, by using a technique called differential GPS, civilian GPS systems now rival, or even exceed, military precision in those areas where the requisite DGPS ground reference stations are available. If you are navigating in a large city, DGPS is probably available to you, but if you are 'in the middle of nowhere' chances are you only have single-frequency GPS to guide you.

We have already discussed this briefly in the context of GPS signal propa-

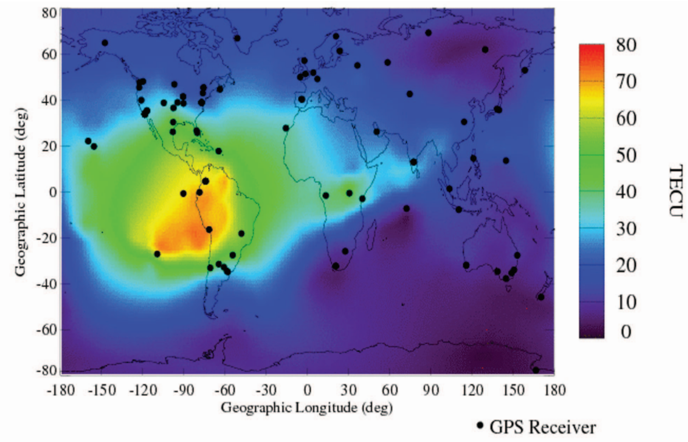

Fig. 2.5. Total electron content (TEC) calculation for April 20, 2012 for 19:00 UT developed by JPL.

gation and ionospheric scintillation. Because many space weather phenomena couple efficiently to the ionosphere, it is unsurprising that space weather issues have always been foremost in the discussion of GPS accuracy and reliability even apart from the fact that the GPS satellites themselves are frequently located in one of the most 'radio-active' regions of the magnetosphere. One of the first unclassified studies to quantitatively assess GPS behavior under solar storm conditions was conducted, inadvertently, by NOAA in 2001. They had set up a network of 70 GPS receivers from Alaska to Florida to test a new weather observation and climate monitoring system called the GPS-MET Demonstration Network. A major geomagnetic storm between March 30 and 31 caused significant changes in the GPS formal error, and was correlated with the published Kp index during the course of the event (NOAA, 2001). Since then, a variety of anomalous changes in GPS precision have been definitively traced to, and found to be correlated with, geomagnetic storms and solar flare events. This also means that systems that rely on GPS for high-precision positioning have almost routinely reported operational upsets of one kind or another. For example (NOAA, 2004):

- On October 29, 2003, the FAA's GPS-based Wide Area Augmentation System (WAAS) was severely affected. The ionosphere was so disturbed that the vertical error limit was exceeded, rendering WAAS unusable. The drillship GSF C.R. Luigs encountered significant differential GPS (DGPS) interruptions because of solar activity. These interruptions made the DGPS solutions unreliable. The drillship ended up using its acoustic array at the seabed as the primary solution for positioning when the DGPS solutions were affected by space weather.

- On December 6, 2006, the largest solar radio burst ever recorded affected GPS receivers over the entire sunlit side of the Earth. There was a widespread loss of GPS in the mountain states region, specifically around the four corners region of New Mexico and Colorado. Several aircraft reported losing lock on GPS. This event was the first of its kind to be detected on the FAA, WAAS network.

Apart from changes in ionospheric propagation, we have the problem that, if the GPS signal cannot be detected by the ground station, and the minimum of 4 satellites is not detected, a position solution will not be available at any accuracy. This situation can arise if the GPS signal is actively blocked or jammed, or if the natural background radio noise level at the L1 and L2 frequencies is too high. This can easily happen during radio outbursts that accompany solar flare events. This happened the day after the December 5, 2006, solar flare, and was intensively studied by Kintner at Cornell, and presented at the Space Weather Enterprise Forum in Washington, DC on April 4, 2007 (NOAA, 2007).

2.1.7 Electrical power grids

The issue of space weather impacts to the electrical power grid is covered more extensively in Chapter 4, we review the main points of this vulnerability, provide concrete examples, and review briefly the impacts and consequences of future large geomagnetic storms.

It has been well known for decades that geomagnetic storms causes changes in the terrestrial ground current. The most dramatic examples of this effect are in the many reports of telegraph system failures during the 1800s. So long as a system requires an 'earth ground', its circuit is vulnerable to the intrusion of geomagnetically-induced currents (GICs). For the electric power grid, these DC currents do not need to exceed much above 100 amperes in order to do damage (Odenwald,1999, Kappenmann, 2010).

When GICs enter a transformer, the added DC current causes the relationship between the AC voltage and current to change. It only takes a hundred amperes of GIC current or less to cause a transformer to overload during one-half of its 60-cycle operation. As the transformer switches 120 times a second between being saturated and unsaturated, the normal hum of a transformer becomes a raucous, crackling whine physicists call magnetostriction. Magnetostriction generates hot spots inside the transformer where temperatures can increase very rapidly to hundreds of degrees in only a few minutes, and last for many hours at a time. During the March 1989 storm, a transformer at a nuclear plant in New Jersey was damaged beyond repair as its insulation gave way after years of cumulative GIC damage. During the 1972 storm, Allegheny Power detected transformer temperature of more than 340 F (171 C). Other transformers have reached temperatures as high as 750 F (400 C). Insulation damage is a cumulative process over the course of many GICs, and it is easy to see how cumulative solar storm and geomagnetic effects were overlooked in the past.

Outright transformer failures are much more frequent in geographic regions where GICs are common. The Northeastern US with the highest rate of detected geomagnetic activity led the pack with 60% more failures. Not only that, but the average working lifetimes of transformers is also shorter in regions with greater geomagnetic storm activity. The rise and fall of these transformer failures even follows a solar activity pattern of roughly 11 years.

The connection between space weather events and terrestrial electrical systems has been documented a number of times. Some of these examples are legendary (1989, 2003) while others are obscure (1903, 1921). Given the great number of geomagnetic storms that have occurred during the last 100 years, and the infrequency of major power outages, this suggests that blackouts following a major geomagnetic storm are actually quite rare events. Consider the following historical cases:

- November 1, 1903: The first public mention that electrical power systems could be disrupted by solar storms appeared in the New York Times, November 2, 1903 "Electric Phenomena in Parts of Europe". The article described the, by now, usual details of how communication channels in France were badly affected by the magnetic storm, but the article then mentions how in Geneva Switzerland (New York Times, 1903). “All the electrical streetcars were brought to a sudden standstill, and the unexpected cessation of the electrical current caused consternation at the generating works where all efforts to discover the cause were fruitless”.
- May 15, 1921: The entire signal and switching system of the New York Central Railroad below 125th street was put out of operation, followed by a fire in the control tower at 57th Street and Park Avenue. The cause of the outage was later ascribed to a “ground current” that had invaded the electrical system. Brewster New York, railroad officials formally assigned blame for a fire destroyed the Central New England Railroad station, to the aurora (New York Times, 1921).
- August 2, 1972: The Bureau of Reclamation power station in Watertown, South Dakota experienced 25,000-volt swings in its power lines. Similar disruptions were reported by Wisconsin Power and Light, Madison Gas and

Electric, and Wisconsin Public Service Corporation. The calamity from this one storm didn't end in Wisconsin. In Newfoundland, induced ground currents activated protective relays at the Bowater Power Company. A 230,000-volt transformer at the British Columbia Hydro and Power Authority actually exploded. The Manitoba Hydro Company recorded 120-megawatt power drops in a matter of a few minutes in the power it was supplying to Minnesota.

- March 13, 1989: The Quebec Blackout Storm - Most newspapers that reported this event considered the spectacular aurora to be the most newsworthy aspect of the storm. Seen as far south as Florida and Cuba, the vast majority of people in the Northern Hemisphere had never seen such a spectacle in recent memory. At 2:45 AM on March 13, electrical ground currents created by the magnetic storm found their way into the power grid of the Hydro-Quebec Power Authority. Network regulation failed within a few seconds as automatic protective systems took them off-line one by one. The entire 9,500 megawatt output from Hydro-Quebec's La Grande Hydroelectric Complex found itself without proper regulation. Power swings tripped the supply lines from the 2000 megawatt Churchill Falls generation complex, and 18 seconds later, the entire Quebec power grid collapsed. Six million people were affected as they woke to find no electricity to see them through a cold Quebec wintry night. People were trapped in darkened office buildings and elevators, stumbling around to find their way out. Traffic lights stopped working, Engineers from the major North American power companies were worried too. Some would later conclude that this could easily have been a $6 billion catastrophe affecting most US East Coast cities. All that prevented the cascade from affecting the United States were a few dozen capacitors on the Allegheny Network (Odenwald, 1999).
- October 30, 2003: Malmö, Sweden, population 50,000 lost electrical power for 50 minutes (Pulkkinen et al., 2005). The blackout was caused by the tripping of a 130 kV line. It resulted from the operation of a relay that had a higher sensitivity to the third harmonic (=150 Hz) than to the fundamental frequency (=50 Hz). The excessive amount of the third harmonics in the system has been concluded to have resulted from transformer saturation caused by GIC. Currents as high as 330 Amperes were recorded on the Simpevarp-1 transformer (Wik et al., 2009).
- October, 2003: South Africa Transformer Damage. The ESKOM Network reported that 15 transformers were damaged by high GIC currents.

Extensive studies have already been conducted on the most cost-effective means for reducing or eliminating GICs in electric power grid components (Kappenman, 2010). The strategies generally include adding individual capacitors to each of the transformer HV lines, or adding a blocking resistor or capacitor to the ground lines in all transformers. Blocking capacitors were, for example, installed on the entire Hydro-Quebec power grid following the March 1989 blackout, as well as the WECC region in the western US. Although this strategy seemed to be successful in reducing GICS and reactive power on some of the lines, the impact was deemed only modest, 12% to 20% for the WECC network with 50% penetration, given the cost expended. Adding blocking capacitors to the transformer neutral ground connector is the simplest and most direct method for achieving a 100% reduction in DC GICs from transformer primaries, but this method is known to alter the impedance of the network in unpredictable ways as the devices are selectively deployed rather than universally adopted.

The next most direct, and also the most cost-effective method is by adding a low-ohmage and low-voltage resistor to the neutral ground of each 3-phase transformer (see red boxes in figure). Preliminary studies (Kappenman, 2010) suggest that this method could achieve a 60% reduction in GIC amperages to transformer primaries. The cost would be at most $100,000 per transformer in the US power grid, which contains some 5000 transformers, for a total cost of about $500 million. A simulation of the Hydro-Quebec power grid during the 1989 failure, but with neutral ground resistors installed reveals a dramatic reduction in the GICs to which the 45 transformers in the 735 kV grid were subjected, with hypothetical 10-ohm blocking resistors reducing the GICs from 550 amps to only about 75 amps.

The maximum storm time disturbance was about 450 nT/min., but even with proper mitigation, the US grid may not be immune from the largest known geomagnetic events, although the severity of the impact could be reduced by 60% from the case where no such mitigation is implemented. During the 1921 storm, a disturbance field of 4800 nT/min. was estimated. Without mitigation, over 500 transformers would be damaged, but with mitigation only about 40 would be damaged according to these simulations (Kappenman, 2010). This tenfold reduction is not inconsequential.

It is also worth mentioning that, although blackouts are a dramatic consequence of severe GICs caused by space weather, economic consequences also flow from the on-going stresses to the power grid during non-black out conditions. For example, Forbes and St. Cyr (2004) note that the constant impacts of minor space weather events over a long period of time disrupts the system that transmits the power from where it is generated to where it is distributed to customers. In examining the determinants of the real-time electricity market price over the period June 1, 2000, through December 31, 2001, they concluded that solar storms (over this period) increased the wholesale price of electricity by approximately 3.7 percent or approximately $500 million. Kappenmann (2012) has recently shown that in the months following the March 1989 Quebec event, a statistically significant number of transformers in the United States had to be prematurely replaced, with the greater number of replacements found in proximity to the Quebec power grid.

Of course, not all electrical power blackouts have anything to do with space weather. Most of us have experienced at least on "outage", and in some regions like Washington, DC, it is typical to have 3-5 outages every year lasting from hours to days. Hamachi-LaCommare and Eto (2004) have studied the economic costs of annual power outages and power "sags" and have found that they cost as much as $130 billion annually to the GDP. We are accustomed to electrical blackouts and quietly absorb them into our economy, with some grumbling about lost food and time. The long term trends for normal blackouts also points to the progressive failures inherent to an ageing domestic power grid (Karn, 2007). The over use of this resource is highlighted by the dramatic growth in bulk power transactions on, for example, the Tennessee Valley Authority system which exploded from less than 20,000 such transactions in 1996 to more than 250,000 by the end of 2001 (Dept. of Energy, 2005).

Increased bulk power transactions have led to a substantial drop in capacity margin, which provides little room either for growth or to maneuver in times of crisis. By some accounts (Patterson, 2010) there were 41 blackouts nationwide between 1991-1995, and 92 between 2001-2005. In 2011 alone, there were 109 affecting communities of 50,000 or more people. The Eaton Corporation, an agrigator of news and industry reports of blackouts across the US states, finds that between 2009 and 2011, the number of power outages rose from 2,169 to 3,041 and the number of people impacted climbed from 26 million to 42 million (Eaton Corporation, 2011).

A "typical" person comes into contact with the following technologies each day: cell phones, portable computing, credit card verification (ATM), navigation (GPS), electrical utilities (water pumps, gasoline pumps, hospital facilities, home lighting, city electrification, cell-phone recharging). All of these "essential" systems rely on electricity either at the point of creation (satellite GPS and ATM verification) or at the point of delivery (cell phone, gas pump, water, etc). All are expected to be ready when needed with 100% reliability. In recent human history, we have been successful in delivering these services even in the face of a number of space weather events. The lynchpin technology is, of course the electric power grid which citizens use to "tap into" essential communication and utility resources. It is unlikely that even a Superstorm event will dramatically impact the number of satellites operational, and backup transponders are readily available in case of emergencies. The ubiquitous cell phone would not fail if satellites failed, but satellites do carry the bulk of financial transactions, GPS and military CCC (command, control, and communications) traffic. The loss of key satellites, or a critical number, would render these services reduced in capability.en made of the cascading problems involved in "re-booting" such a large grid, especially in the event of component failures and burn-outs which would necessitate replacement, not on a local scale, but quite possibly on a global scale, with only a few key manufacturers able to service these needs.

2.1.8 Airline travel

Generally, the known routes for space weather impacts to aviation are through passenger safety (radiation), flight avionics (computer/system glitches), communications (radio interference) and scheduling (delays, route changes). Historically there have been anecdotal instances of each of these being identified. For example, July 19, 1947 - Sunspots delay planes (New York Times, July 18, 1947 p. 15).

Although earlier flight navigation methods involved compass bearings and LORAN-C, which in principle could be affected by geomagnetic storms and shortwave interference, there are actually no known instances where space weather events caused significant disruptions to these navigation technologies. Today, however, the airline industry is adopting GPS navigation as the standard, and its implementation in the Wide-Area Augmentation System (WAAS). WAAS is a combination of GPS and local, ground based metrology reference stations that provides 1-meter lateral and 1.5-meter vertical position resolution every 6 seconds for aircraft flying over the continental US, Canada and Alaska. There are also several satellites, such as Galaxy-15 that are involved in the WAAS system as part of its space-leg. Because its uses GPS satellites, the WAAS system is not immune to space weather effects that impact the ionosphere. Consequently during a severe storm event, WAAS may not be available for several minutes, or even hours, in some regions of the normal coverage area (Doherty, 2011). Studies have shown that approaches with vertical guidance (APV) are restricted during times of geomagnetic activity in terms of the APV coverage with Dst during the period from July 2003 to March 2004. For airports in which APV coverage is not available, flights must revert to IFR or VFR landing regulations within a vertical distance of 200 meters of the runway.

A number of national and international studies have been conducted to assess the radiation load on passengers during active space weather conditions and under otherwise normal circumstances. Normal background radiation doses are typically 0.3 microSv/hr or 3 mSv/year. For passengers and flight crews, the actual cabin exposure varies with the geographic latitude of the flight, the altitude of the flight, and the combined GCR and solar fluxes of particles. For example, Bottollier-Depois et al. (2000) determined from direct measurements at maximum solar activity in 1991-1992 and at minimum activity in 1996-1998. The lowest mean dose rate measured was 3 microSv/hr during a Paris-Buenos Aires flight in 1991. The highest rates were 6.6 microSv/hr during a Paris-Tokyo flight on a Siberian route and 9.7 microSv/hr on Concorde in 1996-1997. A number of similar studies since then have supported the idea that there is in fact some additional passenger and flight crew radiation exposure caused by space weather. However, the levels are cumulatively very low for the vast majority of passengers who travel infrequently during the year.

Nevertheless, some airlines that fly polar routes, such as United Airlines, are sensitive to solar storm events, not necessarily for the added radiation load, but for the disruption of emergency high-frequency communications with ground controller, which violates FAA safety regulations. For example, on January 20, 2005 a severe solar storm event caused 26 United Airlines flights to be detoured to lower altitudes and latitudes. The steady increase in the number of polar routes suggests a larger number of people will be affected by such events as time goes on. As Figure 2.6 indicates, currently, 1.7 million passengers travel these routes each year (Murtagh, 2010).

The most recent, well publicized event where airlines were diverted to other routes came with the powerful January 23, 2012, solar flare and CME. Hailed as the biggest solar storm since 2003, Delta Airlines chose to divert its flights to more southerly routes, while American Airlines took no operational action (Waugh, 2012). Despite the impacts to airline communications and flight safety, there are no instances where space weather has affected passengers or airline flight crews, so in some sense, the issue of flight safety and space weather presents a very modest health risk, but ironically a significant cost in flight time and fuel to airline companies that choose to apply mitigation.

2.2 Forecasting strategies

The specific components of the space weather environment that are known to cause human impacts are solar x-ray flares, coronal mass ejections, solar proton events, geomagnetic storms, galactic cosmic rays, electrostatic discharges, and energetic particles in the magnetosphere. X-ray flares cause ionospheric changes and upper atmosphere heating, which cause problems for LEO satellites and GPS-based systems. During the impulsive phase of a CME, shocks also form that lead to the acceleration of particles and solar proton events.

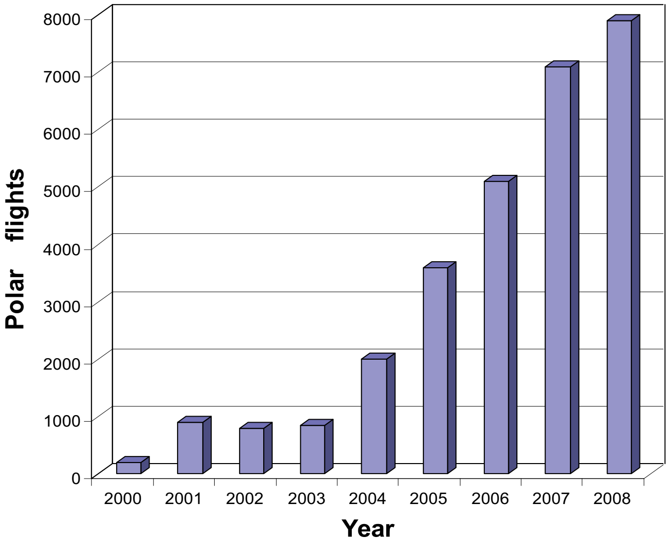

Fig. 2.6. The number of passengers flying "polar routes" continues to sharply increase each year to a current level of nearly 1.7 million passengers each year (Murtagh, 2010).

Some of these energetic particles can also arise from the site of the solar flare itself. Solar proton events and galactic cosmic rays can have energies of 10s of MeV, which lead to SEUs in computer circuitry, or to enhanced radiation exposure by airline passengers and astronauts.

The complexity of the space weather environment, and the many ways in which it can invade our technology to cause problems, almost precludes that we will ever be able to start from an initial set of solar or geophysical data and use this to determine whether a specific transformer or satellite system will fail. Consider that a satellite can be rendered inoperable if a single energetic particle causes a permanent failure of a critical gate, and the vast number of these particles that pass through a satellite's volume during its operational lifetime. In quantum electrodynamics, arguably the most precise physical theory known, precisions of 1 part per 10 billion are routine, but in terms of that fatal energetic particle, we would need a predictive algorithm of nearly the same caliber, otherwise we are forced to always deal in probabilities.

We also know that, in the space sector with over 2000 operating satellites, it is a rare event for satellites to actually fail during solar storms. How is it that, in the most recent Galaxy-15 or Anik-F6 outages, other satellites nearby were not similarly disturbed? We see this curious paradox again and again in reports of satellite outages related to space weather events. The industrial response is that satellite failures have much less to do with external space weather events, which reasonably should have affected more than one satellite simultaneously, than with manufacturing or software problems internal to the satellite itself such as tin wisker growth (e.g. Galaxy-7 in about 2000), solar panels designs (e.g. Tempo-2 around 1997), or software errors (e.g. Galaxy-15 around 2010). This raises another important issue that has been a much-discussed topic among space weather forecasters. How can you predict which space weather events will be important if the various industries that control the vulnerable assets are not transparent with their anomaly data?

For most types of space weather events, by the time they are detected it is already too late to mitigate. This is the case for all of the events that flow from solar or cosmic energetic particles, or solar X-ray flares. Phenomena such as CMEs, on the other hand, provide us with 1 to 3 days notice of arrival near earth once they are spotted leaving the solar vicinity. To keep forecasting costs low, what we would like to do is to come up with a small number of inexpensive measurement indices of the solar and geophysical environment, and through some yet to be developed algorithm, convert these into a statement about whether a particular resource or asset is in eminent danger and what the nature of that danger might be so that industry can take action. We also want to minimize the number of false alarms which can be costly and result in progressive lack of confidence in the forecasts themselves.

2.2.1 Solar storms: flares and CMEs

The simplest correlations we can search for involve solar flares, CMEs, SPEs and the sunspot cycle, because sunspots can be inexpensively counted and studied with ground-based instrumentation. We are reasonably certain that solar flares and magnetic reconnection events require concentrated photospheric magnetic fields, which manifest themselves as sunspots, so we expect more of these events during times of sunspot maximum than sunspot minimum. Yet even during sunspot minimum the number of significant flare events is not zero as for example the M6.4 flare on February 7, 2010, or the X2.6 flare on July 9, 1996. So if you are operating a GPS system or a WAAS system and require 100% coverage for safety, you will need a back up plan even during sunspot minimum. What about coronal mass ejections and the effect they can have on electric power grids?

According to an analysis of 314 halo CMEs during the 23rd cycle by Tripathi and Mishra (2005), about 3 occurred during sunspot minimum (1996) and 61 during sunspot maximum (2001), yet fewer than 1 in 6 halo CMEs led to significant geomagnetic storms.

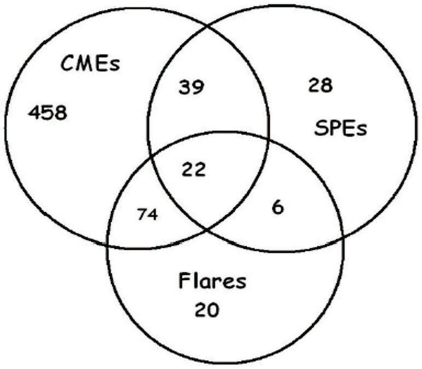

Fig. 2.7. A simple Venn diagram showing the frequency of halo CMEs, X-class flares and Solar Proton Events during Cycle 23. There were a total of 598 Earth-directed CMEs, 95 Solar Proton Events, and 122 X-class flares. It is clear from the intersection statistics that the vast majority of CMEs do not result in SPEs, or are associated with X-class flare events. However, it is also true that the majority of X-class flares are associated with CMEs, and that the majority of SPEs are associated with CMEs as well.

Odenwald (2007) created a combined data base of X-class flares, halo CMEs and Solar Proton Events (SPE) for the period January-1996 to June-2006, during a time in which SOHO/LASCO detected 11,031 coronal mass ejections. Of these, 1186 were nominally 'halo' events including back-side ejections, however, only 598 were actually directed towards Earth. During the same period of time, 95 solar proton events were recorded by the GOES satellite network orbiting Earth. Of these SPEs, 61 coincided with halo CME events. Solar flares were also recorded by the GOES satellites. During this time period, 21,886 flares were detected, of which 122 were X-class flares. Of the X-class flares, 96 coincided with halo CMEs, and 22 X-class flares also coincided with 22 combined SPE+halo CME events. There were 6 X-flares associated with SPEs but not associated with halo CMEs. A total of 28 SPEs were not associated with either halo CMEs or with X-class solar flares. The result can be summarized in the Venn diagram shown in Figure 2.7.

What this simple statistical exercise shows is that many of X-class flares (20 of 122), halo CMEs (458 of 598) and SPEs (28 of 95) are maveric events not associated in general with the other two types of phenomena. One cannot use halo CMEs to predict if an SPE will result (only (39+22)/598 = 10% of the time). One cannot use X-class flares to determine whether an SPE will result (28/122 = 23% of the time), or using halo CMEs to predict X-class flares, (96/598 = 16% of the time), but if an X-class flare is seen, then you have a (96/122 =) 79% chance that a halo CME results, or a (28/95 =) 29% chance that an SPE results. These statistical results demonstrate that the path to any sensible form of space weather prediction will probably always be fraught with the shear uniqueness of each and every space weather event, and the way that it is then "transduced" by a myriad of technological platforms whose properties and succeptabilities are often out of the public domain. This is far less like the tornado that rumbles through a state and wreaks havoc than a lightning storm whose strike points are utterly random on the landscape and damage or death is literally a matter of bad luck. This may well be the situation for global forecasting, but at the individual active region-scale, the situation is fortunately much more optimistic.

2.2.2 Reliability of X-class forecasts

Solar flares and CMEs are the most dramatic precursors of transient changes in space weather conditions, and considerable effort has been expended in developing predictive schemes for them (for a review see Forbes, 2010). In all cases, the ability to predict whether a flare or a CME will occur depends on the quality of the data gathered through the deployment of sophisticated, and expensive equipment. CMEs cannot be studied without space-based coronagraphic equipment (e.g. SOHO/LASCO), sensitive photometers that detect scattered light from them in transit (STEREO), or in-situ particle and field measurements made at L1 (e.g. ACE). Presently, there are no ground-based techniques for studying CMEs that could lead to significant cost savings over using space-based assets that need to be replaced every 10 years or so. Similarly for solar flares and the solar X-ray emission, only space-based sensors provide the data required to detect and quantify their severity, with no ground-based analogues to the X-ray technology. The good news is that both CMEs and X-ray flares produce distinctive "Type-II" bursts of radio-wavelength radiation that ground-based radio telescopes can profitably detect (e.g. Gopalswamy et al., 2005). By combining ground-based and space-based data over the last 20 years, considerable progress has been made in developing reliable forecasting algorithms which can provide nearly 100% certainty over the next 24-hour period for significant CME or flare activity. Most rely on a description of the topology and morphology of sunspots and their precursor magnetic fields.

Solar flares have been studied extensively since the 1930s since they are historically known to cause shortwave outages. It has been understood for some time that sunspots with complex field topologies are prime candidates for flaring activity (Hudson, 2010). Modern-day analyses that incorporate precursor information about changes in sunspot field topology such as rotation and shear, and past time history of activity (Nunez et al, 2005), the McIntosh classification of the sunspot group (Gallagher et al., 2002) lead to forecasts of X-class flares in the next 24-hour period that are more than 90% reliable with few false-positives. More recently, Colak and Oahwaji (2007) use a neural network approach to achieve prediction accuracies of 92% for occurrence and 88% for classification (M or X-class). None of these statistical methods actually employ any physics-based knowledge of the underlying flaring process, but merely search for correlations among a diverse ensemble of parameters available in various data bases and archives.

Detailed measurements of the 3-d shape of active region, surface magnetic fields and their classification (e.g., the Wilson classifications), has led to most of the advances in flare forecasting during the previous sunspot cycle. The legacy of this surface-field approach is best shown in the research by Steward et al. (2011) who used data from the National Solar Observatory's, Global Oscillation Network Group (GONG) to investigate strong-gradient polarity inversion lines, and neutral lines in maps of solar magnetic fields near active regions. By classifying each active region seen in 2003 according to a 5-parameter scheme (e.g., field gradient strength, curvature, neutral line length), they statistically compared 44 combinations of these 5 indices and found an optimal set that maximized the accuracy of predicting a flaring event. The resulting, optimized algorithm, called FlareCast, can predict a flare with 88% confidence within a 24-hour period, with a 10% false-positive rate (i.e. 1 prediction out of 10 will turn out not to occur).

But events transpire so rapidly that, by the time surface fields start to become twisted into the pre-flare state, it is already too late to prepare Earth-based systems for the resulting burst of X-rays and energetic particles, which arrive in less than an hour. During the first decade of the 21st century, a steadily accumulating archive of sub-surface active region images from the GONG program, has allowed changes in the plasma flows some 65,000 km below the surface to be studied at a 10-minute cadence, for over 1,000 active regions. In a path-breaking paper by Reinard et al. (2010), reoccurring sequences of magnetic topology change were discovered that presaged surface flaring events, with a lead time of 2 to 3 days before the surface eruption occurred. A vorticity-based index allows active regions to be classified in terms of its future activity, and can discriminate between regions producing C, M and X-class flares. The study supports the idea that rotational kinetic energy twists sub-surface fields into unstable configurations, which are then involved in explosive magnetic reconnection at the surface.

Coronal mass ejections have a shorter history in the space weather community since their role in the dynamics of the geomagnetic field was only deduced in the 1980s, and actual space weather effects were not observed until the Quebec Blackout in 1989. X-class flares have been studied as precursors for CMEs (Wang et al., 2002) as have sigmoidal features revealed in soft X-ray imaging of the corona (Sterling, 2000; Canfield et al., 1999), and vector magnetogram studies of the length of the "strong-field, strong-shear" main neutral line in sunspot groups (Falconer, 2001). Actual predictive schemes remain lacking due to the failure to bridge the gap between the changes in the small-scale surface fields near measurable active regions, and the often hidden large-scale magnetic field rearrangements that occur before the CME is launched.

Once evidence is available for a CME launch, the arrival time at Earth and the geoeffectivness become important predictive issues. Some progress has been made in this area. It has been known for several decades that the most efficient energy transfer into the geomagnetic field occurs for "south directed" CME field orientations. This can currently only be determined by in situ measurements made by satellites at L1 such as ACE. Also, the speed of the CME and the quantity of entrained plasma determine the ram pressure of the arriving CME at the magnetosphere boundary. The effect of aerodynamic drag by the interplanetary medium has also been studied (Song, 2010), and the results provide significantly improved determinations of the initial CME speed, its speed at 1 AU, and the transit time. The advent of STEREO spacecraft imaging of CMEs at large angles from the sun-earth axis have verified the deceleration of fast-moving CMEs in the interplanetary medium, and that CMEs need to be tracked at least 30 degrees from the Sun in order to obtain arrival time accuracies less than about 6 hours.

Solar proton events, often associated with shock acceleration in the initial stages of CME ejection, have also entered the domain of forecasting through the same statistical studies that proved successful with X-ray forecasting. For instance, Chin (2005) used an archive of 28 SPE between 1997-2000 and compared these events with solar radio bursts recorded between 245 MHz and 15,400 MHz to find a strong correlation between Type-III radio bursts at 245 MHz and the appearance of an SPE observed some 1-2 days later.

2.3 Modeling the economic and societal impacts

Cycle 23 will be seen by historians, no doubt, as a watershed moment in space weather history. Prior to Cycle 23, there was little or no public discussion about space weather vulnerability during the Space Age, although our grandparents surely knew all about the practical consequences of space weather and the insufferable short wave outages. With Cycle 23, we had SOHO providing the public with dazzling and ominous movies of solar storms, and many popularizers, including myself, who went on the stump to sort out for the public all the ways in which we could be affected. Then, just before the famous Halloween Storm of 2003, we had the first high-profile Congressional hearing about space weather in the context of why NOAAs Space Environment Center (SEC) budget should not be halved. Once Homeland Security became involved, we then had a new round of hearings about our infrastructure ulnerability to space weather events. The Space Weather Forum was held in Washington, DC on Capitol Hill in June 2008 to educate Capitol Hill about space weather issues. Meanwhile, researchers began the difficult task of trying to quantify what these impacts could cost us and the social disruption that might follow.

Kappenmann (1997) has an extensive record of modeling the US power grid with increasingly more sophisticated models of the electrodynamics of GICs and exhaustive studies of the North American electric grid network at the component level. Currently, his efforts use historical geomagnetic storms (e.g. 1921 event) and their impact on the contemporary electric power grid. Among the scenarios are four year-long recovery periods costing over $1 trillion in GDP.

Teisberg and Weiher (2000) estimated that the economic benefits of providing reliable warnings of geomagnetic storms to the electric power industry (alone) would be approximately $450 million over three years (note that this does not include any other impacted industries). This is well above the $100 million cost of a new operational satellite that would provide such warnings (ACE, Triana)

Odenwald and Green (2007) modeled the economic losses to commercial satellites in LEO, MEO and GEO orbits and deduced that an 1859-scale "superstorm" arriving near sunspot maximum could cost $50 billion in lost revenue and assets.

In August 1988, Oak Ridge National Laboratory and the NRC published "Evaluation of the Reliability for the Offsite Power Supply as a Contributor to the Risk of Nuclear Plants". This set the stage for considering the impact of space weather-related GICs on the reliability and safety of nuclear power plants (Kirby et al., 1988).

In April 1989, Northwest Power Coordinating Council (NPCC) approved the document "Procedures for Solar Magnetic Disturbances Which Affect Electric Power Systems" which has been updated several time.(NPCC, 1989) October 2003 What is Space weather and who should forecast it? Congressional Hearing on Space Weather held before the Subcommittee on Environment, Technology, and Standards, Committee on Science, House of Representatives, One Hundred Eighth Congress, first session, October 30, 2003, (Congress, 2003)

In December 2005, Idaho National Laboratory and NRC published “Reevaluation of Station Blackout Risk at Nuclear Power Plants–Analysis of Station Blackout Risk.” The executive summary from this report reads in part: The availability of alternating current (ac) power is essential for safe operations and accident recovery at commercial nuclear power plants (INL, 2005).

April 2008 saw the publication of the ”Report of the Commission to Assess the Threat to the United States from Electromagnetic Pulse (EMP) Attack: Critical Infrastructures”. The US Congress funded a vulnerability assessment research under the National Defense Authorization Act to evaluate the impact of an electromagnetic pulse (EMP) from a high altitude nuclear detonation by a terrorist event on the nation’s critical infrastructure including the electric grid. The same study also discussed geomagnetically-induced currents (EMP Commission, 2008).

In 2008 “Severe Space Weather EventsUnderstanding Societal and Economic Impacts Workshop Report”. The National Academy of Sciences determined that severe geomagnetic storms have the potential to cause long-duration outages to widespread areas of the North American grid (NAS, 2008).

In June 2010, the report entitled ”High-Impact, Low-Frequency Event Risk to the North American Bulk Power System” was published, jointly sponsored by NERC and the Department of Energy. NERC now concedes that the North American power grids have significant reliability issues in regard to High-Impact, Low-Frequency events such as severe space weather. The NERC report explains commercial grid vulnerability to space weather (NERC, 2010)

In October 2010 a report entitled “Electromagnetic Pulse: Effects on the U.S. Power Grid” appeared. In relation to that, Oak Ridge National Laboratory released a series of comprehensive technical reports for the Federal Energy Regulatory Commission (FERC) in joint sponsorship with the Department of Energy and the Department of Homeland Security. These reports disclose that the commercial power grids in two large areas of the continental United States are vulnerable to severe space weather. The reports conclude that solar activity and resulting large earthbound CME, occurring on average once every one hundred years, would induce a geomagnetic disturbance and cause probable collapse of the commercial grid in these vulnerable areas. The replacement lead time for extra high voltage transformers is approximately 1-2 years. As a result, about two-thirds of nuclear power plants and their associated spent fuel pools would likely be without commercial grid power for a period of 1-2 years (Oak Ridge Labs, 2010).

Armed with all this bad news, and with the storms of Cycle 24 now beginning, it has become commonplace for Reporters to quote these studies and offer titles such as “A big solar storm could cost $2 trillion, could be a global Katrina” or “Solar storm buffets Earth: How protected is the US power grid?”’. The danger is that, through constant repetition of this Doomsday theme, the public will become inured to the message in the face of the inevitable false alarms such as the January 2012 storm. While it is certainly important to keep the preparation message alive given the consequences to our infrastructure, as scientists and space weather forecasters, we need to be more careful with delivering this complex message to a public increasingly eager for a simple “yes or no” answer to their safety.