Impacts of Severe Space Weather on the Electric Grid

Contact: Dan McMorrow — dmcmorrow@mitre.org

November 2011

JSR-11-320

Approved for public release. Distribution unlimited.
### Abstract

Tasked by the Department of Homeland Security, the 2011 JASON Summer Study focused on the impact of space weather on the electric grid, seeking to understand 1) the current status of solar observations, warnings, and predictions, 2) the plausibility of Mr. Kappenman's worst-case scenario, 3) how previous solar storms have affected some power grids, and 4) what can be done at reasonable cost to protect our grid. This report builds on two previous JASON studies of different aspects of the U.S. electric grid.

### Subject Terms

Plutonium properties, ambient condition studies, equation of state, QMU framework,
## Contents

1 EXECUTIVE SUMMARY ................................................. 1  
1.1 Background ...................................................... 1  
1.2 Principal Findings ............................................... 1  
1.3 Principal Recommendations ................................. 3  

2 INTRODUCTION ....................................................... 7  
2.1 Study Charge .................................................... 7  
2.2 Briefers and Correspondents ................................. 9  

3 SPACE WEATHER AND ITS IMPACTS .......................... 11  
3.1 Space Weather ..................................................... 11  
3.2 Coronal Mass Ejections ......................................... 15  
3.3 CME and the Magnetosphere ................................. 17  
3.3.1 Probabilistic predictions of space weather .......... 18  
3.4 Geomagnetically-Induced Currents (GICs) ............... 21  
3.5 Solar Storms and Variability ............................... 23  
3.5.1 The Carrington Event ....................................... 23  
3.5.2 Electric field variability ................................... 25  
3.5.3 Statistics of extreme events ............................. 27  
3.5.4 Solar cycles .................................................... 29  
3.5.5 Frequency of past solar proton events (SPEs) from 

   nitrate in ice cores ......................................... 30  
3.6 Summary .......................................................... 34  
3.6.1 Findings ....................................................... 34  
3.6.2 Recommendations .......................................... 35  

4 ELECTRICAL GRIDS AND THEIR VULNERABILITIES ....... 37  
4.1 Structure and Regulation .................................... 37  
4.2 Components and Operation ................................. 39  
4.3 Kapppenman (2010) Worst-Case Scenario for Grid Damage 

   by Geomagnetic Storms ..................................... 51  
4.4 Experiences of Some Grids .................................. 53  
4.5 GIC Mitigation .................................................. 56  
4.6 Monitoring and Simulation .................................. 62  
4.7 Findings and Recommendations ............................ 64  
4.7.1 Findings ....................................................... 64
4.7.2 Recommendations ........................................... 66

5 WARNINGS, MONITORING, AND SIMULATION 69
  5.1 Warnings ....................................................... 69
    5.1.1 Space weather prediction center .................. 69
    5.1.2 Community coordinated modeling center .......... 71
  5.2 Monitoring and Simulation ............................... 72
    5.2.1 Monitoring .............................................. 73
    5.2.2 Simulation .............................................. 74
  5.3 Summary .................................................... 75
    5.3.1 Findings ................................................ 75
    5.3.2 Recommendations ..................................... 77

6 CRITICAL OBSERVATIONS ................................. 79
  6.1 Ground-Based Magnetometers ............................ 79
  6.2 Critical Space Systems for Forecasting .............. 80
  6.3 A Minimum Constellation of Spacecraft for CME Space Weather Forecasting ....................... 84
  6.4 A Near-Term Replacement for the ACE Spacecraft .... 84
  6.5 Orbits for Solar Observations .......................... 85
  6.6 Artificial Lagrange Orbits via Station Keeping .... 86
    6.6.1 Ion propulsion for station keeping? ............. 87
    6.6.2 Solar sails .......................................... 89
  6.7 Quasi-satellite Orbits ................................... 90
  6.8 Summary .................................................... 93
    6.8.1 Findings .............................................. 93
    6.8.2 Recommendations ................................... 95

iv


1 EXECUTIVE SUMMARY

1.1 Background

Concerns about the vulnerabilities of technical infrastructure to space weather have been growing since the sun entered the early stages of the current sunspot cycle in 2009, increasing prospects for severe solar storms. The primary issue is not whether these storms will occur but the risks they pose to power grids, satellite communications and GPS. A worst-case scenario by John Kappenman [26] suggests the possibility of catastrophic damage to the U.S. electric grid, leaving millions without power for months to years while damaged major grid components, such as high-power, high-voltage transformers, are slowly replaced.

Tasked by the Department of Homeland Security, the 2011 JASON Summer Study focused on the impact of space weather on the electric grid, seeking to understand 1) the current status of solar observations, warnings, and predictions, 2) the plausibility of Mr. Kappenman’s worst-case scenario, 3) how previous solar storms have affected some power grids, and 4) what can be done at reasonable cost to protect our grid. This report builds on two previous JASON studies of different aspects of the U.S. electric grid [39, 22].

1.2 Principal Findings

1. Technical means exist to mitigate vulnerability of the power grid to severe solar weather resulting from Coronal Mass Ejections (CMEs), and this mitigation could be rapidly implemented if a decision were made to do so. Although exposed to strong solar forcing and large geomagnetically induced currents (GICs), Finland has avoided serious grid problems by specialized transformer design (J. Elovaara, personal communication, 2011). Following a grid collapse in 1989, Hydro Quebec
adjusted control relays, installed series capacitors in transmission lines and blocking capacitors on transformers to minimize chances of further problems [6].

2. Because mitigation has not been widely applied to the U.S. electric grid, severe damage is a possibility, but a rigorous risk assessment has not been done. We are not convinced that the worst-case scenario of [26] is plausible. Nor is the analysis it is based on, using proprietary algorithms, suitable for deciding national policy.

3. Observations vital for space weather warnings are presently supplied by SOHO, ACE and STEREO satellites. Launched for research in 1997 and eight years past its design life, ACE provides the only direct warnings that CMEs are about to hit Earth. Also launched for research, the two STEREO satellites provide the only three-dimensional observations of CME structure used to initialize MHD models propagating the events through the heliosphere.

4. Possible widespread and sustained grid damage is within the broader view of national security issues taken after 9/11, and severe space weather could be one of the causes. The federal response, however, is poorly organized; no one is in charge, resulting in duplications and omissions between agencies.

5. Warning times of imminent CME impacts based on direct observations may be extended from 30-60 minutes to 5-10 hours by placing operational satellites closer to the sun than L1, which is one million miles from Earth. Owing to the large inertia of power grids, the additional warning time would permit more fundamental mitigation measures, some requiring enough lead time to defer work such as maintenance. Research to obtain accurate predictions from models within a few hours of CME formation is also promising, but, although fundamental understanding of solar physics is rapidly improving, operational prediction
of future CME is a distant goal. Predicting CME will remain research efforts for the foreseeable future.

1.3 Principal Recommendations

1. Protecting vital grid components should have higher priority than avoiding temporary blackouts. The safety of vital components, e.g., transformers, should be insured by:

   (a) Using relays set to trip before grid equipment is seriously damaged while avoiding being tripped inappropriately by GIC harmonics,
   (b) Mandating component design standards,
   (c) Blocking GIC with capacitors in transformer neutrals together with short protection against ground faults, and
   (d) Using small series-blocking capacitors in transmission lines where neutral-current blocking is not feasible, e.g., with autotransformers.

2. A program should be established to insure that robust operational satellites for monitoring space weather are installed and optimally positioned. This program should include:

   (a) Funding for the ACE replacement called DSCOVR should be restored in the FY 2012 budget to minimize chances of losing CME observations at L1 before long-term monitoring satellites can be designed and launched,
   (b) Exploring the utility of a small constellation of low-cost satellites in quasi-satellite orbits near 0.1 A.U. from earth as eventual replacements of DSCOVR, such as the constellation outlined in the
SWx_Diamond proposal [44]. These would increase direct warning times tenfold and provide the first three dimensional CME observations, and

(c) Developing STEREO replacements for insertion at L4 and L5 or other suitable orbits.

3. Federal space weather efforts should be centrally directed by an official with authority to insure that all important aspects are covered efficiently. Identified issues needing attention include:

(a) Predictions of severe space weather on earth by NOAA and NASA should be treated as one coordinated effort that includes transitioning proven research predictions to improve operational warnings.

(b) Air Force and NOAA space weather programs would both benefit by being more tightly coupled, as they were in the past. In particular, ways should be explored to use Air Force satellites and sensors to improve civilian warnings, and the Air Force center in Colorado Springs should be able to backup NOAA’s Space Weather Prediction Center in Boulder if it goes down.

(c) DOE space weather and grid developments and programs should be coordinated with work outside DOE to obtain the best results at minimum cost. For instance, DOE wide-area grid monitoring at Oak Ridge should be available to the National Electricity Reliability Council (NERC) as it develops an operational monitoring center for the U.S. grid.

4. A ‘first-principles’ simulation of the U.S. grid should be developed to perform a rigorous risk analysis that can be repeated to uncover vulnerabilities, to test control algorithms and to evaluate proposed grid changes. This effort should begin with a thorough review of [26] in a forum that can protect the proprietary methods and compare them
with approaches used to evaluate other grids, e.g., in Finland. Such a simulation would benefit grid security generally and not be restricted to space weather vulnerabilities.

5. Development and basic research on understanding CME and predicting their effects on the electric grid are progressing well, meriting continued funding.

   (a) Development funding should be increased to transition research-grade forecasts into operational warnings.

   (b) Research into solar physics is improving understanding, but it is not yet sufficiently advanced to warrant funding operational predictions of future CME. Rather, these should still be considered as research efforts until useful predictions can be demonstrated.
2 INTRODUCTION

2.1 Study Charge

The Department of Homeland Security defined five task areas:

1. Improved Warning:

   - What can be done to increase geomagnetic storm warning time from 30 minutes to several hours? Should improvements be made jointly with international partners who also need the increased warning time and should they contribute to this effort?
   - What can be done to clarify policy and procedures to be followed for ground systems if we had several hours warning of an incoming CME? Today it is likely that almost nothing would be done or that we could even alert critical US and international activities in time to react.
   - GPS satellites are reportedly hardened against extreme space weather effects, but what critical space-based systems might be damaged or lost in a severe but credible space weather events?

2. How real and serious is the threat?

   - What is the probability risk to critical infrastructure such as the power grid, computers, aircraft, and communication systems?
   - U.S. critical infrastructure can be defined as assets, systems, and networks, whether physical or virtual, so vital to the United States that their incapacitation or destruction would have a debilitating effect on security, national economic security, public health or safety. How susceptible is our infrastructure to the risk?
3. Mitigation Strategies:
Can we harden our infrastructure against solar storms? What are the hardening protocols used by public and private sector? Can we harden critical systems against extraordinary solar activity by?

- Faraday cage protection (protective metallic shielding)
- Surge Protectors
- Special wire termination procedures
- Screened isolated transformers
- Spark gaps
- Electronic filters
- Backup units
- Satellites
- Smart grids

4. Recovery:
If a major CME event happened what recovery strategies are realistic and technically plausible? Does the private sector have recovery plans?

5. Research:
Should DHS invest in R&D in this area? In what domains? Do we need a national research effort to address warning, protection and reconstitution?

Recognizing that the charge was too broad for the Summer Study, we agreed with the sponsor to concentrate on the first three topics and focus on the electric grid with brief consideration of impacts on GPS availability and how that affects aviation.
2.2 Briefers and Correspondents

We are indebted to many briefers and correspondents for the quality and depth of their experience and the information they shared so freely with us.

On June 8 several of us (Brenner, Gregg, Max, McMorrow, and Prince) visited NOAA’s Space Weather Prediction Center in Boulder, where we were briefed about the center and its work by Tom Bogdan, the director, and Bill Murtagh, the program coordinator. Ed Baker, and B.C. Low gave shorter discussions about specific topics. The next day Brenner, Gregg, and McMorrow met with Capt. Josh Warner and Kelly Hand at Peterson Air Force Base in Colorado Springs.

Briefings held in La Jolla on June 16 and 17 were given by representatives from industry, government, and academia:

John Kappenman - Storm Analysis Consultants
Eric Rollison - North American Electric Reliability Corp. (NERC)
Tom Overbye - University of Illinois
Ed Schweitzer - Schweitzer Engineering
Craig Steigemeir - ABB
Jagadeesh Pamulapti - Office of Science and Technology Policy (OSTP)
Marcelo Elizondo - Pacific Northwest Laboratory (PNNL)

Very helpful telephone conversations were held with: Michael Hesse and Antti Pulkkinen (NASA Space Weather Laboratory), Mark Lauby and Eric Rollison (NERC), and Frank Koza (PJM, Pennsylvania, New Jersey, and Maryland) Regional Transmission Organization. John Kappenman answered many questions and supplied references to numerous relevant publications. In addition, Antti Pulkkinen, Jarmo Elovaara (Fingrid), and Leonard
Bolduc (Hydro Quebec) patiently provided extensive and detailed responses to endless questions. In addition to correspondence with the briefers above, additional help was given by B.C. Low (NCAR), Doug Biesecker (NOAA/SWPC), and Chris McFee (U.K. Science Office).

Finally, our sponsor, Scott Pugh of DHS, introduced us to the briefers and helped convince them to journey to La Jolla.
3 SPACE WEATHER AND ITS IMPACTS

The main phenomena of space weather impacting human infrastructure are examined, particularly in how they generate geomagnetically induced currents (GICs) in electric grids. The solar origins of space weather are reviewed, and histories of past solar storms and statistical measures are used to examine the likelihood of extreme solar storms many times more intense than those experienced in recent times.

3.1 Space Weather

Space weather refers to conditions in the solar system produced by radiation and particles ejected from the sun. Major phenomena producing space weather include:

The Solar Wind: The stream of magnetic flux and hot charged particles, plasma, flowing outward from the sun is known as the solar wind. Some particles seem to originate in the photosphere and others in the corona, particularly in equatorial streamers. The coronal contribution has typical velocities of 400 km/s and temperatures of 1.5 MK. Escaping through coronal holes, photospheric particles tend to be cooler, 0.8 MK, and faster, 750 km/s, flowing outward along open magnetic field lines. Due to the combination of their outward radial velocities and the sun’s rotation, both components move outward along curves known as Parker Spirals [34], like water jets from a lawn sprinkler. Owing to variations in density, speed, and direction, at any time several spiral streams dominate the structure of the solar wind (Figure 1).

Solar Flares and Coronal Mass Ejections: Sudden brightenings in the photosphere and corona above active regions associated with sunspots
are termed flares. By converting energy in the solar magnetic field, flares heat solar plasma to 10 MK or greater [19], emit bursts of electromagnetic waves and energetic particles (SEPs), and eject magnetized plasma known as coronal mass ejections (CMEs). Spanning the spectrum from radio waves to x-rays, their radiation can disrupt communications and GPS signals. Ejected particles sometimes damage satellite electronics. CMEs are discussed in more detail in section 3.2.

**Eruptive Prominences:** Quiescent prominences are arc- or cloud-like plasma structures suspended by magnetic fields above the photosphere. Containing plasma much ($\approx 100$ times) cooler and denser than the surrounding atmosphere, they extend to about 0.15 solar radii above the solar photosphere. When observed from above against the photosphere, these structures appear as dark filaments, but when observed on the solar limb against the corona, they appear as bright arches. More than half of CMEs are associated with eruptive prominences.

**Corotating Interacting Regions:** Corotating interaction regions (CIRs) are formed by adjacent low-speed and high-speed solar wind streams. As high-speed solar wind overtakes low-speed solar wind, an interaction region forms and then moves at a speed that is between the speeds of the two streams [18]. Interplanetary shock waves form at the front and back of the interacting region, moving away and toward the sun respectively. Typically these shocks are formed beyond the earths orbit. Energetic particles accelerated in the shocks can propagate to the earths neighborhood. When a CIR interacts with the earths magnetosphere, a geomagnetic substorm with a storm sudden commencement (SSC) may result, but they are infrequent.

**Solar Energetic Particle Events:** Bursts of solar energetic particles (SEPs) occur when elements of solar wind are accelerated to high velocities, either during ejection or while propagating through the heliosphere.
Solar Proton Events (SPEs) are an important subclass that can penetrate the magnetosphere when sufficiently energetic. Figure 2 shows proton fluxes associated with a CME. NOAA identifies an SPE when protons with energies $\geq 10$ MeV have fluxes exceeding $10$ particles per second per cm$^2$ per steradian.

Figure 1: NASA simulation of 7 June 2011, showing three solar wind spirals and a CME released two days earlier. The slow tail of the CME grazed Earth. Left) Plan view of the inner solar system with earth (yellow circle) fixed to the right of the sun. Middle) Altitude versus range looking from sun to earth shows mostly the solar wind extending across the plane of earth’s orbit. Right) Mercator projection of particle density at earth’s distance, rotated 90° to fit on the page. (Michael Hesse teleconference with JASON, June 2011.

Many large-scale disruptions on earth and in space can be traced to one of these solar phenomena. Affected infrastructure includes:

- Long electrical conductors
  - Electric grids are destabilized and some components are damaged by geomagnetically-induced currents (GICs)
  - Pipelines experience enhanced corrosion due to GIC
Figure 2: The intensity of energetic protons as a function of time for a solar energetic particle event associated with a CME on October 19, 1989 [42, 18]. NOAA’s threshold for defining an SPE is 10 MeV protons with fluxes of $10^3$ protons / cm$^2$ s str.

- Railroad signals and electrical systems are interrupted or damaged by GIC

- Satellites
  - Gamma-rays and fast particles damage satellite electronics
  - Radio noise at GPS frequencies causes receivers to lose lock

- Aviation
  - Aircraft on polar routes are re-routed during intense solar storms to maintain communications during radio storms and avoid radiation hazards to passengers and crew
Navigation systems built around GPS are disrupted when solar radio noise interferes with reception of satellite signals

- Communications
  - Radio and electrical interference compromise HF radio and telephone land lines and cell links

### 3.2 Coronal Mass Ejections

Normal CME begin as streamers that brighten about a day before erupting as massive releases of plasmas carrying intense magnetic fields. Uncertainties remain about most aspects of CME formation, but one important line of interpretation holds that they are generated when lower magnetic flux loops twist and shear an overlying loop, causing magnetic lines to reconnect, expelling the plasma and its magnetic field, usually at high velocity. Figure 3 illustrate a common sequence, beginning with a bright streamer evolving into into a characteristic three-part structure consisting of a bright frontal loop followed by a cavity and then a bright core.

CME are massive \((10^{11} - 4 \times 10^{13} \text{ kg})\), energetic \((10^{22} - 10^{25} \text{ J})\), and wide \((1^\circ - 100^\circ\) at the sun and 0.1 - 0.8 A.U. at earth’s orbit, where the Astronomical Unit (A.U.) is the average distance of earth from the sun, about \(1.5 \times 10^8\) km) [13]. Their initial energy is overwhelmingly magnetic. Though some are slow, CME are often fast, up to 2,500 km/s (W. Murtagh, NOAA/SWPC, personal communication, 2011), and reach earth in less than a day. Typical transits, however, are 2-3 days.

Figure 1 is a snapshot of a NASA simulation of a CME on 7 June 2011 showing its estimated position on 10 June. Three bursts of the solar wind and one CME are shown in the plan view at the left. Because the solar wind bursts were emitted continuously for days and are much slower than
the CME, they followed Parker spirals. Earlier stages of the simulation show CME originating as an arc stretched across the first quadrant. The model propagated it outward in an expanding cone, leaving it in approximately the same direction as it started in relative to earth. The CME’s core missed earth, which was brushed by the slow, low-density tail of the ejection. Because no in-situ three-dimensional observations have been made of CMEs, many details of simulations like this cannot be tested.

CME, from ejection to propagation through the heliosphere, are very difficult to predict in detail, owing to the complexity and multifaceted nature of their interaction with the sun, the solar wind and the magnetosphere. Currently, forecasting accurately where and when a CME will be emitted is not possible, but research is being done and may eventually succeed in understanding the mechanisms producing CMEs or and related phenomena. Several correlations indicate when CME are more likely and what what can
signal their release. For instance, [8] argue that all seven major GIC events since 1930 were due to CME during peak or declining phases of the sunspot cycle, when flares are more likely. Moreover, almost all very large, class X (based on X-ray flux and size) flares are associated with CMEs [23]. This line of analysis seeks statistical predictions of the likelihood that CME will be released and is exploring what detail is needed to be useful.

Once ejected, predicting propagation through the solar wind with high accuracy is challenging, though it is helped significantly by obtaining three-dimensional initial conditions from Stereo satellites. There is not now, however, a reliable way to measure the direction of the magnetic field in the initial ejection, and the interaction of the CME with the solar wind and the interplanetary magnetic field is sufficiently complex that even if the initial field direction were known, substantial uncertainty would still result when it collided with the magnetosphere (M. Hesse, NASA, telecon 2011).

3.3 CME and the Magnetosphere

The most dangerous CMEs to hit Earth are dense and fast, with magnetic fields anti-parallel to earth’s field. One measure of CME intensity is the solar wind convective electric field, $E_{sw}$,

$$E_{sw} = -V_x B_z \ [\text{V/m}] \quad (3-1)$$

where $V_x$ is the radial velocity, and $B_z$ is the strength of its magnetic field in geomagnetic coordinates. For example, the solar wind might have $B_z = 2 \text{ nT}$ and $V_x = 500 \text{ km/s}$, yielding $E_{sw} = -1 \text{ mV/m}$. By comparison, a CME could have $B_z = -20 \text{ nT}$ and $V_x = 1,000 \text{ km/s}$, giving $E_{sw} = 20 \text{ mV/m}$.

Regardless of the orientation of its magnetic field, a fast, dense CME will produce an electromagnetic disturbance at the surface by compressing the magnetospheres’s bow wave (Figure 4). The magnetosphere, however,
repels fields parallel to it, and the disturbance created is minor compared to that produced by CME carrying anti-parallel fields. Anti-parallel magnetic fields ‘reconnect’ with field lines in the bow wave and carry them toward the tail, allowing CME plasma to propagate into the magnetosphere. Aurora and electromagnetic fields are generated during the reconnection which can be sustained as field lines continue to reconnect in the tail while it is compressed by the passing CME. Owing to seasonal changes in the orientation of earth’s axis relative to the radial direction from the sun, ground-level magnetic disturbances recorded between 1868 and 1996 were twice as frequent when the sun was near equinoxes as during summer and winter [10]. CMEs that do reconnect, however, tend to impact electric grids more strongly during winter and summer, when electrical demand is highest and the grids are close to capacity.

3.3.1 Probabilistic predictions of space weather

Although we seem to be very far from accurately predicting the magnetic consequences of a CME from first-principles computer codes, software can and does assimilate and combine diverse sets of data into sound statistically-based predictive tools, with confidence intervals on the predictions. For the GIC consequences of CMEs, this has been carried out especially effectively by A. Pulkkinen and his collaborators in the context of Finnish GIC studies, e.g., [41]. Popular magnetospheric indices used in statistical analyses include:

**Dst:** A measure of equatorial magnetic storm strength, Dst is an hourly average, in nano Teslas, of horizontal field strength measured by four magnetometers located near the equator around the globe.

**Kp:** A 3-hour average of the range of horizontal field variations at thirteen magnetometers, expressed as an integer between 0 and 9. Power companies usually begin monitoring activity when Kp = 7.
Figure 4: Schematic of CME impacting the magnetosphere, adapted from [17]. In the upper panel, the CME’s magnetic field vector is parallel to earth’s which repels it magnetically, although the impact compresses the magnetosphere’s bow wave. In the lower panel, the anti-parallel CME magnetic field connects with magnetosphere field lines in the bow wave, opening it to plasma in the CME.

**AE, AU, AL:** Measures of auroral activity, these are field intensities at magnetometers close to the auroral ellipse. Beginning with 1-minute horizontal field strengths, \( AE = AU - AL \), where \( AU \) and \( AL \) are the highest and lowest values, in nano Teslas.

A recent study [33] correlated space-based metrics of CME strength with ground-based magnetometer measurements. The paper experimented with 20 empirical coupling functions motivated on physical grounds. These ranged from pressure in the solar wind to \( B_z \) and \( V_z \) and the kinetic energy of the wind, as well as a host of other formulae involving combinations of the velocity, density, intensity, and direction of the incoming magnetic field. The
authors find good correlations for functions coupling the velocity of the wind and the strength and direction of the field. The best index is $d\Phi_{MP}/dt \equiv V_x^{4/3}B_T^{2/3}\sin(\theta_c/2)^{8/3}$, where $B_T$ is the total magnetic field, and $\theta_c$ is the angle of the magnetic field relative to that of the earth’s field. Accounting for $r^2 = 0.69$ of the AE variance, where $r$ is the correlation coefficient, $d\Phi_{MP}/dt$ was derived empirically to represent the rate of flux opening at the magnetopause. A few other indices also have similar correlation coefficients, most notably $E_{WAV} = V_x B_T \sin(\theta_c/2)^4$.

Figure 5: Scatter plot of AE, an index of auroral activity, with $d\Phi_{MP}/dt$ [33]. $r$ is the correlation coefficient between AE and $d\Phi_{MP}/dt$.

This shows that at least some success can be found in correlating earth-based measurements with CME properties in space. Of course, the bulk of the data in this study involved normal variations in the solar wind rather than the most violent type of CMEs that lead to potentially the largest GICs. But, nonetheless, this is an excellent starting point. To our knowledge, it has not yet been directly demonstrated whether the distribution of GIC magnitudes can be explained by a reduced set of solar wind parameters. After all, the GICs depend on the time variability of the magnetic field at the surface of the earth and could be used as measures of CME characteristics. Although
there are complications with GIC-based approaches that do not occur with magnetic parameters, e.g., GIC depend on the ground conductivity profile. These are constants at each site and can be parameterized. Work of [41] demonstrates that such parameterizations can be quite simple, with far fewer parameters than might be expected.

Given this, and the success of finding reasonable correlates for solar wind parameters with magnetospheric indices, it seems likely to us that a parameterization can be found that would work equally well with GIC. Although such a correlate may not accurately predict large GICs, it would expose the physical principles producing them and be a valuable constraint for developing more sophisticated predictive tools.

3.4 Geomagnetically-Induced Currents (GICs)

Disturbances to the magnetosphere propagating to ground level can induce large currents in long electrical conductors mounted perpendicular to the magnetic field vectors. For a current loop with area $A_{\text{loop}}$ perpendicular to a changing magnetic field, the induced current is

$$i_{\text{GIC}} = \frac{dB}{dt} \frac{A_{\text{loop}}}{\Omega} \ [A],$$

(3-2)

where $\Omega$ is the total resistance of the loop. For the schematic electric grid in Figure 6, $d = 50$ m is a representative height of high-voltage towers, and $D$, the depth of the return flow, can vary between nearly zero for towers in sea water to 100 km or more for resistive ground, as in Finland.

Taking $dB/dt = 10$ nT/s for a strong magnetic storm, $\Delta x = 100$ km, and $\Omega = 1$ ohm, $i_{\text{GIC}} = 10^{-8} \times 10^5 \times 50 = 50$ mA when the return current is so shallow that $D \approx 0$. GIC of this magnitude are too small to damage grids or their components.Telegraphers discovered the importance of small loop areas when they avoided significant GIC by replacing the ground with a
Figure 6: Schematic of a long transmission line with grounds separated by distance $\Delta x$, towers holding transmission lines $d$ meters above the surface, and a return GIC at depth $D$. The magnetic field inducing the GIC is perpendicular to the page.

second telegraph wire close to the first [43]. In highly resistive ground, return flows are deep, greatly enlarging loop area but not necessarily increasing loop resistance. For $D = 100$ km, $i_{\text{GIC}} = 10^{-8} \times 10^5 \times 10^5 = 100$ A, far more than enough to cause serious problems.

During geomagnetic storms significant GIC can persist for hours, punctuated by short bursts of intense flow. Figure 7 shows an average S-transform of GIC during six geomagnetic storms near midnight. (A Fourier transform with a Gaussian time delay kernel, the S-transform refers displays the temporal evolution of frequency content [40].) The example in Figure 7 exhibits a GIC of 1-10 A lasting six hours at periodicities of 20 minutes to 2 hours. In addition, there were frequent bursts with periods of one to 10 minutes, also with amplitudes of 1-10 A. The Fourier transform, obtained by integrating the S-transform in time, is nearly flat over the slower range and has a power-law decrease at periods less than 4 minutes. Because the short bursts can be damaging, GIC measurements must resolve them, and data are often analyzed as averages over 10-20 s.

\[ S(\tau, f) = \int_{-\infty}^{\infty} h(g(t - \tau)) e^{-i2\pi f t} \, dt, \] where $g(t - \tau, f) = (|f|/\sqrt{2\pi}) e^{-f^2(t-\tau)^2/2}$, $f$ is frequency, and $\tau$ is the central time of the Gaussian window.
Figure 7: Time-frequency content of GIC [40]. a) Average S-transform spectra of GIC during six storms that peaked near midnight. (b) Fourier spectra obtained by integrating the Fourier transform spectra versus time. Black is the average Fourier transform; red and blue spectra were integrated during and after the storm, over times marked in (a) by red and blue lines. The straight line indicates the approximate power law scaling of the spectra.

3.5 Solar Storms and Variability

3.5.1 The Carrington Event

The largest recorded magnetic storm occurred between August 27 and September 7, 1859 and is known as the ‘Carrington Event’ after the amateur astronomer who associated the storm with a solar flare 17.1 hours earlier. The magnetic storm disrupted telegraphs worldwide, in some cases generating fires. Some operators disconnected their batteries and keyed using only the GIC flowing in the lines. Observed in northern South America,
the Aurora Borealis was so bright that it awakened hikers in the Rockies. Fortunately, magnetic deflections were recorded in England and elsewhere, allowing estimates of the storm’s magnitude. At Kew Observatory, this was done on photographic paper (Figure 8) advanced by a watch escapement. Colaba Observatory, India, reported a magnetic deflection of ≈ −1,600 nT, although the maximum at Bombay was only half that, ≈ −850 nT.

Figure 8: Horizontal magnetic field strength (upper) and direction (lower) for the full day of 1 September 1859 at Kew Observatory, during the Carrington Event. The arrow represents a magnitude of ≈ 100 nT. Just left of the arrow, the field strength appeared to go far off scale.

The observation of −1,600 nT at Colaba was modeled [29] by characterizing the CME by a density of ρ = 1,800 particles/cm$^3$, $V_z = 1,500 − 2,000$ km/s, and a field strength of $B_z = −60$ nT. The latter two yield $E_{sw} = 90 − 120$ mV/m for the CME’s electric field strength. By comparison, the quiet background preceeding the event was modeled with $B_z = 2$ nT, $V_x = 450$ km/s, and 5 particles/cm$^3$, giving $E_{sw} = 0.9$ mV/m. The authors noted that the rapid recovery from Carrington required a solar wind faster than any observed, making the event doubly unusual. They also note that estimated properties of a 1972 CME were as strong as Carrington, but apparently its magnetic field did not have a significant component anti-parallel to earth’s field.
3.5.2 Electric field variability

Probability densities of horizontal electric fields produced by CME were constructed using 10-s averages of GIC and simultaneous magnetic field measurements in Finland [41]. At each location, GIC and horizontal magnetic field intensity were used with ground resistivity profiles to estimate the magnitude of the surface electric field, $|E|$. Thirteen years of data were available, 1993-2006, but the histograms in Figure 9 are scaled as the number per century to examine the frequency of 100-year storms. Over low-resistivity ground at high latitudes, e.g., British Columbia, $|E| = 1$ V/km occurs $10^1 - 10^2$ times per year (obtained relative to $10^2$ values per century in panel a). Over high-resistivity ground, e.g., Quebec in panel b, electric fields of this intensity are likely to be 10 times more frequent. These events could occur during the same storm or be months apart. Extending the trend to the axis predicts 10 V/km at most once per century over low-resistivity ground and 10 - 100 times per century over high-resistivity ground. These frequencies are probably upper bounds, as the curves appear to be changing to more rapid descents at high field magnitude.

Estimated electric field magnitudes correlate well with $E_{sw}$ and $D_{st}$, permitting estimates of conditional probability densities $p(E|E_{sw})$ and $p(E|D_{st})$. These were used to estimate 4 V/km as the probable peak magnitude of the electric field during the Carrington Event. Means of these distributions tend to saturate and their variances decrease sharply with increasing $E_{sw}$ and $D_{st}$ (not shown here), indicating that CME forcing becomes increasingly less effective in generating ever larger $E$. Tails of distributions of intermediate forcing events extend to larger $E$ than do tails of larger $E_{sw}$ events. Consequently, the most extreme $E$, and GIC, may be associated with intermediate $E_{sw}$ forcing.
Figure 9: Number of 10 s geoelectric field magnitudes, $|E|$, produced by CME per 100 years, adapted from [41]. The estimates were made using simultaneous observations of GIC and geomagnetic fields in Finland with low and high-resistivity ground models characteristic of British Columbia (a) and Quebec (b). Each curve represents data from a different site. The arrow at 4 V/km marks the estimated magnitude of the Carrington Event, and the red line approximates power-law slopes at high magnitudes.
3.5.3 Statistics of extreme events

How do probabilities of geomagnetic storms, e.g., those examined by [41], compare with frequencies of other natural hazards? A different astronomical hazard offers a useful contrast, namely the hazard from impacts of asteroids or comets on the Earth. Even though such impacts have been negligible over recorded history, the geological and paleontological records show cataclysmic impacts, such as the one that arguably killed all the dinosaurs 65 million years ago, which was perhaps 10 million times as energetic as any in recorded history. The sobering result is that, during any period of time, most of the damage is expected to be done by one or a few largest impacts, not the many small ones. Estimates of this hazard have motivated significant programs of astronomical observation to warn of large, threatening objects.

In this section, to estimate the hazard due to possible geomagnetic storms larger than any recorded, we review what is known from long-term observations about the probability distribution of geomagnetic disturbances. We find that this distribution falls off steeply with size of disturbance. For instance, the largest geomagnetic storm of the last 1,000 years was not likely to have been more than about 3 times as intense (in induced EMF) as those in the last 100 years. In particular we will attempt to estimate the cumulative distribution $N(> Dst)$ of numbers of disturbances exceeding a given threshold of the geomagnetic index $Dst$, in a given period of time; for the greatest events it can be well modeled as a power law

$$N(> Dst) \propto (Dst)^{-\alpha}.$$  \hspace{1cm} (3-3)

Moreover, similar power laws, with the same exponent $\alpha$, seem to hold for other geomagnetic and geoelectric indices of interest. It is reasonable to suppose that this power law can be extrapolated beyond the historical record.

Studying the largest geomagnetic storms of the past $\sim 150$ years, [14] rank-ordered by several different geomagnetic indices. They conclude that
the 1859 “Carrington event” was perhaps the largest, but had several near-
competitors; it was not greatly larger than any other event. From their tables
we have estimated

\[ N(> Dst) \propto (Dst)^{-\alpha}, \quad \alpha \approx 3.3 \quad (3-4) \]

From the analysis of [41] discussed above, we observe that the largest
observed events (roughly 1-10 V/km) are approximately fit by a similar power
law

\[ N(> E) \propto (E)^{-\alpha}, \quad \alpha \approx 3.2 \quad (3-5) \]

where \( E \) is the horizontal geomagnetic field, entirely consistent with above.
This is a steeply falling distribution.

Damage to the grid from extreme events is a threshold effect. From
these distributions, we expect that improving the robustness of the grid by
increasing the damage threshold in \( E \) by a factor of 2 will decrease the rate
of damaging events by a factor of 8 or more. This is an optimistic conclusion.

This is in sharp contrast to asteroids, for which the observed distribution
in diameter \( D \) is shallower

\[ N(> D) \propto D^{-\beta}, \quad \beta \approx 2 \quad (3-6) \]

[15, 12]. Similar conclusions apply to earthquakes [2]. Moreover, impact
energy is proportional to the mass of the asteroid \( M \propto D^3 \), so that the
energy integral over expected impacts in a given time interval

\[ \text{Energy} \propto \int D_{\text{largest}}^3 D^3 \frac{dN}{dD} dD \propto D_{\text{largest}}^{3-\beta} \quad (3-7) \]

diverges at the high end if \( \beta < 3 \); here \( D_{\text{largest}} \) is the diameter of the largest
asteroid under consideration, e.g., \( D \approx 10 \) km over 100 million years. So the
aggregate impact energy is strongly dominated by the largest impacts for
the observed value \( \beta \approx 2 \). In terms of damage, 1000-year earthquakes and
asteroids have roughly 30 times the energy of 100-year events. Assuming that
damage from geomagnetic storms is proportional to $E^2$, a 1000-year storm would likely be only four times as damaging as a 100-year event.

Our conclusion is that we are unlikely to experience geomagnetic storms an order-of-magnitude more intense than those observed to date.

3.5.4 Solar cycles

Another approach to examining the likelihood of very large storms is to consider the solar cycle. Although large CME can occur throughout the solar cycle, their average frequency tracks the sunspot number (Figure 10). Demonstrating the uncertainty in CME definitions, daily occurrence varied two-fold using two different definitions, rising to 8/day with one definition and almost 4/day with the other. For reference, NOAA’s Space Weather Prediction Center (NOAA/SWPC) sees 3-5 CME/day at solar maximum (W. Murtagh, personal communication, 2011). During minima, the lower definition yielded about one CME every two days. As seen in Figure 1, only some CME hit earth and even fewer have strong magnetic fields anti-parallel to earth’s field. We did not find statistics showing these fractions and recommend that they be computed.

![Figure 10: CME and sunspot frequency during solar cycle 23. The official sunspot number is plotted in black for comparison with two estimates of CME per day [13].](image)
Although the current solar cycle, 24, began in 2009, precursors expected for Cycle 25 should have started a year earlier but have not been seen. In particular, solar activity appears to form over deep zonal jets that gradually migrate to the equator as cycles progress. Though not observed directly, these currents are inferred from observations of helioseismicity. Jets for Cycle 25 should have formed at high latitudes during 2008, but no traces of them have been observed (Figure 11). Two other precursors, magnetic field intensity in sunspot umbra and the latitudinal structure of the corona, have also failed to perform as expected to signal Cycle 25, leading to speculation that the cycle may not develop normally. At its most extreme, this could indicate that the sun will enter a prolonged minimum at the end of Cycle 24, similar to Maunder (1645-1715) or Dalton (1790-1830) minima in the sunspot record. Very weak cycles were observed during the Dalton minimum, but only a few sunspots were noted during the Maunder minimum. The rate of CME formation should also decrease if the sun enters a prolonged minimum, but the energy of individual CME will not necessarily be smaller.

3.5.5 Frequency of past solar proton events (SPEs) from nitrate in ice cores

Another avenue to assessing the frequency of extreme solar events comes from ice cores. CMEs and solar flares accelerate solar chromospheric protons and much less abundant heavier nuclei and electrons into the solar wind. When these particles hit the earth’s magnetosphere, protons with energies > 30 MeV can penetrate the magnetosphere at high latitudes and enter the atmosphere, producing nitric oxides that can be included in precipitation onto polar ice caps. Single-nitrogen oxides NO and NO\textsubscript{2} are referred to jointly as NO\textsubscript{x}. NO\textsubscript{y}, known as reactive odd nitrogen, refers to NO\textsubscript{x} plus compounds from their oxidation.
Figure 11: Zonal velocity 7,000 km below the sun’s surface, demonstrating jets that begin at high latitudes in preparation for a new solar cycle and migrate equatorward during the cycle. Jets for Cycle 25 should have formed at high latitudes three years ago, but no trace has been detected \cite{35} reporting data presented by Frank Hill of the National Solar Observatory at a meeting of the Solar Physics Division of the American Astronomical Society in Los Cruces, N.M. during June 2011.

Frozen into ice cores, NO$_x$ and NO$_y$ concentrations can be used to infer fluxes of energetic protons before these fluxes could be measured directly.

Nitrate in one ice core from Greenland and two from Antarctica was analyzed by \cite{31}. The Greenland core was 126 m long and sampled at 15 mm intervals, providing about 20 readings per year between 1561 and 1950. Known dates of volcanic eruptions were used to anchor the time reference. The records revealed seasonal and long-term changes in nitrate concentration in addition to spikes 1-2 months long. For more recent times, when solar records are available, some of these spikes corresponded closely to times of large solar storms. Until the snow has compacted for about 30 years, accurate records cannot be read. Moreover, since 1950 increasing anthropogenic nitrate deposits compete with those produced naturally, reducing the signal-to-noise of solar nitrate signatures by factors of 2 to 3. Cores from both ice caps are needed to compensate for seasonal variations in rates of precipitation.
and photoionization which increase and reduce the nitrate deposition rate, respectively. For instance, a large event during late summer may not appear in the Greenland record due to the lack of precipitation.

The Greenland core yielded 156 strongly impulsive events, each consisting of at least 27 ng/g of nitrate. Applying a linear model to estimate proton fluences yielded 70 samples with fluences exceeding $2 \times 10^9$ cm$^{-2}$ of protons with energies above 30 MeV, enough to penetrate the lower atmosphere. [30] estimates the uncertainty of conversion from nitrate to fluence as ±50%, but they argue that the conversion error will be the same for all samples and not affect the relative magnitudes of events. These fluences are plotted in Figure 12 from Table 1 in the paper, along with fluences from modern measurements of solar cosmic rays given in Table 2.

![Figure 12: Fluence of protons with energies > 30 MeV, plotted from data in [31]. The ordinate is plotted as $10^{-9}$ times fluence expressed in units of cm$^{-2}$. Black values were estimated from nitrate measured in one ice core from Greenland, red ones were obtained from solar cosmic ray observations. Nitrate-derived fluences have a threshold of $2 \times 10^9$ cm$^{-2}$.](image)

The Carrington Event stands out as the largest event, 70% larger than the next. Several large fluences correspond to identified storms, e.g. 1895,
1960, and 1972, and events are sparse during both Maunder and Dalton
sunspot minima, confirming that concentrations of nitrate spikes in ice cores
are related to solar activity. Going further, the text states that nitrate levels
are highly correlated with independent observations of large flares, geomag-
netic storms, and particularly bright aurora. No details are given, however,
to support this statement.

To the best of our understanding, these nitrate observations indeed
record past solar activity, but the record is neither a complete or a propor-
tionate indicator of CME intensity. Instead, recent work shows that intense
SEP are generated when strong CME propagate through regions of the helio-
sphere greatly disturbed by a preceding CME [21]. Thus, large nitrate spikes
are evidence of a strong CME, but other strong CME can occur without
producing proportionate nitrate signatures. For example, the March 1989
geomagnetic storm was intense while the accompanying SEP was not (W.
Murtaugh, personal communication, 2011).

The principal significance of these observations for assessing vulnerabil-
ity of the electric grid is the lack of evidence for events as strong or stronger
than Carrington. The opposite, i.e. discovery of nitrate peaks indicating
past events much more intense than the 1859 event, would be very signifi-
cant, and alarming. The lack of such signatures seems to be good news, but
it must be recognized that the statistics are inadequate for rare events. A
histogram of the fluences reveals a long tail with too few samples to form
realistic probabilities of rare events like Carrington (Figure 13). [31] state
that a broadened program of nitrate analysis could extend estimates of so-
lar activity backward many thousands of years. The increased confidence in
estimates of large events would more than justify the cost of the work.
Figure 13: Histogram of fluences plotted in Figure 12, exhibiting a long, poorly-populated tail at high amplitudes.

3.6 Summary

3.6.1 Findings

1. Solar research is progressing rapidly, and physics-based modeling is impressive, but we are not aware of rigorous comparisons of predictions with observations. The absence of three-dimensional observations within CME is particularly challenging and must be overcome before robust models can be developed, either for CME propagation to earth or their subsequent interaction with the magnetosphere.

2. Statistical predictions of CME characteristics and of GIC are promising approaches. In addition to providing useful probabilistic predictions while physics-based modeling proceeds, statistical approaches may also help identify key parameters and couplings in a manner similar to the successes of similarity turbulent scaling before high-resolution modeling was developed.
3. No evidence has been found of past magnetic storms many times more intense than the Carrington Event. Data are too sparse to be confident that such events have not happened, but available statistics and ice cores suggest that events more than 2-3 times Carrington are very unlikely. Damage likely from large magnetic storms appears significantly more modest than potential consequences of some other rare events, such as earthquakes and asteroid impacts.

4. The lack of precursors for Solar Cycle 25 suggests that the incidence and severity of magnetic storms may decline after Cycle 24 ends in the early 2020s and remain low for at least several decades.

3.6.2 Recommendations

1. As understanding and sensors improve, statistical forecasting should be pursued to determine whether predictors can be optimized and used for practical forecasts with quantifiable uncertainties, e.g., there is a 70% chance of geomagnetic Kp=7 disturbance at a particular site between 24 and 36 hours from now. This would be akin to earlier approaches to weather forecasting based on relating patterns on weather maps to earlier sequences, except that it should be easier to quantify predictions.

2. In addition to considering new satellite configurations, discussed later, adding small, low-power proton and electron sensors developed by the Air Force to all spacecraft should be considered to expand observations of the solar wind in general and CME, particularly where they interact with earth’s magnetosphere.

3. Though not unique indicators of CME, or even of SEP, the nitrate records offer invaluable information about past solar events. More rigorous statistical analysis of existing data in conjunction with independent records of solar activity should be pursued. Even more useful
would be extension of the technique backward in time if that is feasible. Continued absence of events many times more intense than Carrington would be particularly significant.
4 ELECTRICAL GRIDS AND THEIR VULNERABILITIES

4.1 Structure and Regulation

The electric grid is arguably the most important part of U.S. infrastructure threatened by severe space weather. After examining the structure, regulation, and components of the grid, a published worst-case scenario is described and compared with experiences of some high-latitude grids exposed to intense geomagnetic storms. Mitigation, monitoring and simulation are then considered before presenting findings and recommendations about the grid.

Incorporating 3,300 utilities and 1,700 non-utility power producers, the U.S. electric grid consists of 18,000 power plants ranging from nuclear, coal, and hydroelectric to solar panels and wind turbines plus their associated switches, transformers, and high-voltage transmission lines as well as distribution lines to industrial, commercial, and residential customers. It would take us too far afield to describe the tremendously varied aspects of this system, but see [16] for more information. The grid is organized into three major interconnects: Western, Eastern, and Texas (Figure 14). Ties between interconnects can permit power sharing, notably the DC link between Eastern and Western Interconnects that allows independent management of the two systems while avoiding the need to synchronize frequencies. The U.S. grid is also tightly linked with the entire Canadian system as well as the grid in northern Baja California, Mexico, widening the range of power sharing. In particular, Hydro Quebec regularly supplies power to New York and New England.

As part of The Department of Energy Organization Act in 1977, the U.S. Congress established the Federal Energy Regulatory Commission (FERC) as

an independent agency to regulate interstate transmission of natural gas, oil, and electricity. One of FERC’s missions is ‘to protect reliability of the high-voltage transmission system through mandatory reliability standards’. It does this by giving the North American Electric Reliability Corporation (NERC) legal authority to establish and enforce reliability standards for the bulk power system. NERC’s domain includes the entire Canadian grid and the grid in northern Baja California, though it necessarily operates differently in those countries than in the U.S. Standards are developed cooperatively with power companies and regional operators.
4.2 Components and Operation

The primary impact of solar activity on the electrical grid results from geomagnetically-induced current (GICs) in long transmission lines. Indeed, long conductors such as telegraph lines or pipelines can be a significant hazard to personnel and equipment. GIC caused problems for telegraph systems in the mid-19th century, and advances in the technology of electric power amplified vulnerability to solar storms by increasing the importance of distribution over long-distance transmission lines.

Household voltage in the United States is typically 110 volts (110 V) and the average power usage of an American citizen is about 1.6 kilowatt (kW), distributed among residential (38%), commercial (37%) and industrial (25%) uses. During the summer of 2009, on average the U.S. grid produced about 1 TW of power. (For this and other statistics about the grid see the Energy Information Agency web site at http://www.eia.gov/energyexplained.../index.cfm?page=electricity_home#tab2.) As was recognized early in the electrical power industry, resistive losses are much too large to transmit power at household voltages of 110 V. Rather, resistive losses, which go as $I^2R = (P/V)^2R$, are minimized by transmitting a desired power, $P$, at high voltages. To this end, high voltage alternating current (AC) is used because it is simple to transform it from one voltage to another, and transformers are, with suitable capital expenditure, 99.9% efficient.

The steam turbine in a fossil-fueled or nuclear power plant drives an electrical generator (alternator) that typically produces about 10,000 volts (10 kV) at its terminals, and an electrical transformer with a turns ratio of 50:1 is used to step that up to hundreds of kV for transmission along the lines, as shown in Figure 14. For reasons of further economy, generators are typically 3-phase, with either three terminals and a neutral ground connection (for a so-called Y connection) or three terminals without the neutral
(delta connected). In either case, at these very high transmission voltages and high powers, it is common to use three single-phase transformers either in a Y-Y or a Delta-Y connection with the generator, to provide 500 kV or 765 kV for transmission over many hundreds of kilometers.

At the consuming end of the line, there is another transformer to step down voltage for distribution throughout a city, for instance to 35 kV and 11 kV and ultimately to commercial and residential users at 110/220 V. Naturally, the generation of electrical power must closely match in time its consumption, and this is still accomplished largely by the signals provided by the power line currents, voltages, and phases. A sudden increase in consumption by flipping on a large number of toasters, for instance, will tend to slow the generator, as its rotational energy is converted to electrical energy. This signal is then used to provide more steam to the turbine, and thus to restore the generator to its desired frequency.

Because thousands of generators operate at any one time on the large interconnected networks of Figure 14, an increased load in one portion of a municipality will draw power from the interconnected lines, and could be compensated by increased power generation at any one of a number of points. Which generators are preferred at any time is a complex matter of fuel costs, power line capacity, environmental limits, and the like. It is the job of system operators, e.g. the California Independent System Operator (California ISO, http://www.caiso.com/), to oversee and optimize the use of the generation and transmission resources within the constraints existing at the time (as in economic dispatch or environmental dispatch of power).

Maintaining service and regulating voltage are the two highest priorities of grid operators. We were told that ‘Keeping the grid up is in the DNA of all operators,’ who are loathe to deny service. Needed to insure satisfactory operation of customer’s electrical equipment, regulation usually means keeping voltages within a narrow range of +6% to -13%. This requires care-
ful control of the balance between real and reactive power. Real power is resistive and dissipates energy. Reactive power results from inductive and capacitive elements on the grid which are inherently $90^\circ$ out of phase and do not consume power in loads other than by heating wires. Grid components, however, must be designed to handle total current, not just the resistive component. Static VAR (volts-amperes-reactive) compensators (SVCs) are automated impedance matching devices, adding inductive or capacitive elements as needed to keep reactive power below a set fraction of the resistive power. VAR measures the magnitude of the reactive power. Because transmission lines are inherently inductive, series capacitors are often installed in addition.

Beside normal operations of increasing and decreasing power generation and flow there are emergency conditions so extreme that equipment must be protected from damage at the cost of taking it out of service. In these situations, an extensive series of relays ultimately trip to turn off generators, open switches, etc., in order to preserve the system against overload of power or voltage. In substations there are special instrument transformers that reduce the very large voltages and currents in the power system to values that can be connected to instruments called protection relays (mechanical or solid state) set to trip when sufficiently abnormal conditions occur (usually over-voltage or over-current). When the relay trips it throws a circuit breaker and the equipment (transformers, transmission lines, generators, etc.) being protected is disconnected from the grid and thus protected from damage. Lightning strikes and short circuits from downed power lines are common faults that trip a protection relay. Typically, a protection system will try to reconnect very soon after the initial trip in the hopes that the fault was temporary. If the fault has cleared, normal operation will resume.

The grid is extraordinarily complicated and remarkably reliable under normal circumstances that include many daily incidents, such as lightning and fires. There are at least 3,550 extra-high-voltage (EHV) transformers
(230 kV and higher) on the U.S. grid [27], and they cost several million dollars apiece. They are difficult to transport, and transformers are made individually to order. Recognizing the vulnerability of the grid to transformer outage, there have been efforts in recent years to provide modular transformers or even temporary replacements that are smaller and less costly, although less efficient.

A large geomagnetic storm can endanger many transformers, leading to an extremely difficult decision as to whether to take down a substantial portion of the grid for hours or risk permanent damage that could result in substantial portions being down for months or years. It would seem that the decision would make itself, but with real people involved, and complex priorities, that decision is by no means simple. While protection is a good thing, the quality of service declines if relays are tripped when they should not be, as discussed later in this section.

Power transformers make use of the properties of engineered iron cores (actually, transformer steel) which provide an easy path for the magnetic field that is produced by the coils of the transformer when alternating voltage is applied to them. A closed path of laminated steel with layers only a few thousands of an inch thick can provide 10,000 times more magnetic flux at low magnetizing current than would be available in an air-core coil. This allows a primary coil connected to the generator and a secondary coil connected to the transmission line to be spaced from one another to withstand hundreds of kilovolts in the secondary and even higher voltages during a fault or lightning strike. A transformer, in principle, is a very simple concept. A large transformer naturally needs structure, insulation, cooling, and insulating bushings so that the simple concept of Figure 15 becomes the picture of Figure 16.

The very efficiency of the transformer makes it peculiarly vulnerable to the quasi-DC that is produced in a geomagnetic storm. This situation could
be mimicked by connecting a welding generator between the neutral lead of the transformer and ground, thus giving rise to GIC flowing through the
secondary coil of the transformer into three sets of overhead lines extending hundreds of kilometers before returning to earth ground via the neutral lead of the receiving transformers. From there, the GIC travels through the earth to the grounding mat of the originating transformer.

Each of the three phases of a high-power transmission line carrying 3,000 MW of electrical power at a voltage (rms) of 500 kV must carry about 2,000 A of 60 Hz current. It seems incredible that the secondary loop, consisting of the high-voltage winding on the sending transformer and the high-voltage winding on the receiving transformer hundreds of kilometers away, would be sensitive to an automobile battery of 12 V attached in the transformer ground leg. Yet that is the case, because these high voltage power lines use such large conductors that overall loop resistance can be as little as a few ohms. Hence a 12 V battery or other generator would produce some 4 A in this loop, which is not long resisted by the magnetic impedance (inductance) of the cores of the transformer.

To be specific, if a receiving transformer without a load attached to its secondary coil draws an excitation current of 1 A from the 500 kV line, its impedance is of magnitude 500,000 Ω, although it is purely reactive, so that the zero-load heat generation in the transformer is not $VI = 500$ kW, but on the order of 200 kW, produced by hysteresis in the core and proportional core’s mass. At 60 Hz the transformer inductance, $L$, needed to provide reactance of $X = 500,000$ Ω at $f = 60$ Hz is $L = X/2\pi f = 1,326$ H (Henry, the unit of inductance), and the two transformers at each end of the line contribute 2,652 H.

Because high-voltage transformers are built to be very efficient, even relatively small GIC can saturate their cores during one half of the power cycle, as shown schematically in Figure 17. Adding the GIC shifts the magnetic flux out of the linear range of the transformer during half of the power cycle, producing large current spikes in the secondary during that half of the cycle.
The simulation in Figure 18 shows current magnitudes increasing 100-fold during half of the cycle as a result of 5 A of GIC. These spikes are rich in harmonics that can confuse relays, causing them to trip accidentally.

Figure 17: Schematic of transformer half-cycle saturation [4]. Adding a DC offset to a normal flux (B) versus time relation (shown in blue in the upper left) shifts the relation upward (shown in red) into the strongly nonlinear region of B versus current (I), producing large current spikes during one half of the cycle, the positive half in this example.

Sudden application of a $V_{\text{GIC}} = 1,000$ V would increase the DC magnetizing current at a rate of $V_{\text{GIC}}/(2L) = 0.38$ A/s. After 30 seconds, the DC current would be about 12 A, more than enough to produce substantial half-cycle saturation of the transformer cores and allow large half-cycle AC currents to flow in the transformers, as illustrated in Figure 18. Ultimately, with an assumed 4 Ω loop resistance, and with an inductance of the saturated transformer of $L \approx 13$ H (1% that of a transformer with an unsaturated core and giving 26 H for the two transformers), the 1000 $V_{\text{GIC}}$ (if applied instantaneously) would rise at a rate of $1000/26 = 40$ A/s.
The relatively small DC current (≈ 50 A) of a large GIC is impressed on a transformer carrying much larger AC currents. (This discussion follows a simulation presented to JASON by Kappenman in 2011.)

Figure 18: Simulation of excitation currents in a 500 kV transformer under normal conditions (upper) and with half-cycle saturation produced by 5 A GIC [26]. Maximum amplitudes are ≈ ±0.58 A in the upper panel and ≈ +85 A in the lower.

The impact of transformer saturation by GICs was studied in some detail by [38], who combined simulation with some experimental verification. In Figure 19 he shows the stray magnetic flux leaking from the transformer core to the surrounding structure and the tank housing the transformer and its oil bath. Depending on the transformer type and design, different parts heat more than others. For example, in this simulated 5-limb autotransformer, the shunt shown in Figure 19 reached a temperature of 500°C after 30 minutes GIC at 100 A/phase. Because autotransformers are constructed using taps on a single winding, mitigating GIC in them is more difficult than it is for
three-phase transformers. For example, equipment connected to the portion of the winding used as the secondary ‘floats’ at the voltage of the neutral-current-blocking-capacitor.

Figure 19: Stray flux leaking from the core of a 5-limb 240-MVAR autotransformer when a GIC current of 100 Amps/phase is present along with a 300 A line current [38].

Temperatures over 110°C damage cellulose insulation commonly used in transformers. The ABB Service Handbook for Transformers states that ‘Essentially all GE Mark II designs (1965-1971), Mark I designs (1959-1964) and pre-Mark 1 designs (pre 1959) are at risk. These units have shown a tendency for shrinkage of the low density pressboard used in their construction, resulting in loose coil blocking.’ This reinforces the need for a detailed census of all large transformers, important to the transformer operator, but also to FERC and NERC in their mission to enhance greater reliability. To this end, NERC is conducting a census of these transformers, but considerable work will then be needed to assess vulnerabilities of old transformers. For instance, when briefing JASON, Craig Steigemeier estimated that just finding drawings might cost about $1,000 per transformer while each detailed engineering analysis could be about $50,000.
Also of interest here is the protection system hardware of the power grid. Although protection of personnel is important, the focus here is on hardware protection. An overview of transmission system protection is given in Figure 20, where we see protection of generators, buses, transformers and transmission lines. Within each dashed-line box are a number of circuit breakers that can be tripped to protect the hardware in the box, generators, transformers, etc. Protective relays perform the essential function of controlling these breakers to prevent or limit damage to expensive hardware. Of interest here is the impact of GICs on these relays. Protective relays might trip unnecessarily as a result of the large harmonic content of geomagnetically induced currents, although some relays might be desensitized and not trip.

Figure 20: Transmission protection system. The stop sign octagons are generators, the small squares are circuit breakers, the dark lines are either buses or transmission lines and the transformers are indicated by pairs of spirals. The dashed lines enclose breakers controlled by a protection relay system [5].

The operation of a protective relay is illustrated in Figure 21. The relay in the blue box collects information from small instrumentation transformers
to sense current (CT for current transformer) and voltage (PT for potential transformer) as well as other information, such as temperature (RTD) to drive a protection algorithm to actuate the breakers (squares) as necessary to protect the transformer. When the relay trips the circuit breakers, the transformer is isolated and protected from damage by power system currents and voltages. Protective relays operate to provide a variety of functions, e.g., over-current protection, reclosing after a fault, and under-frequency load shedding. Older relays are typically mechanical and modern ones are computer controlled, digital relays.

Figure 21: Computer relay for power system protection. The relay in the blue box senses relevant information and actuates the circuit breakers (white squares) to protect the transformer in the middle. Adapted from Schweitzer presentation to JASON in 2011).

The issue of interest here is the impact of GICs on the protection relays themselves. There are multiple modes of impact: relatively high harmonics in power lines due to transformer saturation and over-temperature and/or gas. GIC impacts on relays is of particular interest considering that the two instances known to us where GICs have caused significant power outages were both due to improper tripping of protection relays (Section 4.4).
Schweitzer (2011) notes that power line currents affected by transformer saturation in a GIC event can be up to 100% more than normal. The higher harmonic content during GIC events could be used to sense the presence of transformer saturation and prevent damage. On the other hand, in unprepared relays the harmonics might cause either improper tripping or failure to trip properly. While advanced computer relay devices, such as noted by Schweitzer (2011), can be aware of GIC-related harmonics and not be deceived by them, there are many different types and vendors of protection relays with unknown vulnerability. Further, there are reported to be some 50,000 mechanical relays still in operation in the U.S. power system [37]. Several sources refer to concerns over the impact of elevated harmonic content on protection relays. For example, ‘The large second-harmonic levels produced by GIC may restrain differential relays during fault conditions that occur during geomagnetic disturbances’ [9], and 'Multilevels of protection for individual apparatus such as generators, transformers, capacitors and AC and HVDC transmission lines all have tripped in prior storms due to relay mis-operations’ (J. Kappenman, Jason Presentation 2011).

Over-temperatures and gas accumulation in transformer casings are another means by which GICs can cause improper relay tripping. GIC related over-temperature operation, as analyzed by [38], does not occur in the core but rather in the transformer case and other metal parts (Figure 19). Such a mode of temperature increase and resultant gas accumulation may be misinterpreted by protection devices, such as Buckholtz relays, that respond to gas accumulation in a oil filled transformer. Thus, a transformer may be tripped out of service when it is unnecessary or vice versa.

In summary, improper relay tripping (too much or too little) is likely to be a significant player in the impact of GIC events on power grids. High quality computer relays can be a help if properly implemented, but older or improperly implemented relays (computer or mechanical) have been and can be a vulnerability.
4.3 Kapppenman (2010) Worst-Case Scenario for Grid Damage by Geomagnetic Storms

A report prepared for Oak Ridge National Laboratory by John Kapppenman [26], an independent consultant, examines the vulnerability of the U.S. electric grid to severe solar storms. The report has generated great interest in government and the public, largely owing to the projected scenario of catastrophic damage, including:

- More than 300 EHV transformers destroyed
- 130 million people lacking power for several years until damaged high-voltage transformers are replaced
- 1-2 trillion dollars of economic loss.

Based on estimates of magnetic disturbances of past major storms, Kapppenman’s worst-case scenario for a 100-year event assumes an accompanying ionospheric electrojet centered at 50°N. A geomagnetic storm environmental model with magnetic fields measured on the ground during solar storms was used to estimate the shape and intensity of the electrojet. Its intensity was adjusted to produce horizontal magnetic fields at ground level changing by 4800 nT per minute east of the Mississippi River and half that west of it. Assuming that these fields propagate downward as plane waves, models of ground resistivity, expressed by nine profiles in different geographic regions, were used to calculate horizontal electric fields at ground level. Finally, a model incorporating the electric grid’s topology and resistive impedances, mostly produced by transformers, lines, and substation ground connections, was used to estimate GIC in lines carrying 345 kV and higher voltages. Technical details and algorithms are not included in the report.
The appendix to [25] gives some specifics, but others are considered proprietary (J. Kappenman, email, 20 June 2011) and were not available to JASON. The model was validated by other simulations of recent storms comparing observed and measured GIC.

Figure 22: Summary of worst-case scenario disruption of the U.S. electric grid by a severe geomagnetic storm centered at 50°N (J. Kappenman presentation to JASON, 16 June 2011). Red and green circles indicate magnitudes of GIC into and out of transformers from the ground.

This scenario shows large GIC in most of the Eastern Interconnect and the Pacific Northwest (Figure 22), inducing unprecedented reactive demands that could knock out over 70% of U.S. generating capacity affecting more than 130 million people. More serious than the blackout, at least 300 EHV transformers likely would be permanently damaged or destroyed (Figure 23). Because current production rates quote delivery times of 15 months, Kappenman estimates that 4 or more years would be needed to restore normal service.
4.4 Experiences of Some Grids

There are reports that five transformers were destroyed and ten more damaged in South Africa by the October 2003 CME. We did not look into the details, but instead read multiple reports and exchanged emails with Leonard Bolduc of Hydro Quebec and Jarmo Elovaara (retired from Finn Grid) and Antti Pulkkinen (from Finland, but now at NASA Goddard) about experiences of their systems.

Figure 23: High-voltage transformers estimated to be a risk in Kappenman’s worst-case scenario (J. Kappenman presentation to JASON, 16 June 2011).
Figure 24: Hydro Quebec Grid with hydroelectric generation in the north and customers in the south, including New York and New England. Series compensation was added in 1996 to minimize chances of the grid collapsing again [6].

The world’s largest hydroelectric system, Hydro Quebec is capable of generating 36.6 GW from 60 hydroelectric and 1 nuclear plant (Figure 24). Most of the hydroelectric generators are about 900 miles north of the Canadian customers and farther from those in New York and New England (24), whom it also serves. Most of the transmission lines carry 735 kV.

In March 1989 a strong solar storm with ground-level magnetic deflections of 500-1,000 nT [10] caused half-cycle transformer saturation to generate harmonics that improperly tripped five power lines, knocking out nearly 10 GW of generating capacity and collapsing the entire grid within a minute.
This blackout was traced to improper relay tripping associated with the concurrent GIC event; seven large static VAR compensators were improperly tripped offline by relays influenced by the GIC-induced harmonics [11, 4]. Two transformers were destroyed not by the GIC but by high voltages accompanying the tripping (Ed Schweitzer, 2011 presentation to JASON). In spite of the inconvenience to millions of customers, most power was restored within 9 hours, and the cost to the utility was only $13M, half of which was for equipment, including replacing two destroyed transformers. We did not find an estimate of the economic cost to customers.

Subsequently, Hydro Quebec made changes to reduce their vulnerability, including reprogramming the malfunctioning relays. The modifications were soon tested when a strong storm on March 13, 1991 generated measured horizontal electric fields up to 1.7 V/km [6]. GIC of 110 A were measured in one of two transformers connected in parallel, indicating peak GIC of 220 A. Further improvements in 1996 added series capacitors to improve VAR regulation in major lines (These are shown in Figure 24.) Although not their purpose, these capacitors also block GIC in those lines. As a result of these changes, [6] states ‘So, we now are confident that our network will survive the anticipated worst-case GIC.’

Another case of protection relays being tripped improperly occurred in Finland in the 1990s when a new computer relay configured incorrectly was installed in a 220 V line (Jarmo Elovaara, personal communication, 2011). Finland, however, has experienced no blackouts or major component damage due to magnetic storms, in spite of measured GIC of 200 A produced by strong forcing and highly resistive soil. This record has been achieved by (1) using fully-wound 3-phase core transformers rather than cheaper auto-transformers favored in the U.S., (2) designing and testing transformers to tolerate high leakage fluxes and elevated temperatures, (3) using coils with typical reactive impedances of 500 Ω to limit earth faults, also reducing GIC as a side benefit (by virtue of the coil resistance of some 2.5 Ω added to the
loop resistance), and (4) installing series capacitors on some lines to balance reactive loads, also blocking GIC, though that is not their purpose. The Finns believe that chances of major transformer damage are small but are willing to accept one temporary grid collapse per year.

In the U.S., a transformer at the Salem nuclear plant in New Jersey was badly damaged by a magnetic storm in March 1989. The damage, however, was only discovered during an inspection a week later, and the transformer was not taken off line for several more weeks. Salem is in the PJM (Pennsylvania, New Jersey, and Maryland) Interconnection, which loses a transformer catastrophically about once a year (Frank Koza, personal communication, 2011). The only loss in PJM attributed to GIC, the Salem transformer, was found to have a design defect responsible for the failure. It, however, has become the ‘poster child’ for GIC transformer damage in the U.S.

4.5 GIC Mitigation

Simple and well-understood mitigation has been demonstrated to block GIC. To be specific, consider a large CME anti-parallel to earth’s magnetic field producing ground-level fluxes changing as fast as 1000 nT/min. When perpendicular to 500 kV transmission lines, these fluxes can induce quasi-DC voltages of 5 V/km or more, resulting in a 2500 V potential drop over a distance of 500 km. Figure 25(a) schematically shows the resulting GIC in a three-phase transmission line.

Transmission lines are inherently inductive, e.g., on the order of 0.4 H over a distance of 500 km. At 60 Hz (ω ≈ 400 rad/s), the reactance of a 0.4 H line inductance is 160 Ω. For a line carrying 500 MW per phase, the current is about 1000 A, so that the drop across the line inductance is 160 kV. Because the inductive load is in quadrature with the resistive load, it reduces the carrying capacity of the line. The long-known solution is to
Figure 25: Schematic of a three-phase transmission line terminated by grounded wye connections at the secondary sides of extra-high-voltage (EHV) transformers, adapted from [28]. A horizontal magnetic field induces GIC in an unprotected transformer (a) but not when series capacitors are added to the three lines (b) or, more simply, when a single neutral-current blocking capacitor (NCBD) is inserted in one of the neutrals (c). Each, however, must be accompanied by bypass protection against ground faults.
add a series capacitor tuned to reactance similar in magnitude but opposite in phase to the inductive load at 60 Hz. Though added for other reasons, these capacitors block GIC in the lines (Figure 25(b)).

Series capacitors, compensating the inductance of the line, increase its carrying capacity. To be specific, the line’s series inductance consumes ‘reactive power’ which is just the product of voltage and current, in this case 160 MVAR and the series capacitor would thus also need a reactance magnitude of 160 MVAR to supply the reactive power to the line inductance. Since this energy flows into and out of the capacitance every half cycle, the series capacitor stores 0.4 MJ of energy, precisely the peak energy storage in the line’s inductance. Furthermore, the series capacitor must be supported on an insulated structure because the whole thing floats at the ac voltage of 500 kV, and it is subject to impulsive voltages considerably higher than that in the guise of switching transients, lightning strokes, and the like. Its dissipation is of the order of 0.4 kW/MVAR, or 64 kW. In view of these requirements, series capacitors are large (Figure 26) and cost about $10M per substation.

To withstand 160 kV or more across them, series capacitors usually contain about 100 elementary capacitors arranged in series, requiring each capacitor to have a capacitance 100 times larger than the entire bank. All
carry the same phase current of 1,000 A. Only one of these in each phase, however, is needed to block GIC. Instead of a large bank of capacitors floating at 500 kV AC, a series-line-blocking capacitor would have 1% of that bank on a much smaller structure. This capacitance standing alone would be too large to balance the inductance of the line, but it would block GIC. It would need to withstand the same energy and voltage drop as a single member of the capacitor series, hence much less than the whole series. To protect the protection circuitry, the series-line-blocking capacitor needs over-voltage protection in the form of a paralleled varistor and spark gap as well as a closing switch that could be remotely operated to short out the capacitor in case of a persistent fault. The series-line-blocking capacitor itself does not need insulating bushings rated for EHV of 500 or 765 kV; instead it would have a metal case and a single small bushing like the cans on power distribution poles. It would be mounted on a tall insulated platform, or even on each line from a power pylon in the drop to the transformer at the substation.

For most three-phase EHV lines, the single three-phase transformer at one end (or the set of three single-phase transformers) can all be protected from GIC with a single capacitor introduced between the Earth ground and the neutral of the Y-connected transformers, as shown in Figure 25c. This neutral-current blocking device (NCBD) must withstand the same quasi-DC GIC voltage as the series-line-blocking capacitor, but it does not need to carry the full AC line current even of a single phase; it normally would be subject only to the unbalance current, typically limited to 10% of the current of a single line. The NCBD capacitor is a much more economical solution where it can be used, but its application to autotransformers is problematic and not possible in many cases.

Reference [7] described a year-long test of an NCBD by Hydro Quebec. Designed to withstand 2500 V induced by GIC, the main capacitor was specified as 2650 µF (1 Ω at 60 Hz) and a power of 384 kVAR at 620 V. This value
was selected after a compulsory test to verify that the device could withstand a 60 Hz AC voltage of $2.15U_N$ (1333 V; $U_N$ is the unit rating) for 10 s and a DC voltage of $4.3U_N$ (2666 V) also for 10 s. In contrast, a GIC series-line mitigation capacitor would be rated at 1600 kVAR at 1600 V and, by the same standards, should survive testing at AC voltage of $2.15 \times 1600$ V (3440 V) and a DC voltage of 6880 V. According to personal correspondence with Lonard Bolduc, such an NCBD capacitor box (Figure 27), with complete and approved protection could be reproduced for about $100,000.

Each approach requires a self-healing capacitor that can safely block the GIC voltage for the duration of peak currents during a magnetic storm, typically less than 10 minutes. (Made with very thin aluminum layers, self-healing capacitors do not fail catastrophically because short-circuits across the dielectric vaporize locally without significantly affecting the capacitor’s overall performance.) In this example, we have taken a GIC voltage of 2500 V DC. Even with three single-phase transformers at the powering and receiving ends of the 3-phase line, the neutral current blocking capacitor (NCBC)

---

2In addition to the design to block GIC of 420 V, Bolduc et al. indicate that the NCBD would serve also to block 465 V$_{DC}$ present at the station in unipolar operation of a parallel EHV line carrying 3, 700 A$_{DC}$
under normal circumstances and in the GIC-blocking mode would be carrying an AC current of magnitude equal to the unbalance among the three lines, typically less than 10% of the current in a single line. For 1000 A per line, the NCBC would carry less than 100 A, so that limiting its AC voltage to 1600 V for comparison, would represent a capacitor rating of $1600 \times 100 = 160$ kVAR in fact, given the low duty cycle of the GIC blocking voltage, a capacitor rated for steady AC at 1000 V rms would probably be adequate to block $2500 \ V_{DC}$. The 1.6-fold increase in capacitance would be accompanied by a corresponding 1.6-fold decrease in VAR rating and hence in the already negligible capacitor cost.

In contrast, for the Series-Line (SL) approach to GIC mitigation, each of the three SL capacitors must carry 1000 A, as well as block the assumed 2.5 kV GIC. The kVAR rating of each capacitor is thus 1600 kVAR, which together with the fact that there are three SL capacitors means an overall SL kVAR rating 30 times that of the NCBC. In the instance of the SL capacitor, increasing its capacitance by a factor 1.6 would reduce the AC voltage by that similar factor, with a corresponding factor 1.6 reduction in kVAR and hence capacitor cost. The dissipation of such capacitors at 0.4 W/kVAR thus corresponds to about 640 W for each of the three SL capacitors – about one millionth of the line’s transmitted power.

The NCBC approach would be a clear winner for transmission lines that use standard transformers, as contrasted with autotransformers at both terminals. In fact, if one end has a standard 3-phase transformer or three single-phase transformers, the NCBC should be used there. For autotransformer configurations, much of the substation may be connected to the lower-voltage portion of the transformer winding, and that portion (perhaps including the generator) would float at the potential of the NCB capacitor. Under line-fault conditions, this would be limited (in the Bolduc design) to about 1500 V for his 120 kV line (perhaps 4,000 V for the 500 kV system).
Detailed analysis would be necessary to ensure safety of personnel and equipment in this autotransformer approach.

As for the Series-Line capacitors, they could be located at either terminal, with an investment in capacitors on the order of 1% that needed for series-capacitor compensation of long-line inductance. Bolduc, however, also notes that awareness of the problem and mitigation steps are not widely appreciated, nor is action usually taken.

4.6 Monitoring and Simulation

At its most basic, monitoring begins with sensors on key components, such as transformers. For example, the ABB Transformer Handbook [3] describes in considerable detail how large transformers are monitored to ensure their design life. On p. 121 it indicates that the normal life expectancy at a continuous hottest-spot temperature of 110°C is 20.55 years. However, older transformers are not fully equipped with the sensors required for full online monitoring, and, in any case, gas evolution and water content of the insulating oil are lagging indicators of problems in the transformer. An alternative to monitoring the hottest point in the transformer (which in the presence of GIC may well not be a coil conductor but a structural member or the transformer case) would be a real time digital modeling of the transformer design, together with the measured GIC and AC voltage for that transformer.

Presently, local and regional operators monitor only their portions of the grid. Recognizing the need to monitor the entire U.S. grid, NERC and the operators are developing a system to monitor all grid facilities carrying 230 kV and/or generating at least 500 MW. Common data formats will enable comprehensive graphic displays and allow ‘drilling down’ several layers to obtain more detail. The data, however, will be available only to U.S. government employees and they cannot keep it more than seven days owing
to commercial and legal concerns of the operators. One concern is that in the
deregulated power market other operators could gain commercial advantage
by observing how competitors operate their systems. Should only one or
a few competitors be able to access the information, companies would be
liable under disclosure laws in situations similar to insider trading. Although
NERC has heard that Oak Ridge has developed a wide-area grid monitoring
program, they know little about it, indicating that some of these efforts may
be duplicative.

Sophisticated simulation software exists, as demonstrated to us by Tom
Overbye (Univ. Illinois), but we are not aware of any efforts to simulate the
entire U.S. grid, say in parallel with or as part of comprehensive monitor-
ing. Indeed, simulation will be severely constrained until the data issues just
discussed are resolved because real data will be needed to do useful simula-
tions. For instance, when large-scale or subtle grid problems arise, forensics
are likely to take weeks to months to complete, and additional time will be
needed develop and test fixes.

Outside the U.S., we learned of a DC model of the electrical grid in
Finland (personal communication from Antti A. Pulkkinen, 2011) that will
provide currents and response to a specified GIC voltage at any point in the
network. Presumably this model contains all of the resistances associated
with the ground current spreading in the poor-conductivity soil of Finland,
together with resistances of transformers, lines, etc., as affected by switchgear
and the like. Some aspects of GIC current modeling in the Finnish power grid
are discussed by [45]. The inadequacy of DC models is indicated by the one
disconnect on the Finnish grid related to GIC (Jarmo Elovaara, personal
communication, 2011). It was caused by a misconfigured digital relay and
was likely tripped by GIC harmonics. The relay has given no problems since
it was reconfigured. A DC model would allow appropriate cross-correlation
of observed GIC in transformer neutrals against the $dB/dt$ of appropriate
sensors in the vicinity of the line. Routine data from small fluctuations
would then provide data to enable real-time alerts on the basis of $dB/dt$, independent of the monitoring and reporting of neutral currents.

We have not learned of a full AC model for the Finnish grid, but do note the strong assertion that Finland has never had major damage to a transformer from GIC, despite many incidents of severe space weather. The Finns are aware of the danger of GICs and take prudent measures to prevent damage both in the grid itself, e.g., through appropriate transformer specification and testing during acquisition, and through research centered in the Finnish Meteorological Institute, Geophysical Research [36]. In fact there is a Nordic GIC Network linking research in the four Scandinavian nations (http://www.lund.irf.se/HeliosHome/nordicgicnetwork.html).

Several studies have pointed out the need for interconnect-scale simulation of the North American power grid for applications from planning, operational support and optimization, market analysis as well as vulnerability and resiliency [1, 24]. As far as we could determine, none of the existing models provides a comprehensive approach to the U.S. grid by itself, let alone its connections to Canadian and Baja California grids. In addition to representing the entire system, focused work is needed on new issues arising. For instance, intermittent power sources, e.g., solar and wind, can destabilize grids when they exceed 2-3% of total power.

4.7 Findings and Recommendations

4.7.1 Findings

1. We agree that the U.S. electric grid remains vulnerable but are not convinced that Kappenman’s worst-case scenario [26] is plausible, i.e.
that a severe solar storm will probably destroy up to 300 EHV transformers, leaving as many as 130 million people without power for years while replacement transformers are manufactured and installed. Our reservations stem from:

(a) The nature and characteristics of the EHV transformers are not well known. Consequently, there is no way to quantify how many will fail at maximum GIC.

(b) As far as we know, there is no comprehensive simulation of the entire grid with quantitative uncertainties on GIC estimates throughout the grid. Several groups show good agreement in specific cases, but general agreement over the grid is quite another matter and has not been demonstrated as far as we know.

(c) Some of Kappenman’s algorithms are proprietary and hence cannot be evaluated.

(d) Experiences of high-latitude grids such as Hydro Quebec and Finland do not suggest such catastrophic damage even for storms several times larger than any experienced to date.

2. Well-understood mitigation is available at a cost per transformer comparable to that of the engineering analysis needed to analyze the vulnerability of each transformer. Though exposed to severe geomagnetic storms owing to their high latitudes, operators of Hydro Quebec and Finnish grids are confident that mitigation undertaken will prevent catastrophic damage.

3. Identifying transformers in which GIC can flow, i.e. those lacking series protection) can remove power before damage occurs, and modest investments to monitor GIC in neutrals will allow rapid restoration of power that is cut off. Understanding GIC flows at the grid monitoring center being developed by NERC should allow safe operation of
the system at higher levels of magnetic storm intensity than otherwise possible.

4. U.S. efforts badly need coordination and central direction, as evidenced by NERC developing comprehensive grid monitoring without having access to wide-area monitoring reportedly developed at Oak Ridge.

5. Restrictions of grid data, placed by power companies for commercial and legal reasons, must be resolved before adequate simulations can be done to assess grid vulnerabilities, optimize its performance under new situations, and analyze damage that does occur during solar storms.

4.7.2 Recommendations

1. Mitigation should be undertaken as soon as possible to reduce the vulnerability of the U.S. grid. The cost appears modest compared to just the economic impact of a single storm, e.g., $8B in August 2003. Specific steps should include:

   (a) Adding real-time GIC monitors to each vulnerable transformer, with procedures to cut AC power when needed to avoid permanent damage.

   (b) Using digital relays to avoid false tripping when GIC harmonics are present and to provide data for essential tripping.

   (c) Add neutral-current-blocking-capacitors (NCBCs) with shunt protection, following the experience of Hydro Quebec. At $100k per transformer, the 1,000 most vulnerable transformers could be protected for perhaps $100M including installation.

   (d) For autotransformers on vulnerable lines, deploy small series-line capacitors with shunt protection.
2. A rigorous and fully transparent risk analysis should be done of the U.S. grid. It should begin with analyzing Kappenman’s worst-case scenario while protecting his proprietary methods and proceed to a full-up simulation along the lines envisioned by [1]. National policy should not be based on methods not fully available to the government.
5 WARNINGS, MONITORING, AND SIMULATION

Official and research-grade warnings of impending magnetic storms allow operators of electrical grids to take actions mitigating the impact of disruptions. By indicating present grid conditions, monitoring allows choosing appropriate actions. Simulations, however, are needed to understand past disruptions and prevent future ones.

5.1 Warnings

Predictions and warnings of impending magnetic storms are being issued by NOAA’s Space Weather Prediction Center (SWPC) and the Community Coordinated Modeling Center (CCMC) of NASA’s Space Weather Laboratory.

5.1.1 Space weather prediction center

Warnings of impending geomagnetic storms allow operators of electric grids to minimize disruptions by adjusting operations. Grids are most vulnerable when operating at peak capacity, with little reserve to maintain voltages during power surges accompanying shifting GIC patterns. Largely seasonal, the most dangerous times generally match seasonal needs for air conditioning, as in Texas in August, or for heating, e.g., Quebec in February. Deferring planned maintenance and repairs requires warning times of a day or two, but shorter lead times, from minutes to an hour or so, are useful for allowing some adjustments, such as rebalancing load patterns and power flows.
Based on a wide array of observations and modeling, NOAA’s Space Weather Prediction Center (SWPC) issues three types of warnings related to geomagnetic storms:

**Watches:** Originating with observation of a CME by the coronagraph on the SOHO satellite, watches are issued after a model predicts that the CME will hit earth in 1-4 days.

**Warnings:** The only direct evidence of an impending geomagnetic storm comes when a CME with a magnetic field anti-parallel to earth’s is detected at the Advanced Composition Explorer (ACE) satellite, located 1.5 million kilometers from earth on the earth-sun line. Depending on the speed of the CME, warnings based on magnetic fields at ACE provide lead times of 15-60 minutes.

**Alerts:** Based on deflections of horizontal magnetic field strength, usually at the Boulder magnetometer, Alerts indicate that a magnetic storm is taking place.

Corresponding to Kp = 5 to 9, Alerts are categorized as G1 to G5, with 5 being the most severe, roughly corresponding to a Category 5 hurricane. SWPC issues a G5 Alert when the Boulder magnetometer has a negative horizontal deflection exceeding -500 nT. During severe storms, Warnings and Alerts often come in ascending sequences, such as those shown in Figure 28 for the second half of October 2003. Begun with a series of Watches, the increasing sequence of Warnings and Alerts culminated in three G5 Alerts marking events known as the Halloween Storm.
Figure 28: Summary of geomagnetic alerts, watches, and warnings issued by NOAA’s Space Weather Prediction Center during the second half of October 2003. Elevated activity culminated in the Halloween Storms at the end of the month.

Official SWPC warnings go the major interconnects comprising the U.S. electric grid which then distribute them to their regional operators and individual companies. In addition, data and warnings are passed to Air Force space weather as well as many organizations in other countries.

5.1.2 Community coordinated modeling center

NASA’s Space Weather Laboratory has the mission of minimizing damage to astronauts and satellites by predicting space weather conditions throughout the heliosphere. Timely warnings of impending events allow protective measures, such as shutting down electronics, delaying space walks, or having astronauts take shelter in shielded spaces. Partly funded by the National Science Foundation (NSF), the Community Coordinated Modeling Center (CCMC) of the Space Weather Laboratory does extensive research and modeling to improve the predictions.
As part of a pilot program termed Solar Shield [32], CCMC worked with the Electric Power Research Institute (EPRI) to predict GIC at defined places in grids where these currents were measured. These research-grade forecasts were made in two levels:

**Level One:** Initialized with observations from SOHO and STEREO satellites, magnetohydrodynamic models forecast CME velocity and magnetic intensity which in turn predicted the electric field strength impacting the magnetosphere, as in equation 3-1. Models of the magnetosphere and ionosphere then estimated GIC using resistivity profiles determined from earlier observations. Posted on EPRI’s Sunburst server, these predictions gave a range for the probable GIC magnitude and its start and end times. The example in Figure 29 predicted GIC of 11 to 82 A for ten hours during the Halloween Storm.

**Level Two:** As CME effects are detected on ACE, the data initialize a three-dimensional model of the magnetosphere that estimates GIC time series. Figure 29 compares the prediction with recorded GIC. In this section of the data, peak currents were -20 A, within the predicted range.

### 5.2 Monitoring and Simulation

Power companies, regional operators, and some interconnects monitor flows of power and voltages on their grids, and the effort is being extended to the entire U.S. grid. Simulation, however, is presently much more restricted.
Figure 29: EPRI/NASA space weather warning during the 2003 Halloween solar storms. The Level One warning, in text boxes on the left, estimates a magnitude range of the GIC and its start and end times. The Level Two warning, on the right, compares predicted (red) and observed (black) GIC at a particular grid location.

5.2.1 Monitoring

Monitoring begins with sensors on individual components. For instance, modern transformer designs provide real-time digital monitoring, particularly of known or suspected hot spots to ensure that transformers meet their design lives. Page 121 of the ABB Transformer Handbook \cite{ABBTransformerHandbook} notes that the normal life expectancy at a continuous hottest-spot temperature of 110°C is 20.55 years, before proceeding to point out that older transformers are not fully equipped with the sensors required for full online-monitoring. Because overheating has usually resulted from over-voltages accompanying faulty relay tripping, digital relays detect GIC by observing persistent elevated harmonics on all three phases. Tripping can be restricted to transmission lines lacking series capacitors, and trip levels can be adjusted in response to NOAA space weather warning levels (Ed Schweitzer 2011 presentation to JASON). Data from these relays are available for monitoring grids and diagnosing problems.

Some signals from equipment monitors are sent to utility monitoring
centers which in turn transmit data to regional centers. NERC and the regional utilities are developing monitoring for the entire U.S. grid, defined as facilities operating at 230 kV and above with generation of at least 500 MW. Data from power companies will have a common format, facilitating graphical displays and making it relatively easy to inspect more detailed data when desired. Due to commercial and legal concerns, power companies will allow only U.S. government personnel to access the data and then only for seven days. Companies are concerned that data showing how they operate their systems would give competitors an edge in selling power. Legal concerns are similar to those about insider trading, i.e. giving privileged information to one or a few competitors could lead to federal prosecution. NERC has heard that Oak Ridge National Laboratory has developed a wide-area grid monitoring program but knows little about it (E. Rollison, personal communication, 2011).

5.2.2 Simulation

Grid simulations are needed to probe responses to magnetic storms and other disruptions as well as to test proposed changes in grid configuration. We learned of a DC model of the electrical grid in Finland that will provide currents and response to a specified GIC voltage at any point in the network (A. Pulkkinen, personal communication, 2011). Presumably this model contains all resistances associated with GIC spreading in the poor-conductivity soil of Finland, together with those of transformers, lines, etc. Some aspects of GIC current modeling in the Finnish power grid are discussed by [45]. A DC model would allow cross-correlation of observed GIC in transformer neutrals against $dB/dt$ measured near the line. Routine data from small fluctuations would then enable real-time alerts based on $dB/dt$, independent of neutral current monitoring and reporting. The limitation of DC models is indicated by the one disconnect on the Finnish grid related to GIC (Jarmo
Elovaara, personal communication, 2011). It was caused by a misconfigured digital relay and was likely tripped by GIC harmonics. The relay has given no problems since it was reconfigured.

Several studies have pointed out the need for interconnect-scale simulation of the North American power grid for planning, operational support and optimization, market analysis, and vulnerability and resiliency [1, 24]. Presently there are good modeling projects, but none provides a comprehensive approach to the U.S. grid by itself, let alone its connections to Canadian and Baja California grids. In addition to representing the entire system, focused work is needed on new issues arising. For instance, intermittent power sources, e.g., solar and wind, can destabilize grids when they exceed 2-3% of total power, and studies are needed to understand how these effects can be minimized. Data, however, are essential for realistic simulations. Consequently, limitations being placed on data to be supplied to the NERC grid monitoring must be solved to allow retention of these data for future studies. For instance, how could a grid collapse during a magnetic storm be diagnosed and prevented in the future without full data during the collapse as well as during normal operation?

5.3 Summary

5.3.1 Findings

1. The U.S. has two excellent groups predicting unclassified aspects of space weather, NOAA’s Space Weather Prediction Center and the Coordinated Community Modeling Center (CCMC) of NASA’s Space Laboratory. As we understand the situation, SWPC is charged with issuing official warnings of impending space weather disturbances to
U.S. civilians, while under the Solar Shield test program CCMC extended the approach on a research basis to predict GIC at several sites. In addition, the Air Force issues warnings, some classified, to its assets based on inputs from SWPC.

2. The Solar Shield test program proved successful, but grid operators are confused when research-grade forecasts and official predictions do not coincide.

3. SWPC and the Air Force space weather center at Peterson AFB in Colorado Springs were formerly more tightly linked than at present, and both feel that they would benefit from being better connected.

4. NASA’s modeling of ground-based effects is presently limited to north of 50°N, i.e. to Canada. The limitation is not conceptual or technical but simply a lack of funding and priorities needed to do the work.

5. Due to much higher funding levels, CCMC develops most of the models likely to be valuable for predicting space weather and its terrestrial impacts. Although the ENLIL model, exhibited in Figure 1, is being transitioned from NASA to NOAA, transitions are not considered part of normal operations at these centers, limiting the speed of improvements in official warnings.

6. NERC is developing monitoring for the entire U.S. grid, which is essential for understanding how the complex system interacts, but NERC is not privy to wide-area monitoring already developed at Oak Ridge National Laboratory. Nor will NERC be able to retain data supplied to by power companies, greatly limiting the ability to understand future grid failures.
5.3.2 Recommendations

1. NOAA and NASA should resolve space weather issues falling between their responsibilities. These issues center on treating transitions from research predictions by NASA to official warnings by NOAA as part of normal operations and on coordinating predictions so the user community clearly distinguished between research forecasts and official warnings. This should include deciding whether GIC predictions made by NASA should eventually be part of official warnings or remain advisory special products and extending CCMC’s modeling domain to include the entire U.S.

2. Air Force and NOAA should increase links between their space weather operations. SWPC lacks a backup in case it goes down; the Air Force space weather center could easily serve that purpose. Lacking the long technical backgrounds in space weather common at SWPC, Air Force personnel would benefit from closer technical support in training their users how to interpret and use space weather products. Also, Air Force has some sensors and platforms that could aide civilian efforts, provided that classified aspects are protected. Similar efforts have been successful in other aspects of environmental measurements made by the intelligence community and should be possible here.

3. NERC should consider extending its monitoring effort to include the entire North American grid in their domain, i.e. including Canada and northern Baja California, because this system is operated jointly and shares problems and opportunities. To make full use of the effort, NERC should resolve the stringent limitation on data retention now envisioned, even if this requires seeking new legislation from Congress.
6 CRITICAL OBSERVATIONS

Presently, critical space weather observations applicable to electric grids consist of magnetic field strengths observed on Earth and CME characteristics detected in space.

6.1 Ground-Based Magnetometers

In addition to triggering SWPC Alerts, data from ground-based magnetometers are used with nearby GIC observations to estimate profiles of ground resistivity. Resistivity can change over horizontal scales of kilometers and smaller, but because GIC typically depend on average resistivity over horizontal scales of 100 km, that is the rough resolution needed along transmission lines (A. Pulkkinen, personal communication, 2011).

The Geological Survey operates official magnetometers in the United States. Boulder (BOU) is one of their six stations in the continental U.S. (Figure 30). Fredericksburg, MD (FRD) is often the backup when Boulder is down. Relative to the needs for estimating GIC in the power grid, the U.S.G.S. network is very sparse in the lower-48 states.

In contrast to the six USGS magnetometers in the continental U.S., the Sino Magnetic Array at Low Latitudes (SMALL) consists of 24 installations to understand better how solar magnetic storms couple to ground at low latitudes, particularly when influenced by the equatorial electrojet (Figure 31). Begun in 1999, SMALL is a cooperative program between the University of California at Los Angeles (UCLA) and China’s Seismological Bureau. UCLA and NASA funded development of the magnetometers, and NSF supported their manufacture at UCLA. The goal was to make each magnetometer and its data system for $6,000 to $15,000.
Continental coverage much closer to SMALL than to the U.S.G.S. array is needed for U.S. space weather studies and predictions. Using inexpensive research magnetometers appears to be an affordable approach that should be considered if the pilot program being tried by NASA/CCMC and a utility is successful. Without this component, improvements in understanding and predicting CME impacts on the magnetosphere will not lead to commensurate increases in accuracy of GIC predictions which are the measures needed by power companies when deciding how to respond to warnings.

6.2 Critical Space Systems for Forecasting

Space systems are critical for the forecasting of Space Weather [32]. Coupled with models of the heliosphere and the Earth’s magnetosphere, data from these space systems are the primary means for prediction of the probabilistic impact of energetic solar events on the Earth. Here we concentrate
Figure 31: The Sino Magnetic Array at Low Latitudes (SMALL) begun in 1999 as a joint project between China and UCLA with NSF funding to study magnetic storms at low latitudes.

on Coronal Mass Ejection (CME) events [20] because of the importance of these events for operation of the electrical power grid. Key space weather parameters that need to be measured routinely by operational space systems include:

- **Location and size of solar flares** (if not determined by Earth based observatories). These are provided by white light, X-ray, and UV imagers.

- **3-D Spatial extent and velocity of CME events**. These are provided by coronagraphs at two or more separated locations.

- **3-D magnetic field components associated with CME or high-speed streams, upstream of the Earth**. These are provided by in-situ satellite measurements.

These measurements can be made from a minimum constellation of
satellites at Lagrangian Points, a set of saddle points in the combined gravitational fields of Sun and Earth that move with Earth around the Sun. As shown in Figure 32, there are five Lagrangian Points, L1 through L5. Directly in line with the Sun, L1 is $1.5 \times 10^6$ km from earth. At $60^\circ$ ahead of and behind the Earth in its orbit, L4 and L5 are well-located for obtaining three-dimensional observations of CME.

![Figure 32: Geopotential potential showing the five Lagrange Points produced by balances between gravitational attractions of the sun and the earth and centrifugal forces. Phase-locked to earth as it revolves around the sun, the Lagrange Points are saddles in geopotential with red and blue arrows showing attraction to and away from the points. From ‘Lagrangian Point’ in Wikipedia at http://en.wikipedia.org/wiki/Lagrangian_point.](image)

Presently, the most important space weather data comes from four research satellites:

**SOHO:** A joint project of NASA and the European Space Agency, the Solar and Heliospheric Observatory carries three coronagraphs that collectively image the corona from 1.1 to 32 solar radii. Launched in 1995 and operational at L1 the next year, SOHO is long past its 2-year intended mission. The mission is now extended to December 2012.
ACE: Launched into a halo orbit around L1 (Figure 33) in 1997, the Advanced Composition Explorer carries a magnetometer that provides the only direct evidence that a CME is about to hit earth as well as the orientation of the CME’s magnetic field. Although eight years past its design life, has enough propellant to operate until 2024.

STEREO A and B: Two Solar and TErrestrial RElations Observatory satellites were launched in 2006. One, A, is in an orbit slightly inside earth’s, giving it a period of 346 days, and B is slightly outside for a 388-day period. Due to their orbits, A moves ahead of Earth at 21.7°/year as B falls behind at 22.0°/year. Presently, they are slightly past being 90° ahead and behind. Obtaining stereoscopic views of the Sun is their principal mission, which is done with cameras imaging the solar disk and inner and outer coronas. Fourteen years past their design life, the STEREO satellites will be on the far side of the Sun, out of direct contact with Earth, for several months in 2015. In 2023, they will converge near Earth.

Figure 33: Advanced Composition Explorer (ACE) launched in 1997 into a halo orbit about L1.

In addition to needing spacecraft at three positions, near L1, L4, and L5, we also emphasize that all three being used now for operational forecast-
ing were developed by NASA primarily as scientific research satellites, rather
than as operational space weather satellites. These spacecraft are aging and
no replacements are currently under development. Given the large poten-
tial economic impact of severe solar events and the importance of accurate
forecasting in mitigating impacts, we recommend that an operational sys-
tem for space weather forecasting be developed, possibly with international
cooperation.

6.3 A Minimum Constellation of Spacecraft for CME
Space Weather Forecasting

Given the preceding discussion the following constellation of spacecraft is
needed:

- A spacecraft at L1 or other orbit between the Earth and the Sun with a
  3-axis magnetometer to provide warnings of CME with high magnetic
  field gradients in directions likely to cause significant disturbance of the
  Earth’s magnetosphere. Additional instruments should be minimized
to hold down costs. Possible orbits are discussed below.

- Two coronagraphs off the Sun-Earth line to provide three-dimensional
  information warnings of CME size, direction, and velocity. The two
  coronagraphs need to be in orbits that allow reconstruction of the
  Earth-directed CMEs. Possible orbits are discussed below, with L4
  and L5 being among the candidate orbits.

6.4 A Near-Term Replacement for the ACE Spacecraft

The highest near-term priority for space weather prediction is measurement
of magnetic fields at L1 or other suitable location, a capability currently pro-
vided by the aging ACE spacecraft. Although a spacecraft with the required
capability has been built and is available, it has not been funded for launch. This is the Deep Space Climate Observatory (DSCOVR) spacecraft, which has undergone extensive testing and is currently at the NASA Goddard Space Flight Center (see figure).

The spacecraft began as the Triana Earth-observation mission, but NOAA hopes to refurbish the spacecraft bus and payload for use as a space weather satellite, including the Plasma-Mag instrument which would make magnetometer measurements at L1, thereby providing a backup/replacement for ACE. A possible coronagraph instrument is also being considered. Funding was requested in the Administration’s FY12 budget for the refurbishment. This appears to be an excellent opportunity, allowing a replacement for ACE to be developed relatively quickly and at low cost.

6.5 Orbits for Solar Observations

Several types of orbits are particularly useful for making observations important to space weather forecasting, in particular those providing stereoscopic views to characterize CME in three dimensions, and those on the Sun-Earth line, but closer to the Sun to verify that CMEs are on track to hit Earth. We described several of these now.

- **Lagrange Points.** L1 is of particular interest for verifying that CME are heading for Earth, providing a warning time of about 25 minutes for a CME moving at 1000 km/s. Sixty degrees ahead of and behind the Earth, L4 and L5 are good candidates for stereoscopic views, but they are not the only possibilities. Spacecraft can be positioned anywhere along the Earth orbit with a fixed angle relative to the Earth-Sun axis by using trajectories having a low total ∆v requirement.

- **“Artificial Lagrange Orbits” using Station Keeping.** It is possible in principle to position a spacecraft on the Earth-Sun axis, closer
to the Sun than the L1 point, using thrust to keep the spacecraft positioned (“station-keeping”). Possibilities include ion thrusters and solar sails. We evaluate these possibilities below and conclude that they are not attractive as near-term options for orbits for an ACE replacement.

- **“Quasi-satellite” Orbits.** An attractive possibility, both from an operational space weather perspective and from a research perspective, is a constellation of small spacecraft in so-called quasi-satellite orbits. For the Earth, quasi-satellite orbits have the same orbital period and phase as the Earth around the Sun, but a larger eccentricity. Consequently, in the coordinate frame centered on the Earth, a spacecraft in a quasi-satellite orbit appears to orbit the Earth even though it is actually in heliocentric orbit around the Sun. The orbits are called quasi-satellite orbits because many large bodies, including the Earth, have small “companion” objects in quasi-spacecraft orbits around them (e.g. the asteroid Cruithne “orbiting” the Earth).

Next we discuss non-standard orbits that might have the advantage of earlier warning than a spacecraft at L1.

### 6.6 Artificial Lagrange Orbits via Station Keeping.

Eventually, the ACE spacecraft will need to be replaced if accurate predictions of the impact of CMEs on the Earth are to continue, particularly for critical magnetic field measurements. One possible orbit we investigated for an ACE spacecraft replacement is a ‘artificial Lagrange (L1) ’ orbit, i.e. an orbit which keeps the spacecraft on the Sun-Earth line, but further from the Earth than L1, accomplished by station keeping.

A spacecraft in interplanetary space can detect, *in situ*, CME on their way to Earth, and can measure their magnetic field strength and direction
directly. The ACE spacecraft, stationed at the L1 Lagrange point between the Earth and Sun (a point of unstable equilibrium at which station-keeping requires only small thrust) about 0.010 AU from the Earth has been collecting these data for many years. Unfortunately, it is far past its design life, and no replacement is scheduled. At its distance it provides 25 minutes or longer warning for a typical CME.

Could a future spacecraft be placed closer to the Sun while maintaining its station along the Earth-Sun line? Unfortunately, this is difficult. Station-keeping with the Earth’s orbital period at distance \( R \) from the Sun may easily be shown, in the approximation of circular orbits, to require a radial thrust

\[
F = \frac{GM_{\odot}m}{R_0^2} \left[ \left( \frac{R_0}{R} \right)^2 - \frac{R}{R_0} \right],
\]

(6-8)

where \( M_{\odot} \) is the Solar mass, \( m \) is the mass of the spacecraft, \( R_0 \equiv 1 \) AU and \( R \) is the radius of the spacecraft’s orbit; away from L1 the force of the Earth’s gravity may be neglected. (The Earth’s orbit is not precisely circular, having an eccentricity of 0.0167 which implies that the spacecraft will lead or lag the Earth by as much as 1.44° in azimuth unless this is corrected for by also putting the spacecraft in an orbit with the same eccentricity).

Two possibilities for station keeping are ion propulsion and solar sails.

6.6.1 Ion propulsion for station keeping?

Ion propulsion is a proven technology for interplanetary spacecraft, used for example on the NASA DAWN mission currently in orbit around the asteroid Vesta.
Could ion thrusters be used for station keeping at a location significantly further from the Earth than L1? Providing thrust requires the expenditure of propellant mass. Ideally, the combined mass of spacecraft and propellant decays exponentially with an $e$-folding time

$$
\tau = \frac{I_{sp}g R_0^2}{GM_\odot} \left[ \left( \frac{R_0}{R} \right)^2 - \frac{R}{R_0} \right]^{-1},
$$

where $I_{sp}$ is the specific impulse of the propellant and $g$ is the Earth’s acceleration of gravity (its average exhaust speed opposite to the thrust direction is $I_{sp}g$; this is the definition of $I_{sp}$).

For $R = R_0/\sqrt{2}$ the factor in brackets is 1.293. The Solar acceleration at Earth’s orbit is 0.581 cm/s$^2$ so the required incremental acceleration is 0.751 cm/s$^2$. For a 100 kg spacecraft this requires a force 0.751 N. For (impulsive) firings of solid fuel with specific impulse 300 s (mean exhaust velocity along the thrust direction of 3 km/s) the mean rate of fuel expenditure is 0.25 gm/s. Independent of the satellite’s mass, the implied exponential decay time of its mass is $4 \times 10^5$ s, about five days. The asymptotic value ($R \to R_0$) of the factor in parentheses is $3(R_0 - R)/R_0$. For $R = 0.9R_0$ the exponential decay time is 20 days.

An alternative to solid fuel rockets might be ion thrusters. Solar electric propulsion ions thrusters developed for interplanetary space probes (NASA-NSTAR program) have 0.1 Newton thrust capability, require 2.3 kW of power, and have a specific impulse of 3300 seconds (Wikipedia, Ion Thruster http://en.wikipedia.org/wiki/Solar-electric-propulsion). Multiple thrusters can be mounted on a single spacecraft, as on DAWN.

Use of a ion engine with larger specific impulse increases the decay time in proportion to the specific impulse. For a specific impulse of 3000 s and $R = 0.95R_0$ decay times of about one year are feasible. Unfortunately, this is still too short to be useful; one would like a satellite lifetime of a decade or more.
We also note that ion engines with high specific impulse are energetically inefficient. At $R = 0.95R_0$, the area of photovoltaic cells required to power the engine is about $4.2 \text{ m}^2/\text{kg}$.

We conclude that ion propulsion is not an attractive approach for positioning a satellite significantly closer to the Sun than L1 on the Earth-Sun axis.

### 6.6.2 Solar sails

Station keeping with solar sails is technically feasible and using sails for station-keeping at an artificial Lagrange point might be an ideal application of solar sailing.

That is, for an insolation of 1 kW/m$^2$ at 1 AU, the energy density is $(1 \text{ kJ/s/m}^2)/c$ (where $c$ is the speed of light) implying a radiation pressure of $\sim 3 \times 10^{-6} \text{ N/m}^2$, and the total radial force on an 80% reflecting solar sail perpendicular to the radius is $\sim 6 \times 10^{-6} \text{ N/m}^2$. For a foil of mass density $0.0075 \text{ kg/m}^2$ (corresponding to 0.3 mil-thick aluminized polyester foil), the acceleration is $\sim 8 \times 10^{-4} \text{ m/s}^2$. We note that sails with 0.3 mil-thick foils have been deployed on the Japanese IKAROS spacecraft.

The acceleration needed for station-keeping twice as far from Earth as L1 (0.98 AU) is $\sim 3.5 \times 10^{-4} \text{ m/s}^2$ compared to the solar acceleration of $8 \times 10^{-4} \text{ m/s}^2$ for 0.3 mil foil. Therefore, a spacecraft propelled by a 0.3 mil-thick solar sail at 0.98 AU need have less than half its mass in solar sail. Thinner/lighter sail material has been proposed, which would allow positioning even closer to the Sun. For instance, a spacecraft five times further from Earth than L1 needs an acceleration of $\sim 9.4 \times 10^{-4} \text{ m/s}^2$ for station keeping, thus requiring a thinner foil (e.g. $\leq 0.2 \text{ mil thick}$).
Solar sails are technically feasible, their state of technical readiness is a major issue. Currently solar sailing spacecraft are still considered technology demonstration missions. Therefore, solar sails are unfortunately not appropriate for an operational space weather satellite in the near term. The first successful demonstration in space occurred in 2010 with the Japanese IKAROS mission which used a 200 m$^2$ solar sail. However, the acceleration demonstrated by IKAROS was only $10^{-3}$ Newton on a 310 kg spacecraft, or an acceleration of about $3 \times 10^{-6}$ m/s$^2$, about 1% of what would be needed for station-keeping at 0.98 AU, and only 0.3% of that required for 0.95 AU.

### 6.7 Quasi-satellite Orbits

An attractive option for positioning spacecraft closer to the Sun than L1, but still relatively close to the Earth-Sun line, is to establish a constellation of small-spacecraft in quasi-satellite orbits (also called “distant retrograde orbits (DRO)”). Here we will consider quasi-satellite orbits that have the same period as the Earth (1 yr), but are in orbits with significantly larger eccentricity, e.g. $e_{sat} \sim 0.1$ compared to the Earth ($e_{earth} \sim 0.017$). Provided that the phase of the orbit is properly chosen, the spacecraft will appear to “orbit” the Earth when seen in a frame that is rotating about the Sun with a 1-year period. That is, for part of its orbit, the spacecraft will be closer to the Sun than the Earth and for the other part of its orbit, the spacecraft will be further from the Sun than the Earth (Figure 34). Thus even though the quasi-satellite is in heliocentric orbit (with perturbations from Earth gravity) it will appear to orbit the Earth. The quasi-satellite orbits would exist even if the Earth had only a small mass, but they would not be stable over eons.

Now consider a constellation of such spacecraft all in orbits with similar eccentricity, but different orientation and phase. If the number, orientation, and the phase of the orbits of the spacecraft are properly chosen, then the
orbits can be arranged so that at least one spacecraft is always significantly closer to the Sun than the Earth. Such a constellation could provide space weather warnings of particle densities and magnetic fields an order of magnitude earlier than ACE at L1. In addition, the constellation would have considerable scientific research value by sampling CMEs and other interplanetary phenomena at multiple points in space on spatial scales not previously monitored.

The velocity increment required to place a satellite in a solar orbit of eccentricity, $e$ (quasi-orbit with perihelion $e\text{AU}$) can be estimated in several ways. The final orbit must have a period of one year and therefore a semi-major axis of 1 AU. It is straightforward to show that such a Keplerian orbit requires a velocity at 1 AU transverse to the Earth’s orbit i.e. a radial velocity, $v_{rad}$ of:

$$v_{rad} \sim \sqrt{\frac{GM_{\text{sol}}}{\text{AU}}} \ e \sim 30 \text{ km/sec} \times e \quad \text{[for small } e\text{]} \quad (6-10)$$
For example, $e = 0.1$ requires a radial velocity of $3 \text{ km/sec}$. If this radial velocity is acquired as the result of an impulsive thrust, e.g. by transfer from low-earth orbit, then taking into account that the spacecraft must also achieve escape velocity, by energy conservation, $V_{\text{thrust}}^2 - V_{\text{esc}}^2 = v_{\text{rad}}^2$. Because $V_{\text{esc}} \sim 11 \text{ km/s}$ is about a factor of 3.6 larger than $v_{\text{rad}} \sim 3 \text{ km/sec}$, the additional velocity increment needed in low-earth orbit to achieve the required $v_{\text{rad}}$ is less than 10% of that needed to achieve escape. In this scenario, an additional maneuver is needed to adjust the angular momentum and phase of the orbit, but these can be fairly small.

A specific constellation of spacecraft in quasi-satellite orbits was proposed in 2000 by [44], who called the concept the “Space Weather Diamond”. They estimated the required “characteristic energy (C3)”, corresponding to the excess orbital velocity over the escape velocity, to be $\sim 5 \text{ km}^2/\text{sec}^2 = (1/2)v_{\text{excess}}^2$. This result for $v_{\text{excess}}$ is very close to the above estimate of $3 \text{ km/sec}$ for the required radial velocity. In addition, the St. Cyr et al. quasi-orbit required a small additional injection maneuver having $\Delta v_{\text{inj}} \sim 350 \text{ m/sec}$. The $\Delta v$ requirements for placing spacecraft in quasi-satellite orbits 0.1 AU from the Earth seem therefore not severe. The constellation proposed by St. Cyr et al. consists of four spacecraft with orbits such that the distance from Earth ranges from $\pm 0.1$ AU along the Earth-Sun line to 0.2 AU east and west of Earth. For this constellation, the closest spacecraft to the Earth-Sun line can be as much as $\sim 11^\circ$ off axis.

Our own preference would be for a constellation of “small-sats” with payloads consisting only of a 3-axis magnetometer and particle sensors. Such a package could be very light weight and have low power consumption, possibly allowing launch of multiple spacecraft with a single rocket. For a distance of 0.1 AU from the Earth, it appears that a minimum of 4 spacecraft is needed so that at least one is always significantly closer to the Sun than L1 and not too far off the Earth-Sun line; more than 4 spacecraft would be even better. The required number is in part dictated by the coherence length of magnetic
fields over the CME. To provide magnetic field measurements that allow useful GIC predictions, the magnetic field measured at the spacecraft must be representative of the magnetic field of the part of the CME that impacts the Earth. CMEs are believed to have coherent magnetic fields on scales of 0.1 AU.

We also remark that there is no absolute requirement for the orbits of the spacecraft to be strictly in the ecliptic plane. There may be scientific value in orbits which go above and below the ecliptic. However, this requires additional $\Delta v$ and may therefore be undesirable from the perspective of propellant weight.

6.8 Summary

Observations, monitoring, and simulation are essential to predicting how space weather will impact electric grids and developing mitigation strategies.

6.8.1 Findings

1. Observations from four research satellites - SOHO, ACE, STEREO A, and STEREO B - have demonstrated their value to space weather predictions. Now dependent on these satellites, these predictions are in jeopardy because all four of the satellites are years past their design lives and no replacements are ready or even funded for development. Loss of ACE would be the most damaging, as it provides the only confirmation that a CME is going to hit Earth and triggers actions by power companies to mitigate its impact.
2. Three operational satellites are needed for long-term solar monitoring, one in line with the Sun and the others off this line at positions optimized for stereoscopic viewing of CME.

3. The DSCOVR satellite being stored at NASA Goddard offers the quickest and cheapest means of backing up or replacing ACE.

4. After initial replacement of ACE with DSCOVR, a constellation of simple satellites in quasi-satellite orbits would provide robust monitoring in line with the Sun and could be positioned to increase warnings times to as much as ten hours by increasing distances from earth to 0.1 A.U.

5. The present number of magnetometers in the U.S. is inadequate for space weather predictions. A larger array is needed to map spatial and temporal structures of electrojets and to infer ground resistivity affecting GIC in transmission lines.

6. Organizational weaknesses in the space weather enterprise tend to lie in gaps between agencies. For instance, Air Force space weather would benefit from closer coupling with the greater technical experience at SWPC, and SWPC would be aided by access to some Air Force assets and having a backup capability in Colorado Springs. Overlaps between official and research-grade space weather predictions by NOAA and NASA need to be worked out, as does development of a regular path for selecting successful research developments and transitioning them to official predictions. In another area, to aid their monitoring program, NERC should have access to DOE work on the topic.
6.8.2 Recommendations

1. Owing to the importance of direct confirmation that a CME will hit Earth, DSCOVR should be readied and launched as soon as possible to backup ACE.

2. Based on recent successes in predicting space weather using research satellites, the U.S. should commit to a long-term set of operational satellites for monitoring space weather. A minimum constellation could be 3-4 platforms in quasi-satellite orbits, as in the SWx Diamond proposal of [44] and two more at L4 and L5, or in other appropriate orbits.

3. NERC should consider extending monitoring being developed for the U.S. electric grid to the entire North American interconnected system. To obtain maximum impact, this project should find a way to retain data supplied by power companies to use in post-event analysis and simulation.

4. Consideration should be given to coupling a simulation capability to full-grid monitoring, to test hypotheses explaining grid problems and well as developing strategies for new developments, such as significant reliance on intermittent power sources.

5. Because present coordination between agencies dealing with space weather is far less than optimum, the entire effort should have someone in charge with authority to resolve inter-agency problems.
6. A project should be developed to optimize magnetometers available for space weather observations. One aspect should involve an objective analysis to determine locations where sensors are needed to infer the electrojet structure in the ionosphere. Another issue is determining how finely to map the resistivity structure of the ground and how much of this can be done using temporary installations, as only one good data set is needed per station.
References


[28] J. Kappenman, Low-frequency protection concepts for the electric power grid: Geomagnetically induced current (gic) and E3 HEMP mitigation.


