

The risk of geomagnetic storms to the grid

A preliminary review

David Roodman

GiveWell

June 29, 2015

Acknowledgements: Thanks to Dan Baker, Emanuel Bernabeu, Ed Cliver, Trevor Gaunt, Ramsis Girgis, Nat Gopalswamy, David Jacobson, John Kappenman, Stephen Parrott, and Antti Pulkkinen for their time and wisdom, and their patience to engage in conversations. Thanks to Gaunt, Baker, Kappenman, and Sébastien Guillon for comments on an earlier draft. All views expressed are attributable only to me. Uncredited photos are hyperlinked to their Wikimedia Commons source pages.

Contents

Introduction and overview.....	1
Background.....	5
Electromagnetism.....	5
Solar activity and geomagnetic storms.....	10
Geomagnetic storms and power grids.....	11
What is the probability per unit time of a storm at least as extreme as the Carrington event?	17
Carrington comparisons.....	17
The July 2012 near-miss.....	18
Kappenman's factor of 10.....	19
Extrapolating statistically from the historical record.....	23
Probability densities and cumulative probability densities.....	23
Extreme value theory.....	26
Applying EVT to geomagnetic storms.....	32
CME speeds.....	32
The D_{st} index.....	36
Published studies of the historical record.....	38
Tsubouchi and Omura (2007), "Long-term occurrence probabilities of intense geomagnetic storm events," <i>Space Weather</i>	38
Ruzmaikin, Feynman, and Stoev (2011), Distribution and clustering of fast coronal mass ejections, <i>Journal of Geophysical Research</i>	39
Thomson, Dawson, and Reay (2011), "Quantifying Extreme Behavior in Geomagnetic Activity," <i>Space Weather</i>	39
Riley (2012), "On the probability of occurrence of extreme space weather events," <i>Space Weather</i>	40
Love (2012), "Credible occurrence probabilities for extreme geophysical events: Earthquakes, volcanic eruptions, magnetic storms," <i>Geophysical Research Letters</i>	41
Kataoka (2013), "Probability of occurrence of extreme magnetic storms," <i>Space Weather</i>	41
Summary: What is the probability per unit time of a storm at least as extreme as the Carrington event?	42
Are magnetic disturbance extremes localized?.....	42
How vulnerable are high-voltage transformers to geomagnetically induced currents?.....	48
Conclusion.....	51
Sources.....	52

Introduction and overview

Life on earth evolved to exploit the flow of energy from the sun—and to withstand its extremes, from ultraviolet radiation to bombardment by magnetically charged plasma clouds. As the name of NASA's "Living with a Star" mission aptly suggests, the sun is a source of both sustenance and danger.

But if life on earth writ large has adapted to its home star, perhaps civilization has not. Perhaps modern societies are unprepared for what the sun can be expected to deliver even on the fleeting time scale of human history. In particular, the concern motivating this document is that a cataclysm on the sun could trigger a "geomagnetic storm" that would knock out so many satellites and high-voltage transformers that advanced societies would lose electricity for months or years while waiting for replacements. Loss of power that long could compromise hospitals, water treatment plants, pipelines, and food transport, creating an economic and humanitarian disaster (NRC 2008, pp. 11–12).

Does the risk of a "perfect geomagnetic storm" deserve more attention than it is receiving? My initial assessment is that it almost certainly does, for the attention has been minimal relative to the stakes. I am not at this point convinced that the *probabilities* are as high as some have suggested. (For example, Riley's (2012) oft-cited 12%/decade probability estimate for an extreme storm looks like an unrepresentative extrapolation from the historical record.) But the present inquiry is layered in uncertainty. Scientific understanding of the sun's behavior is limited. Likewise for the response of power systems to geomagnetic storms. *My* understanding of the state of knowledge is itself limited. Significant "tail risk"—of events extreme enough to cause great suffering—should not be ruled out.

A distinctive feature of the geomagnetic storm issue is the sequential, probabilistic nature of the phenomenon of concern. A preliminary assessment of the risk, as performed here, has to touch on each step in the sequence. Cataclysmic explosions with the power of a billion hydrogen bombs occur on face of the sun. Each event may throw off some amount of magnetically charged plasma, producing a coronal mass ejection (CME). In the abstract, a CME has some probability of hitting the earth, which depends on its angular breadth. If it hits, it will do so at some speed, perhaps as high as 1% of the speed of light, meaning 3,000 kilometers per second. The CME's magnetic field may point substantially in the same direction as the earth's, producing a magnetic collision (Gopalswamy 2006, p. 248) rather like slamming together two magnetized toy trains the way they don't want to go. Sometimes several CMEs come over a few days, the first one clearing a path through interstellar matter that speeds the transit of its successors. Each magnetic blast will, over hours or days, bend the earth's magnetic field and accelerate electrical currents that flow at great heights above the planet, including the electrojets that cause the Aurora Borealis and Aurora Australis. The gusts of "solar weather" will also strew turbulence in the earth's magnetic field, like a strong wind over water (Kappenman 2005, p. 6), producing even sharper, if more localized, magnetic oscillations.

According to the laws of electromagnetism, when the magnetic field fluctuates in a given spot, it induces a voltage there. The faster the magnetic change, the greater the voltage. Before the Industrial Revolution, electrical pressures induced by magnetic storms along the surface of the earth could only be relieved by the flow of electric charge through air, sea, or land. But now people have laced the planet with less resistive conduits: long-distance power lines. Especially when crossing terrain whose (igneous) mineralogy resists electrical current or when terminating near conductive seawater, and especially when the wires happen to align with the induced electrical force, these cables can become geomagnetic lightning rods.

Like lightning rods, high-voltage power lines are grounded: for safety, they are connected to the earth at either end. But at each end of most of these power lines, interposed between them and the earth, are transformers,

garage-sized or bigger. They put the “high-voltage” in “high-voltage power line.” In preparation for long-distance transmission from a generating source, the transformers step the voltage up—to as high as 765,000 volts in the US. At the receiving end, transformers symmetrically step the voltage back down for distribution to factories, offices, and homes. (Boosting the voltage for long-distance transmission cuts energy losses from the electrical resistance of the power lines.)

Transformers exploit the symmetry of electromagnetism: just as a changing magnetic field induces a voltage, so does the movement of electrical charge (electricity) produce a magnetic field. Inside each transformer, two wires, one connected to the input line and one to the output, coil hundreds of times within or around a shared core of magnetically permeable material such as silicon steel. The normal input is alternating current (AC), like that in an ordinary home, its voltage flipping from positive to negative and back 50 or 60 times a second. The oscillating electricity in the wire produces an oscillating magnetic field in the transformer’s core. That in turn induces an oscillating current in the output wire, typically at a different voltage. The capacity of AC to be transformed in this way for long-distance transmission is precisely why at the dawn of the electrical age AC beat out DC (constant, “direct” current) as the standard for power systems.

Under design conditions, a transformer’s core is magnetically capacious enough to carry the entirety of the field produced by the input wire. But if too large a current enters, the core will *saturate*. Magnetic force fields will stray out of the core and into the surrounding wires, where they can exact invisible mayhem: random currents in both the input and output wires and “hotspots” of burnt insulation. Possibly, the transformer will fail immediately. Or it may continue operating while the hot spots cool into something analogous to dots of rust: they escape attention at first, but initiate degradation that spreads over weeks or months. Eventually a failure may be triggered, which engineers may not even recognize as storm damage (Albertson et al. 1973, p. 475; Gaunt and Coetzee 2007, p. 444).

Geomagnetic storms can send such damaging currents into transformers in two ways. The storms can directly induce them, as just described. Or the storms can disrupt currents, voltages, and frequencies in an operating grid enough to overwhelm the equipment meant to counteract such distortions, and thus trigger sudden shutdowns of power plants or disconnections between sections of the grid. These automatic responses are designed to protect the grid, and may largely do so—but perhaps not completely in extreme cases. In Québec during the great storm of March 1989, the sudden disconnection of the La Grande hydroelectric dam complex from the rest of the grid caused an “overvoltage” that damaged two big transformers, part of a larger cascade of events that led to a widespread blackout (NERC 1990, p. 42). A wildcard that has emerged since 1989—but which is beyond the scope of this report—is that a storm might damage GPS and communications satellites, which utilities have increasingly used to coordinate components of the grid. (Giant generators spinning at 50 or 60 times per second, hundreds of miles apart, must be precisely synchronized if serving the same grid.)

In the worst case, some analysts believe, a geomagnetic storm would take out hundreds of high-voltage transformers across a continent-scale area. High-voltage transformers are large, expensive, custom industrial products. There are not a lot of spares around. New ones would take months each to manufacture and deliver since under normal circumstances, it takes 5–12 months to produce and deliver a large transformer in the US, and 6–16 if it is imported (USITC 2011, p. II-7); and limited global production capacity could produce a backlog of years. The blackout would be measured in months. The failures would cascade to all corners of industrial societies because of the interdependence of systems—power, pipelines, sewage treatment, police, air traffic control, hospitals. The scariest potential consequence is the loss of cooling at storage facilities for spent nuclear fuel, as at Fukushima in 2011 (Foundation for Resilient Societies 2011).

Offsetting such risks is the paradoxical resilience built into grids, as seen in Québec. If a geomagnetic storm sufficiently distorts the current entering or exiting a major transformer, safety equipment trips, shutting it down. Large areas may be blacked out within seconds. But the system may become immune to more permanent damage. Short-term fragility bestows long-term resilience. In Québec, power was largely restored after nine hours (NERC 1990, p. 42), and life went on.

In addition, the power system is arguably more prepared for electrical storm surges today. Satellite-based warning systems are more sophisticated (“GoreSat” was launched on February 11 to strengthen capacity to monitor solar activity); utility officials are wiser to the danger and so are perhaps more ready to preemptively shut down grids to protect them; and some systems have been modified to make them more robust (NERC 2010, p. 63). That is not to counsel complacency, but to highlight the complexity of this issue.

In past reviews for the Open Philanthropy Project, I have laced my conclusions with caveats about how well researchers have been able to answer various empirical questions—above all because of the difficulty of determining cause and effect in social systems. This time, I must offer similar warnings but for different reasons. One reason is a sort of good news: some of the limits to collective knowledge on this issue arise as much from lack of study as from any deep barriers to human understanding. The impact of nonstandard currents on large transformers, for example, could be much better researched. There may lie an opportunity for philanthropy, perhaps. But just as important, because of the limitations of my expertise, the limit on time, the trans-disciplinary complexity of the topic, and the sharp disagreements among experts, I am less confident that I have reached the frontier of knowledge.

So I offer the following assessment with tentativeness, as a snapshot of Open Phil’s current understanding. Key points:

- Solar activity, as measured by the number of sunspots, follows an 11-year cycle, with the number of sunspots rising and then falling. (The first sunspots of the current cycle, defining its start, appeared January 2008 (NASA 2008).) Coronal mass ejections capable of causing storms are more common in the high-sunspot-number phase but the correlation is not absolute. Fast CMEs occur in the declining part of the cycle too. There are also dynamics on longer time scales, which are not well understood. Some cycles are more active than others. Trends appear in century-scale data. At present, we have little basis for forecasting the evolution of storm frequency, beyond the observation that a major one occurs about once a decade.
- Geomagnetic storms are not unusual. Major ones occurred in 1859, 1872, 1909, 1921, 1960, 1972, 1982, 1989, and 2003, among other years (Kappenman 2006; Silverman 2006).
- In a storm of any given extent, higher-latitude regions feel greater magnetic distortions—notably Scandinavia, Canada, and the northern US.
- Studies extrapolating from historical data to estimate the per-decade probability of giant storms like the ones that hit in 1859 (the “Carrington event”) have tended to err on the high side. In particular, the 12%/decade figure cited by the *Washington Post* (Washington Post 2014; Riley 2012), appears based on a model that, roughly speaking, fits a straight line to the curved tail of the storm distribution. My own estimates suggest a risk of 0.33%/decade, with a 95% confidence interval of 0.0–4.0%.
- Yet the past in this case—the historical record—is short. We should not attain confidence by extrapolating from this limited record.
- Some geomagnetic storms have taken out high-voltage transformers (Gaunt and Coetzee 2007; Moodley and Gaunt 2012; NERC 1990). But none has done enough damage to warrant substantial economic or humanitarian concern.

- Three questions seem central to the analysis of the threat posed by extreme geomagnetic storms to transformers:
 - How much stronger the worst-case—say, a 100-year storm—than the storms that have hit since modern high-voltage grids were built.
 - How widespread would be the extremes in minute-to-minute, second-to-second, magnetic field change. If the extremes are confined to a few hot spots, as is plausible under a chaotic model, then the risk of long-term and widespread blackouts may be low. If a few transformers go, compensation will not be too hard. If one region loses power, its neighbors can deliver many kinds of aid, from spare transformers to a share of their power.
 - How widespread would transformer failures would be, during the storm or in the weeks and months after. In principle, the vulnerability of a given transformer depends on many factors: its design (there are many types); its construction; its age; how quickly safety equipment trips to shut it down (Gaunt and Coetzee 2007; Girgis and Vedante 2012). In practice, these dependencies are not well understood, in the sense of being empirically verified under realistic conditions. In assessing these empirical issues, it is worth distinguishing between two sources of stress: an extreme current that quickly disables a transformer; and a smaller one that does not trigger a protective shutdown, and yet starts a process of decay leading to failure.
- My best estimate at this writing is that the probability of catastrophe is well under 1%/decade, but is nevertheless uncertain enough, given the immense stakes, to warrant more serious attention. In particular:
 - Most measures suggest that what appears to have been the largest storm since the industrial revolution, the 1859 Carrington event, was less than two times as strong as recent storms, which civilization has shrugged off. In a review of storm strength indicators, Cliver and Svalgaard (2004) put the Carrington event near the top of the list of great storms of the last 150 years on every dimension of strength for which data are available—but never in a class by itself. Cliver and Dietrich (2013) describe the Carrington event as 50–100% larger than more-recent storms.
 - It is hard to imagine how a doubling in storm intensity could make the difference between a handful and hundreds of transformers destroyed. It is not *impossible* to imagine though: Perhaps the perfect storm, with the most damaging combination of speed, magnetic strength and orientation, and tight sequencing of several CMEs, has yet to occur. Perhaps civilization has become more vulnerable because of rising dependence on vulnerable satellites. Perhaps there are engineering thresholds, which, once crossed, lead to exponentially more damage.
 - Recent tree ring analysis has revealed jumps in the atmospheric concentration of radioactive carbon in the years 774–775 and 992–993 (Miyake et al. 2012; Miyake, Masuda, and Nakamura 2013). This evidence *could* point to solar flares 10 times brighter than any seen in more recent centuries (Cliver et al. 2014, p. 3). But whether it does and whether geomagnetic disruption would have been proportionally large are at the moment points of dispute and uncertainty (Miyake et al. 2012; Usokin et al. 2013; Cliver et al. 2014; Neuhäuser and Hambaryan 2014).
 - Thanks to automatic shutdowns, the high-voltage transformer fleet may not be prone to immediate, permanent, and widespread damage during a storm (Girgis and Vedante 2012). However, emerging evidence suggests that transformers suffer more than commonly realized, perhaps especially from currents not quite large enough to trip safeties. Gaunt and Coetzee (2007) document slow-motion degradation in eight transformers in the low-latitude nation of South Africa beginning right after the Halloween storms of 2003. These permanently disabled transformers months, not moments, later. SAC (2013, p. 3-2) provides intriguing graphical statistical evidence that geomagnetic disturbance is *the* major cause of transformer failure in the

US, its role obscured by time delays. On the other hand, the delaying and spreading of failures over time may buffer society against the risk of exhaustion of spares.

And yet we should not be complacent about the threat. The true probability of something even more severe than Carrington is unknown. And the effects of storms weak and strong on transformers is poorly understood, at least in public-domain science. (Perhaps the military and industry actors know more than they share.) The long manufacturing times make a nation's high-voltage transformer fleet an Achilles Heel if enough damage occurs at once. And in many countries, electric industry regulation is heavily influenced by utilities and equipment manufacturers, who out of professional pride and institutional interests may resist efforts to adequately assess and address the risk.

Background

Electromagnetism

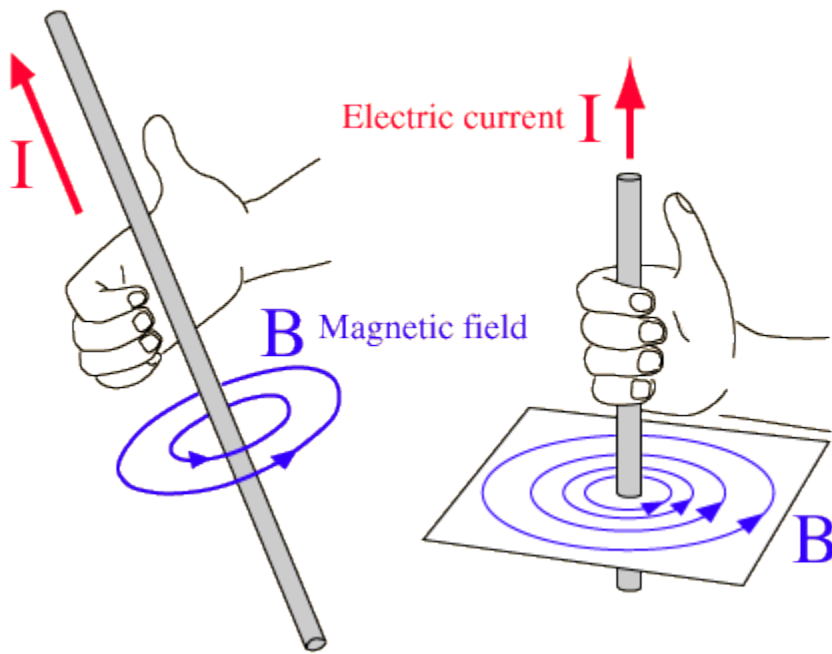
Current is the aggregate movement of charged particles. It is measured in amount of charge per unit time; a standard unit is the ampere, or “amp.”

By convention, electrons have negative charge even as, by convention, the direction of current is the direction of movement of positive charge. Thus electric current is thought of as moving in the direction opposite that of the actual electrons involved. This is a matter of semantics, not physics.

The end of a magnetic object that is drawn to the north is naturally called its north pole. Since with magnets, as with electric charges, opposites attract, the earth's northerly magnetic pole is actually a *south* pole under the usual labelling convention of physics. And with magnets too, there is a sign convention: northerly magnetic force is positive.

Ampère's law observes that electric current produces a magnetic field that is geometrically dual to the path of movement.¹ That is, if you stick out the thumb of your right hand, curl the fingers as if you were trying to hitch a ride, and place the line of your thumb parallel to a wire so that the thumb points in the direction of current, then your fingers will follow the induced magnetic field that encircles the wire all along its length:

¹An astute reader will note that since movement is relative, so is magnetism. How large a magnetic field an observer perceives depends on her velocity relative to the moving charge.



If the sign convention for electricity or magnetism were flipped, then you'd need to use your left hand.

This means that the needle of a compass placed next to a wire from a battery will deflect when the switch to the wire is turned on.

Rather symmetrically, just as movement of electrical charge creates a magnetic field, a *changing* magnetic field induces an electrical field that is geometrically dual to it. And one way for the local magnetic field in some place to change is for the object generating the field to move closer or farther. This is Faraday's law. It is governed by a *left-hand rule*. That is, if we stand at the north end of a giant bar magnet and looking toward its south end, and imagine the bar's magnetism suddenly rising, the process of change induces what we perceive as a clockwise electric field around the bar. If a wire were coiled around the bar, electricity inside it would move clockwise, producing a current. But since the electrical field is induced by a *changing* magnetic field, when the magnetic field stopped strengthening, the current would stop.

The more sudden the change in the local magnetic field, the greater the electrical force created, however momentarily, and the larger the currents induced in any conducting media in the field.

The strength of an electrical field at any given point is measured in volts per meter. How much current a voltage induces at a point depends on the voltage, the electrical conductivity of the medium, and the length of conductor subjected to the field.

To recap, movement of charged particles relative to some point induces a magnetic field there. Movement of a magnet toward some point induces an electric field there. The duality is the heart of Maxwell's equations, which are the unifying mathematical description of electromagnetism.

One consequence of the duality is a negative feedback loop called *reactance*. When you flip on a light switch, current starts to run through the wire. This causes an encircling magnetic field to materialize all along the wire. As just asserted, the sudden *change* in magnetic field strength momentarily induces voltages all along the wire that work out, if you use the right and left hand rules, to oppose the direction of the original current. This *reactance* momentarily delays the current from reaching its full strength. But the current and its magnetic field quickly stabilize, and the opposing voltage disappears since it only arises from changes in the magnetic field.

In other physical contexts the negative feedback is strong enough to cause permanent oscillation—producing electromagnetic waves, including light.

The modern world is built on devices that exploit the duality between electricity and magnetism—electric motors and generators, transformers, radio transmitters and receivers, etc.

Most grid-based electric power is alternating current (AC). Its strength and direction is always changing, cycling 50 or 60 times a second. AC is naturally produced by any generator that rotates, notably the steam turbines in nuclear and fossil fuel plants and the water turbines in dams. Non-alternating or direct current (DC) is naturally produced by non-mechanical processes such as the chemical reactions in batteries and the photoelectric effect in solar cells.²

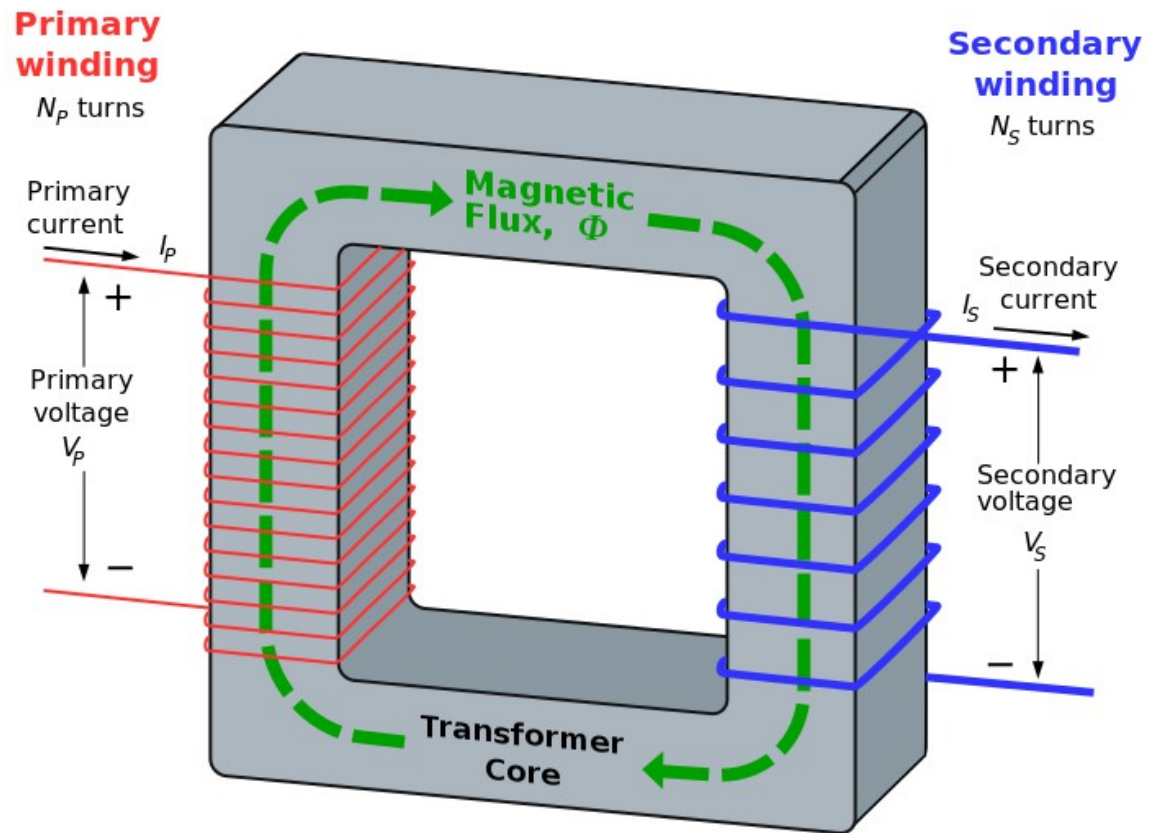
In fact, reactance has quite different consequences for AC than DC. Since AC is constantly changing, reactive current is too. The effect is not only to delay the achievement of equilibrium as in the simple light switch example above, but to permanently shift the waveform within its 50 or 60 Hertz cycle. Much of the design and operation of electric grids is shaped by the need to control this effect in order to synchronize alternating currents from various sources and keep the AC rhythm perfectly stable. These days, one source of precise timing information is GPS-type satellite networks.

Several kinds of electrical components exploit the ability of electrical and magnetic fields to influence each other. These typically contain large *coils* of wire. Why? If a single strand of wire creates a weak magnetic field in its vicinity, 1000 strands packed together produce one 1000 times stronger. By the same token, if the magnetic field in the component suddenly strengthens, it induces an electric field in every winding near it, so the more wire subject to the electric field, the greater the total force created in that wire.

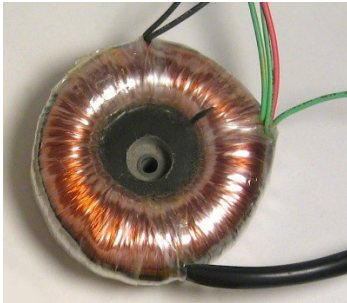
A *transformer* is made by coiling two wires around or within the same metallic core, typically a different number of times. The ends of one wire, the primary winding in the diagram below, might connect to a power source such as a dam or wind farm. The ends of the other might tie to a long-distance transmission line linking to a distant city. If the primary wire carries AC then the windings induce a constantly oscillating magnetic field, ideally confined to the magnetically permeable core. The alternating magnetic field in the core in turn creates an oscillating voltage along each winding of the secondary wire.³ Since the total voltage created along the secondary wire depends on how many times it is wound around the core, the output from the input can differ in voltage. Electrical energy is *transformed* from one voltage to another. This does not violate the law of conservation of energy; it is rather like using the energy from two balls falling one meter to lift one ball two meters.

²A minority of high-voltage transmission lines carry DC, which further reduces losses to electrical resistance. Since the transformers to which they are tied are designed for sustained currents in one direction, they are less vulnerable to geomagnetically induced currents.

³In fact, many high-voltage transformers are *shell form*. These reverse the placement of electrical and magnetic conductors in the schematic diagram above. The magnetic material wraps around the electrical wires. They operate on the same principles.



Some transformers are as small as coins:



...while others are much bigger:



Utilities use *step-up* transformers to raise the voltage of generated power for long distance transmission. The higher the voltage, the less current is moved in order to transmit a given amount of energy and the lower the energy losses from electrical resistance in the long-distance power lines. Step-down transformers at the receiving end reduce the voltage back to a level appropriate for distribution to households and businesses. Today, power lines in the US run as high as 765 kilovolts. Some lines in China are built for 1 megavolt.

A major challenge in engineering transformers is that the materials they use, such as copper and iron, are not perfectly conductive of electricity nor infinitely permeable to magnetism. To a degree, they resist, and in the process generate heat. Power lines resist and make heat too, but their high-surface-to-volume ratios let them dissipate it easily, so that the added heat from a geomagnetically induced current (GIC) will not do lasting damage.

In contrast, transformers coil huge lengths of wire into small volumes, making heat harder to dissipate (NERC 2012, p. 25). Similarly, and most crucially, temporarily increased magnetic forces may saturate the ability of the transformer's core to carry the magnetic field.⁴ This will push the field out of the core and into the surrounding

⁴Technically as the regular alternating current oscillates, the geomagnetically induced direct current will increase the total current to potentially dangerous levels during half the regular current's cycle and reduce it

coils, where its constant oscillations will distort the currents on both sides of the transformer, potentially creating damaging hot spots.

Solar activity and geomagnetic storms

Sun spots are relatively cool spots that occasionally appear on the surface of the sun. They are magnetic phenomena. Through an appropriately strong light filter, they look like black dots.

Solar flares are cataclysms on the surface of the sun that cause sudden bursts of radiation, including visible light.

Coronal mass ejections (CMEs) are what they sound like—expulsions into space of coronal matter. The ejections vary in speed, mass, breadth, and orientation and strength of embedded magnetic field. The fastest CME observed by the SOHO satellite since it began monitoring in 1996 left the sun at 3000–3500 kilometers per second, about 1% of the speed of light ([data](#), [video](#)). CMEs do not go in all directions at once: angular widths are typically 45–60 degrees (Riley et al. 2006, pp. 648, 652), making for a one-in-eight to one-in-six chance of earth impact. CMEs are now understood to be the most energetic solar phenomena (Gopalswamy 2006, p. 252). CMEs are an extreme form of solar wind, which is an ongoing flow of particles away from the sun in all directions.

Solar particle events (SPEs) cause large numbers of electrically charged particles, notably protons, bombard the earth.

Sun spots, solar flares, CMEs, and SPEs are distinct but related. For example, solar flares can cause SPEs. CMEs can generate them too by accelerating the interstellar matter they plow through, like a motorboat sending a shockwave before it.

Since the early 19th century, it has been understood that the frequency of sun spots rises and falls in a cycle of about 11 years (Lakhina et al. 2005, p. 3). The sun is said to oscillate between *solar minima* and *solar maxima*, and the strength of solar activity is often still indexed by the sunspot number:

during the other half (NERC 2012, p. 25). Think of raising the graph of a sine wave so its oscillations no longer center on zero.

Sunspot number, 1610–1995



After the existence of CMEs was confirmed in the early 1970s, it became apparent that they too occur more during solar maxima—several times per earth day on average, as opposed to once every two days during minima (Gopalswamy 2006, p. 246).

CMEs often launch within hours or days of solar flares, which is why the more easily observed solar flares were long thought to be the cause of geomagnetic storms.

It is now understood that CMEs are the primary cause of the most intense geomagnetic storms, which are transient disruptions of the earth's magnetic field (Gosling 1993).

Solar activity exhibits dynamics at cadences longer than the 11-year solar cycle, which are poorly understood. During the Maunder Minimum, between 1645 and 1715, few sunspots were observed. The second half of the 20th century, the main baseline for projections of future activity, was more active in sunspot terms than any 50-year period since 1750. On the other hand, the sun has gone unusually quiet in last few years, at least in sunspot terms. The solar minimum between the previous and current sunspot cycles, running approximately 2005–10, was the quietest and longest of the space age (Lockwood et al. 2011, p. 1). And the solar maximum now being experienced looks to be the lowest since 1906 (NASA 2014).

The relationship between sunspot activity and CMEs is not well understood. Despite the recent sunspot quietude, in July 2012 the sun threw off one of the fastest CMEs in the modern record (Baker et al. 2013; it missed the earth). And as shown above, the sunspot peak associated with the Carrington event of 1859 was low. As a result, physicists do not have a good model of solar dynamics with which to predict future activity. That uncertainty invites the use of statistical methods to extrapolate from the past, discussed below.

Geomagnetic storms and power grids

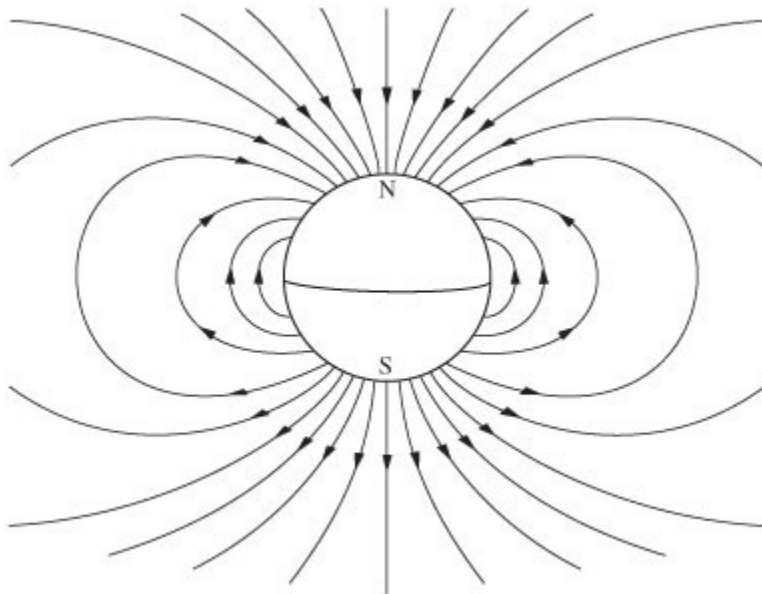
The physics of the arrival of CMEs at earth are perhaps better understood than the physics of their origin in the sun.

The earth's geomagnetic poles flip and shift over time. Today, the northern geomagnetic pole deviates from the northern spin pole by about 10 degrees, roughly toward New York (wdc.kugi.kyoto-u.ac.jp/poles/polesexp.html). So the US-Canada border has about the same geomagnetic latitude as Stockholm, despite being farther south in the conventional sense. This matters because it is the earth's magnetic field that guides the incoming "missiles" of CMEs.

Broadly, the impacts of CMEs are intuitive. Magnetically "charged" matter hurtles towards the earth. If the matter is magnetically oriented the same way as the earth, the greater the magnetic collision, since like magnetic poles repel. By the same token, the more opposed a CME's magnetic field is to the earth's, the less disruptive its arrival (Gopalswamy 2006, p. 248).

At higher resolution, the effects ricochet in complex ways via the dualism of electromagnetism. The sudden arrival of magnetically charged material affects the speed and direction of electric currents above the earth—*electrojets*—which in turn affect magnetic fields at the surface, which in turn induce electrical currents there too. One electrojet, the *ring current*, encircles the earth 10,000–20,000 miles above the equator, running east to west. By the right hand rule, its intensification during a storm creates a stronger southerly magnetic field beneath it. Since this opposes the earth's magnetic field, the effect is a net reduction in the measured field along the equator. This is why some measures of storm strength are in negative nanotesla, the tesla being a unit of magnetic field strength. This predominantly equatorial effect of geomagnetic storms receives less attention in the literature I read. (But see Ngwira et al. 2013a.)

Then there are the Birkeland currents, which are best known for causing the Auroras Borealis and Australis as they intersect the upper atmosphere. Where the ring current orbits the earth, the Birkeland flow to and from the earth, spiraling along magnetic field lines. Disproportionally often, those field lines will arrive at the earth near the geomagnetic poles. To see why, consider this diagram of the earth's magnetic field:

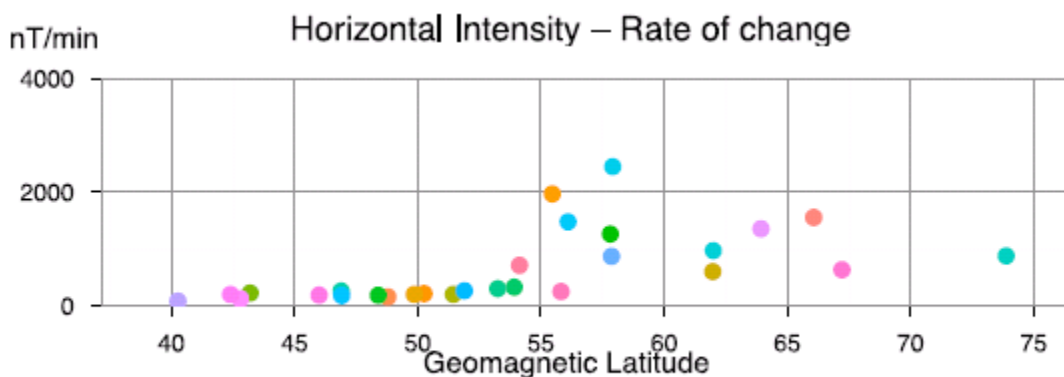


If you put your finger on a point far from the earth—say, at least one earth diameter away—and then figure out what field line you are on and trace it toward the earth, the odds are you will end up near a pole. That is where most of the far-reaching field lines puncture the earth's surface. But unless you start due north or south of the earth, you won't end up *at* a pole. This is why the auroras and geomagnetic storms are strongest at high latitudes but taper toward the poles. (See Pulkkinen et al. 2012, pp. 5–10, on geomagnetic storms.)

Of course, the stronger the storm, the stronger the effect felt at any given geomagnetic latitude. This is why in the biggest storms, the Birkeland currents reach farthest toward the equators. It is why in 1859 auroras were visible within 23 degrees of the magnetic equator: San Salvador in the northern hemisphere and Santiago in the southern (Cliver and Svalgaard 2004, p. 417).

Geomagnetic storms last hours or sometimes days. Some of the biggest are triggered by a succession of CMEs, as in 1859 and 2003. The first CME can accentuate the impact of its successors by clearing the transit path of interstellar dust and saturating certain regions of the earth's magnetosphere with particles. But the proximate cause of harm to grids is dynamics on the scale of minutes, since the faster the magnetic change, the larger the induced voltage. To understand the potential for such rapid changes, we can draw on the concept of turbulence. Kappenman (2005, p. 6) refers to Kelvin-Helmholtz shearing, which is a model for what happens at the boundary between two fluids moving at different velocities. One can imagine that a high wind over a perfectly flat sea would make no waves. But such a state turns out to be an unstable equilibrium, like a pin balanced on its point. The slightest deviation from balance is self-reinforcing. If a few molecules of sea water happen to rise above the rest, the wind catches them, creating ripples that raise other molecules. Bigger waves give the wind more purchase, and turbulence develops.⁵ At any given moment, some molecules are moving much faster than the wind. CMEs are apparently capable of inducing analogous turbulence in the earth's magnetic field. These chaotic magnetic shudders are what can most easily damage electronics on earth.

The graph below provides evidence on where magnetic volatility is most common. It shows the magnitude of the biggest one-minute change in the horizontal magnetic field ever recorded at each of 28 selected magnetic observatories across Europe (Thomson, Dawson, and Reay 2011, fig. 6). The observatories began operating at different times, mostly between 1980 and 2000, so not all captured the big 1989 storm. Despite being only partially comparable, the observations suggest that geomagnetic disturbances are largely confined to territory above 55° geomagnetic latitude, which includes Canada and most of Europe and the United States.

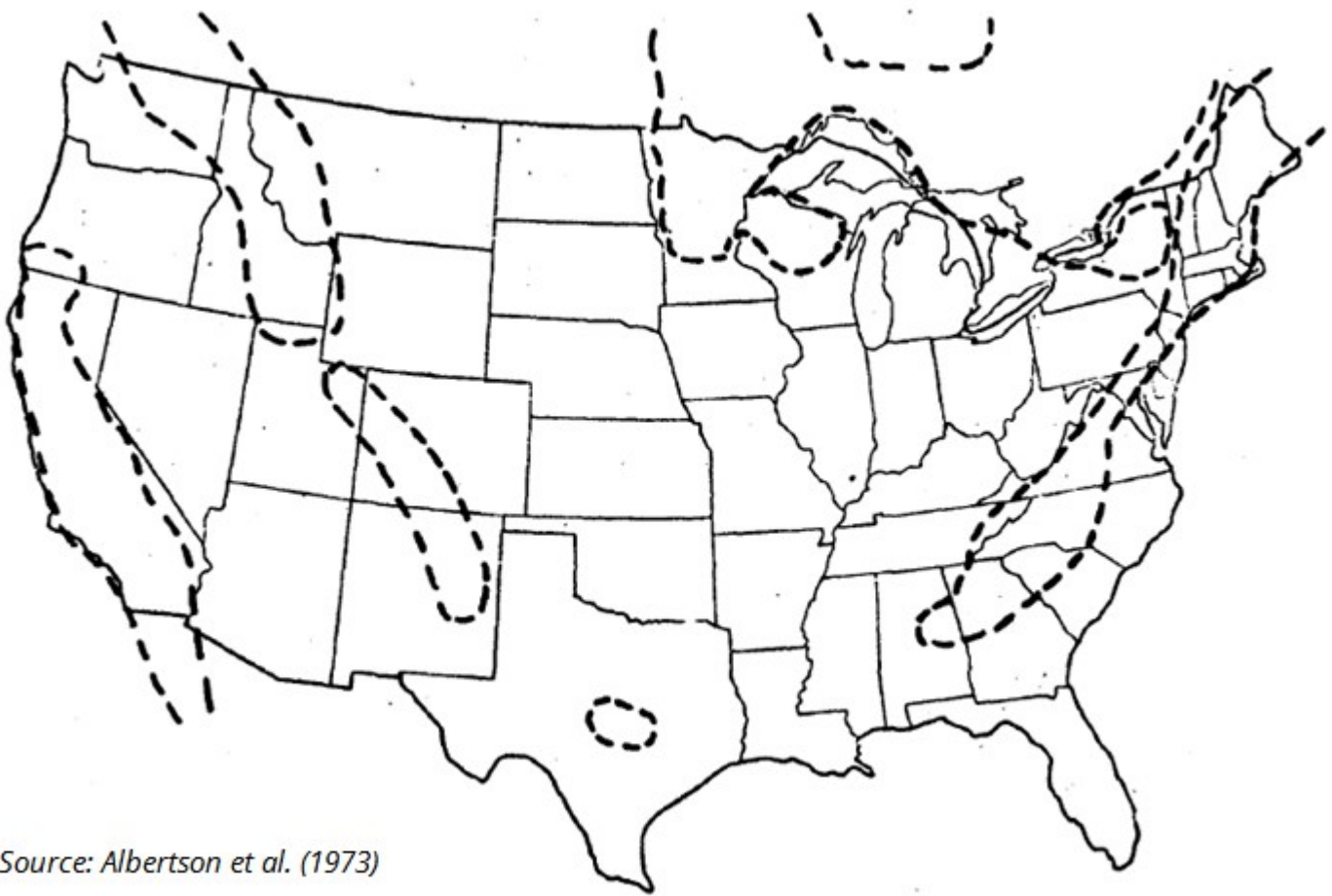


Source: Thomson, Dawson, and Reay (2011)

Before modernity, geomagnetic storms induced currents mainly in seawater and the earth itself. But by stringing high-voltage power lines across the continents, humanity has created a new path for electrons. We have built the space weather equivalent of lightning rods. How attractive a power line is to GICs (geomagnetically induced currents) depends on its length; on the electrical conductivity of the rock beneath it; and on the proximity of either end to the sea, salt water being a good conductor. The map below, taken from a pioneering investigation of the geomagnetic storm risk (Albertson et al. 1973, fig 1) shows which parts of the United States lie on igneous rock, which particularly resists electricity. In these areas—notably along the populous coasts—power lines are

⁵ [wikipedia.org/wiki/Kelvin%E2%80%93Helmholtz_instability](https://en.wikipedia.org/wiki/Kelvin%E2%80%93Helmholtz_instability) features a nice graphical simulation.

particularly attractive conduits for geomagnetically induced currents. Modern modelling by Pulkkinen et al. (2012, p. 11) suggests that being over resistive ground quintuples the electrical forces at play.



Source: Albertson et al. (1973)

Power lines can withstand the large, transient currents induced by geomagnetic storms. But the transformers at either end can overheat, potentially crippling the grid. The human mind is drawn to stories of catastrophic failure, and there are examples of that, as in the 1989 which permanently disabled a large transformer at the Salem nuclear plant in New Jersey (Kappenman 2010, p. 2-29). Through simulations, Kappenman (2010, pp. 1-14, 4-14, 4-15) concludes that 100-year storm over the continental US could put 368–1003 high-voltage transformers at risk of permanent damage, out of some 2146 in service. Manufacturing replacements can take months—and requires electricity. In an interview with GiveWell’s Ben Rachbach⁶, John Kappenman stated that “One factory could make 30–50 transformers per year.” This raises the specter of very long-term outages over wide areas. Kappenman (2008, p. 10) estimates that full recovery could take 4–10 years and economic costs would be \$1–2 trillion in the first year alone. However, these slides provide no specifics for the economic calculation.

⁶files.givewell.org/files/conversations/Kappenman%208-6-13.pdf

But in a few pages, I will question one basis for this scenario.⁷ And it increasingly seems that the dominant mode of transformer destruction has been subtler. In this mode, storms do not immediately disable transformers on a large scale. Rather, they cause hotspots within transformers, large enough to do local damage. Like untreated rust, the flaws then spread in the following months until the transformer fails. Or perhaps the transformer holds until another storm delivers the coup de grâce. Gaunt and Coetzee (2007) document such slow-motion destruction in eight transformers in South Africa after the great Halloween geomagnetic storm of 2003.⁸ This is one of them:



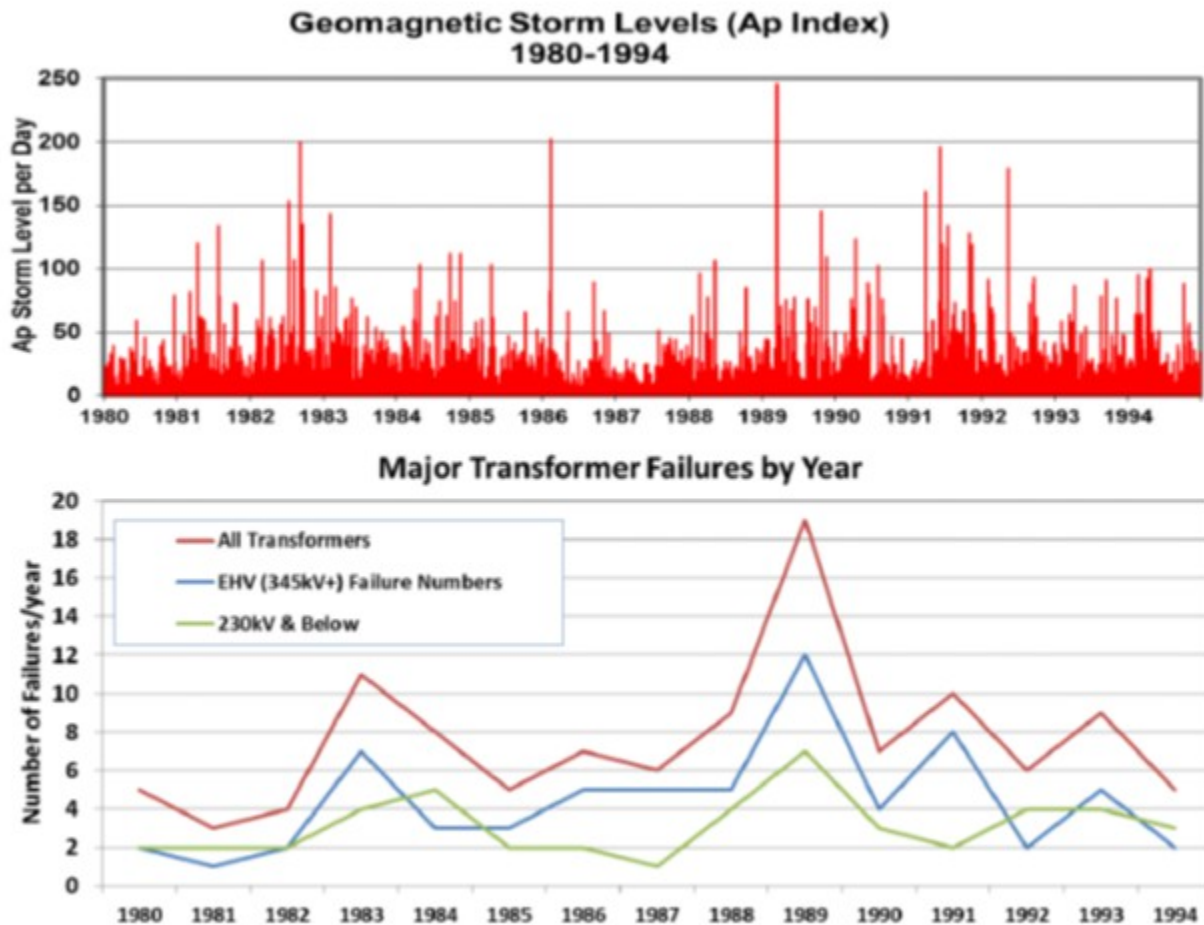
Source: Gaunt and Coetzee (2007)

Gaunt and Coetzee suggest that this failure mode is more common than appreciated. Because of the time lag, when a damaged transformer finally fails, engineers may not recognize a storm as the true cause. Indeed, Storm Analysis Consultants (2013, p. 3-2) has gathered statistical evidence suggesting that storms were indeed a major cause, if not *the* major cause, of transformer failure in the United States between 1980 and 1994. Below, the first graph shows the intensity of global geomagnetic disturbance using something call the Ap index. The bottom graph shows the number of failures of major US transformers reported in an incomplete, voluntary survey of utilities that used to be conducted by the Institute of Electrical and Electronics Engineers (IEEE). The correlation appears strong:⁹

⁷The scenario assumes that magnetic changes of 4800 nT/min (2400 nT/min west of the Mississippi) would occur across the US in a 5-degree band centered on 50° N geomagnetic latitude (Kappenman 2010, p. 4-11), a premise that I will challenge below as unrepresentative of the historical data. The value of 4800, as an actual historical reading, appears to misconstrue the primary source by about a factor of two; it comes from 55° N, not 50° N; and it is for an isolated location, not a continent-wide region.

⁸Moodley and Gaunt (2012, §V.E) also links damage of one of the transformers to smaller geomagnetic disturbance in 2001.

⁹It would be interesting to carry out formal hazard modelling, incorporating time lags and storm strength.



Source: SAC (2013)

In general, geomagnetic storms pose several risks to society: damaging communication and global positioning satellites, accelerating corrosion of pipelines, inducing disruptive currents in electrical grids. This last concern is greatest. As the NRC (2008, p. 3) put it:

Electric power is modern society's cornerstone technology, the technology on which virtually all other infrastructures and services depend. Although the probability of a wide-area electric power blackout resulting from an extreme space weather event is low, the consequences of such an event could be very high, as its effects would cascade through other, dependent systems. Collateral effects of a longer-term outage would likely include, for example, disruption of the transportation, communication, banking, and finance systems, and government services; the breakdown of the distribution of potable water owing to pump failure; and the loss of perishable foods and medications because of lack of refrigeration. The resulting loss of services for a significant period of time in even one region of the country could affect the entire nation and have international impacts as well.

Citing the presentation of R. James Caverly of the US Department of Homeland Security, the NRC (2008, p. 31) continues with examples of risks:

- *Loss of key infrastructure for extended periods due to the cascading effects from a space weather event (or other disturbance) could lead to a lack of food, given low inventories and*

reliance on just-in-time delivery, loss of basic transportation, inability to pump fuel, and loss of refrigeration.

- *Emergency services would be strained, and command and control might be lost.*
- *Medical care systems would be seriously challenged.*
- *Home dependency on electrically operated medical devices would be jeopardized.*

In addition, prolonged lack of external power and diesel fuel delivery might even compromise cooling systems for spent fuel pools at nuclear installations, as at Fukushima (Foundation for Resilient Societies 2011).

What is the probability per unit time of a storm at least as extreme as the Carrington event?

Carrington comparisons

Geomagnetic storms are not rare. The literature mentions major events in 1847, 1859, 1872, 1909, 1921, 1960, 1972, 1982, 1989, and 2003, among others (Kappenman 2006; Silverman 2006, Cliver and Dietrich 2013). Since postwar society has survived many storms without difficulty, a key question is whether something much bigger lurks around the corner, which could wreak havoc of a different order. In this regard, the Carrington CMEs of 1859 are often taken as a benchmark. When the Carrington storm hit, the main consequences were spectacular auroras and fires at a few telegraph stations (Green 2008). Today the consequences might be far worse.

This concern raises a question: how much stronger was the Carrington storm than recent ones? Though low-quality by modern standards, data are available to partially answer this question. This table shows some indicators along with corresponding values for modern comparators:

Storm strength indicator	Carrington	Modern comparators	Sources
Associated solar flare intensity (soft X-ray emissions)	0.0045 W/m ²	0.0035 W/m ² , Nov. 2003	Cliver and Dietrich (2013), pp. 2–3
Transit time of CME to earth	17.6 h	14.6 h, Aug. 1972 20.3 h, Oct. 2003	Cliver and Svalgaard (2004), Table III
D_{st} (low-latitude magnetic field depression)	–850 nT	–589 nT, Mar. 1989	Siscoe, Crooker, and Clauer (2006); wdc.kugi.kyoto-u.ac.jp/dst_final/198903
Lowest magnetic latitude where aurora visible	23°	29°, Mar. 1989	Cliver and Svalgaard (2004), p. 417; Silverman (2006), p. 141

W/m² = watts/square meter; h = hours; nT = nanotesla

Cliver and Svalgaard (2004) observe that the Carrington event consistently appears near or at the top in rankings of storms by various indicators. Yet “various lines of evidence indicate that the intensity of the geomagnetic storm beginning 2 September 1859 was not markedly larger (if it was larger at all) than that of the top tier of subsequent great storms” (p. 419).

These comparisons suggest, conservatively, that the Carrington event was at most twice as strong as anything yet experienced in the postwar era. The roughly estimated D_{st} of –850 nT is smaller than twice the –589 nT of 1989. Likewise for the solar flare intensity of 0.0045 W/m², against the 0.0035 of 2003.

The July 2012 near-miss

Another important comparator is the major CME of July 23, 2012. Despite an angular width estimated at 160° (Baker et al. 2013, p. 587), the CME missed the earth. Indeed it left from what was then the far side of the sun (Baker et al. 2013, fig. 2). However, the NASA satellite STEREO-A was travelling along earth's orbit about 4 months ahead of the planet, and lay in the CME's path, while STEREO-B, trailing four months behind earth, was also positioned to observe. The twin probes produced the best measurements ever of a Carrington-class solar event (Baker et al. 2013). Since the sun rotates about its axis in less than a month, had the CME come a couple of weeks sooner or later, it could well have smashed into our planet.

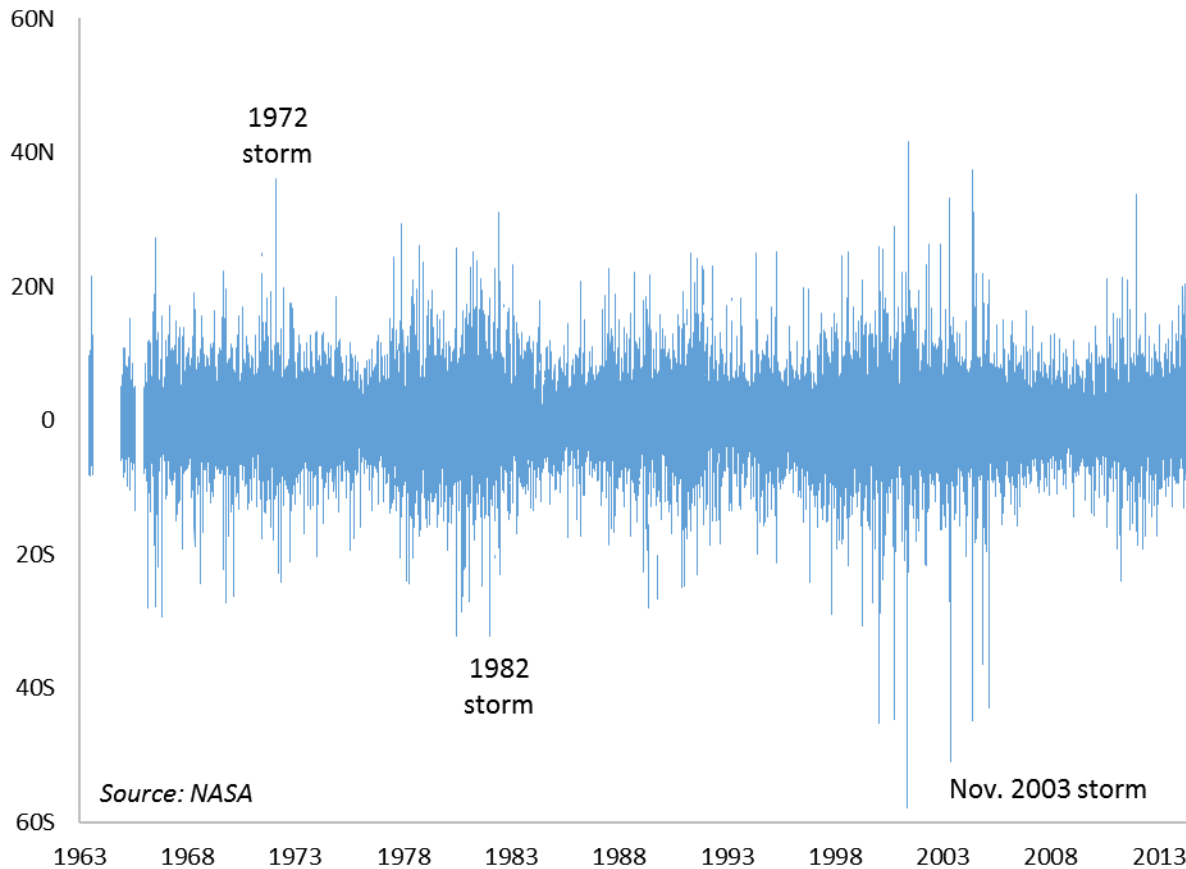
Two numbers convey the power of the CME. First is its transit time to earth orbit: at just under 18 hours, almost exactly the same as in the Carrington event.¹⁰ A slower CME on July 19 appears to have cleared the interplanetary medium of solar plasma, resulting in minimal slowdown of the big one on July 23 (Liu et al. 2014). Second is the strength of the component of the CME's magnetic field running parallel to earth's. Recall that a CME strews the most magnetic chaos when its field parallels the earth's (meaning that both point south) and leaves the least imprint when oriented oppositely. The magnetic field of the great July 2012 CME was measured at 50 nT south at its strongest point (Baker et al. 2013, fig 3, panel 1). Here, however, "south" means perpendicular to the earth's orbital plane. Since the earth's spin axis is tilted 23.5° and its magnetic poles deviate from the spin poles by another 10° , the southerly magnetic force of the near miss CME *had it hit the earth* could have been more or less than 50 nT. Baker et al. (2013, p. 590) estimate the worst case as 70 nT south, relative to earth's magnetic orientation.

For comparison, the graph below shows the north-south component of the interplanetary magnetic field near earth since 1963, where north and south are also defined by the earth's magnetic poles. Unfortunately, data are missing for the largest storm in the time range, the one of March 1989.¹¹ But the graph does reveal a large northerly spike in 1972, which explains why that year's great CME caused minimal disruption despite its record speed (Tsurutani et al. 2003, pp. 6–7). Also shown are large southerly magnetic forces in storms of 1982 and 2003, the latter reaching 50 nT.

¹⁰The eruption occurred at about 2:05 universal time on July 23, 2012. STEREO-A began to sense it around 21:00. (Baker et al. 2013, pp. 587–88.)

¹¹Downloaded from cdaweb.sci.gsfc.nasa.gov/cdaweb/sp_phys, data set OMNI2_H0_MRG1HR, variable "1AU IP Bz (nT), GSM" (meaning 1 astronomical unit from sun, interplanetary magnetic field Z component, geocentric solar magnetospheric coordinates, nanotesla). Readings are hourly, with gaps.

North-south component of interplanetary magnetic field near earth (nT)



Given the July 2012 CME's speed, magnetic field, and density, how big a storm could it have caused had it hit earth? Baker et al. (2013, pp. 589–90) estimate that it would have rated between -480 and -1182 on the D_{st} index, depending on how much the CME's magnetic field paralleled the earth's at the moment of collision. Separately, Liu et al. (2014, p. 5), estimated the range as between -600 and 1150 nT.

As the authors note, the higher number is somewhat more conjectural because it is produced by a model that has not been calibrated to real data on such extremes, for lack of instances. Nevertheless, taking the high-end D_{st} at face value and comparing to the actual modern record of -589 , for March 1989, again points to a realistic worst-case storm as being twice as strong as anything experienced since the construction of modern grids.

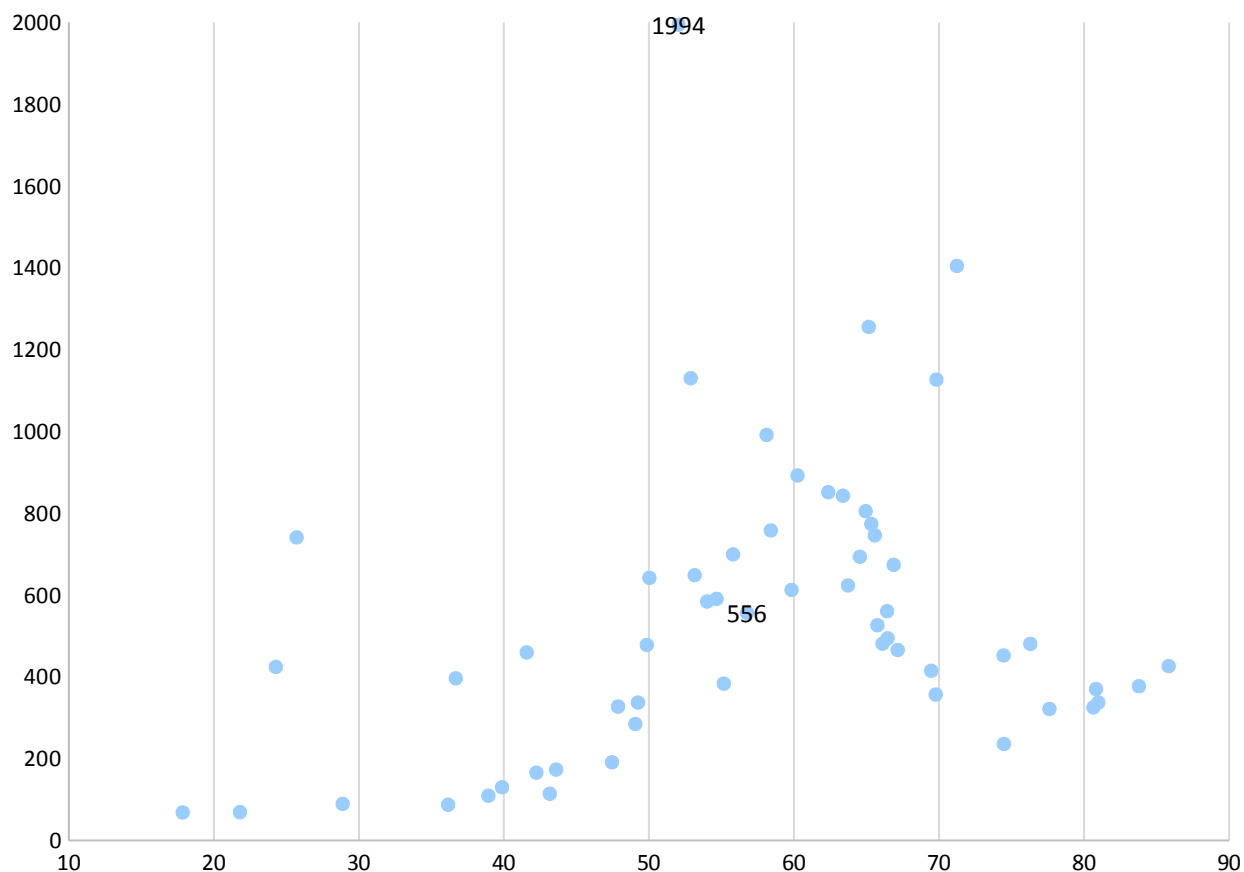
In a companion paper, the authors of Baker et al. (2013) run computer simulations to develop a more sophisticated understanding of what would have happened if earth had been in STEREO-A's place on July 23. Their results do not point to a counterfactual catastrophe. "Had the 23 July CME hit Earth, there is a possibility that it could have produced comparable or slightly larger geomagnetically induced electric fields to those produced by previously observed Earth directed events such as the March 1989 storm or the Halloween 2003 storms." (Ngwira et al. 2013b, p. 677)

Kappenman's factor of 10

In contrast, the prominent analyst John Kappenman has favored a factor of 10 for the once-in-a-century scenario.¹² Recognizing that this difference begs explanation, I investigated the basis for the factor of 10. Readings and correspondence with Kappenman lead me to understand that the factor of 10 is the ratio of two numbers. One represents the worst disruption that geomagnetic storms have wrought in the modern age: “the regional disturbance intensity that triggered the Hydro Quebec collapse during the 13 March 1989 storm only reached an intensity of 479 nT/min” (Kappenman 2004; see also Kappenman 2006, p. 188; Kappenman 2010, pp. 1–30; Kappenman 2012b, p. 17-3). While I did not find a clear citation of source for this statistic, it looks highly plausible. The graph below, based on my own extracts of magnetic observatory data, shows the maximum one-minute horizontal field changes at 58 stations on that day in 1989. Each 3-letter code represents an observatory; e.g., FRD is Fredericksburg, VA, and BFE is Brorfelde, Denmark.¹³

¹²“Because the [1-in-100-year scenario] 4800nT/min threat environment is ~10 times larger than the peak March 1989 storm environment, this comparison also indicates that resulting GIC peaks will also in general be nearly 10 times larger as well” (Kappenman 2010, p. 4–12). “Historical evidence indicates that even more severe storm levels could reach an intensity of as much as 5000 nT/min, ~10 times larger than the March 1989 storm” (SAC 2011a, p. Exec-1). “This disturbance level is nearly 10 times larger than the levels that precipitated the North American power system impacts of 13 March 1989” (Kappenman 2004).

¹³Plotted are all stations with data for the period in NOAA’s SPIDR system, at spidr.ngdc.noaa.gov, or in the Nordic IMAGE network, at www.geo.fmi.fi/image. Geomagnetic latitudes are from the calculator at omniweb.gsfc.nasa.gov/vitmo/cgm_vitmo.html. For a list and maps of observatories, see Rasson (2005).



Ottawa (OTT, in red) recorded a peak change of 556 nT/min, between 9:50 and 9:51pm universal time, which is compatible with Kappenman’s 479. BFE recorded the highest value, 1994 nT/min.

The other number in the factor-of-10 ratio represents the highest estimate we have of any magnetic field change before World War II, at least at a latitude low enough to represent a major concern for Europe or North America. It comes from Karlstad, in southern Sweden, during the storm of May 13–15, 1921. The rate of change of the magnetic field was not measured there, but the electric field induced in a telegraph line coming into the town was estimated at 20 volts/kilometer (V/km). Calibrating to modern observations, “the 20 V/km observation... suggests the possibility that the disturbance intensity approached a level of 5000 nT/min” (Kappenman 2006, p. 195). Kappenman (2010, p. 3-22) suggests 4800 nT/min. And $4800/479 \approx 10$.

I have two concerns about the estimate of this ratio. First, the top number appears to have been unintentionally increased by a scholarly game of telephone. As a source for the 20 V/km observation, Kappenman (2006) cites—and correctly represents—Elovaara et al. (1992, p. 2), who write, “The earth surface potentials produced are typically characterized by the value 1 V/km, but in extreme cases much higher values has been recorded like 20 V/km in a wire communication system in Sweden in May 1922.” No source is given there; but Jarmo Elovaara pointed to Sanders (1961) as likely (correspondence, October 28, 2014, citing aid from Risto Pirjola). Indeed, in Sanders (1961), we read, “In May, 1921, during an outstanding magnetic storm, the largest earth-current

voltages measured on wirelines in Sweden ranged from 6.3 to 20 v/km" (p. 371). The source for *that* range is the "Earth Currents" article of the 1943 *Encyclopedia Britannica*, which states: "In May 1921, during an outstanding magnetic storm, Stenquist calculated from the fusing of some copper wires and the non-fusing of others that the largest earth current voltage in Sweden lay between 6.3 and 20 volts per kilometre" (Britannica 1943). "Stenquist" is David Stenquist, a Swedish telegraph engineer who in 1925 published *Étude des Courants Telluriques* (*Study of Earth Currents*, Stenquist 1925). The pertinent passage thereof comes on page 54:

Neanmoins j'ai essayé de calculer la valeur la plus grande des courants telluriques. Jusqu'ici l'opinion ordinaire est que les différences de potentiel les plus grandes dans la terre à cause des courants telluriques sont de deux volts par kilomètre. Pendant les nuits du 13—14 et 14—15 mai, cette valeur fut de beaucoup dépassée. En plusieurs cas les courants ont été si forts dans les lignes de cuivre (3 mm), que les tubes de fusion ont fondu, c'est à dire l'intensité du courant a dépassée 2,5 ampères. Parceque le fil de cuivre maintenant mentionné a une résistance de 2,5 ohms par kilomètre, on reçoit une différence de potentiel de 6,3 volts par kilomètre. Par contre les tubes de fusion placés sur les lignes de fer (4 mm) n'ont pas fondu. Ces lignes de fer ont une résistance de 8 ohms par kilomètre. Par cela on sait, que 20 volts n'ont pas été dé-passés. Avec une assez grande sûreté on peut dire, qu'une différence de 10 volts par kilomètre s'est trouvée. Sur

Nevertheless I tried to calculate the largest value of telluric [earth] currents. Until now, standard opinion was that the largest potential differences in the earth because of telluric currents are two volts per kilometer. During the nights of May 13–14 and 14–15, this value was greatly exceeded. In many cases the currents were so strong in the lines of copper (3 mm [millimeters]), the conduits melted, i.e. the current exceeded 2.5 amps. Because the copper wire just mentioned had a resistance of 2.5 ohms per kilometer, we get a potential difference of 6.3 volts per kilometer. In contrast, the [fusion tubes?] placed on the iron lines (4 mm) did not melt. These iron lines have a resistance of 8 ohms per kilometer. So it is known that 20 volts did not occur. With a large enough security to speak, a difference of 10 volts per kilometer was found.

Stenquist believed the electric force field reached 10 V/km but explicitly rejected 20. Yet through the chain of citations, "20 volts n'ont pas été dé-passés" became "higher values has been recorded like 20 V/km." Using Kappenman's rule of thumb, Stenquist's 10 V/km electrical force field suggests peaks of 2500 rather than 5000 nT/min of magnetic change on that night in Karlstad.

The second concern I have about the estimated ratio of 10 between distant and recent past is that it appears to compare apples to oranges—an isolated, global peak value in one storm to a wide-area value in another. As we have already seen, the highest value observed in 1989 was not 479 but 1994 nT/min, in Brorfelde, 500 kilometers south of Karlstad. And back in July 13–14, 1982, the Lovo observatory, at the same latitude as Karlstad, experienced 2688 nT/min (Kappenman 2006, p. 193, concurs). At nearly the same moment, some 300 kilometers to the southeast in the town of Töreboda, 9.1 V/km was observed on a 0.921 kilometer Swedish Rail monitoring

line¹⁴; this lines up reasonably with Stenquist's rough estimate of 10 V/km from 1921. It is therefore not clear that the 1921 storm was more intense than those of the 1980s, let alone 10 times more so. Maximum magnetic changes and voltages may have been the same.

If this is correct, a factor of two for the worst-case extrapolation from history, relative to recent experience, still looks reasonable.

A deep question here is about the nature of geomagnetic disturbances. Are they uniform in peak intensity across thousands of kilometers? Or does their turbulent nature create isolated hot spots? In other words, if 2500 nT/min hit Karlstad in May 1921 is it likely that all of Scandinavia, or even Canada and the northern US, also underwent such geomagnetic stress? Or was Karlstad just unlucky and unrepresentative? The distinction matters greatly, for the real fear about geomagnetic storms is that they could disable grids over very large areas. Isolated hot spots, on the other hand, might take out a handful of transformers: enough to make blackouts widespread but not long-term. I return to this question below.

Of course, none of this means that a storm 10 times as intense as recent ones is impossible, only that to contemplate it requires more than extrapolation from the limited historical record.

And it should be said that in the last few years a potentially far more fearsome event has appeared in the historical record. Chemical analysis of tree rings has revealed a jump in the atmospheric concentration of radioactive carbon—carbon 14—between the years 774 and 775, 20 times normal variation (Miyake et al. 2012). Another spike, 60% as big, was found between 992 and 993 (Miyake, Masuda, and Nakamura 2013). Scientists seem agreed that the proximate cause was a jump in extraterrestrial radiation, which converted more atmospheric carbon 12 to its radioactive isotope. They are intensely divided as to the source—the sun, another star, or another galaxy (Miyake et al. 2012; Usokin et al. 2013; Cliver et al. 2014; Neuhäuser and Hambaryan 2014). If the source of either event was a flare from our own star, it must have been far larger than any modern event, perhaps ten times so (Cliver et al. 2014, p. 3). Compounding the uncertainty about the implications for our inquiry is the lack of knowledge about the scale of any concomitant magnetic disruption. Solar flares do not cause geomagnetic storms; CMEs do. Sometimes the two go hand in hand, sometimes not. Whether any CMEs would have been proportional is not known.

At this point, as the scientific debate is hot, it is hard to know what to make of the tree ring findings.

Extrapolating statistically from the historical record

Probability densities and cumulative probability densities

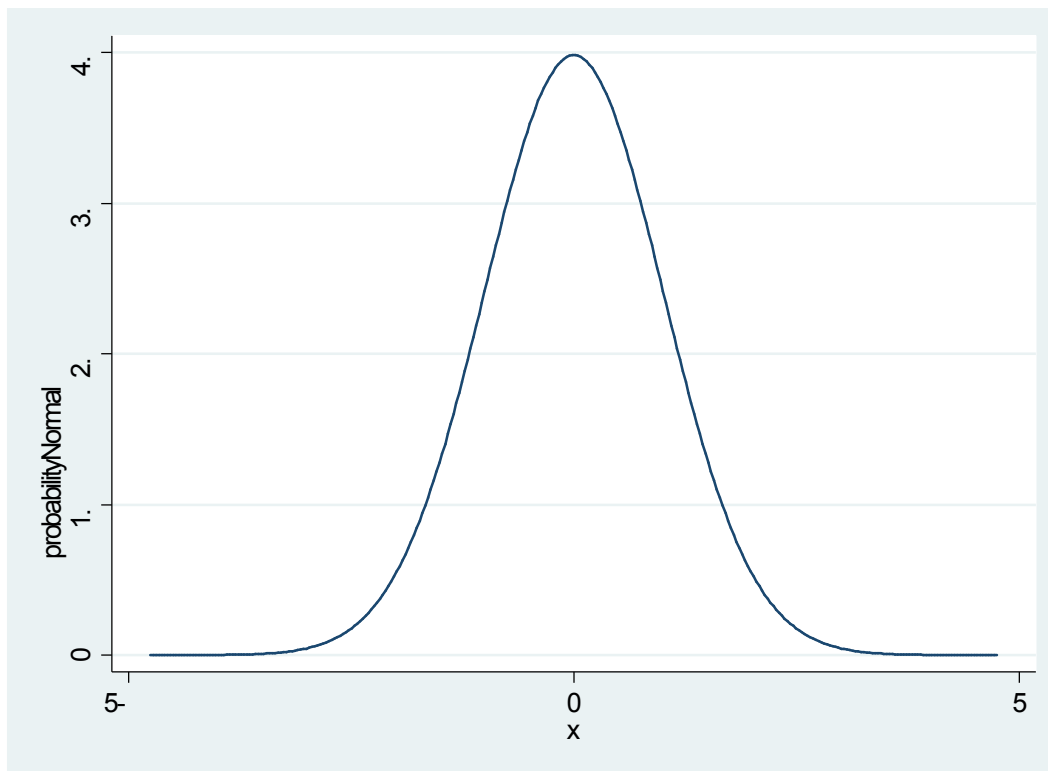
Another approach to estimating the probability of extreme events is to compile (more recent) historical data on indicators such as the storm-time disturbance index (D_{st}) and then use statistical methods to extrapolate probabilities to or beyond the edge of what has so far been observed. This strategy makes fuller use of available data. One result in this vein has also reached the popular press, that the risk of another Carrington event is

¹⁴Kappenman (2006, p. 192) reports this voltage as occurring along a "communication circuit [with] length ~100 km" between Töreboda and Stockholm, which are actually about 300 km apart. However, close inspection of the primary source reproduced in Kappenman—a magnetograph printout—reveals that 8.42V was measured across a line of just 0.921 km, for the reported average of 9.1 V/km. This means that the observation should not be taken as evidence of such a high voltage over a large area. Otterberg (1982, p. 2), confirms that Swedish rail (SJ) maintained equipment in, but not necessarily between, Töreboda and Stockholm to monitor ground potentials created by geomagnetic storms. Artelius (1982, pp. 2–3) also contains this magnetograph printout, alongside two more from ~28 km lines, which show contemporaneous peaks of ~3 V/km over these longer distances. Documents courtesy of Sture Lindahl, Gothia Power, June 28, 2015.

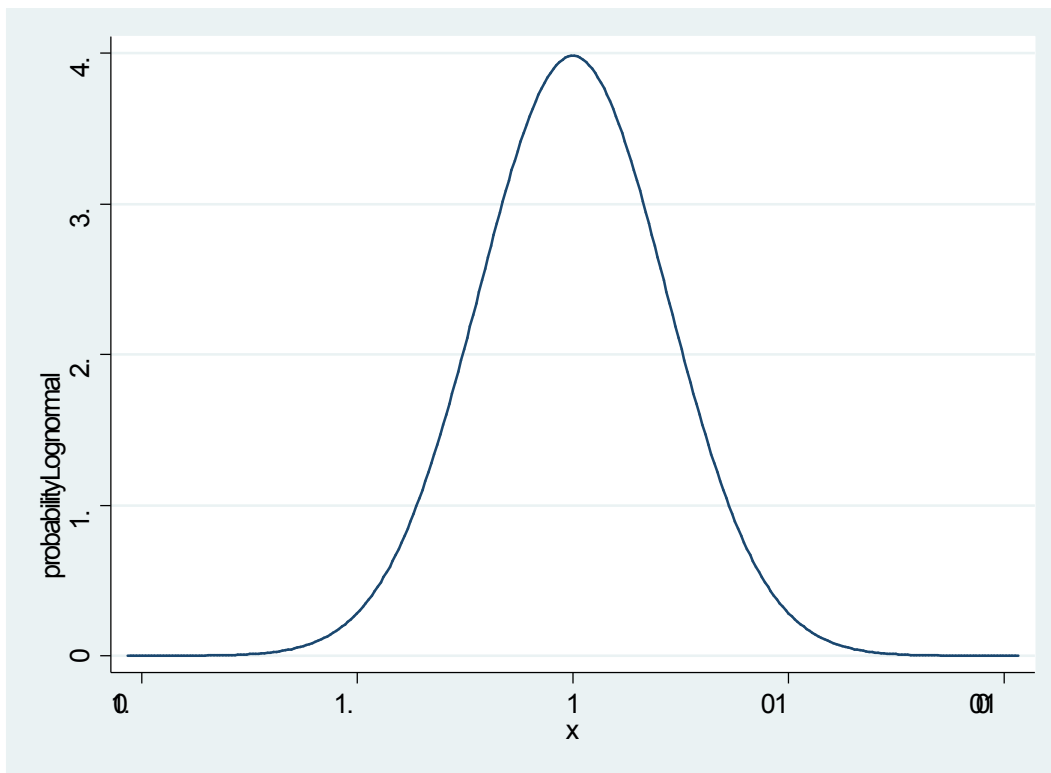
12%/decade (Riley 2012). The rest of this section is devoted to explaining and applying the statistical approach, and explaining why the 12% rate looks too high as an extrapolation from the recent past.

A fundamental notion in statistics is the *distribution*. A distribution is a graph that represents the probabilities of all possible outcomes of a process, such as the roll of a die. Much of the academic discussion over the probability of extreme geomagnetic storms revolves around which mathematical family of distributions best represents the actual distribution of storms. A chosen distribution is fit to the data, and then it is used to estimate probabilities per year or decade of events of various strengths.

The most famous distribution is the normal density, or “bell curve”:

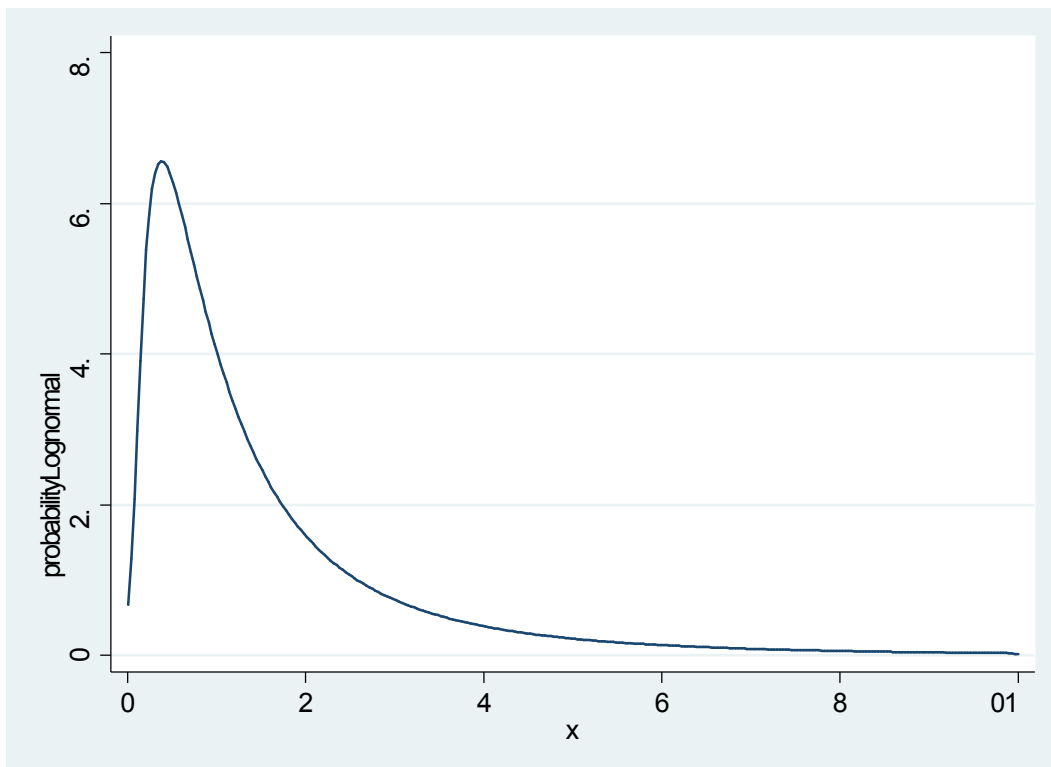


More relevant for us is the lognormal density, which arises when the *order* of magnitude of some variable, such as the population of towns and cities, is normal. E.g., maybe cities of size 1 million are most common, so they form the peak of the distribution. On either side, cities of size 0.1 million and 10 million are equally common. The lognormal distribution can be drawn this way, just by changing the labels on the horizontal axis:



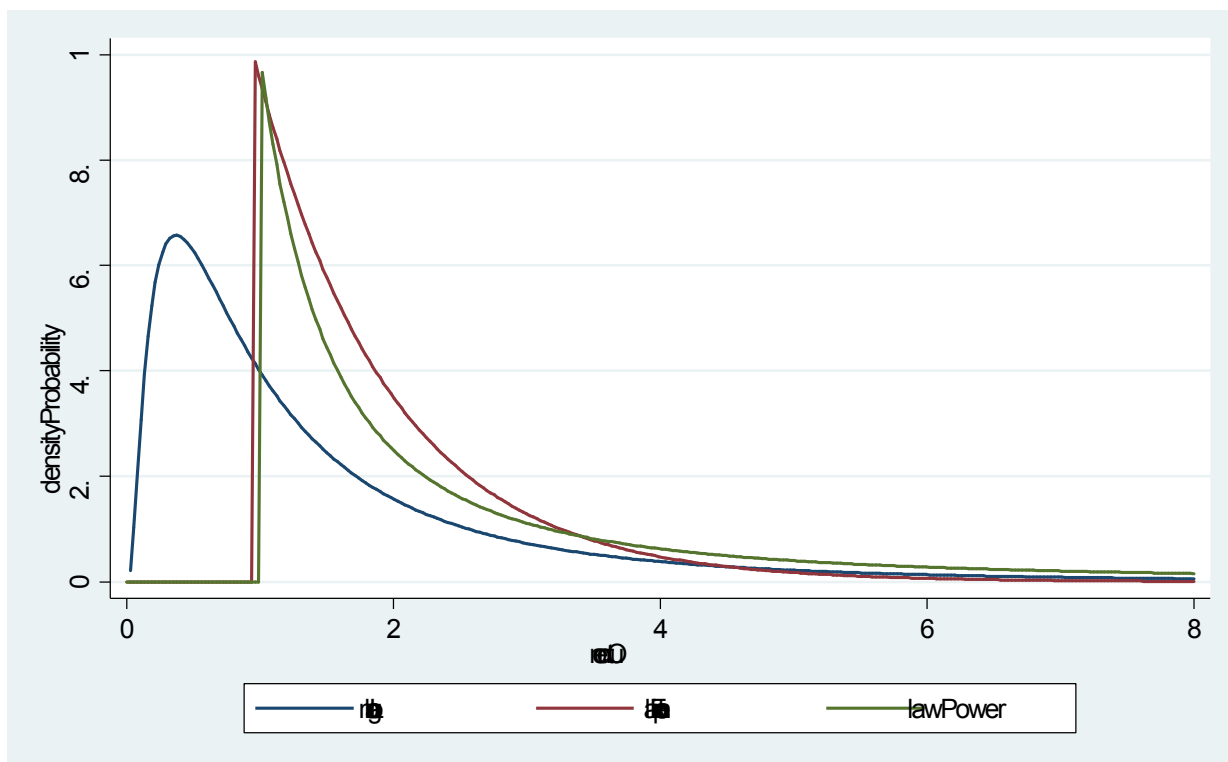
(Even spacing of 1, 10, 100 on the horizontal axis is called a logarithmic scale.) Under this distribution, negative values are impossible, while large positive values are more probable than in the standard normal distribution.

If we rescale the horizontal axis so that 1, 2, 3, ... rather than 1, 10, 100, ... are evenly spaced, the lognormal looks like this:



Two more distributions that figure in this discussion are the exponential and power law distributions. Both put zero probability on values below some minimum; they start high at the minimum and decay for larger values. They differ in the pattern of decay. For example, if the populations of the world's large cities were exponential distributed, then it could happen that 50% of cities have populations between 1 and 2 million, 25% between 2 and 3 million, 12.5% between 3 and 4 million, and so on, halving the share for each increment of 1 million. In contrast, under the power law distribution, the decay could manifest this way: 50% between 1 and 2 million, 25% between 2 and 4 million, 12.5% between 4 and 8 million, and so on, halving the share for each doubling of population. Notice how the power law has a fatter tail, assigning more probability to very large cities—and indeed, urban populations are found to follow a power law (Newman 2006, p. 323).

This graph compares examples of the lognormal, exponential, and power law distributions.¹⁵ The latter two are zero below the chosen cut-offs. Not far above these thresholds, the power law curve is lower than the exponential; but farther to the right, it is higher. Thus the power law predicts fewer low outcomes and more high ones.¹⁶

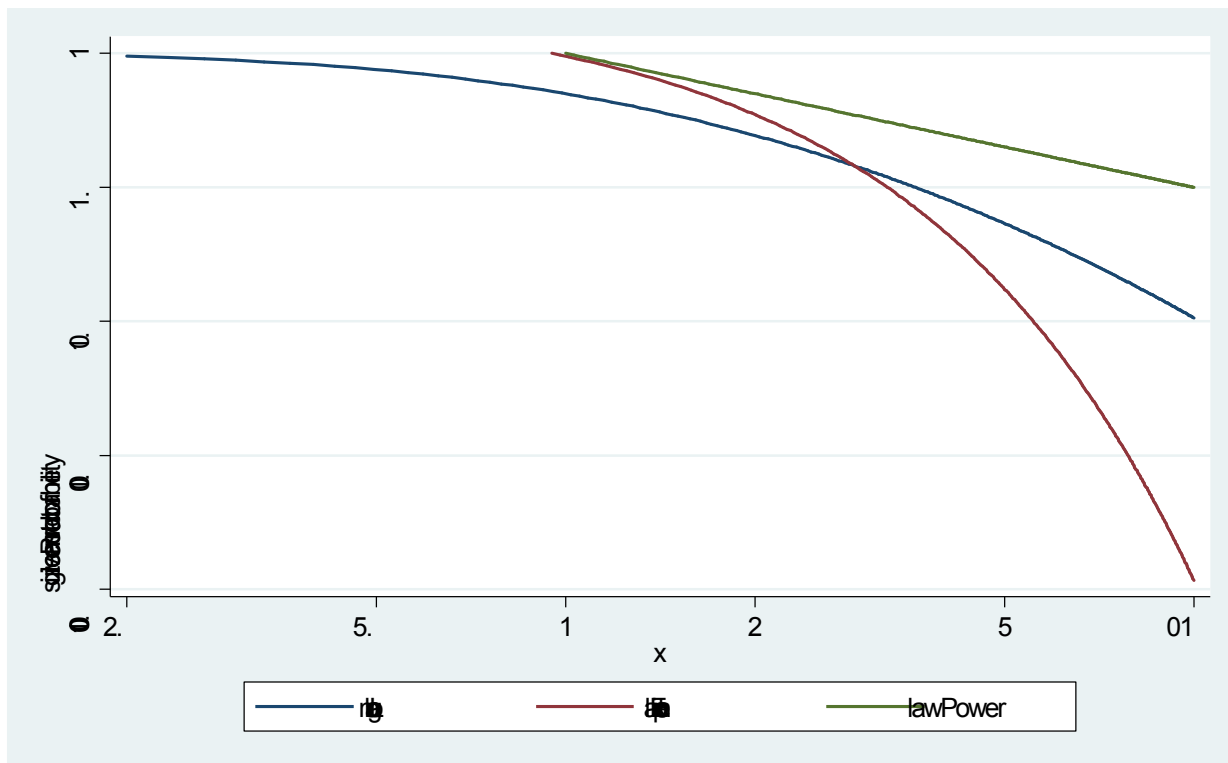


These graphs show probabilities of events of a given size. Another way of graphing distributions that serves our interest in right tails is to show the implied probability of an outcome of *at least* a given size—such as a storm of Carrington size or larger—and do so with both axes on logarithmic scales, which magnifies the tail region for inspection. The next graph redraws the same three distributions in this way. We see that there is a 10% (0.1)

¹⁵One way of organizing one's thinking about these distributions is to observe that the exponential is linear on a log-linear plot, the power law is linear on a log-log plot, and the lognormal is parabolic on a log-log plot.

¹⁶The densities are $\frac{1}{\sqrt{2\pi}x} e^{-(\ln x)^2/2}$ for $x > 0$, $e^{-(x-0.95)}$ for $x > 0.95$, and $1/x^2$ for $x > 1$.

chance of an outcome above 10 according to the power law distribution in the previous graph, but only a 1% (0.01) chance according to the chosen lognormal distribution, and only 0.01% under the exponential:



Even though this graph technically contains no more information than the previous one, to the human eye it reveals something new. The power law distribution chosen here predicts that an event of at least size 10 is 1,000 times more likely than according to the exponential distribution. This gap is remarkable given how similar the two distributions appear in the previous graph. And it illustrates how the choice of distribution one fits to real data can drastically affect the extrapolated probabilities of extremes.

Extreme value theory

All three distributions graphed above have been fit to historical data on CMEs, geomagnetic field disturbances, and related data sets. Yurchyshyn et al. (2005) fit the lognormal to CMEs recorded by NASA's SOHO satellite during 1996–2001. Love and Gannon (2009) fit a power law distribution to the bulk of an equatorial geomagnetic disturbance (D_{st}) series for 1958–2007, except they find the right tail to decay faster, closer to an exponential. Riley (2012), whose work has reached the popular press (Washington Post 2014), uses the power law, as does Kataoka (2013).

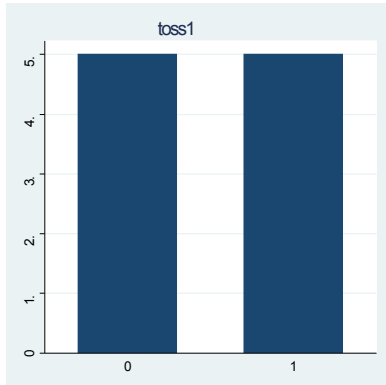
As was just suggested, and as will be illustrated, results are sensitive to choices of distribution for fitting. So which distribution is best?

A branch of statistics called extreme value theory addresses this question. Roughly, its answer is: none of the above. The deep reason is that it may be unrealistic to assume that much or all of a distribution obeys a single, known probability law. For examples, CMEs in different speed ranges may be generated by distinct physical processes (Ruzmaikin, Feynman, and Stoev 2011).

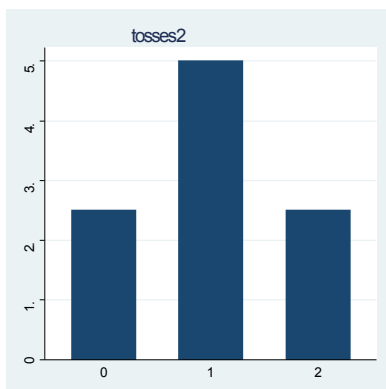
And the choice of distributions is remarkably avoidable. Using extreme value theory, researchers can infer probability ranges for extreme events while remaining agnostic as to the underlying distribution.

One way to explain this is to again start with the normal distribution. A fundamental result in statistics is the Central Limit Theorem, which says that adding or averaging together unrelated random variables almost always tends to produce that normal bell curve.

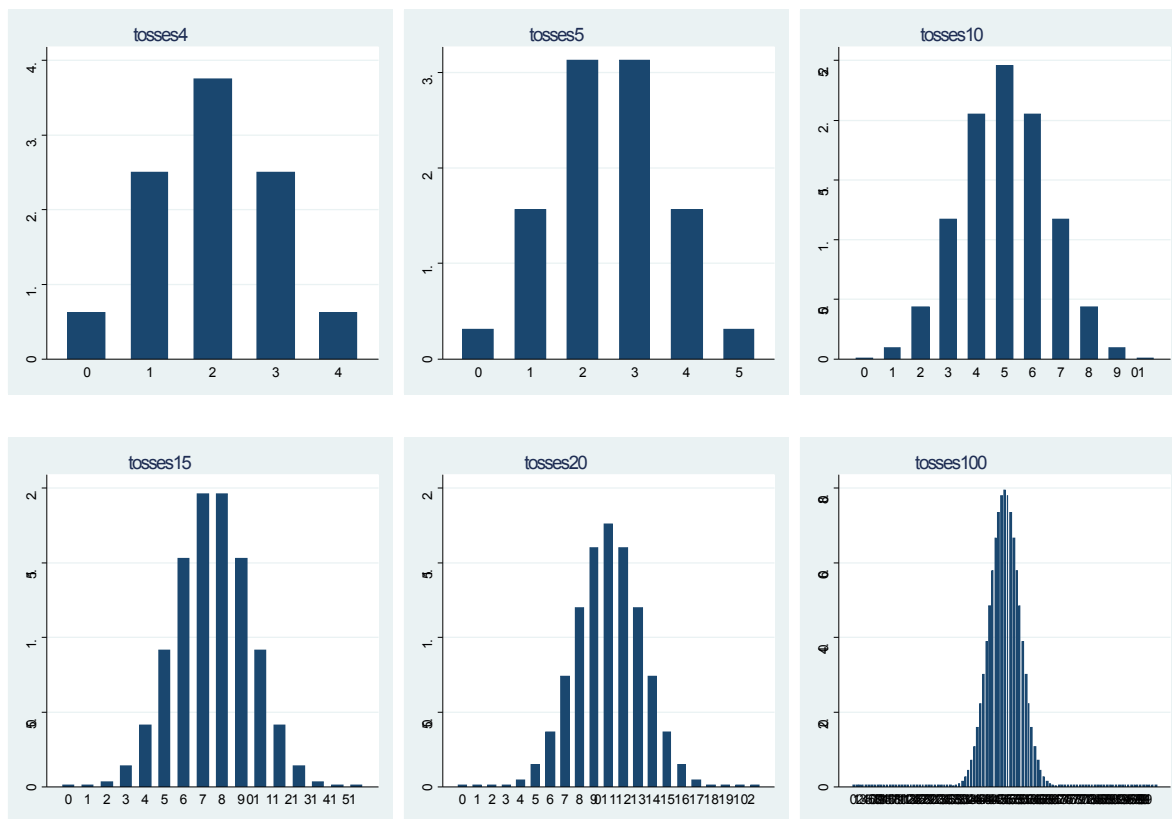
As an example, imagine coin tosses. This graph below shows the probability distribution for the number of heads in a single toss of a fair coin: 50/50 zero-one:



If we flip two coins, four sequences could occur, with equal probability: HH, HT, TH, and TT, where the letters symbolize heads and tails. Two of these sequences would yield a total of one head, so that outcome has a 50% chance. Zero and no heads each have a 25% probability. The distribution goes from rectangular to triangular:



And here are the probability graphs of the number of heads for 4, 5, 10, 15, 20, and 100 tosses:



The more tosses, the closer the distribution comes to a bell curve.

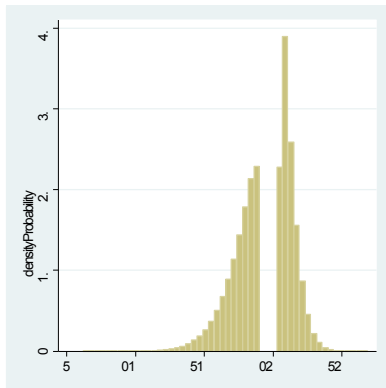
The same thing happens when starting with just about any process you can imagine—rolling dice, polling voters, measuring travel time to work. No matter how many times the underlying distribution zigs and zags, if you sample it enough times and sum or average the results, the bell curve will emerge like a phoenix. Its center will be at the overall average—in this case at 50% heads.

Moreover, as one increases the number of samples that are summed or averaged, the curve narrows in a predictable and universal way. With two tosses, achieving heads 0% of the time is not unexpected: it can happen 25% of the time. With 100 tosses, getting 0% heads is astronomically unlikely. It works out that for every quadrupling of sample size, such as from 25 to 100 coin tosses, the bell curve narrows by half. This square root law is what allows pollsters to compute margins of error. They know that if they repeated the same poll at the same national moment, they wouldn't get precisely the same average answers, since they would randomly call different people. But if many otherwise identical polls are taken at the same time, their individual results will cluster around the true average of citizen sentiment according to a normal distribution whose spread is determined by the number of people polled.

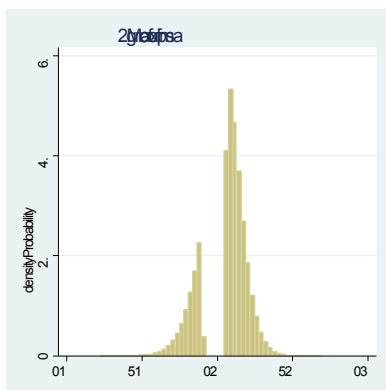
If we sample a distribution many times and *multiply* rather than add or average the results, we will typically get a lognormal distribution.¹⁷ So when lognormal distributions are observed in nature, as in the study of CME speeds by Yurchyshyn et al. (2005), it is reasonable to hypothesize that the underlying physical process is a multiplicative interaction of several erratically varying forces.

¹⁷An added requirement is that all possible outcomes are positive—unlike in the coin toss example, in which tails is treated as 0.

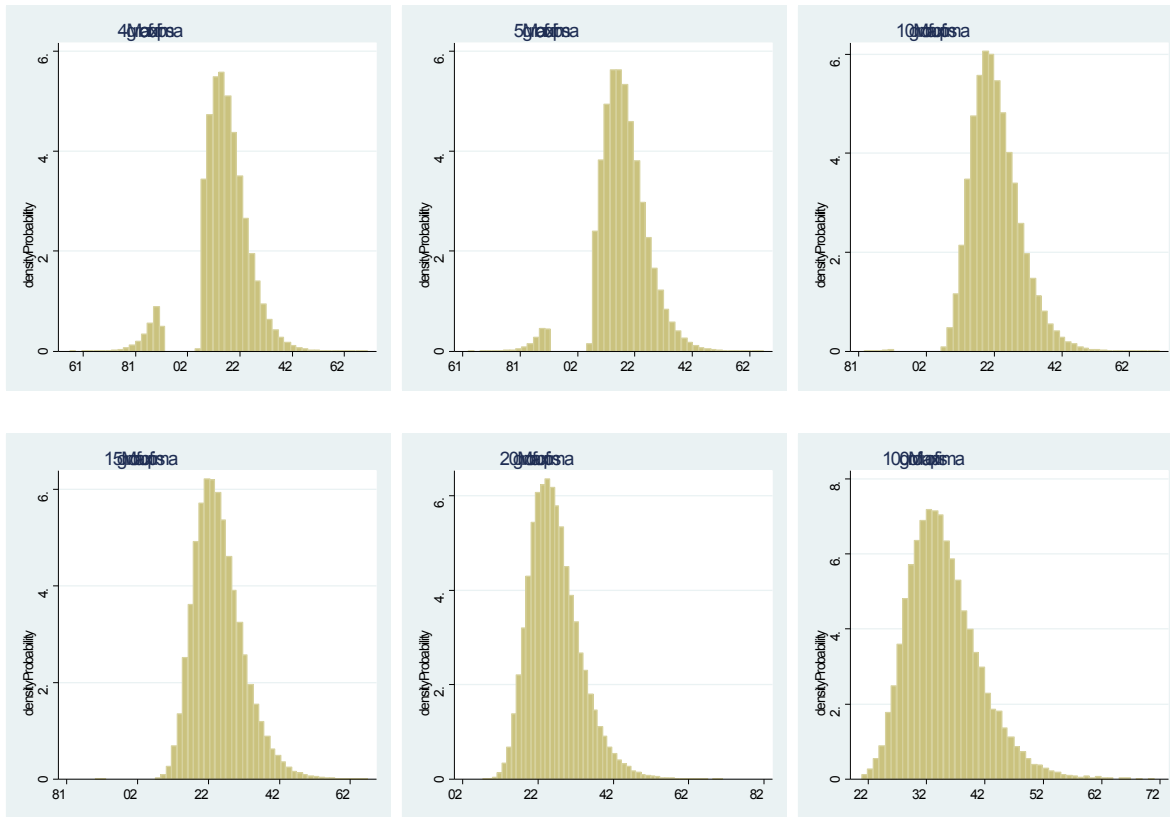
One extreme value theory (EVT) method is quite analogous in motivation to the pollster's reliance on the bell curve. It involves taking *block maxima* (Coles 2001, ch. 3). Imagine that we have data on millimeters of rainfall at an airport for each day for 30 years. The daily data happen have this peculiar distribution, for which I generated a million data points:



Now imagine that we divide the data set into pairs of days. For each pair, instead of adding the results, as we did with coin tosses, we take the maximum: we keep the higher rainfall value and throw away the lower one. In my simulation, these two-day maxima are distributed like this:



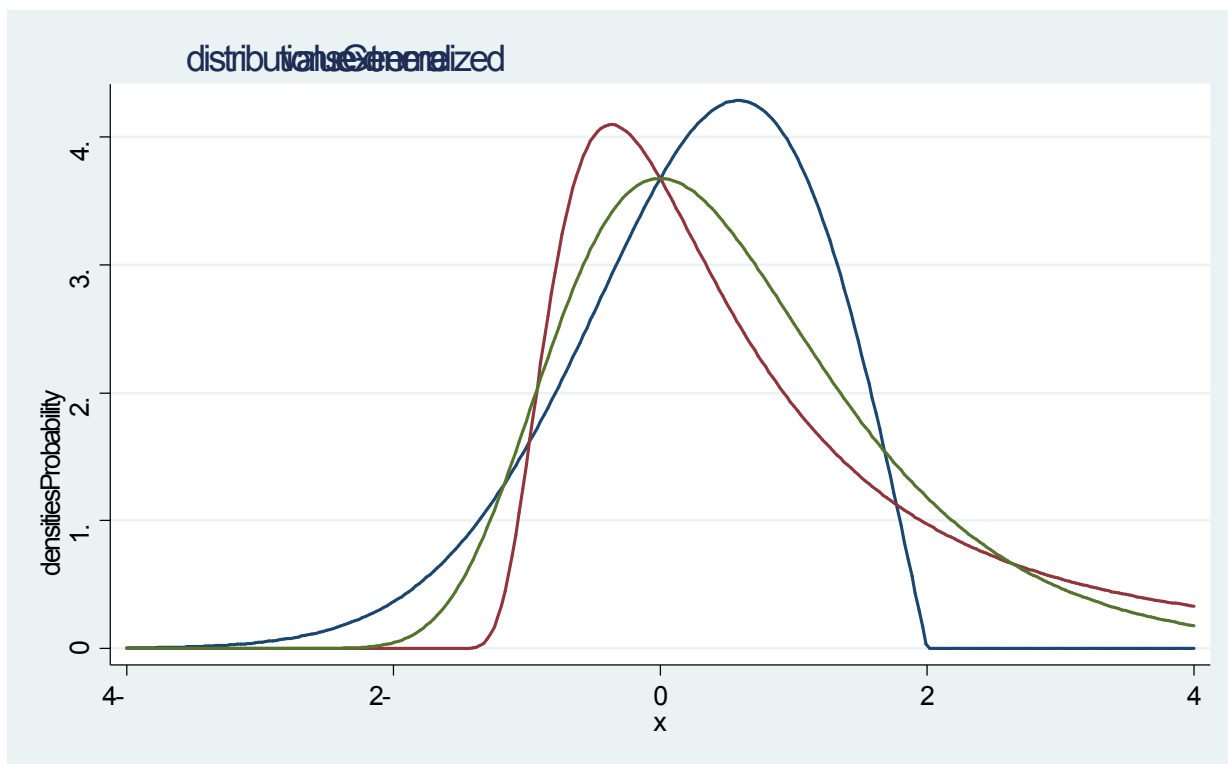
When I instead take maxima over groups of 4, 5, 10, 15, 20, or 100 days, I get:



Again, an elegant curve emerges. But it cannot be the bell curve because it is asymmetric, with a long right tail. Instead, theory typically assures, it is a member of the *generalized extreme distribution* (GEV) family. Three members of this family are depicted here¹⁸:

¹⁸The GEV is $g(x) = \frac{\partial}{\partial x} e^{-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi}} = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi - 1} e^{-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi}}$, with $\lim_{\xi \rightarrow 0} g(x) = \frac{1}{\sigma} e^{-e^{-(x - \mu)/\sigma}}$. The

members graphed here have $\mu = 0, \sigma = 1, \xi = -0.5, 0, 0.5$.

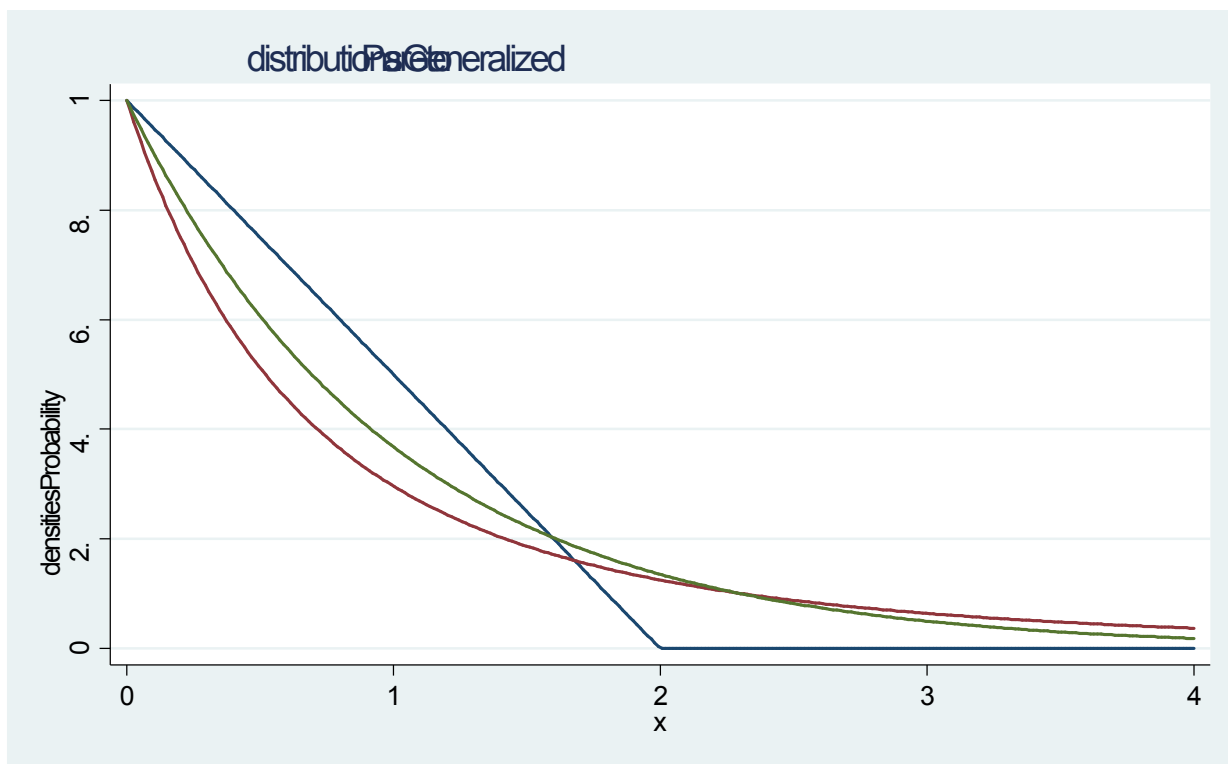


As with the Central Limit Theorem, the key point is that the GEV forms emerge almost regardless of the distribution of the original data. The forms differ from the normal curve because instead of averaging or summing groups of data points, we are taking maxima.

So without making strong and perhaps debatable claims about the pattern of *daily* rainfall data, we can group the rainfall data into 100-day blocks, take maxima, find the member of the GEV family that best fits the maxima, then follow the contour of this member's rightward tail to estimate the probability of say, at least 10 centimeters of rain falling in a single day within any 100-day period. Tsubouchi and Omura (2007, Table 1) do the analogous for daily geomagnetic storm D_{st} statistics for 1957–2001, taking one maximum for each year.

Of course rainfall patterns, like geomagnetic storm patterns, could change, defying predictions. Past need not be prologue. But that challenge applies to any method of extrapolating from historical data. The virtue of EVT methods is that they are grounded in rigorous statistical theory and reduce the need for *a priori* assumptions. EVT methods provide the firmest basis for extrapolating from the past.

A distinct but closely related EVT method focusses more exclusively on extreme data points (Coles 2001, ch. 4). It turns out that for all the diversity in probability distributions, their tails tend to be pretty much alike in how they decay toward zero. In form, they too converge to members of a particular family of distributions, called generalized Pareto (GP) distributions. Some members of this family are graphed below. To repeat, the idea is that pretty much all extreme event distributions look like one of the curves below.



Remarkably, there is a correspondence between block maxima and tails. For example, if a distribution's block maxima follow the red contour in the previous graph, rather like my made-up rainfall data, then its right tail will look like the red curve just above.¹⁹

This provides another way to estimate extreme probabilities while avoiding strong and potentially debatable assumptions about the overall distribution of events in question. If a particular generalized Pareto distribution well fits the tail above some high threshold, such as 25 millimeters/day in the rainfall example, then we can reasonably use it to project probabilities at even higher levels. We can also use standard methods to compute confidence intervals. Tsubouchi and Omura (2007, Table 1) apply this technique too to D_{st} statistics for 1957–2001. Thomson, Dawson, and Reay (2011, Table 1) do the same for readings from a selection of European magnetic observatories for 1979–2010.

Applying EVT to geomagnetic storms

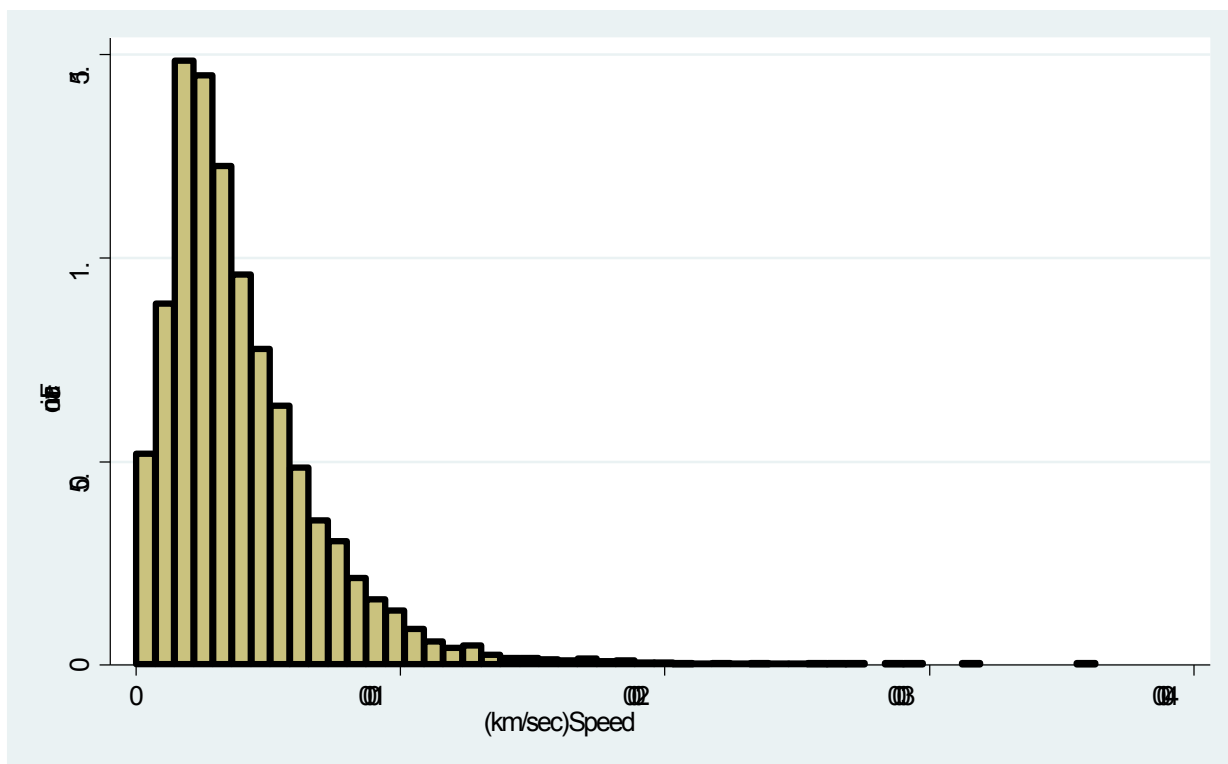
To better understand the probability estimates for extreme geomagnetic storms, I applied extreme value theory to two kinds of data: speeds of coronal mass ejections (CMEs) from the sun, and values of the storm-time disturbance index (D_{st}), which, recall, measures the average equatorial deviation in the magnetic field. Both are correlated, if imperfectly, with the destructive potential of a CME.

CME speeds

NASA data on the speeds of CMEs are accessible at cdaw.gsfc.nasa.gov/CME_list.²⁰ This graph presents the distribution of speeds, in km/sec, of the more than 22,000 CMEs detected since 1996:

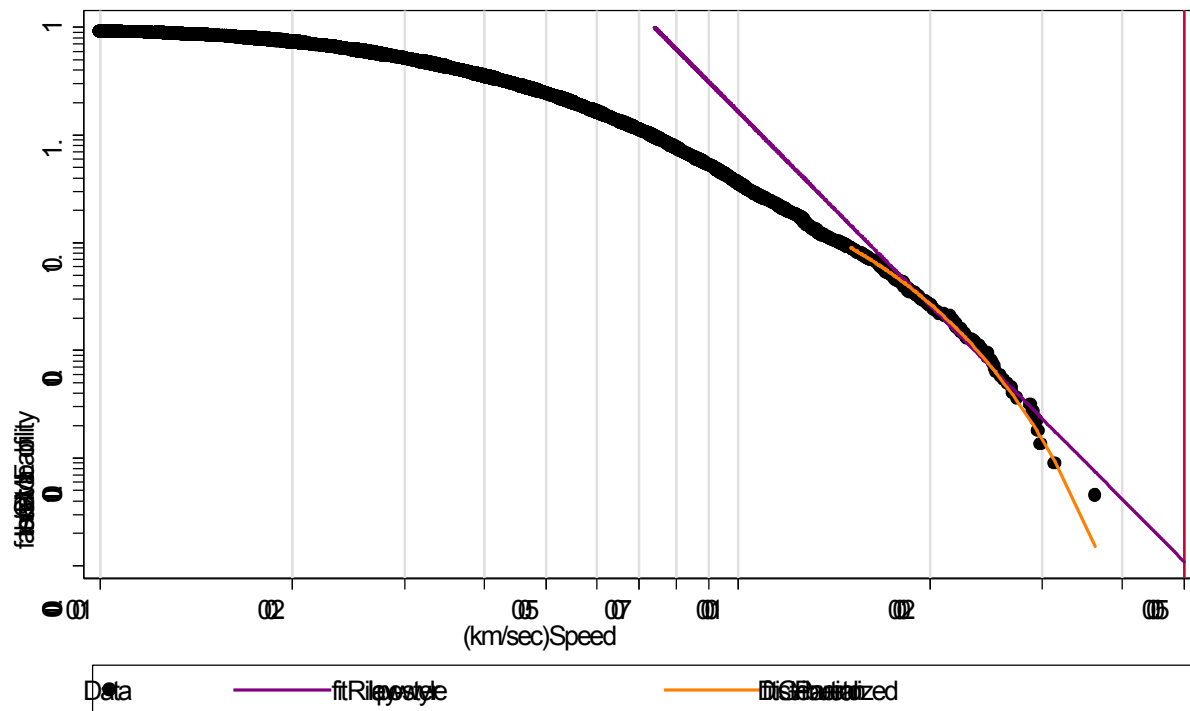
¹⁹The GP is $g(x) = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-1/\xi - 1}$, with $\lim_{\xi \rightarrow 0} g(x) = \frac{1}{\sigma} e^{-(x - \mu)/\sigma}$. The members graphed here again

have $\mu = 0, \sigma = 1, \xi = -0.5, 0, 0.5$.



As explained in the previous section, instead of counting how many CMEs occur in each speed bracket, it is useful to graph the probability of a CME being of a given speed or higher, and to do so with both axes logarithmic (“log-log scales”). The black dots below do that. Where packed together, the dots look like a solid curve:

²⁰These are “plane-of-sky” speeds, meaning that an extremely narrow CME coming straight toward the earth would have a speed of zero, because it would not appear to be moving. But CMEs can be 45 or more degrees wide, so that even when directed straight at earth, their perimeters are moving quickly across the plane of the sky. “Quadratic speeds” at initial reading—based on parabolic fits to at least three observations of the CME—are taken from the NASA data set where provided. Linear speeds are used otherwise. CMEs with no speed data are assigned a zero speed.



We see for example that about 10% (0.1) of CMEs since 1996 left the sun faster than 700 km/sec.

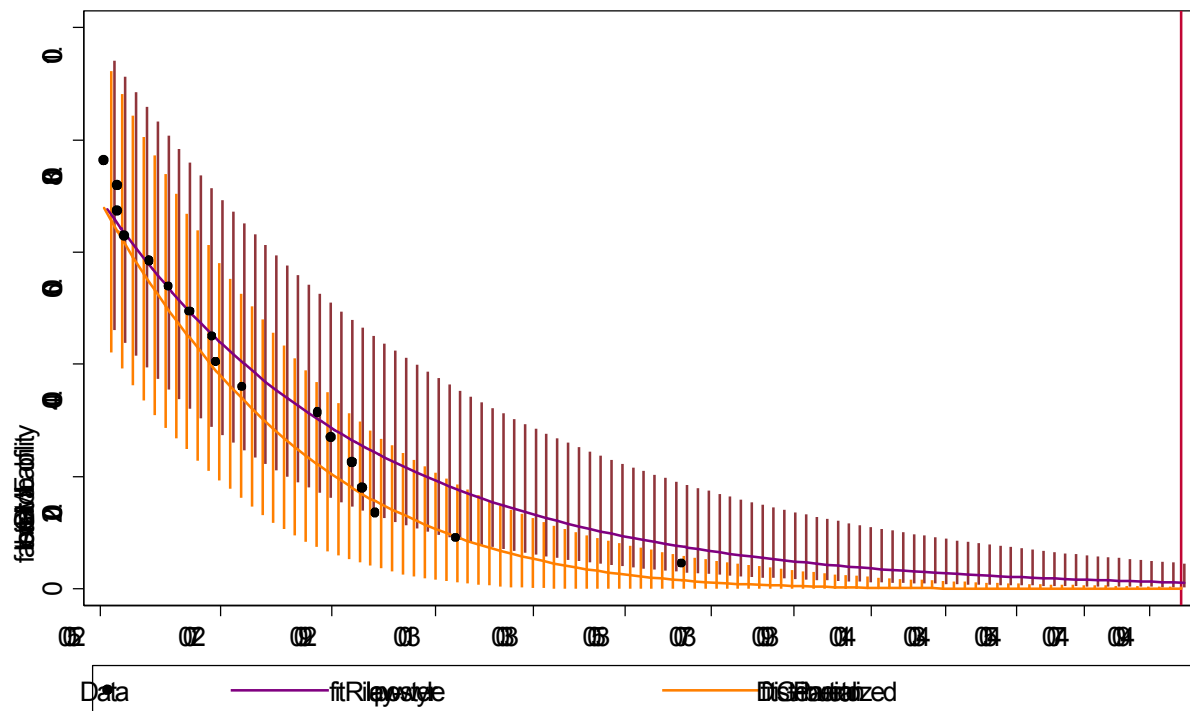
Superimposed on this graph are two fitted distributions. One is a straight line, in purple, which is fit to the CMEs above 2000 km/sec, as in Riley (2012, p. 6). A straight-line fit on this graph corresponds to a power law, which has a fat tail. The other fitted distribution, in orange, is a generalized Pareto curve, also fit to the 201 CMEs above 1500 km/sec.²¹ The red vertical line marks 5000 km/sec, the speed at which Riley (2012, p. 6) estimates the first Carrington CME left the sun in 1859.

The graph helps us think about the probability that another Carrington-speed CME could be generated today. Extending the power law line to 5000 suggests a probability *per CME* of 0.0011%: that is the vertical coordinate where the purple power law line meets the red Carrington line. That probability may seem low, but the data set reports 22,267 CMEs in 18 years, of which 0.0011% works out to 0.135 Carrington CMEs/decade—i.e., we should expect about one-sixth of a Carrington CME per decade or, more intuitively, one every 74 years. A more rigorous calculation turns 0.0011% per event into a 12.7% chance of at least one per decade.²² This is close to Riley’s 12% figure.

²¹The threshold of 1500 was chosen with a graphical method described in Coles (2001, §4.3.1). If a population obeys a GP, then the average excess of the data points above a threshold x should be linear in x . This appears to be the case in the CME data at and above 1500 km/sec.

²²The first calculation is $22,267/1.797 \times 0.00001093$. The second uses the Poisson distribution (Riley 2012, eq. 6): $1 - e^{-22,267/1.797 \times 0.00001093}$.

But the GP curve (above, in orange) points to lower probabilities for extreme geomagnetic storms. To see how much so, we need to abandon the logarithmic scaling of the vertical axis and zoom in on the right tail. This graph does that, along the way adding 95% confidence intervals.



Based on the GP fit to the extreme right tail (orange), the central estimate of the probability of a 5000 km/sec CME is *6 in 1 billion* (6×10^{-9}). That is where the orange curve hits the red. However, the 95% confidence interval runs from 0 to 0.00033% per CME, which is 0.0–4.0% per decade.²³

But not all CMEs hit Earth. A big CME on July 23, 2012, missed the planet for example. Scientists measured its launch speed at 2000–3000 km/sec and angular width at $140^\circ \pm 30^\circ$ (Baker et al. 2013, p. 587). If we conservatively take 180° as a representative angular width, then fast CMEs have a 50% chance of hitting the earth. We might divide by two again to account for that hopeful possibility that a CME’s magnetic field will parallel rather than oppose that of earth’s, reducing magnetic disruption. These adjustments would narrow our 95% confidence interval to 0.0–1.0%.

This analysis suggests that based on this data set, the risk is lower than that presented in Riley (2012). Still, the high end of that range represents a serious risk if the result would be a long-term, continent-scale blackout.

In absorbing this finding, bear in mind several caveats. First, we do not know precisely how fast the Carrington CMEs left the sun. More to the point, we do not know precisely how fast a CME would need to launch in order to inflict catastrophic damage on electrical grids.²⁴ 5000 km/sec may not be the right benchmark. Or it could be that what matters less is the speed of any single CME than the tight sequencing of several, as happened in 1859. The

²³The GP was fit with Maximum Likelihood, with a parametric-bootstrap bias correction with 100 replications. Standard errors of the parameter estimates and predicted probabilities were in turn non-parametrically bootstrapped, clustering by calendar half-year to adjust for serial correlation in CME speeds. Confidence intervals are one-tailed, left-anchored at 0. All estimates performed with my “extreme” package for Stata.

earlier CMEs literally clear the way for the later ones. And the sequence may progressively amplify the electrical and magnetic energy flows about the earth. More important, it would be dangerous to extrapolate confidently from 18 years of solar activity data. Evidently a 5000 km/sec CME was unlikely in the last 18 years, as one did not happen. Some future solar cycles will be more energetic.

The D_{st} index

Another historical data set that is central to the study of geomagnetic storms is the record of the “ D_{st} ” index, which is a measure of the strength of the horizontal component of the earth’s magnetic field based on hourly readings from four mostly low-latitude observatories around the world (Love and Gannon 2009, p. 3103). As explained earlier, geomagnetic storms systematically weaken the horizontal component of the earth’s magnetic field at low latitudes. The hourly D_{st} series is not an ideal proxy for the risk to electric grids at higher latitudes—not only because they are at higher latitudes, but also because power systems are most vulnerable to magnetic field oscillations that occur over seconds or minutes, not hours. Total magnetic field depression can be small even as oscillations are large, and vice versa.

Nevertheless, the D_{st} index does broadly track magnetic storm activity on earth. And the index has the virtue of age: the World Data Center for Geomagnetism in Kyoto supplies hourly D_{st} readings back to 1957 (wdc.kugi.kyoto-u.ac.jp/dstdir)—a continuous track record three times as long as for CME speeds.

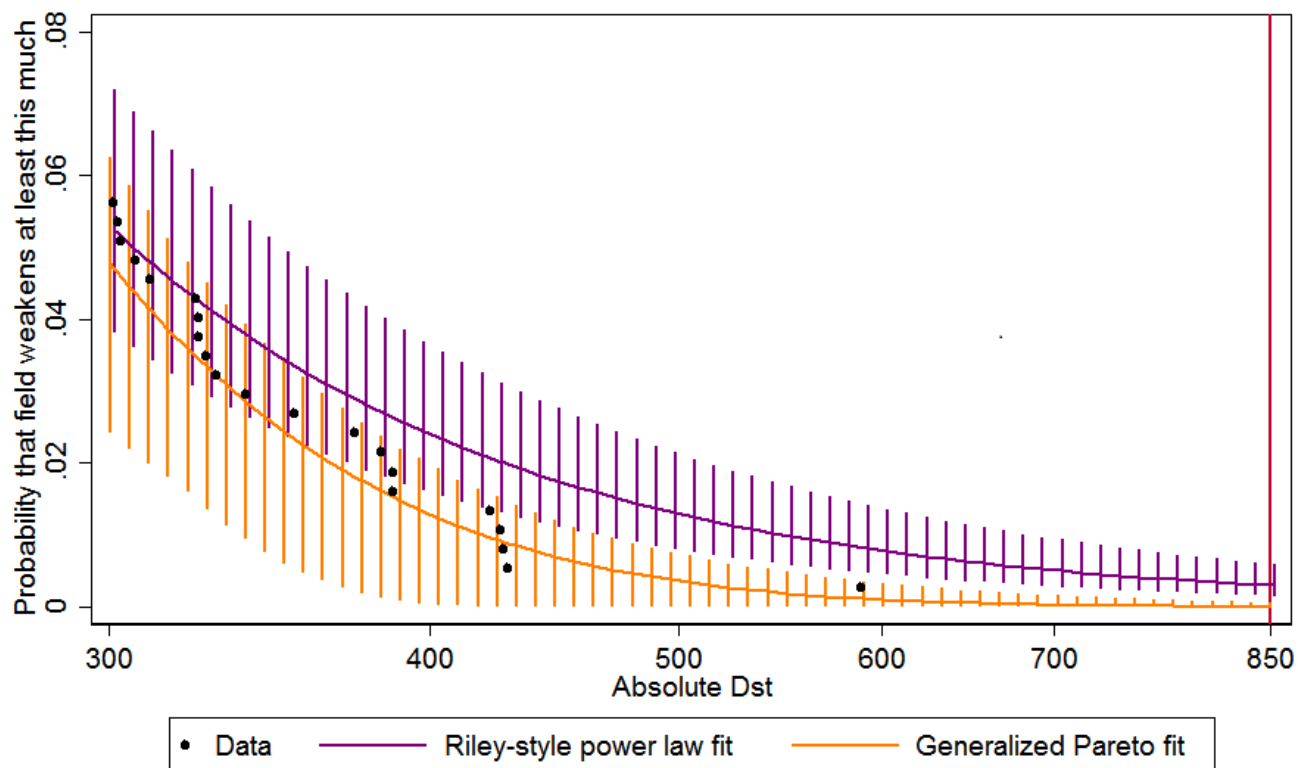
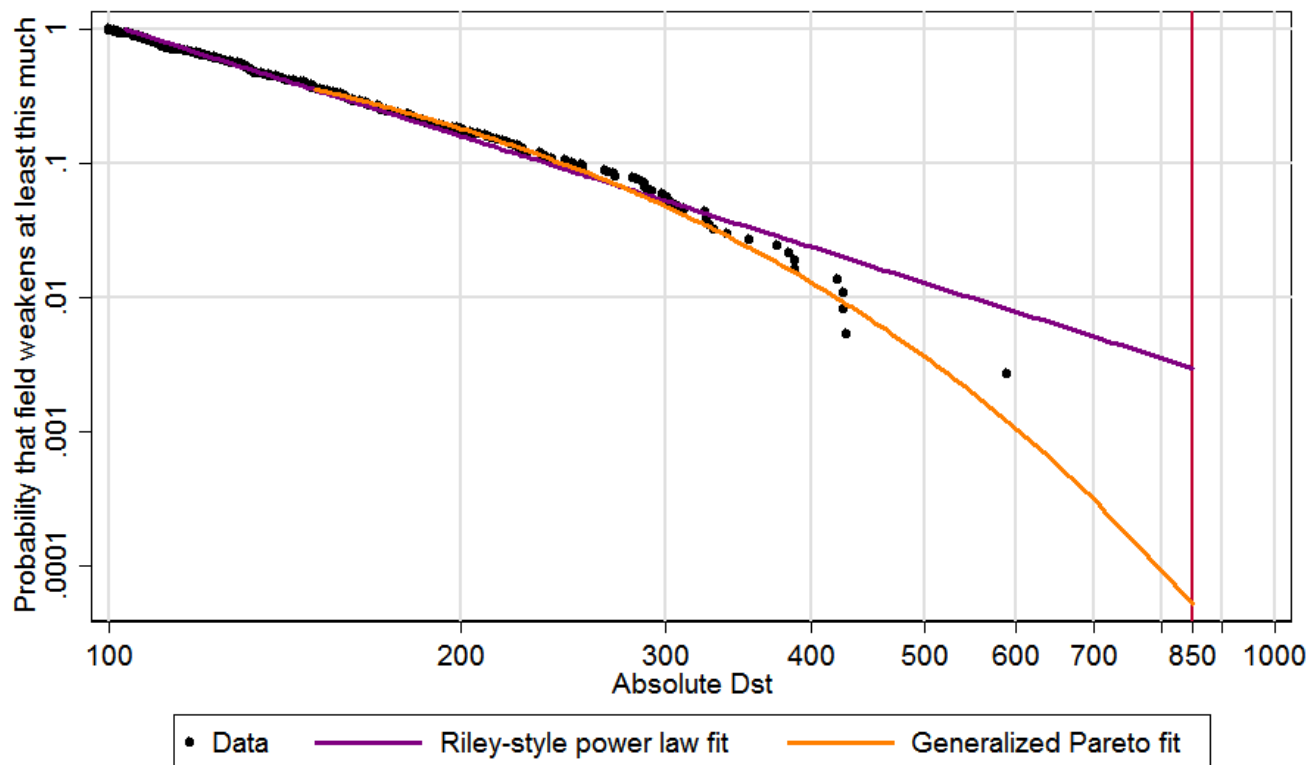
The next two graphs are analogous to the last two, but for D_{st} .²⁵ The data cover January 1, 1957–September 26, 2014. Following Riley, the unit of analysis is not the hour but the geomagnetic *event*, which is defined as one or more consecutive hours with an absolute D_{st} above 100 nanotesla.²⁶ Vertical red lines are drawn at an absolute D_{st} of 850 nT, which Siscoe, Crooker, and Clauer (2006) estimate for the Carrington storm. The rightmost dot is the 1989 storm that knocked out the Québec power grid; it registered at 589. The GP is fit to all events above 150 nT.²⁷

²⁴It is also worth noting that the SOHO speed measurements are imperfect, especially of CMEs heading away from earth. A powerful CME on July 23, 2012, reached the earth’s orbital path in 18 hours, about as fast as the Carrington CME. Its launch speed was 2000–3000 km/sec (Baker et al. 2013, p. 587). The SOHO estimate used here, 2,103 km/sec (cdaw.gsfc.nasa.gov/CME_list/UNIVERSAL/2012_07/univ2012_07.html), is on the low end of that range, suggesting that the SOHO measurement underestimated the true speed. On the other hand the deprecation of this observation is appropriate in a sense, since it was poorly observed precisely because it was not a threat to Earth.

²⁵Since the interesting changes in D_{st} reflect field weakening, its value is usually negative. Thus the references to “absolute Dst” in text.

²⁶Following Tsubouchi and Omura (2007, p. 3), if less than 48 hours separates two episodes above 100 nT, I treat this as one event.

²⁷The threshold was chosen by the method discussed in note 21.



Again, the GP-based estimates of a Carrington-sized event are lower than the power law–based ones: just 0.33%/decade (95% confidence interval 0.0–4.0%) vs. 17.6% (9.4%–31.8%).^{28,29} In the latter graph, the GP fits the extreme data better, suggesting that its lower probabilities are better extrapolations.³⁰

But the implications of the D_{st} series—more proximate than CMEs to our earthly concern with geomagnetic disturbances and covering a longer timeframe—are more worrisome. The 95% confidence interval embraces a substantial chance of another Carrington or worse. Of course earlier caveats apply here too. The most that EVT can do is assure that we extrapolate reasonably from the available data. It cannot banish the legitimate concern that even 45 years is too short a period from which to extrapolate.

Published studies of the historical record

Tsubouchi and Omura (2007), “Long-term occurrence probabilities of intense geomagnetic storm events,” *Space Weather*

Much as I do above, Tsubouchi and Omura (2007) fit the generalized Pareto distribution to the tail of the D_{st} data series, using data through 2001 or 2003, in different variations.

One difference is that despite describing how they extract *events* from the hourly data—sequences of hours or days of high readings to be treated as a single storm—the paper analyzes the data set with one observation per hour rather than per event.³¹ This effectively treats high readings in successive hours as statistically independent, sample-expanding events; my view is that they are not. That said, the thrust of Tsubouchi and Omura’s discussion is to emphasize the uncertainty of their predictions. Perhaps as a result of the seemingly larger sample, Tsubouchi and Omura cut the tail at 280 nT rather than my 150 nT.³² (The higher the cut-off the more accurate is the GP model in theory, but the smaller the sample.)

Tsubouchi and Omura’s preferred estimate is that the largest storm in the data set, the one that was measured at 589 and caused the blackout in Québec in 1989, has a *return rate* of 60 years—i.e., it was a “60-year storm.” Alternate estimates (Tsubouchi and Omura’s Table 1, rows 2–3) peg the 1989 event as a 75- or 100-year storm. Similarly, after adding data through 2014 and collapsing groups of closely spaced observations into single events, I estimate the implied return rate at 99 years.³³

But, like Tsubouchi and Omura, I should emphasize uncertainty. The 95% confidence interval for my estimated return time for 1989-scale storm is 16–605 years.

²⁸Because the event definition treats prolonged episodes as single observations, standard errors are bootstrapped without clustering.

²⁹Using the USGS version of Dst, which removes extraneous cyclical patterns (Love and Gannon 2009), yields a GP estimate for 850 of 0.001%/decade (confidence interval 0–4.1%).

³⁰Tsubouchi and Omura (2007) also fit the GP distribution to the Dst data set. When I restrict my sample to match theirs (up to 2001 or 2003) I closely match the results in the first two rows of their Table 1.

³¹I count 121 hourly observations above 280 through the end of 2001, which coalesce into 26 events. The “data” column of their Table 1 reports 121 observations and I achieve close matches to their coefficient estimates in that table when I used hourly rather than event data.

³²Like Tsubouchi and Omura, I use the mean residual life plot to select a threshold (Coles 2001, §4.3.1).

Ruzmaikin, Feynman, and Stoev (2011), *Distribution and clustering of fast coronal mass ejections, Journal of Geophysical Research*

Ruzmaikin, Feynman, and Stoev (2011) apply an innovative technique within the extreme value theory tradition (Stoev, Michailidis, and Taqqu 2006), to assess whether parts of the CME distribution obey a power law—in other words, whether the CME speed data have a section that follows a straight line on those log-log graphs.

The method is based on the following insight. Think back to the fake rainfall data set I constructed, whose observations I grouped into larger and larger blocks to show how the distribution of the maxima evolved. If, say, 100 observations out of the million are classed as extreme, as the number of blocks shrinks, the fraction of the blocks that happen to contain an extreme event will rise. (When there are half a million blocks of size two, almost none will contain an extreme event.) So the *average* maximum across all blocks will rise as the number of blocks falls and their individual size grows. Stoev, Michailidis, and Taqqu (2006) show that *if* the underlying data follow a power law, then so will the average maxima. Each doubling of the size of the blocks—e.g., taking the maximum rainfall for each fortnight instead of each week—will increase the average maximum by the same percentage. On log-log scales the graph of the average maximum with respect to block size—the “max spectrum”—will be straight. Checking for such straightness becomes a way to detect power law behavior.³⁴

Ruzmaikin, Feynman, and Stoev (2011) find that between 700 and 2000 km/sec, CMEs seem to follow a power law. You can examine the second graph in the “CME speeds” subsection above to see whether you agree that the curve is straight in that range. It is worth noting that Ruzmaikin, Feynman, and Stoev do not formally test the power law hypothesis against competing models such as the lognormal (Yurchyshyn et al. 2005), whose “max spectrum” is only slightly curved.

At any rate, the finding is of scientific interest for what it implies about the physics of solar activity, but it does not quite speak to the odds of the most dangerous CMEs, above 2000 km/sec. For the authors find that above 2000, the probabilities of a CME at a given speed drops off more rapidly than a power law would suggest, and as is evident in the CME graphs above. The authors avoid estimating probabilities of extreme events above 2000 km/sec.

³³The formula is $m = \frac{1}{\zeta_u} \left[1 + \xi \frac{x_m - \mu}{\sigma} \right]^{1/\xi}$, where N_t is the return time in years, ζ_u is the fraction of observations in the region to which the GP distribution is fit, x_m is the storm strength of 589, μ is threshold of 280, and σ and ξ are parameters determining the shape of the GP distribution (Coles 2001, eq 4.12). For Tsubouchi and Omura’s Table 1, rows 2–3, the values are $(\xi, \sigma, \zeta_u) = (.081, 45.8, 139/394464)$ and $(.031, 80.2, 45/45)$. For mine, they are $(.054, 70.2, 134/373)$.

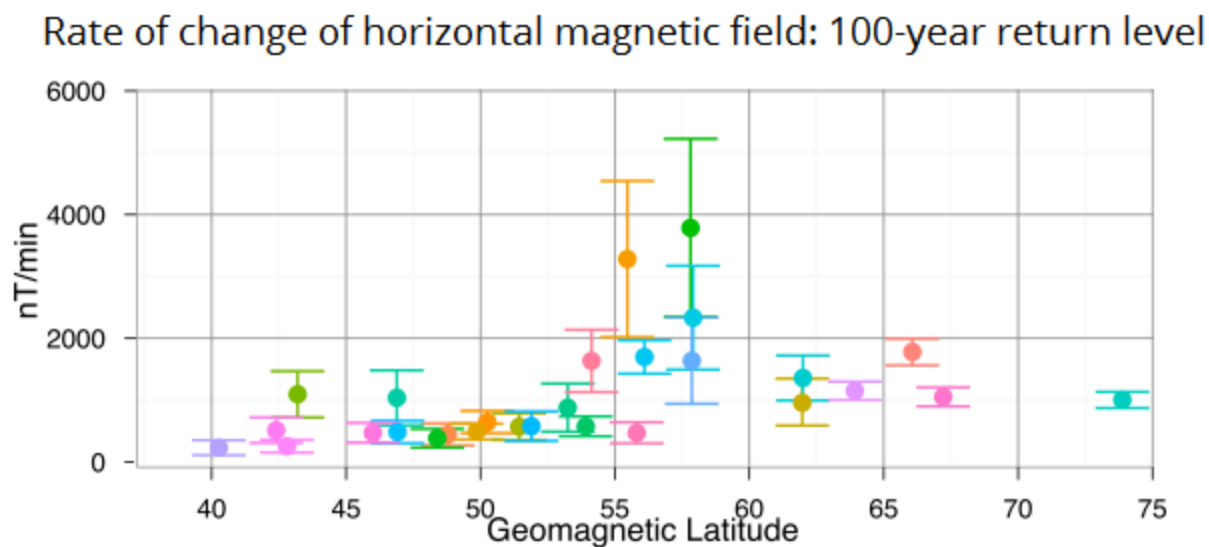
³⁴The effect is weakened if extreme events are clustered in certain time periods, which will happen if there is serial correlation their probability. In this case, fewer time blocks will gain extremes even as block size rises and block count falls. So the average maximum will not rise as fast. On the other hand, if the data are randomly reordered before the process is executed, this weakening will not occur. Doing it both ways provides a measure of the clustering of fast CMEs, a point that Ruzmaikin, Feynman, and Stoev pursue with rigor. They find significant clustering. This finding—and the clear predominance of geomagnetic storms at the equinoxes—is why in my regressions I bootstrap standard errors while clustering by half-year.

Thomson, Dawson, and Reay (2011), “Quantifying Extreme Behavior in Geomagnetic Activity,” *Space Weather*

Like Tsubouchi and Omura (2007), and as in my own analysis above, this paper harnesses the GP distribution. The difference is in the data set, which consists of minute-by-minute magnetic readings from 28 selected European observatories over recent decades. The disadvantage of this data set is its brevity: about half of the observatories began collecting data at the minute cadence after 1990, and none did before 1979 (Thomson, Dawson, and Reay 2011, fig. 2). These are short periods from which to forecast risks over 100 or 200 years, as the paper does. The advantage is that the per-minute magnetic field change is a better measure of the threat to power lines than CME speeds and hourly changes in the equatorially focused D_{st} .

The authors apply the methods well and present the results clearly, not exaggerating certainty. They analyze each observatory’s data separately, in each case taking the top 0.03% of extreme observations as the tail whose shape is approximated by the GP distribution. They also cluster data points into single events to prevent spurious statistical precision, treating extreme observations within 12 hours of each other as part of the same event.

The next graph, from Thomson, Dawson, and Reay (2011, fig. 6) shows their estimates for the 100-year “return level” for one-minute horizontal magnetic change at each observatory—that is, a level of change that would only be expected once a century. The circles show the central estimates and the vertical bars show 95% confidence intervals:



Source: Thomson, Dawson, and Reay (2011)

For reference, Ottawa and Brorfelde, two stations mentioned earlier, are both at about 55 degrees geomagnetic latitude. They experienced peaks of 556 and 1994 nT/min in 1989. The highest two estimates of the 100-year return level are between 3000 and 4000 nT/min and are for Brorfelde (orange) and Eskdalemuir, in southern Scotland (green). Factoring in the confidence intervals, these estimates are roughly 50–100% above the highest change I have found in the modern record south of 70° geomagnetic latitude (2688 at Lovo between 11:59pm and midnight on July 13, 1982). Again, the suggestion is that a worst-case extrapolation from the historical record is something twice as bad as recently experienced.

Riley (2012), “On the probability of occurrence of extreme space weather events,” *Space Weather*

Riley fits power laws to historical data on four phenomena that scientists have connected to geomagnetic storms: CME speeds; D_{st} levels; solar X-ray emissions during the flares; and nitrate deposits in terrestrial ice cores. However, the ice core relationship appears to have been firmly rebutted by Wolff et al. (2012).

Riley’s straight-line power law fits to log-log distribution plots imply probabilities of 3–12%/decade for a Carrington-scale event. However, Parrott (2014, p. 14) points out that the low number, based on ice core data, appears to be miscalculated according to Riley’s stated approach, and apparently should be 18% or perhaps 25%.

I have several reservations about the Riley extrapolations. First, confidence intervals are not reported, and they can be wide, as we have seen. (To be fair, the Riley text does emphasize the uncertainties.) Second, descriptions of methods sometimes seem incomplete or ambiguous (Parrott 2014, 2015). Third, there appear to be mathematical errors (Parrott 2014, 2015). Fourth, the true distributions may be curved everywhere when plotted log-log (e.g., they would appear parabolic if lognormal). Straight-line fits to sections where the curvature is low would serve as good approximations locally, but could lead to inaccurate conclusions when extrapolating to extremes.³⁵ In every case, the most extreme data points fall below Riley’s best-fit lines (Riley 2012, figs 2b, 4b, 8b, and 10b). This observation is compatible with the theory that the distributions curve downward and the straight-line power law fits therefore overestimate the risks predicted by history. As we saw earlier, fits drawing on the branch of statistics developed for the purpose at hand, extreme value theory, allow for this curvature while permitting weaker, thus more plausible, assumptions about the data generating process.

Love (2012), “Credible occurrence probabilities for extreme geophysical events: Earthquakes, volcanic eruptions, magnetic storms,” *Geophysical Research Letters*

Love (2012) takes an exceptionally conservative approach to assessing probabilities of extreme natural disasters, working only with data for the handful of most-extreme historical events. In the case of geomagnetic storms, he judges that three super-storms had occurred in the 153 years since Carrington: the Carrington event itself, a poorly documented storm in 1909, and the 1989 storm. Using a statistical tool called the Poisson distribution, he then asks, What probability per unit time is most compatible with observing this many “superstorms” in this many years? The answer: 17.8%/decade (95.4% confidence interval: 3.4–38.6%) (Love 2012, Table 1, row 5). Redefining “superstorm” to refer only to the Carrington event, Love’s answer becomes 6.3%/decade (95.4% confidence interval: 0.0–23.0%) (Table 1, row 9). “The 10-yr recurrence probability for a Carrington event,” he concludes, “is somewhere between vanishingly unlikely and surprisingly likely.”

Love’s approach has strengths and weaknesses. Its informal definitions of “superstorm” free him from the historical bounds of systematic recordkeeping. He can jump to 1909 and 1859 despite the lack of comparable, regular geomagnetic data back that far. The approach is also extremely conservative in throwing out all information relating to smaller storms, which is correct if the physical processes generating superstorms are distinct.

A weakness is that the choice of start time could lead to bias. Why not start at 1800? Assuming reasonable confidence that no superstorms occurred between 1800 and 1859, this would reduce the apparent per-decade

³⁵E.g., Riley (2012, p. 6), first fits a power law to the CMEs between 700 and 2000 km/sec, the range identified by Ruzmaikin, Feynman, and Stoev (2011). This fitted line is flatter than one I graphed, so it implies a higher probability of a Carrington-speed event: 85%/decade. Riley rejects this as implausible and shifts to fitting to CMEs above 2000. The shift is appropriate in my view, but does not address the underlying issue that the straight-line power law may just be a poor model for curved data distributions. (The 85% figure also appears to be an erroneous estimate under Riley’s approach (Parrott 2014, p. 9).)

probability of superstorms. Using 1859 as the start point, like using 2001 as the start point for an analysis of terrorist attack probabilities, increases the number of dramatic events in the sample while minimally extending the time frame. In addition, the conservatism comes at the cost of throwing away information about all other storms. It is reasonable to think that the frequency of not-quite-superstorms tells us something about that of superstorms, however defined. EVT gives us a systematic way to incorporate that information to the extent it has been collected.

Kataoka (2013), “Probability of occurrence of extreme magnetic storms,” *Space Weather*

Kataoka essentially applies Riley’s method to a different data series, namely the magnetic field readings collected at the Kakioka Magnetic Observatory in Japan since 1924. These readings are one input to the D_{st} index, but reach back farther than D_{st} , which begins in 1957.

From power law fits to the entire data set, subsets for particular 11-year solar cycles or particular phases, Kataoka (2013, p. 1) estimates the chance of another Carrington event at 4–6%/decade. However, just as in Riley (2012), in almost every case the most extreme geomagnetic disturbances are less frequent than the fitted power law lines predict, falling below it. This again suggests that use of the power law leads to overestimation of the risks implied by the historical record.

Summary: What is the probability per unit time of a storm at least as extreme as the Carrington event?

This section has reviewed analyses of the record of geomagnetic storms, with an eye toward extrapolating probabilities of extremes. The analyses of small data sets relating to the handful of most extreme events on record suggest that the largest known, the Carrington event, was at most twice as intense as more recent events. My preferred estimate from the more formalized analysis of larger data sets confined to more recent years remains my own from the D_{st} history: a 0.266%/decade probability of an event of at least 850 nT in magnitude with a 95% confidence interval of 0.0–8.6%/decade. This is the only estimate discussed that is based on appropriate extreme value theory methods and connects to an estimated magnitude for the Carrington event. And since that value, a D_{st} of 850 nT, is less than twice the highest recorded (in 1989), this again points to a storm twice the intensity of anything in the modern record as the outer limit for a worst-case scenario.

However, all such analyses should be treated as inputs to a threat assessment, not outputs. The historical record is short. Indeed the basis for my estimate is a data set that begins a century after Carrington; if it reached back to that remarkable day, to embrace the Carrington event too, the resulting risk estimate might well be higher. Accepting that extreme CMEs are emitted about once per decade, it is not improbable that the sun has dealt us a good hand in the last half century or so, a set of CMEs that lack the most destructive combination of speed, mass, magnetic field strength and orientation, close sequencing, and direct earth collision. Our next hand could be worse.

It appears responsible therefore to emphasize the high end of the 0.0–8.6%/decade confidence interval.

Are magnetic disturbance extremes localized?

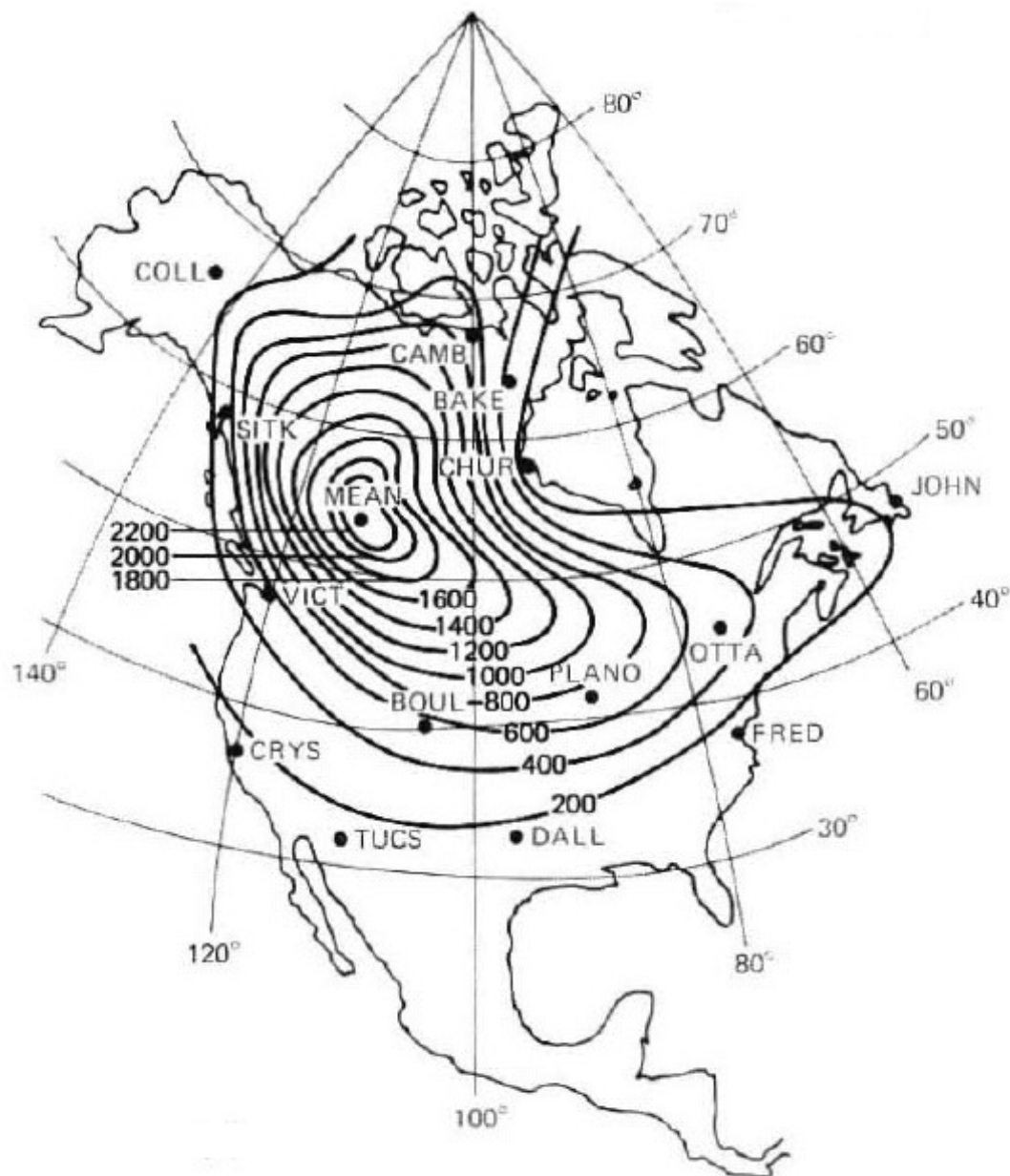
Storm strength has long been measured by global, aggregate indexes such as Dst, which hide local variation. In thinking about the great storms of 1859 and 1989, the mind is drawn to the reports of extremes, such as Victorian magnetometers whose needles went off scale. But it is one thing to observe, say, that magnetic field changes as rapid as 2200 nT/min were recorded during the 1972 storm, and another to suggest that entire national grids might someday undergo such stress. The chance of the latter depends on the spatial patterns of disturbance during geomagnetic storms. To the extent that the humanitarian threat lies in geomagnetic

disturbances over large areas, it becomes important when comparing past storms and formulating worst-case scenarios to distinguish between localized and widespread disturbance levels.³⁶

Here, I present data from two storms which suggest that the extremes are usually localized and confined to high latitudes.

The following map appears in Anderson, Lanzerotti, and MacLennan (1974). It shows their estimates of the contours of the magnetic field rate-of-change over North America between 10:41pm and 10:42pm universal time on August 4, 1972. Superimposed are curved contours, analogous to lines of constant elevation on conventional maps. In this map the highest “mountain” reaches to 2200 nT/min, and is centered over Meanook, Alberta, denoted by “MEAN.” (Other stations include Dallas (DALL), Boulder (BOUL), Ottawa (OTTA), Fredericksburg (FRED), and Tucson (TUCS).)

³⁶In a paper that cannot be shared because it is undergoing peer review, Antti Pulkkinen, Emanuel Bernabeu, and coauthors revise previous work (Pulkkinen et al. 2012) on modeling 100-year storm scenarios in part by factoring in that most places won’t experience the magnetic extremes that characterize the scenario (conversation with Pulkkinen, October 5, 2014).



In reading this map, keep in mind that its rich shapes are derived from just 14 bits of hard evidence. Across the continent at the time, only 14 magnetometers were in operation that were taking readings every minute or so—a speed needed to detect such sudden jumps—and that could later provide their data to the researchers. In mathematically deriving the contours of this transient magnetic mountain, the researchers assumed that the contours were as simple as possible while still fitting the 14 points. Most likely the sparseness of their magnetometer network missed a lot of complexity, perhaps including hotspots further south.

By a minute later, the magnetic Mount Meanook had split into twin peaks about 700 nT/min tall, in Meanook, the other to the east in Fort Churchill, northern Manitoba (Anderson, Lanzerotti, and MacLennan 1974, fig. 3b).

These observations tentatively suggest that the extremes of the 1972 storm were indeed localized in space as well as time.³⁷ Roughly, a good fraction of North America may have experienced 500 nT/min or higher, but probably not 2000 nT/min or higher. As noted earlier, a widespread flash of at least ~500 nT/min probably sufficed in 1989 to knock out the Québec grid, as well as to knock out a transformer serving the Salem nuclear plant in New Jersey (Kappenman 2010, 2-29). But widespread 500 nT/min did not to cause a humanitarian crisis then or in 1972. In 1972, the US had 12,000 miles of power lines capable of operating above 460 kilovolts (kV). By 1989, it had 28,000 miles above 400 kV and by 2012, 38,000 (EEI 1995, Table 86a; EEI 2014, Table 10.6).³⁸ So while the US grid was smaller during those storms than now, its scale was presumably still substantial enough to meaningfully test whether large-scale impulses of ~500 nT/min pose a catastrophic risk—and to produce reassuring evidence about the 500 nT/min disturbance level.

The second data set, along with my analysis thereof, is inspired by the first. The IMAGE network, started in 1982, consists now of 33 magnetic observatories across Scandinavia, whose digital instruments record the local magnetic force field at a typical cadence of 10 seconds:

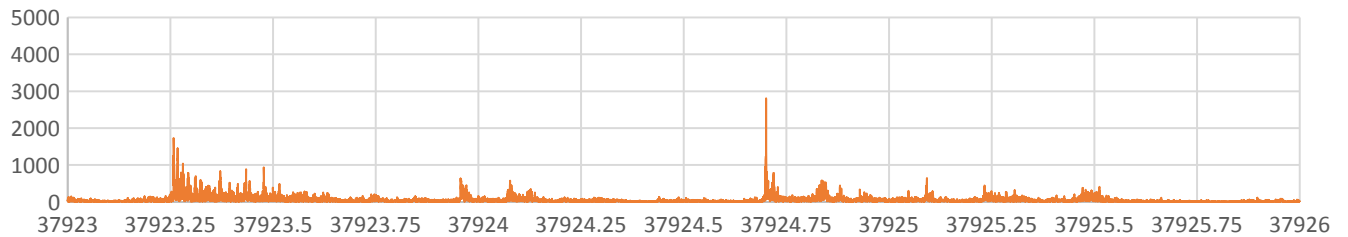


³⁷This reading of the map is in tension with Kappenman's (2006, p. 192): "This disturbance was estimated to be at an intensity of ~2200 nT/min over extensive portions of North America (Anderson et al., 1974)."

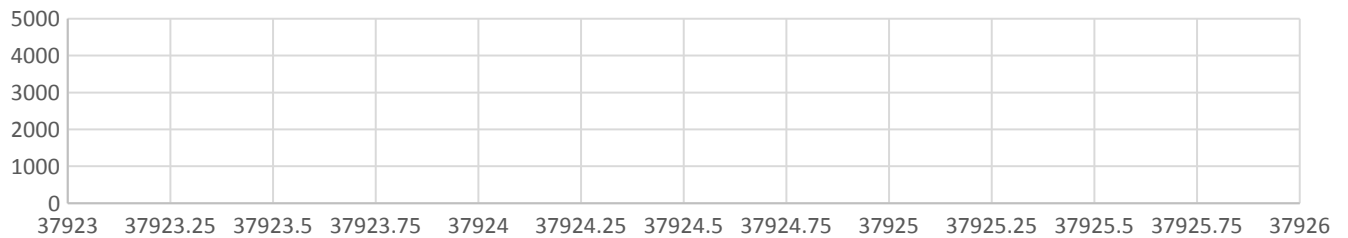
³⁸EEI revised its mileage brackets for reporting power line mileage in 1978, so statistics from before and after that year are not perfectly comparable.

Since the IMAGE network had many more stations by the 2003 Halloween storm than it did during the 1989 one —26 vs. 7—I took data for the 2003 storm. These graphs plot absolute horizontal magnetic changes at four representative stations during October 29–31, 2003, ordered from north to south. Changes are computed for each 10-second interval but expressed in nanotesla per minute:

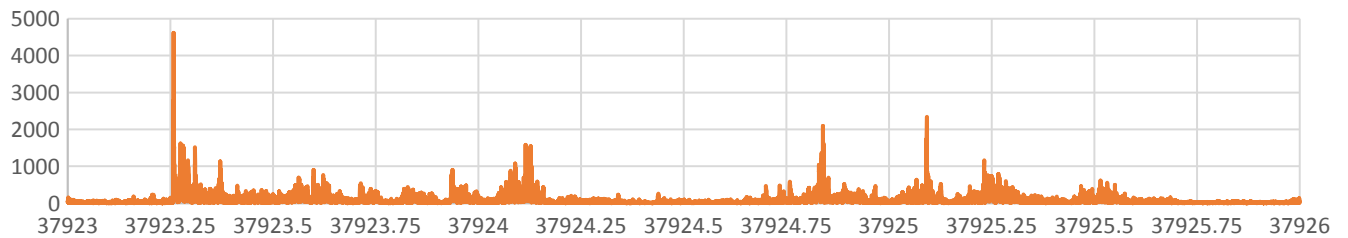
HOP (Hopen Island, well north of Scandinavia)



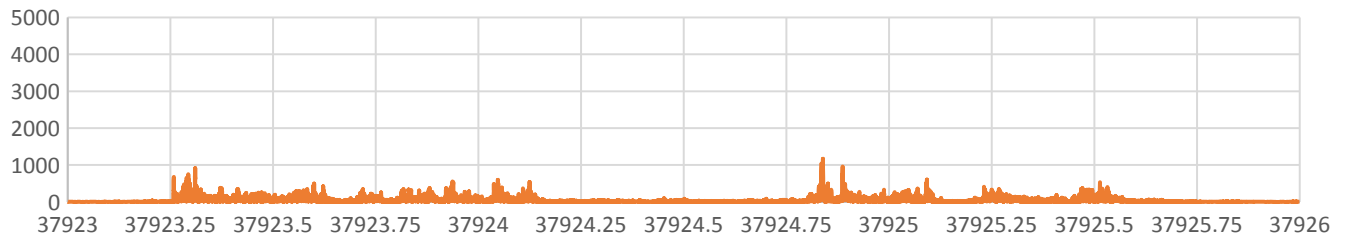
SOD (Sodankylä, northern Finland)



RVK (Rørvik, central Norway)



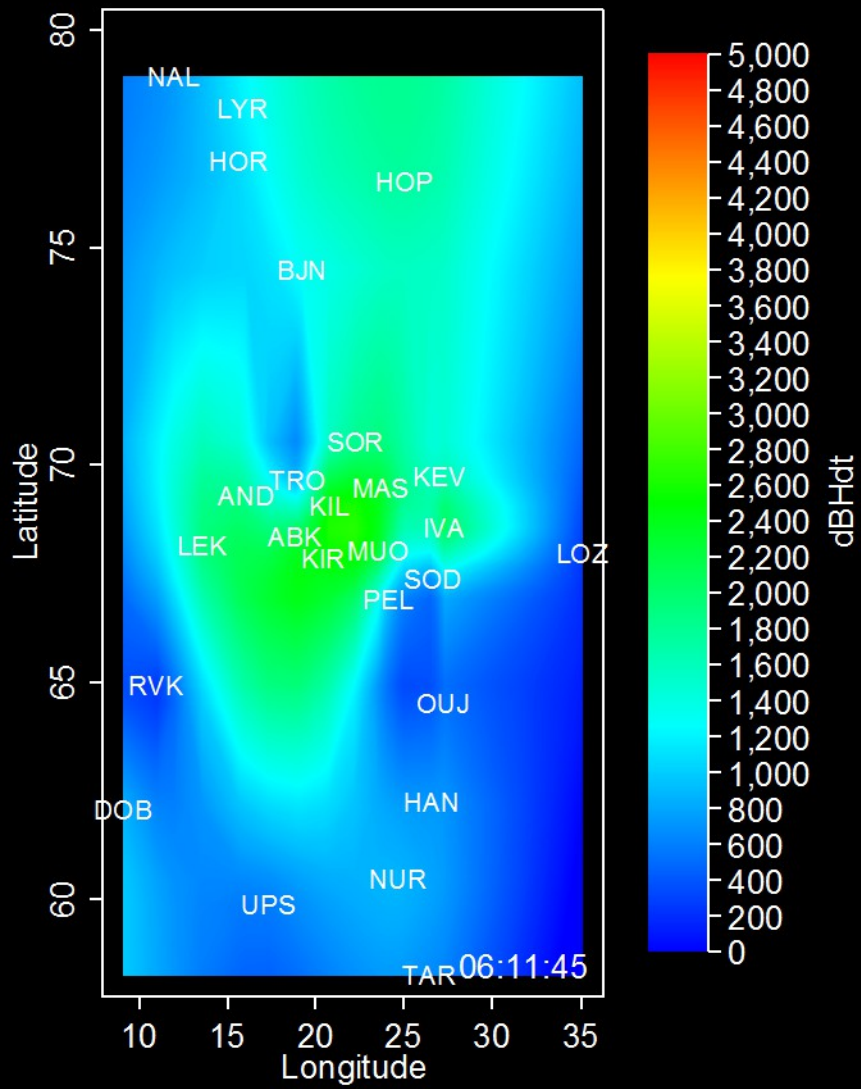
UPS (Upsala, southern Sweden)

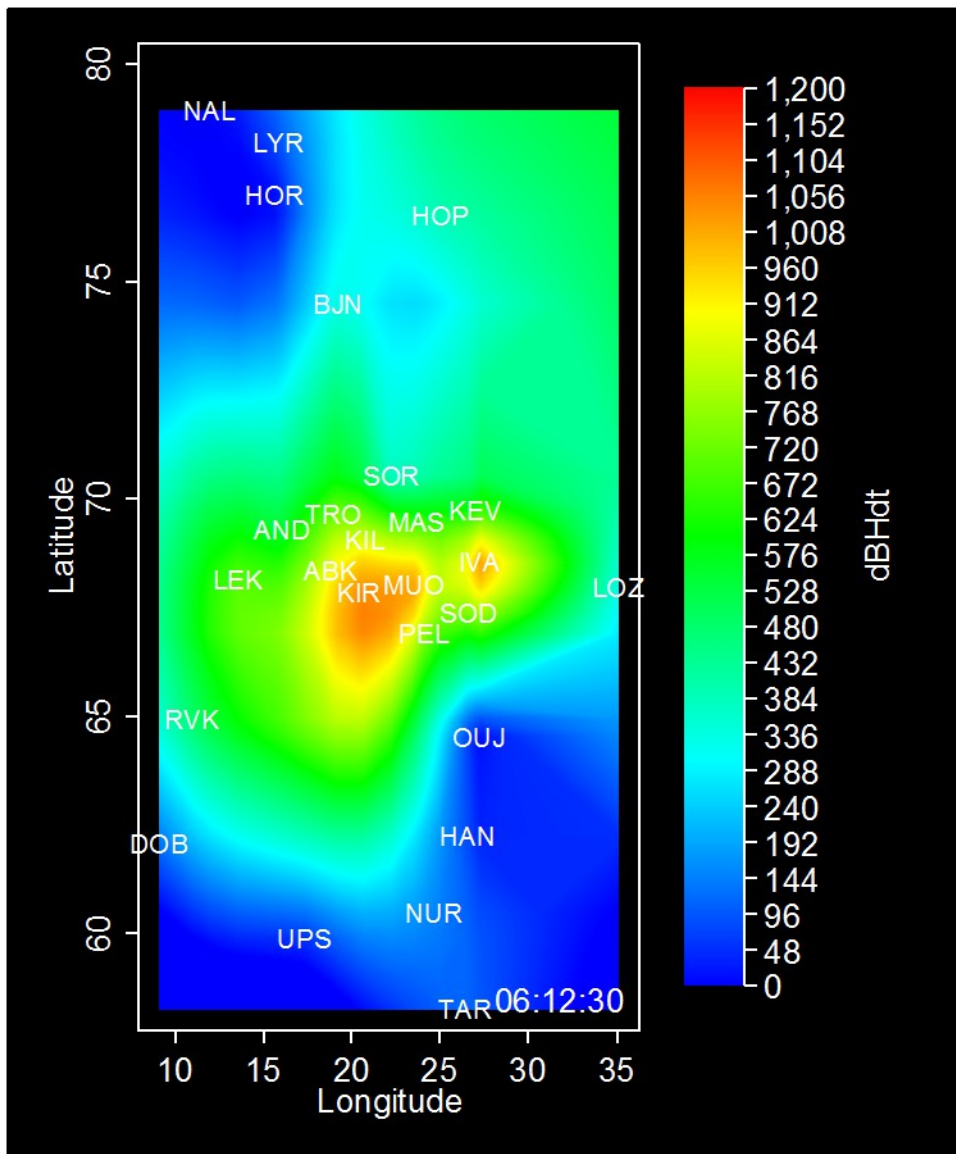


The graphs show strong spatial correlations across the region; that is, to a substantial degree the magnetic field varied at all the stations in concert. This makes sense since the electrojets driving the fluctuations are planetary in scale. But the locations differed greatly in the magnitude of the changes. Just as in North America in 1972, the effects were much lower to the south—here, represented by Uppsala (last graph).

Computing change over 60 instead of 10 seconds, for greater comparability with the 1972 figures, damps the biggest spikes (not shown). The largest rate of change observed falls from 4618 nT/min (at RVK, above) to 1612 (at LOZ, the Russian observatory Lovozero). Either way, the biggest individual changes do not represent the experience of the region as a whole.

To visualize the spatial dimensions of the data, I also computed contours, as shown below. Imagine these graphs superimposed on the map of Scandinavia above. Covering an area roughly 800 miles east-west and 1100 north-south, they capture the initial concussion of the Halloween CME with Scandinavia, on October 29, 2003, at 6:11 universal time. The first contour plot shows the absolute horizontal magnetic field changes between 06:11:40 and 06:11:50, expressed in nT/min. The second shows the total change between 6:12 and 6:13, producing some lower values by smoothing over that longer interval. Notice that the colorings in the two plots are on different scales. Especially the second shows magnetic field changes of at least 500 nT/min occurring over a large area, and being sustained for at least a minute.





Animated versions of these plots are at posted in [10-second](#) and [1-minute-change](#) versions. Each covers about five minutes 10 times faster than real time. Another pair of videos does the same for a period of high activity on the evening of 30 October, also visible in the time series graphs above ([here](#), [here](#)).

We find that geomagnetic storms are phenomena of turbulence, which gives them a fractal quality. The higher the resolution with which we measure them, in time or space, the more complexity and the more potential for extreme points. Turning that around, the more we average over time or space, the lower the extremes we will perceive. The means that in citing extreme values for storms, we need to take care to mention the temporal and spatial coverage of the statistic, and to compare like to like.

In the 1972 and 2003 storms, field changes on the order of 500 nT appear to have occurred over large areas over the time scale of a minute. The practical question then becomes whether some multiple of that observed value, thought to correspond to a reasonable worst case, would cripple the grid. That is an engineering question, which the next section approaches.

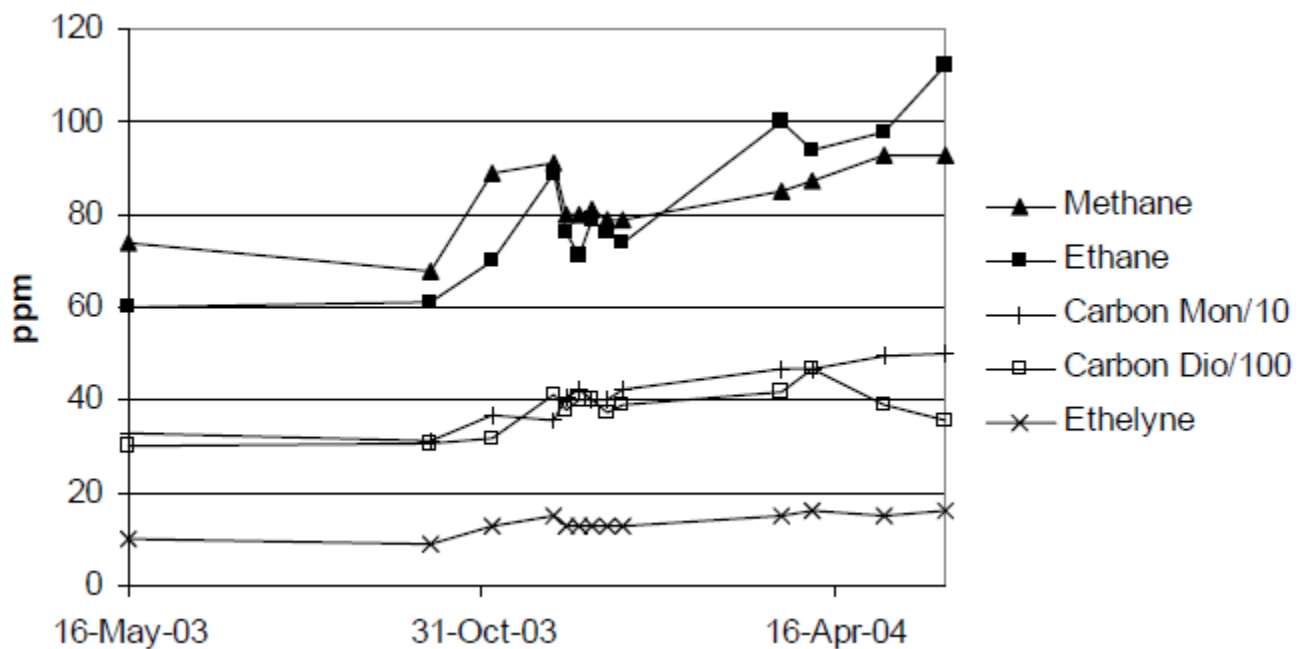
How vulnerable are high-voltage transformers to geomagnetically induced currents?

The discussion to this point has stretched the imagination with sun spots the size of the earth, coronal mass hurtling through space at 1% of the speed of light, and planetary magnetic concussions. But at some point the analysis must literally come down to earth, and when it does, it brings us to objects on a more human scale, including power lines and transformers merely the size of houses.

Yet if power lines and transformers lines are easier to contemplate, their behavior during storms is not necessarily better understood. One challenge for science is that to test a high-voltage transformer under realistic conditions, you need to channel enough power through it to serve a small city. Another is that transformers are diverse, in design, materials, age, quality of manufacture, and operating conditions, all of which impedes generalization. And they are, rather literally, black boxes. Still another barrier to science is that utilities and transformer manufactures probably know more than they say about the nature and consequences of the geomagnetic currents flowing through their systems. Unlike airlines, which must share black box data recorders, utilities are under little compulsion to share such information, despite its relevance to public safety (Kappenman 2012a, p. 9).

I am aware of four lines of analysis on the impacts of storms on grids. First, there are scattered reports of transformer failures—such as in the U.K., US, South Africa, Sweden, and New Zealand. (Girgis and Vedante 2012, pp. 5–6, succinctly list reports while arguing against geomagnetically induced currents as the cause.) And the effects are probably more widespread than the public reports suggest. Kappenman (2012a) testified that after 1989, “The only US electric power company to openly report transformer impacts...(Allegheny Power), reported that deleterious impacts caus[ed] loss-of-life to transformer insulation on 36% of their [high-voltage] transformer infrastructure.”

Particularly important in this vein is the report by Gaunt and Coetzee (2007) on the gradual deterioration of eight transformers in South Africa after the Halloween 2003 storm. Notably, South Africa had generally been seen as protected from geomagnetic storms by its low latitude. (Moodley and Gaunt 2012 refines the story of one of them, suggesting to a link to 2001 activity as well.) As Gaunt and Coetzee explain, high-voltage transformers are bathed in oil for cooling. As the oil circulates, it carries away heat. When magnetic cores saturate and the field strays outside the core, wires and insulation can overheat and in effect burn. Hot spots manifest not as a flames (one hopes) but as chemical decomposition that forces gaseous byproducts into the oil. This graph (Gaunt and Coetzee 2007, fig 5) shows parts per million of dissolved gases at one of the ill-fated transformers over the course of 12 months, at the Matimba power plant (carma.org/plant/detail/27542):



The gas content began to creep upward after the Halloween storm. Gaunt and Coetzee infer that the transformers were more vulnerable to the storm's geomagnetically induced currents precisely because those currents were too weak to trip protective switches. Perhaps during the storm paper insulation around some of the wires was damaged, leading to short circuits and, in time, additional overheating. Or perhaps some other positive feedback mechanism operates. And Gaunt and Coetzee suggest that GICs disable more transformers than is commonly appreciated, the origin of the harm masked by its slow progression.

The second line of evidence on transformer vulnerability is statistical. As noted in the background section above on "geomagnetic storms and power grids," there is a highly suggestive correlation between magnetic storms and transformer failures in the US between 1980 and 1994. Gaunt and Coetzee (2007, fig. 10) show evidence of the same in South Africa between 1984 and 2002. Geomagnetic storms appear to be a major cause of transformer failures, if with lags of months or years.

Third, some engineers have experimentally run large direct currents through high-voltage transformers not designed to accommodate them. The resulting literature is the one I understand least well of all those central to this inquiry.

In the early 1990s, no doubt inspired by the 1989 storm, engineers for Tokyo Electric Power, Toshiba, Hitachi, and Mitsubishi ran heavy currents through scale models of various transformer types. The largest model was of the type that theory and smaller-scale tests indicated was most vulnerable. When receiving 200 amperes of current, this model's temperature rose 110 degrees Celsius over 20–30 minutes, and then stabilized. "The magnitude of this temperature rise corresponds to a fairly large [geomagnetically induced current] level, and since the frequency of occurrence of this level is low and the duration of the effect is short, the effects on transformers from the viewpoint of lifetime reduction are sufficiently small" (Takasu et al. 1994, p. 1177). Still, this experiment took place in a laboratory, not the field. One aspect of the unrealism was the lack of any normal AC transmission load in addition to the DC test current (Takasu et al. 1994, p. 1182, in response to query from reviewer Kappenman).

In Finland, Lahtinen and Elovaara (2002) tested a new, full-scale model of transformer commissioned by the Finnish electric grid operator Finngrid. They too poured large DC currents into a transformer that was carrying no

normal AC load. They too were reassured. “The field tests...show that tested transformers may tolerate very high dc currents for times typical to geomagnetic storms.” In Québec, Picher et al. (1997) comparably tested two production transformers and found comparable reassurance.

The fourth line of evidence is theoretical. Analysts have built computer models to simulate the flow of GICs into transformers. Leading proponents of this approach are Ramsis Girgis and Keran Vedante of ABB, which through acquisitions has become the manufacturer of most high-voltage transformers operating in the US. Girgis and Vedante (2012) argue emphatically that GICs would do no permanent harm to the “great majority” of transformers. Why? GICs affect transformers in two main ways. They distort the currents going in and out of the devices. And they force magnetic field lines out of the core, which can lead to overheating. The second effect, argue the ABB engineers, takes places over minutes, perhaps even hours, while the first is nearly instantaneous—and should trip protective equipment that will shut down the transformer before permanent harm is done. If a large storm shuts down many transformers, a network-wide voltage drop or blackout may ensue. But in that fragility lies resilience: after the storm, the grid can be switched back on.

Yes, concede Girgis and Vedante (2012, §IV), some transformers have failed. But they were of old designs, or poorly constructed.

A problem with the Girgis and Vedante analysis is that it is not open. The paper provides little detail on how its model was specified and tested against empirical data (SAC 2011b). Queries on my part have yielded nothing additional. Evidently the model is proprietary. In effect, the reader is asked to trust the manufacturer’s reassurances about the reliability of its products. Also perplexing is the response to the empirical evidence from transformers operating in the field such as in Gaunt and Coetzee (2007). “The magnitudes of the GIC reported to have been associated with these incidents do not seem to be high enough to cause much winding overheating” (Girgis and Vedante 2012, p. 6). In other words, it seems, the ABB model is not compatible with the evidence from South Africa. This contradiction does not seem to have been explained.

These four lines of evidence combine to produce a strange state of knowledge. Theory and pure DC tests reassure. Yet statistical correlations and evidence from specific cases such as in South Africa suggest that utilities underestimate the rate of storm damage to the transformer fleet.

Further complicating the problematique is the complex behavior of the grid itself when stressed. Running a grid is a high-wire act in more than one sense. Sources and sinks for energy must be balanced. The waveforms of voltages must be finely calibrated across long distances to avoid damaging equipment. Systems such as capacitor banks and static VAR compensators work automatically to maintain the balance. When they are overwhelmed, components of the grid can shut down or disconnect within seconds, cascading disruption. The Québec 1989 example shows that an unplanned shutdown can itself harm transformers. Perhaps the permanent damage can be expected to be limited, so that the lights will soon come back on. How secure we should feel about this is itself a matter for theoretical and empirical research on the responses of grids to geomagnetic storms.

I draw two conclusions—or infer two hypotheses for further exploration. First, when it comes to geomagnetic disturbances, moderate tempests may do more harm over months than massive blasts do over minutes. (Of course, the modest tempests and massive blasts can be the same storms, experienced at different latitudes.) This hypothesis reconciles the year-to-year correlations between geomagnetic disturbances and transformer failures with the apparent loss of only a handful of transformers during the 1989 storm, the biggest of the postwar era. This idea has complex implications for the effects of magnetic superstorms. A really big storm *might* cause hundreds of transformers to fail—perhaps over the following year, perhaps predominately at middle latitudes. The resulting exhaustion of spares and manufacturing capacity *could* be a slow-motion train wreck. Or perhaps,

as may have already happened without anyone realizing it, the degradation will be so gradual as to create no humanitarian risk.

The second conclusion is practical. The empirical questions relating to transformers and grids are easier to answer than most in the social sciences or, for that matter, in astrophysics, because experiments can be run. With money, much more could be learned about the vulnerability of transformers and grids. The effort should be well be worth it, for it would help us better understand a potential Achilles heel of modern civilization.

Conclusion

The humanitarian threat from geomagnetic storm is a subject of vast complexity, slicing across many disciplines. This review by one newcomer cannot do it full justice. However, I think this preliminary analysis suffices to justify more attention to the issue.

Some summary points:

- Some prominent, high-side estimates of the risk appear to be unrepresentative extrapolations from the historical record. Riley's (2012) estimates derive from what are, in log-log space, straight-line extrapolations of a curved distribution. Kappenman's (2004, 2010) multiplier of 10 for a worst case relative to recent experience appears to depend on a comparison of a local extreme in one storm to a widespread disturbance in another. Comparisons of major storms since Carrington based on available data suggest that it was at most twice as strong as any storm in the modern area. And a doubling does not seem worrisome, since past storms have caused little real suffering.
- Nevertheless, the historical record is limited. Major coronal mass ejection events occur about once per 11-year solar cycle. Modern grids have been in place for only a handful of decades. High-frequency magnetometry only began in the 1980s. And several effectively random traits of CMEs—speed, breadth, direction, mass, sequencing, angle and strength of magnetic field—combine multiplicatively to determine the impact on earth. This causal structure gives rise to a right-skewed or fat-tailed distribution for storm impact, whose profile is particularly difficult to estimate from small samples.
- Grids and their management evolve, both in ways that make them more vulnerable and ways that make them less. The storm of 1989 appears to have led to more space weather monitoring, and to spending on making the grid more robust and resilient to geomagnetic disruption.³⁹ On the other hand, grids have become more interconnected, and their management more dependent on satellites that are also vulnerable to CMEs.
- More generally, the response of grid components and grids as systems to geomagnetic storms appears to be a prime subject for research. Much more can be learned—and put in the public domain—that would be of direct use in minimizing this threat. Because grids are complex systems, they may respond to rising storm strength in highly nonlinear ways. This layers yet more uncertainty on our assessment of the risk—but it is an uncertainty that can be reduced.
- Attention should be paid to how utilities and their regulators are responding to the threat, and opportunities should be sought for corrective advocacy, where needed.

³⁹Sébastien Guillon, Hydro-Québec - TransÉnergie, Montréal, personal communication, November 14, 2014, with regard to Québec.

Sources

Document	Source
Albertson et al. 1973	Source
Anderson, Lanzerotti, and MacLennan 1974	Source
Artelius 1982	Source
Baker et al. 2013	Source
Britannica 1943	Source
Cliver et al. 2014	Source
Cliver and Dietrich 2013	Source
Cliver and Svalgaard 2004	Source (ungated)
Coles 2001	Source
EEl 1995	Source
EEl 2014	Source
Einstein 1905	Source
Elovaara et al. 1992	Source
Foundation for Resilient Societies 2011	Source (archive)
Gaunt and Coetzee 2007	Source
Giles, Feng, and Godwin 2011	Source (archive)
Girgis and Vedante 2012	Source
Gopalswamy 2006	Source
Gosling 1993	Source
Green 2008	Archive
Hoyt and Schatten 1998	Source
Kappenman 2004	Source (archive)
Kappenman 2005	Source
Kappenman 2006	Source
Kappenman 2008	Source
Kappenman 2010	Source
Kappenman 2012a	Source (archive)
Kappenman 2012b	Source
Kataoka 2013	Source
Lahtinen and Elovaara 2002	Source
Lakhina et al. 2005	Source
Liu et al. 2014	Source (ungated)
Lockwood et al. 2011	Source
Love 2012	Source
Love and Gannon 2009	Source , data
McCracken 2001	Source (archive)
Miyake et al. 2012	Source
Miyake, Masuda, and Nakamura 2013	Source
Moodley and Gaunt 2012	Source

NASA 2008	Source (archive)
NASA 2014	Source (archive)
NERC 1990	Source
NERC 2010	Source
NERC 2012	Source
NERC 2014	Source (archive)
Neuhäuser and Hambaryan 2014	Source
Newman 2006	Source (ungated)
Ngwira et al. 2013a	Source
Ngwira et al. 2013b	Source
NOAA 2004	Source (archive)
NRC 2008	Source
Otterberg 1982	Source
Parrott 2014	Source (archive)
Parrott 2015	Source (archive)
Picher et al. 1997	Source
Pulkkinen et al. 2012	Source
Rasson 2005	Source (archive)
Riley et al. 2006	Source (archive)
Riley 2012	Source
Ruzmaikin, Feynman, and Stoev 2011	Source
SAC 2011a	["An Analysis of the Equipment Vulnerability from Severe Geomagnetic Storms"]
SAC 2011b	["Technical Comments and Concerns on Draft Chapter 10 by Author Girgis"]
SAC 2013	Source (archive)
Sanders 1961	Source
Silverman 2006	Source
Siscoe, Crooker, and Clauer 2006	Source
Stenquist 1925	Source
Stoev, Michailidis, and Taquu 2006	Source
Takasu et al. 1992	Source
Thomson, Dawson, and Reay 2011	Source
Tsubouchi and Omura 2007	Source
Tsurutani et al. 2003	Source
USITC 2011	Source
Usokin et al. 2013	Source
Washington Post 2014	Source (archive)
Wolff et al. 2012	Source
Yurchyshyn et al. 2005	Source (ungated)