

Los Alamos National Laboratory

Review of the GMD Benchmark Event in TPL-007-1

Michael Rivera and Scott Backhaus

September 2015

This work was funded by the Energy Systems Predictive Capability program in the Office of Electricity Delivery and Energy Reliability of the U.S. Department of Energy under Contract DE-AC52-06NA25396 with Los Alamos National Laboratory.

Contents

Overview	3
Estimating the 100-Year Exceedance Geo-electric Field Magnitude	3
IMAGE magnetometer data and conversion to geo-electric field	3
Definition of an Independent GMD Event	4
Statistics of the Independent Peak Geo-Electric Field Events	4
Extrapolation of the PDFs and Calculation of Exceedance Geo-Electric Field Magnitude	5
Summary of Differences with TPL-007-1	6
Recommendation.....	7
Latitude Scaling of Geo-electric Field for the GMD Benchmark Event	8
Electrojet Properties--Qualitative Discussion	8
Current TPL-007-1 Latitude Scaling of the Benchmark Event.....	8
Reanalysis of the Magnetometer Data Conditioned On the Dst Index	9
Recommendation.....	12
Effects of Earth Conductivity on Uncertainty and Scaling of Geo-electric Calculations	13
Process for Estimating the Geo-Electric Field for the Standard GMD Event in Different Regions	13
Earth Conductivity (Resistivity) Data	14
Effect of Uncertainty In the Earth Conductivity Data On the Peak Geo-Electric Field	14
Recommendation.....	16
Appendix	17
Charge to the reviewers.....	17
Reviewer comments and author response.....	18
Internal Reviewer #1: Reinhard Hans Walter Friedel (Reiner), Los Alamos National Laboratory	18
Internal Reviewer #2: Earl Christopher Lawrence, Los Alamos National Laboratory	20
External Reviewer #1: Pascal Van Hentenryck, Brown University.....	21
External Reviewer #2: Ian Hiskens, University of Michigan.....	24

OVERVIEW

Los Alamos National Laboratory (LANL) examined the approaches suggested in NERC Standard TPL-007-1 for defining the geo-electric field for the Benchmark Geomagnetic Disturbance (GMD) Event. Specifically,

1. Estimating 100-year exceedance geo-electric field magnitude,
2. The scaling of the GMD Benchmark Event to geomagnetic latitudes below 60 degrees north, and
3. The effect of uncertainties in earth conductivity data on the conversion from geomagnetic field to geo-electric field.

The body of this document summarizes the review and presents recommendations for consideration.

The Appendix contains the comments from four independent reviewers of this work and our responses to these comments, including a discussion of any changes made to the content of the body of the document.

ESTIMATING THE 100-YEAR EXCEEDANCE GEO-ELECTRIC FIELD MAGNITUDE

A clear and rigorous definition of the recurrence time and exceedance geo-electric field is critical to the screening and risk assessment process for geomagnetic disturbance (GMD). Using historical data from the International Monitor for Auroral Geomagnetic Effects (IMAGE) magnetometer array, we describe an analysis that addresses the definition of an event, the geo-electric field of that event, and the recurrence time of these fields.

IMAGE MAGNETOMETER DATA AND CONVERSION TO GEO-ELECTRIC FIELD

The IMAGE magnetometer data¹ covers portions of northern Europe and Scandinavia and is available at a 10-second time resolution. Many of the magnetometers have relatively contiguous data sets stretching over 25 years. There are occasional breaks in the magnetometer data because of maintenance or temporary device failure. We linearly interpolate over small time gaps of less than 20 minutes. Data gaps longer than 20 minutes are not interpolated. Instead, we separate the time series into smaller subsets and analyze only subsets longer than two days. Shorter data subsets are excluded from the analysis. There is no other filtering or preprocessing of the data. By excluding gaps larger than 20 minutes, the interpolated points are less than one percent of the available data.

The resulting magnetic field time series from the IMAGE data is converted to a geo-electric field using a one-dimensional plane wave model² and the Quebec earth model.³ Although there is some question concerning the validity of generating an electric field using an earth model from a single location with magnetometer data measured at a different locations, we accept this as a valid technique. Using these data and this approach, we generate up to 20-year geo-electric field time series for each IMAGE magnetometer station.

¹ Tanskanen, E.I. (2009), A comprehensive high-throughput analysis of substorms observed by IMAGE magnetometer network: Years 1993-2003 examined, *J. Geophys. Res.*, 114, A05204, doi:10.1029/2008JA013682.

² Wait, J. R. (1953). Propagation of Radio Waves Over a Stratified Ground. *Geophysics* 18(2), 416-422.

³ NERC Benchmark Geomagnetic Disturbance Event Description, December 5, 2014, pg. 19.

DEFINITION OF AN INDEPENDENT GMD EVENT

The physics of the electrojet during a GMD event creates time correlations in the resulting geo-magnetic and geo-electric fields. In generating geo-electric field magnitude statistics from the time series data, it is important to consider these correlations when defining a GMD event. Figure 1 displays the autocorrelation for the geo-electric time series generated using magnetometer data from the SIT station in the IMAGE data set. Analysis of the trace shown in Figure 1 yields a decorrelation time, τ , of the geo-electric field of about 2 minutes.

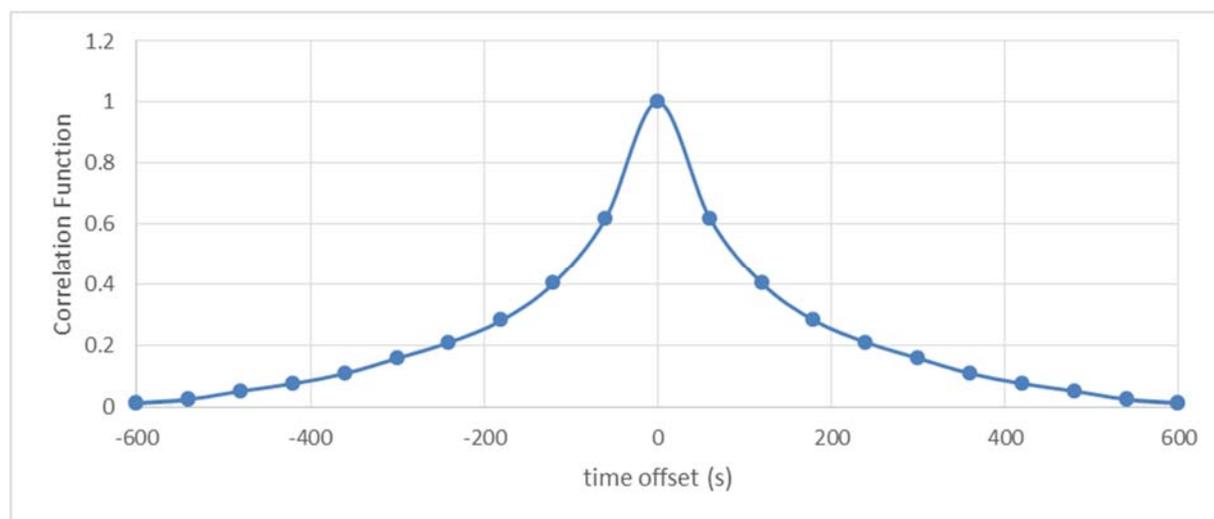


Figure 1. Autocorrelation function of the geo-electric field generated from station SIT's time series data using the plane wave model and Quebec earth conductivity model

At time leads or lags of less than 2 minutes, samples of the geo-electric field magnitude cannot be considered to be independent. Therefore, we define independent geo-electric events as maximum values in the geo-electric field magnitude over a 2-minute lead or lag. That is to say, an event occurs at time t if and only if $|E(t)| > |E(t + \tau)|$ for all τ such that $0 < |\tau| \leq 2min$.

This definition is in contrast to the approach taken in TPL-007-1, where each 10-second sample was considered as an independent event. The current approach in TPL-007-1 results in over-counting of samples in the ramp of the geo-electric field up to a maximum and during the ramp down from the maximum. The over-counting of samples injects a bias into geo-electric field magnitude statistics by adding more counts at geo-electric field magnitudes below the independent peak magnitudes, $|E^{peak}|$, resulting in an artificially faster decline of the event rates. This bias results in an underestimation of the rate of high-magnitude geo-electric field events, i.e., those magnitudes beyond the largest values in the analyzed time series data, which are the relevant field magnitudes in the risk screening process of TPL-007-1.

STATISTICS OF THE INDEPENDENT PEAK GEO-ELECTRIC FIELD EVENTS

Using the definition of independent geo-electric field events discussed above and the 20-year time series data from the IMAGE magnetometer stations, we obtain the set of peak electric field magnitudes that occur during each event (defined as $|E^{peak}|$). That is to say if independent event i (as defined

above) occurs at time t_i then $|E^{peak}|_i = |E(t_i)|$. From the set of event magnitudes, we generate the probability distribution function (PDF), $f(|E^{peak}|)$ for all the events for a given station. Here, we emphasize that we are computing a PDF of field magnitudes, rather than a histogram of the number of events (as was done in TPL-007-1) because the choice of data binning in the histogram affects the scaling and shape of the histogram. The scaling and shape, in turn, affect the extrapolation of the histogram to large geo-electric field magnitudes beyond the largest in the time series data, i.e., to field magnitudes relevant for the risk screening process in TPL-007-1. Figure 2 shows the PDFs generated by normalizing the histogram for total number of events and bin size for the independent geo-electric field amplitude events for a subset of the IMAGE magnetometer stations. We emphasize that the integral of these PDFs over field magnitude is equal to one and each of these PDFs gives the probability of the observation of a single, independent peak geo-electric field magnitude falling between $|E^{peak}|$ and $|E^{peak}| + d|E^{peak}|$, i.e., $f(|E^{peak}|)d|E^{peak}|$. In contrast to the analysis in TPL-007-1, this representation of the geo-electric field magnitude data is independent of the choice of data binning.

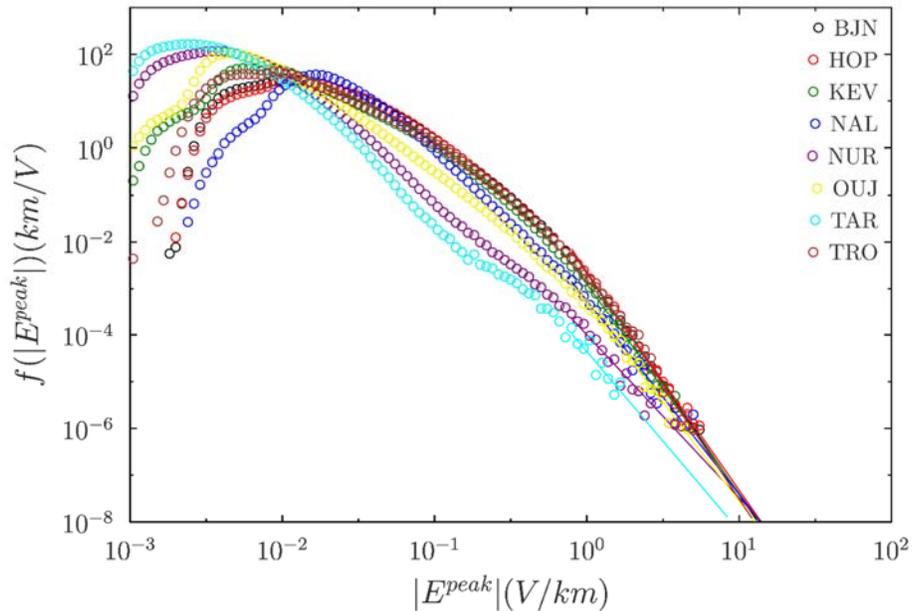


Figure 2. PDF for independent peak geo-electric field magnitude observations for a subset of the IMAGE magnetometer stations. The solid lines are the extrapolations using power law fits to the extreme events.

EXTRAPOLATION OF THE PDFS AND CALCULATION OF EXCEEDANCE GEO-ELECTRIC FIELD MAGNITUDE

The exceedance value for the independent peak geo-electric field observations is known to be larger than the largest geo-electric field observed in the IMAGE time series data; therefore, we are forced to extrapolate the PDFs (Figure 2). These extrapolations are done with power law fits to the larger values of $|E^{peak}|$ and are displayed as solid lines in Figure 2. We note here that these extrapolations assume that the space weather physics for the larger (currently unobserved) events are the same as for the smaller events in the IMAGE data. In the analysis described here, we adopt this assumption but note that the differences in the shapes of the PDFs in Figure 2 from station to station may be an indication of just this sort of change in the physics.

Using the power law extrapolations of the PDFs in Figure 2, the exceedance value E^* for observation time T for independent samples of the peak geo-electric field magnitude $|E^{peak}|$ is given by solving

$$\int_{E^*}^{\infty} f(|E^{peak}|) dE^{peak} = \frac{\tau}{T}$$

for E^* . In this definition, independent observations of $|E^{peak}|$ with decorrelation time τ will exceed E^* once in a total observation time T .

Carrying out this integral over the power law extrapolations of the PDFs in Figure 2 and setting the observation time to $T=100$ years and decorrelation time $\tau=2$ minutes, we solve for the E^* for each IMAGE magnetometer station and list these in Table 1. The average of the E^* is 13.2 volts per kilometer (V/km) and the median of the E^* is 13.5 V/km. These values should be compared to the current value in TPL-007-1 of 8 V/km. The number of stations is not large, and we indicate the uncertainty in E^* via the range of 8.4 V/km to 16.6 V/km.

Table 1. E^* for each IMAGE magnetometer station

Station	E^* (V/km)
BJN	13.1
HOP	14.0
KEV	13.8
NAL	14.6
NUR	16.6
OUJ	12.9
TAR	8.4
TRO	12.2

SUMMARY OF DIFFERENCES WITH TPL-007-1

At each step in our analysis of the IMAGE magnetometer time series, our analysis differs from that in TPL-007-1. Specifically:

1. Instead of considering every sample of the geo-electric field to be independent, we estimate the decorrelation time τ (see Figure 1) and analyze the IMAGE time series data to extract statistically independent observations of the peak geo-electric field magnitude. The effect is to remove a bias in the subsequent statistical analysis resulting in slower decrease in the probability of large, independent observations as compared to TPL-007-1.
2. We cast our analysis of the statistics in terms of PDFs, which removes the arbitrariness of how the IMAGE data is binned.

3. Instead of extrapolating histograms of geo-electric field observations to a rate of one observation in one hundred years, we instead compute the 100-year exceedance geo-electric field magnitude E^* from the PDF via integration.
4. Compared to the 8 V/km given in TPL-007-1, our E^* has an average of 13.2 V/km, a median of 13.5 V/km and a range of 8.4 V/km to 16.6 V/km.

RECOMMENDATION

The historical data used to define the TPL-007-1 Benchmark Event should be reanalyzed with consideration for the time correlations between IMAGE magnetometer samples (or other datasets) to develop a statistically meaningful definition of an independent GMD event. The GMD statistics should be developed in terms of a PDF for a single, statistically independent event with geo-electric field magnitude exceedance values computed from this PDF. These recommendations will help to avoid over counting of non-independent magnetometer samples, the distortion of the GMD statistics, and the arbitrariness of histogram binning—all of which impact the extrapolation to unobserved field values. When we incorporate these recommendations into analysis of the IMAGE magnetometer data, we find a one-in-one-hundred year geo-electric field magnitude for a statistically independent GMD event with an average of 13.2 V/km, a median of 13.5 V/km and a range of 8.4 V/km to 16.6 V/km.

LATITUDE SCALING OF GEO-ELECTRIC FIELD FOR THE GMD BENCHMARK EVENT

The TPL-007-1 Standard defines a GMD Benchmark Event in terms of the geomagnetic field time series generated at 60-degrees north geomagnetic latitude. The standard allows for the amplitude of this geomagnetic field time series to be scaled down for geomagnetic latitudes within the range of 60- to 40-degrees north geomagnetic latitude. The scaling is derived using functional fits to historical magnetometer datasets observed over the last several decades.⁴ The process of deriving this single scaling function for geomagnetic latitude assumes that the scaling is the same for the relatively small events in the historical data and an event of the magnitude of a one in a hundred year event, such as the Carrington event of the late 19th century. Here, we analyze historical magnetometer data with a focus on determining if larger GMD events scale with latitude as the smaller events observed to date.

ELECTROJET PROPERTIES--QUALITATIVE DISCUSSION

During a GMD, the fluctuations of the magnetic field within the Earth are largely caused by current fluctuations in the three electrojets within the E-regions of the Earth's ionosphere, where currents flow relatively easily.⁵ The Earth has two auroral electrojets, typically located at 60-degrees north and south geomagnetic latitude, and an equatorial electrojet around the geomagnetic equator. To a first approximation, these electrojets are ring currents. Fluctuations in their currents will cause the strongest magnetic field fluctuations and geo-electric fields at a location on the Earth directly below the electrojets. The influence of these fluctuations and the geomagnetic field they generate decays farther away from geomagnetic latitude of the electrojet.

CURRENT TPL-007-1 LATITUDE SCALING OF THE BENCHMARK EVENT

The TPL-007-1 Standard proposes to model this decay in the geomagnetic field using a scaling function derived from magnetometer observational data. Using 1-minute-resolution data from many magnetometer stations worldwide, Ngwira et al.⁴ computed the resulting geo-electric field using the one-dimensional plane wave model² and the Quebec Earth conductivity data.⁶ Then, they extract the peak geo-electric field magnitude for 12 large disturbances.⁷ Using 1-minute-resolution SuperMAG⁸ magnetometer data for the same 12 disturbances, we reproduce the analysis of Ngwira, shown in Figure 3. The exponential fit (dotted blue line in Figure 3) to the aggregated data set between 60-degrees and 40-degrees north geomagnetic latitude reproduces the scaling function in TPL-007-1 (solid red line) with sufficient accuracy. We note that this choice of fitting function is not physically motivated, however, given the variance in the data at any given latitude, almost any function that rises with geomagnetic latitude could be fit to the data.

⁴ Ngwira et al., "Extended study of extreme geoelectric field event scenarios for geomagnetically induced current applications," *Space Weather*, Vol. 11, 121-131.

⁵ Potemra, T. A., "Current systems in the Earth's magnetosphere: A Review of U.S. Progress for the 1975–1978 IUGG Quadrennial Report", *Reviews of Geophysics*, vol 17(4).

⁶ NERC Benchmark Geomagnetic Disturbance Event Description, December 5, 2014, pg. 19

⁷ Here, we distinguish between an "event" as defined using the decorrelation time (of ~ 2 minutes) computed in the previous section and a "disturbance" which may last several hours to several days.

⁸ Gjerloev, J. W., The SuperMAG data processing technique, *Geophys. Res.*, 117, A09213, doi:10.1029/2012JA017683, 2012.

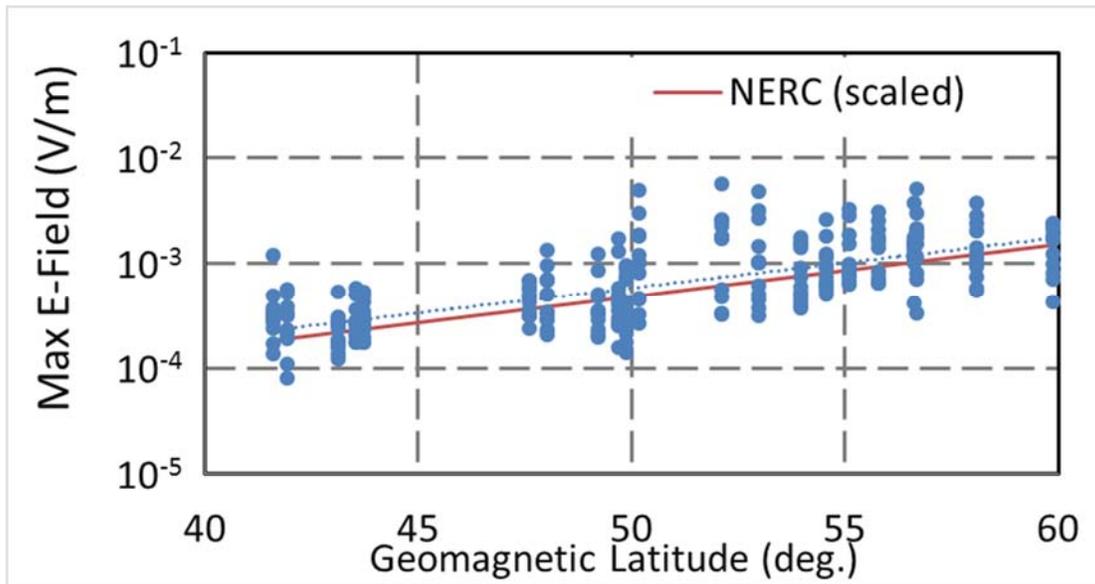


Figure 3. Maximum geo-electric field at different geomagnetic latitudes for 12 of the largest geomagnetic disturbances over the past two decades. The solid red line is the proposed scaling function from TPL-007-1. The dotted blue line is a fit to the SuperMAG data analyzed in this work.

REANALYSIS OF THE MAGNETOMETER DATA CONDITIONED ON THE DST INDEX

A fundamental assumption in the TPL-007-1 analysis is that events as large as the one in one hundred year disturbance that the standard proposes to address scale with geomagnetic latitude in the same way as the weaker events in the Ngwira data set. To investigate this assumption, we consider a somewhat larger set of disturbances from the same SuperMAG magnetometer array. Instead of analyzing the data in aggregate, we condition the data into bins based on both the disturbance storm time index^{9,10} (*Dst* index) and magnetic latitude. We then generate the average value of the disturbance maximum $|dB/dt|$ (the quantity most associated with the electric field) for each bin.

Figure 4 displays the average values of maximum $|dB/dt|$ as a function of geomagnetic latitude for different values of *Dst* (Here, the *Dst* bins are 100 units wide). Also plotted in the figure is the maximum $|dB/dt|$ averaged over all *Dst* values, as well as the proposed NERC scaling. It is important to note that, for the most extreme disturbances, there are a very small number of observations in the reported averages. Indeed for *Dst* = -550, there is only one disturbance, thus, it is not at all an average. The same holds true for the *Dst* = -450. Higher *Dst* bins include many more disturbances. Even though the number of disturbances in the two extreme bins is small, the *Dst*-conditioned data displayed in Figure 4 shows systematic deviations from the NERC scaling function as the *Dst* index increases in magnitude, i.e.,

⁹ Mayaud, P. N., "Derivation, meaning and use of geomagnetic indices," AGU monograph, Washington, DC, 1980.

¹⁰ The *Dst* index is the average value of the horizontal component of the Earth's magnetic field measured at four geomagnetic observatories near the equator. The *Dst* index is frequently used as a measure of geomagnetic activity. A more negative *Dst* index indicates stronger activity.

becomes more negative. As the value of Dst decreases below -300 , there is a significant and systematic deviation of the data from the proposed NERC scaling law. The most notable and relevant feature is the appearance and growth of a peak in dB/dt at geomagnetic latitudes of 52 degrees, in the vicinity of New York City.

From this analysis, it is clear that the NERC-proposed scaling, which was empirically derived from statistics dominated by weaker disturbances, does not accurately represent the stronger events at lower Dst . We speculate that for the disturbances with $Dst < -300$, the Earth's auroral electrojet is moving to lower magnetic latitude, and the appearance and growth of the peak in dB/dt as Dst becomes more negative as a result of this movement. Whether this speculation is correct or not, the results in Figure 4 indicate that the scaling of the geo-electric field in TPL-007-1 does not capture the behavior of large disturbances during periods of very low Dst .

Similar to Figure 3, the NERC scaling law shown in Figure 4 also fits relatively well over the geomagnetic latitude range of interest in TPL-007-1 (40 to 60 degrees north) for the average over all values of Dst . However, the NERC scaling law in Figure 4 is somewhat larger than the measured values for weak values of Dst (i.e., -150) because Ngwira *et al.* did not utilize such weak disturbances in their analysis.

To make this analysis clearer, Figure 5 displays the same data as in Figure 4, but normalized by the proposed NERC scaling function for the geomagnetic latitude range of interest (40- to 60-degrees north). Perfect agreement between the NERC scaling law and observed data would be indicated by a ratio of unity in Figure 5. There is approximate agreement for $Dst = -250$. For less negative Dst , the ratio dips well below 1.0 for the reasons noted above, i.e., these higher Dst values are not included in Ngwira *et al.* data set. For large disturbances ($Dst < -300$), a peak appears at a geomagnetic latitude of 52 degrees north. As Dst becomes more negative, this peak consistently grows and indicates a more severe GMD environment at this latitude (and at most latitudes) than indicated by the NERC scaling law used in TPL-007-1. The trace for $Dst = -550$ only includes one disturbance, as does the $Dst = -450$. One could argue that these data are not representative of all disturbances of this magnitude. However, these events did clearly show a threat 2.5 to 3.2 times higher at 52 degrees north geomagnetic latitude than would have been indicated by the NERC scaling. The data for $Dst = -350$ includes seven samples, making the elevated risk relative to the NERC scaling law (a factor of 2.0) more significant.

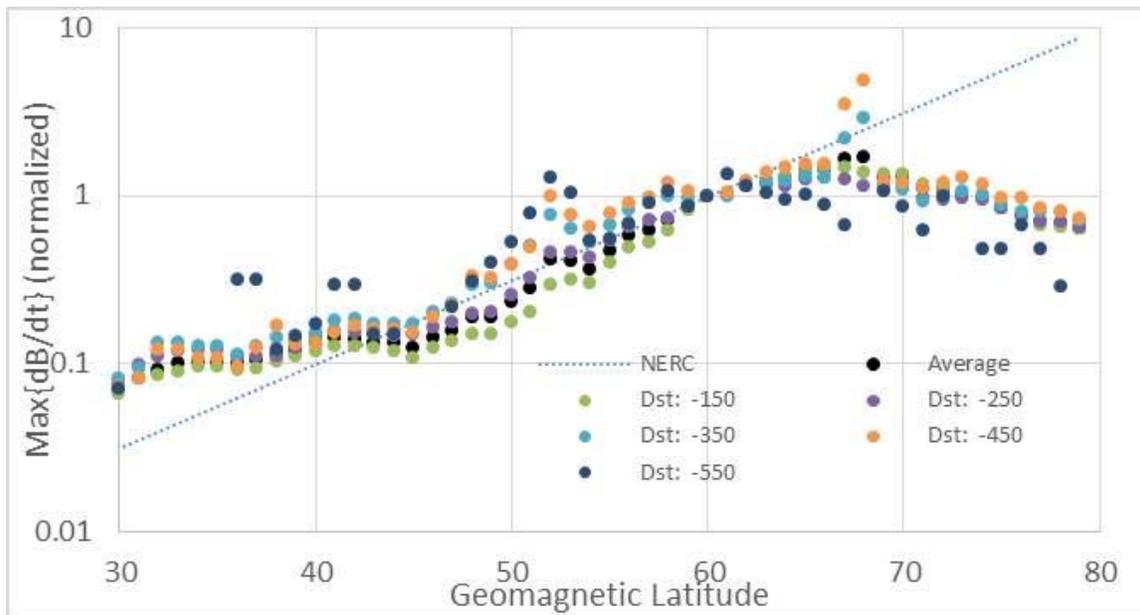


Figure 4. Average maximum $|dB/dt|$ for the largest 122 geomagnetic disturbances from 1980 through 2011, conditioned on the strength of the disturbance (Dst). The dotted line represents NERC's suggested scaling from TPL-007-1. The $Dst=-550$ and -450 bins each contain one disturbance. The $Dst=-350$, -250 , and -150 bins contain 7, 19, and 94 disturbances, respectively. The black dots represent an average over all Dst values.

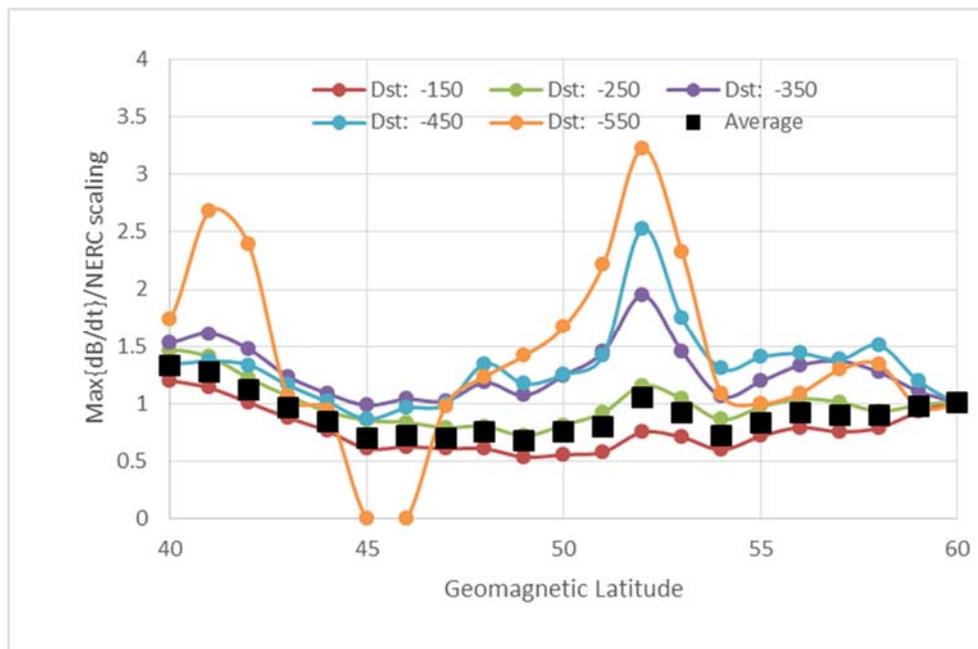


Figure 5. Average maximum $|dB/dt|$ for the largest 122 geomagnetic disturbances from 1980 through 2011, conditioned on the strength of the disturbance (Dst) and normalized by the suggested NERC scaling. The $Dst=-550$ and -450 bins each contain one disturbance. The $Dst=-350$, -250 , and -150 bins contain 7, 19, and 94 disturbances, respectively. The black dots represent an average over all Dst values. A value of 1.0 corresponds to a perfect agreement with NERC scaling predictions.

RECOMMENDATION

The physics and location of the electrojet during large GMD introduces a significant uncertainty into the latitude scaling suggested in TPL-007-1. Data-driven models based on historical observations will not capture the behavior of events like the TPL-007-1 Benchmark Event because there are currently no observations of events this large. Instead, a physically motivated model of the electrojet is needed to effectively extrapolate the small to moderate disturbance data currently in the historical record to disturbances as large as the TPL-007-1 Benchmark Event. Until the time that such a model is developed, an additional degree of conservatism in the mid-geomagnetic latitudes is warranted. Figure 5 gives an estimate of the correction factor. Discarding the strongest *Dst* bins, for which there are not more than a few, or even one, event in the average, we suggest a factor of two as a conservative correction factor.

EFFECTS OF EARTH CONDUCTIVITY ON UNCERTAINTY AND SCALING OF GEO-ELECTRIC CALCULATIONS

The standard GMD event in TPL-007-1 is defined in terms of a time series of magnetic field. Analysis of the impact of the GMD event on power systems requires converting this standard magnetic field time series to a local geo-electric field time series in different geographic regions using a plane wave model² of the Earth and data for the local Earth conductivity. Here, we assess the uncertainty in the geo-electric field caused by uncertainty in the Earth conductivity data and estimate the worst-case geo-electric field magnitude relative to the geo-electric field suggested by TPL-007-1.

PROCESS FOR ESTIMATING THE GEO-ELECTRIC FIELD FOR THE STANDARD GMD EVENT IN DIFFERENT REGIONS

To adapt the standard GMD event in TPL-007-1 to a local area, TPL-007-1 recommends several possible approaches, each of which is ultimately based on a one-dimensional plane wave model² of the Earth and United States Geological Survey (USGS) data for the Earth's conductivity.¹¹ Figure 6 shows typical one-dimensional Earth conductivity data taken from USGS sources for a single geological region. The solid green line is the recommended conductivity as a function of depth into the Earth. To apply the one-dimensional plane wave model, the conductivity is assumed to be spatially uniform and isotropic over a horizontal scale larger than the typical vertical scale of the conductivity variations. The conductivity is also assumed to be uniform within each layer, making abrupt changes in value only between the layers. Under these assumptions, a one-dimensional plane wave model² that accounts for the frequency-dependent reflections of the geo-electric fields at each layer-to-layer interface may be used to compute the resultant geo-electric field near the Earth's surface.

If detailed data for the local Earth conductivity is known, and if the utility in question has the capability, TPL-007-1 allows the use of these data in the plane wave model for calculating the geo-electric field¹² from the benchmark magnetic signal. If such specific data are not available, then TPL-007-1 recommends the use of regionally specific conductivity data from USGS.¹³ Figure 6 is an example of these data. Finally, if the utility chooses not to implement the plane wave model, TPL-007-1 provides for a simplified approach. A standard Earth conductivity model (the Quebec model provided in TPL-007-1) is used to compute a geo-electric field time series for the TPL-007-1 standard magnetic field time series, and regionally-specific factors are provided in TPL-007-1 to directly scale this geo-electric time series to the other regions defined by USGS.

The latter two processes to estimate the geo-electric field time series suffer from uncertainty in the Earth conductivity data. Note the wide range of uncertainty within the USGS data for some of the reported values of ground conductivity, shown as dashed lines in Figure 6. Without specific data for the local Earth conductivity, TPL-007-1 should consider a worst-case scenario for each of the USGS regions.

¹¹ U.S. Geological Survey, "Regional Conductivity Maps," <http://geomag.usgs.gov/conductivity>.

¹² NERC Benchmark Geomagnetic Disturbance Event Description, December 5, 2014, pg. 19.

¹³ NERC Benchmark Geomagnetic Disturbance Event Description, December 5, 2014, pg. 20.

EARTH CONDUCTIVITY (RESISTIVITY) DATA

Figure 6 shows a typical Earth layer resistivity data taken from the USGS website¹⁴ for the Cascade Sierra region. The dark green line in the figure represents the recommended resistivity value for each layer. USGS provides a considerable range of resistivity for many of the layers, denoted by the dashed lines in Figure 6. For example, the third layer below the surface in Figure 6 has a suggested range of resistivity that spans nearly three orders of magnitude. The TPL-007-1 standard only utilizes the recommend value of resistivity to calculate the geo-electric fields and, ultimately the geo-electric field scaling ratios used in the regional scaling coefficients.

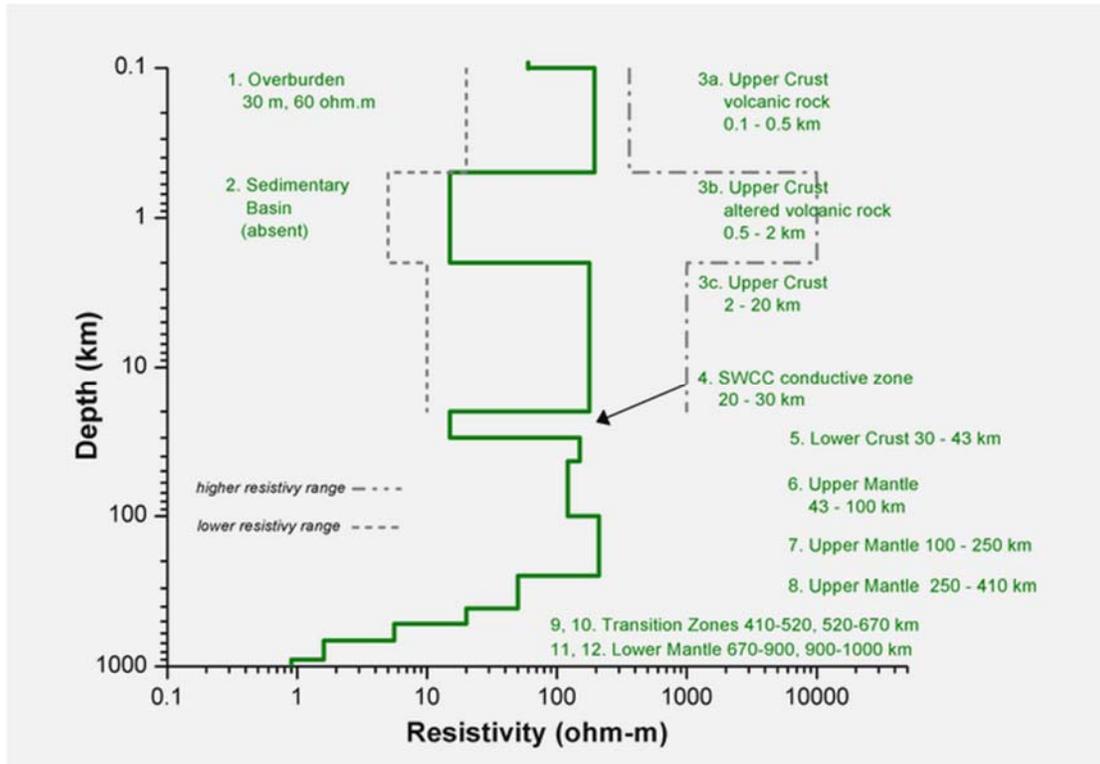


Figure 4. USGS earth layer resistivity data for the Cascade Sierra region

EFFECT OF UNCERTAINTY IN THE EARTH CONDUCTIVITY DATA ON THE PEAK GEO-ELECTRIC FIELD

Using the Cascade Sierra region in Figure 6 as an example, the large range of Earth layer conductivities around the recommended value is expected to lead to uncertainties in the geo-electric field computed from the TPL-007-1 Benchmark Event magnetic field. The complex interaction of reflections near the Earth's surface complicates the estimation of the uncertainty in geo-electric field driven by the Earth conductivity data uncertainty. We take a pragmatic approach to this estimation by varying the top four layers for each resistivity model over their entire resistivity range in the USGS data individually, holding all other layers at the recommended value. We find the largest (smallest) peak geo-electric field is always found when the layer in question is assigned the lowest (highest) conductivity value. We have

¹⁴ U.S. Geological Survey, "Regional Conductivity Maps," <http://geomag.usgs.gov/conductivity>.

extended this study by selecting all pairs of conductivity layers and simultaneously varying their conductivity over the entire two-dimensional space defined by range in the USGS data. For all pairs, we again find that the largest (smallest) peak geo-electric field is obtained with both layers set to their lowest (highest) conductivity

We use the intuition built by these two studies to define worst-case and best-case scenarios for the uncertainty in the Earth conductivity data. For the worst case, all of the conductivity layers are set to the minimum value of the range quoted in the USGS data. For the best case, all of the conductivity layers are set to the maximum value of the range quoted in the USGS data. If no range is given, the recommended conductivity value is used.

For those USGS regions that specify uncertainty ranges, the geo-electric time series (at 10-second resolution) is computed for the worst and best-case configuration of conductivities from the TPL-007-1 Benchmark Event magnetic field time series. The peak geo-electric field is determined for each geo-electric time series. The same procedure is carried out for the recommended conductivity values. We did not complete this procedure for all USGS regions because, for many regions, no resistivity range is reported. Although there is little or no uncertainty in the reported value, in these cases, the USGS reports the conductivity value as “possibly representative,” and, thus, it is impossible to specify a range.

The results for the peak geo-electric values for a subset of the USGS regions that provide a range of conductivity values are shown in Figure 5. For several regions, the difference between the results for the recommended conductivity values and the maximum and minimum values is quite small (e.g., AK-1A), whereas other regions show substantial spread between the results (FL-1). The results obtained by using the recommended values of conductivity tend to the conservative end of the range between worst and best case, although for some regions (e.g., FL-1), the gap in peak geo-electric field between the recommended and worst case conductivity is considerable. It is interesting to note that, within the scope of the one-dimensional plane wave model, the best-case result in Figure 7 provides some guidance as to the reduction in predicted peak geo-electric field that could be obtained by more careful study of the Earth conductivity in the local area of interest.

Before concluding this section, we note that the layers that are expected to have the largest effect on electric field magnitude exist at depths between 10km and 100km for the low frequency fluctuations considered in this study. In many cases, resistivity ranges were not reported for these depths within the USGS layer models; typically, uncertainties are only reported for depths down to 20km. Therefore, it is likely that the range in peak geo-electric field value reported in Figure 7 is an underestimate of the range that would occur if uncertainties were ascribed to the lower layers. More work needs to be done to generate better-defined Earth layer conductivity data sources for use in TPL-007-1

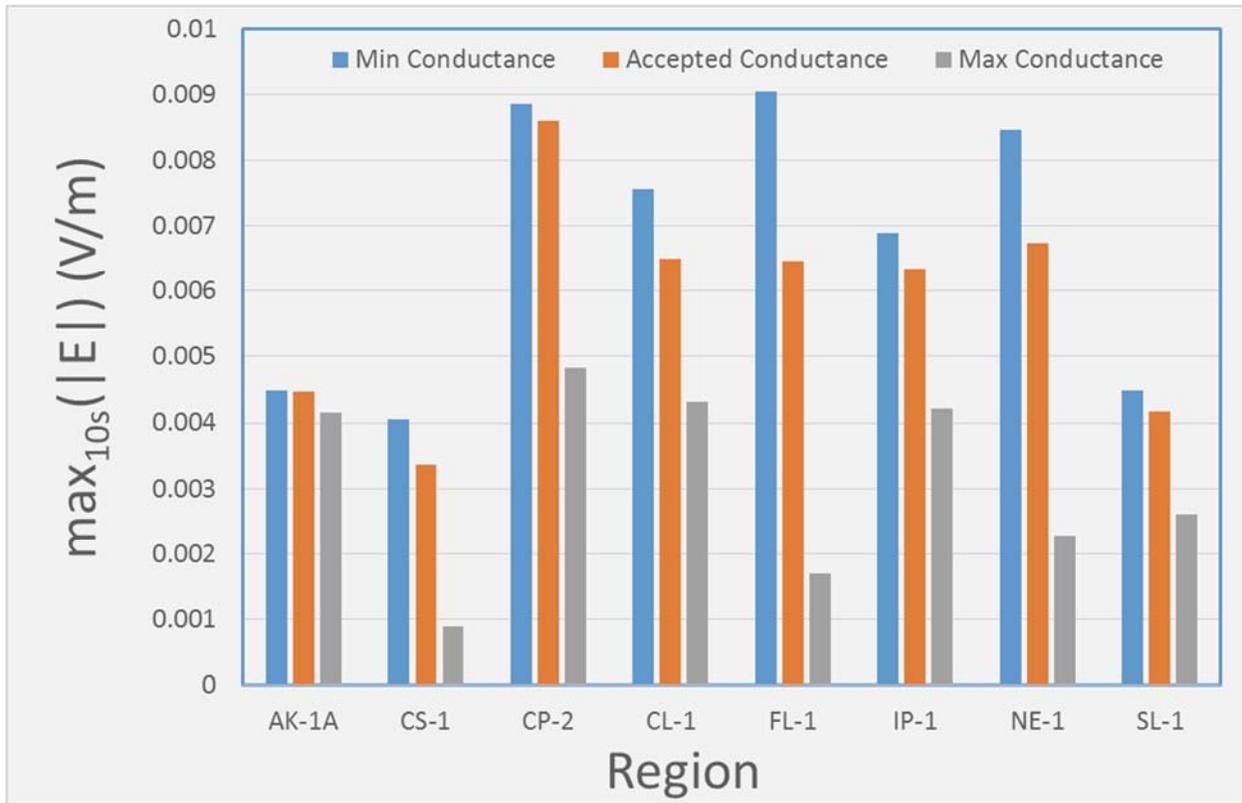


Figure 5. Worst case (blue bars) and best case (grey bars) peak geo-electric field values for a subset of USGS regions that report uncertainty ranges in conductivity values. The orange bars are the values obtained using the USGS recommended value.

RECOMMENDATION

Lacking better data for Earth conductivity, it seems prudent, within the scope of the one-dimensional plane wave model, to use the regionally specific, worst-case Earth conductivity configurations in the TPL-007-1 risk screening process.

APPENDIX

CHARGE TO THE REVIEWERS

The initial draft of this report was circulated to four reviewers for comment and criticism. Two of the reviewers were internal to Los Alamos National Laboratory while the other two reviewers were external. The reviewers were given a simple but broad charge:

Dear Reviewer,

Thank you for agreeing to review our report on our investigation of the North American Electric Reliability Corporation's (NERC) draft standard TPL-007-1 on the threat to the electric power grid from geomagnetic disturbance (GMD). The report discusses our investigation of three main topics related to TPL-007-1:

- 1. Statistical approaches used to determine the exceedance geo-electric field magnitude for a one in one-hundred-year GMD event*
- 2. The geomagnetic latitude scaling of geo-electric field generated by the one-hundred-year event GMD event*
- 3. The effect of the uncertainty in the Earth conductivity on the calculation of the geo-electric field from the NERC benchmark geomagnetic field*

Please review each of the three topics in the report and provide feedback on the following:

- 1. Are the approaches we used appropriate for the analysis performed?*
- 2. Are there alternative analysis approaches that you would recommend?*
- 3. Are the recommendations we provide at the end of each topic scientifically reasonable based on the availability of data at the time of the analysis?*

The topics being discussed in this report are broad and interdisciplinary. If you do not feel qualified to provide certain feedback on a topic, please indicate as such.

After receiving comments and criticism from the reviewers, the report authors took on of two approaches to resolve each issue:

1. For minor issues regarding places where additional text was needed to clarify the intended message, this was added without any further discussion with the reviewer.
2. For more substantial issues, the report authors performed additional analysis (as needed) and then made what they considered to be the relevant changes to the report. These were circulated back to the reviewer to ensure the original issue was resolved.

The remainder of this Appendix is documentation of the review process and of the changes made to the original draft of this report.

REVIEWER COMMENTS AND AUTHOR RESPONSE

Internal Reviewer #1: Reinhard Hans Walter Friedel (Reiner), Los Alamos National Laboratory

I have read through your report and have some comments, see below. I think the report needs some additional work before it is distributed further.

1. Your estimate of the 100-Year Exceedance geo-Electric field is, of course, limited by the time resolution of the original input data you considered - which is 10 seconds. I take it the geo-electric field from this data is at the same cadence? Since the electric field comes from the dB/dt changes in the field, the same dB over faster than 10 sec would give you a bigger electric field.

Have you looked at magnetometer data from a sample station with higher time resolution to see if the 10s resolution is sufficient to capture the fastest rate of change likely to be observed? A figure demonstrating the variation of dB/dt as a function of data time resolution would be needed here to support your result.

Backhaus/Rivera (Response #1): We have worked with SuperMAG data (1-min resolution) and IMAGE data (10-sec resolution). Similar PDFs were obtained from both data sets; we do not expect significant changes by extending the resolution to 1 sec (see below). We do agree that analyzing magnetometer data with higher resolution would be interesting, however, we do not have access to publicly available, online-data that has a faster than 10-sec resolution.

Reiner (Reply to Response #1): Jesse Woodroffe ran a similar analysis using 1 second data and also got similar results, he may be able to provide you a plot of that.

Backhaus/Rivera: We will ask Jesse for his data. We agree that performing the analysis with the higher time-resolution data might lead to a slightly larger 100-year exceedance field. However, the observation that there is little difference in the results obtained with 1-min and 10-sec resolution data coupled with the fact that spectral power in the electrical field fluctuations are falling at roughly the rate of $f^{-2.5}$ (thus, the power in 1-sec fluctuations is 300 times smaller than 100-sec fluctuations), gives us confidence that the analysis in the paper adequately captures the magnitude of the 100-year exceedance peak to within the confidence reported.

2. Regarding the latitude scaling and use of Dst: For the latitudes you consider, Dst may not be the best index. It represents the ring current, while the magnetic disturbances at high latitudes are due to the electrojets, which are more related to the Auroral Electrojet (AE) index. While Dst and AE are also correlated to some degree, using AE may give you a cleaner signature in Figs 4 and 5.

Backhaus/Rivera: We agree. We have done a preliminary analysis using AE index, as well as the Kyoto Sym-H index, however, we have not yet completed this work. Once complete, we will include it in subsequent journal papers. However, our preliminary analysis shows that no change is expected in the analysis results presented in the report. For the large events that contribute to the Dst<-400 plot in Figure 5, AE and Dst are sufficiently correlated that AE selects the same subset of events as Dst. The plot in Figure 5 will remain the same.

Additional comments on this area: Your speculation, however, is correct, as one moves to more active storms the electrojets typically move further south. AE is derived from a fixed set of high latitude stations, Dst from a fixed set of equatorial stations, which means that as the electrojets shift equator-ward, esp. for the more extreme events, AE becomes less representative and Dst “picks up” more effects from the electrojets. This may explain why you get the best “peak” in Fig 5 for the highest activity periods. Data from both the IMAGE, Carisma and SuperMAG chains has been used to derive AE- like indices that track the position of the electrojet, from which you can also get information on the latitude of the electrojet for a given event. These data should be utilized to derive a better scaling: i.e., for a given level of disturbance you first need to pick the max location of the electrojet, and for each of these levels provide a scaling for areas north and south of it. As noted in your discussion of Fig 5 the peak for low activity is likely to be north of 60 degrees, so the study should be extend to higher latitudes. This should be part of your recommendation.

Backhaus/Rivera: We agree that the latitude scaling of the geo-electric field needs to be a function of the magnitude of the event. Our analysis in Figure 5 already shows this effect. We would very much like to perform this analysis to extract a magnitude-dependent latitude scaling function that better represents the spatial distribution of the GMD hazard. Our approach would be very similar to the one you describe above; however, limitations of time and funding keep us from pursuing this path. Until we are able to complete this work, we believe our recommendation (see page 11 of the report) of an additional factor of two safety in the geomagnetic mid-latitudes (50–55 degrees) is a prudent, but not overly alarmist.

3. On the Sensitivity study on earth conductance - your approach is to vary the top 4 layers of the USGS model. It is the only bit I actually disagree with. For the frequency of fluctuations we’re dealing with, the work of Anti Pulkkinen showed that the layers that have the biggest effects are not near the top but rather exist in the 10-100km range of depth - so for your sensitivity study you should vary those layer’s resistivity range. This is kind of hinted at by your result - where up to three orders of magnitude resistivity variation in say the second layer results in only factors of 2-4 and often less variation in the induced field. That rather speaks to the induced electric field strength being rather insensitive to the variations of resistivity in the top layers. I suggest that this part of the work needs to be redone.

Backhaus/Rivera: We agree that the computed geo-electric field is more sensitive to uncertainty in the lower layers. However, the USGS data available to us report resistivity ranges only for the upper layers and not for the lower layers. The intent of this report is to discover and advise on the unresolved uncertainties within the TPL-007-1 screening process. If these uncertainties are overestimated without justification, the end result may be over-application of very expensive GMD mitigation measures on a nationwide basis. Therefore, we have chosen our current approach to avoid alarmism by only considering layers in the geological model where USGS provides ranges, even though those ranges are ill defined, with descriptive phrases like “possibly representative.”

Our conclusion is that more work needs to be done to generate better-defined Earth layer conductivity data sources for use in TPL-007-1. To address your concerns, we have added a

discussion of this matter indicating that the reported ranges in Figure 7 likely underestimate the true range when uncertainty in the resistivity of the lower layers is included.

Internal Reviewer #2: Earl Christopher Lawrence, Los Alamos National Laboratory

1. The data pipeline for the exceedance analysis is not clear to me. How is the interpolation done? Are the 10-sec data measurements at those time points, or something like averages, sums, or maxima over the 10-sec interval?

Backhaus/Rivera: The magnetometer sensor is buried in the ground and produces a signal at several 100-Hz. This signal is low-pass filtered to approximately 1 Hz, usually by electronics attached directly to the magnetometer buried in the ground. The heavy filtering is intended to reject extremely high frequency events. These 1-Hz samples are then averaged over 10 seconds. Therefore, the reported 10-sec resolution IMAGE data is the average ground magnetic field over a 10-sec time span. Similar processing is performed to generate the 1-min resolution SuperMAG data.

2. How are the events/peaks determined?

Backhaus/Rivera: As discussed in the report, we compute a normalized auto-correlation function from the 10-sec time series data and a correlation time of approximately 2 minutes by integrating the central peak of the normalized auto-correlation function (see Figure 1). Peak events are defined as maximum values in the 10-sec data over a time horizon of plus or minus 2 minutes (that is to say, peaks in the data that are larger than any other peaks within a correlation time).

3. I think the exceedance integral assumes that the PDF is describing events that have some notion of frequency. For example, if the PDF is the magnitude every 10s and you assume independent intervals than something like the $1 - 1/(100 \times 365 \times 24 \times 60 \times 6)$ quantile is your E^ (I think).*

Backhaus/Rivera: We agree with this reasoning, however, the frequency in this case is the expected frequency based on the inverse correlation time (1/2 minutes), not the sample frequency (1/10 seconds). Therefore, the appropriate quantile is $1 - (2\text{min}/100 \text{ years})$ rather than $1 - (10\text{s}/100 \text{ years})$.

4. Did you consider parametric forms for the distributions in Figure 2? If something fit well, you wouldn't need to extrapolate and the integral would be easier.

Backhaus/Rivera: Based on our observations of the latitude scaling (and confirmation of this by Reviewer #1), we expect there to be a smooth transition in the tail behavior dependent on the latitude of the sensor. The transition is caused by the relocation of the electrojets as a function of the magnitude of the event. This cannot be represented by a simple parametric form for the distributions in Figure 2. Rather, we believe, as is pointed out in the discussion by Reviewer #1, point #2, that these distributions will ultimately be better represented by the combination of a statistical characterization of the magnitude of the event combined with a physical model of the impact of that event (i.e., the geo-electric field) on the ground. Until we can complete this work,

we believe the power-law extrapolation over less than one order of magnitude is an appropriate approach.

5. Similarly, Generalized Extreme Value distributions are often used for answering questions of this type. Were these considered?

Backhaus/Rivera: We agree that Generalized Extreme Value distributions are another valid approach. Our approach to fitting the tail of the distributions to a power law was motivated by the approach taken during the development of TPL-007-1. Extreme event analysis should be utilized in future work.

6. On page 10, you discuss a peak in dB/dt. What is the argument that this is not spurious?

Backhaus/Rivera: There are physical reasons to expect the electrojet will move to lower latitude during extreme events. This has been confirmed in the above discussion by Reviewer #1, point #2. The peak's existence is not unexpected, however, we agree that determining the magnitude of that peak requires additional work. Please see the extended comments in the previously referenced discussion with Reviewer #1 for a few details. We believe our recommendation (see page 11 of the report) of an additional factor of two safety in the geomagnetic mid latitudes (50-55 degrees) is a prudent, but not overly alarmist. Further work could support the accuracy of the recommendation. In response to this comment, as well as to comments from Office of Electricity, we added more detail to this portion of the paper and presented the results more clearly.

7. My main suggestion is to be more explicit about some of the steps in the data processing. It's a bit too terse currently.

Backhaus/Rivera: We have worked to enhance the discussion in the paper and, based on your input, we add more clarity to the descriptions of procedure.

External Reviewer #1: Pascal Van Hentenryck, Brown University

Here is some feedback on this request. This is certainly an interesting perspective, compared to the original approach.

1. Page 3: "We interpolate over the small time gaps of less than 20 minutes": How is this done? What not use a more pattern-based recognition using the available data? Why is interpolation the right solution here? Can you assure that nothing significant is missed by this interpolation? Can you quantify? - This may not be that important and not something that you can do much about. Would just omitting these parts change the results? Have you measure that?

Backhaus/Rivera: The interpolation is a linear interpolation from the last remaining data point before the gap to the first point after the gap. We tried sample and hold, an iterative averaging process, as well as exponential tapering to interpolate small gaps in the data with little effect on the results, in particular to the high-magnitude events of interest in this study. This is likely

explained by the fact that the number of interpolated points was small compared to the number of actual data points. We tested a subset of the data to see if altering the 20-minute threshold had a significant effect on results; it did not. To address this concern, we added a note in the text indicating the number of interpolated points was less than one percent of data points.

2. Pages 3 and 4: This is the core of the document. What I was wondering here naively is whether there is some physical justification that you can use in arguing for this definition? Are the physical processes understood well enough to get an intuition about the definition of an event?

Backhaus/Rivera: The reviewer is correct that a better understanding of the underlying physics would greatly enhance our ability to define an event. The physics, however, is rather complicated. The fluctuations in the electrojet that ultimately give rise to the electric field depend on fluctuations in the strength of the coronal mass filtered through the magnetic response of the earth. At this time, the authors are not aware of research that could illuminate this physics to the point where it could inform the definition of an event. Absent a physical model, we feel the empirical definition we have developed based on the autocorrelation is the next best approach.

3. A side question here is whether two or more events could actually overlap with your definition

Backhaus/Rivera: Given the definition, two events will not be found within a correlation time of one another. Even if two peaks exist within a correlation time, the maximum of those two peaks would be chosen as defining the event. However, if a strong peak event has an extent larger than a correlation time, it is possible that smaller “sub-events” will be picked up riding the strong peak event so long as they are more than one correlation time removed from the strong peak event.

4. What I would also do personally is to do a sensitivity analysis on the value of this definition? How much do the results change if you change the definition of the event? Say from 30 seconds to 4 minutes in increments of say 10 seconds?

Backhaus/Rivera: As a result of this comment, we continuously changed the time requirement of event definition from 10 seconds to 1 day. We naively expected results to become asymptotic to the event definition time, and, to some degree, they are. The slope of the PDF tail approaches $|E|^{-3}$. What does not asymptote, however, is the magnitude of the PDF because as the event definition time increases, the number of weak events filters out. Because the PDFs must integrate to one, the amplitude of events in the large $|E|$ tail rises. Unfortunately, this also means that E^* continuously increases as event definition time increases.

Although at first this seems problematic, it demonstrates an incompleteness in our approach. We have rightfully understood that there is a lower bound to the even definition time, namely the 2-minute correlation time discussed in the paper. There is also an upper bound, namely a point beyond which the event definition time is filtering actual events. We attempted to approximate this correlation time by considering the average amount of time the signal spends above a threshold value of $|E|$. For large threshold values, the time above threshold was, on average, 1-minute wide, and was fairly insensitive to changes in threshold value. This indicates that the event definition time should be on order of 1 minute, but not significantly larger as this would start to filter weaker real events. We take this as an indication that choosing an event

definition time on the order of the correlation time of 2 minutes to ensure statistical independence, but not much larger to ensure real events are not being filtered, is the appropriate selection.

We appreciate the referee's comments in this matter. The subsequent results we generated while addressing this comment will be included in an IEEE transaction we are currently writing. However, because the work done in response to this comment reaffirmed the selection of a 2-minute even definition time, and therefore does not change the recommendation, we have not included it in the text.

5. Page 4: Your use of autocorrelation is nice. Did you actually try to see if there were outliers over time (either very small or very long events) that could influence the autocorrelation? Probably giving the variance would address this to some extent.

Backhaus/Rivera: In reanalyzing our response to referee comment 4, we also took the time to investigate the amount of time an event stays above a threshold value of $|E|$. If there were significant outliers, one would expect this time to vary quite a bit as threshold level changed; it did not. On average, events were on the order of 1-minute duration for thresholds large enough to have only 200 events in a 10-year period.

6. Page 4: faster decline -> faster decline of

Backhaus/Rivera: This has been altered.

7. Page 4: Why not give the formulation in the last paragraph that explains mathematically the point you are making in English in the paragraph starting with "This definition"

Backhaus/Rivera: We will certainly be presenting a more mathematical treatment in an IEEE transaction on this matter. For the purposes of the audience of this particular document, a formulation would have distracted from the recommendation we were trying to present.

8. Page 5: Before-last paragraph: The point you are making there is unclear (partly because the last sentence does not parse).

Backhaus/Rivera: Apologies. The text has been fixed.

9. Page 5: What power laws did you use? Not sure how they show a long tail but I may be misunderstanding Figure 2. It would be good to give a bit more detail here on how you do this.

Backhaus/Rivera: We did not try a discrete set of power laws. The power law was a best fit power law decay to the tail of the distribution function. We did not consider other canonical forms.

External Reviewer #2: Ian Hiskens, University of Michigan

After reading the 2015-08-14 version of the LANL report on TPL-007-1, I have the following comments/questions:

1. *The data conversion process is not clear. What is meant by “we separate the time series into smaller subsets”? More particularly, how are the “smaller subsets” defined? How is the 20-year time series related to the data subsets?*

Backhaus/Rivera: The 20-year time series is the raw data from the magnetometers. When the magnetometers drop data points, a default value of 99999.9 nT is inserted into the data (well above even the strongest events). Therefore, the data set can be broken apart into a set of contiguous “good” regions, that is, continuous sets of data that do not have a value of 99999.9 nT, and contiguous “bad” regions, which have values of 99999.9 nT. Good and bad regions are complementary and alternate (i.e., every bad region is bounded by either two good regions or the data sets endpoints). We consider each bad region individually. If the bad region in question is less than 20 minutes, we replace the bad region with interpolated values from the good endpoints and merge the bounding good regions into a new larger good region. This is done iteratively until all regions are either good, or bad regions of greater than 20 minutes in extent (selection of 20 minutes ensures that the number of interpolated points is less than 1 percent of the total data). The resultant set of good regions are then filtered so that we only consider those good regions that are at least 2 days long to ensure the Fourier modes of interest are not windowed significantly. The resultant set of good regions are the “smaller subsets” discussed in the text.

2. *In the definition of an event, it would be neater to state, “for all τ such that $0 < \tau < 25$ min.”*

Backhaus/Rivera: This has been altered.

3. *It's not clear why “bias, if not corrected ... results in an under-estimation of the rate ...”. Likewise, the discussion of bias in the summary on page 6, particularly “slower decrease in the probability ...”, is not clear. Some explanation would be helpful. Maybe you could include a figure showing PDFs with and without bias.*

Backhaus/Rivera: The slope of the PDF tail changes continuously even as definition time increases, though it does asymptote as was discussed in an earlier reviewer response. Without an event definition time, the samples are not statistically independent, and, thus, points on the way up to a peak are counted along with the peak itself. This multiple counting is what is meant by biased. As was indicated in the response to the earlier referee comment, we will discuss this in more detail in an IEEE transaction we are writing.

4. *How are the PDFs generated? There are a few ways this could be achieved. Clarification would enable better comparison with the histogram approach.*

Backhaus/Rivera: PDFs are generated by first generating the histogram, but then normalizing the histogram by total number of points and bin-size; thus, the resultant PDF integrates to one. This has been added to the text.

5. *In the reanalysis of magnetometer data, it is currently stated that, "...we condition the data into bins based on the Dst index ..." The data is binned on the basis of both latitude and Dst index.*

Backhaus/Rivera: Thank you. This has been altered.

6. *In the reanalysis of magnetometer data, it is stated, "... it is clear that the physics modeled by the NERC proposed scaling ..." This could be more accurately stated, "... it is clear that the empirical model proposed by NERC ..." In the related recommendation, perhaps it would be better to say, "The poorly understood physics and location of the electrojet during large GMD events ...".*

Backhaus/Rivera: Thank you. The text has been altered.

7. *In the last paragraph on page 13, it is stated "...TPL-007-1 should consider a worst-case scenario for each of the USGS regions." Is this suggesting that TPL-007-1 only consider a worst-case scenario or that a worst-case scenario be considered as part of a robustness assessment?*

Backhaus/Rivera: We are suggesting a worst-case scenario be considered as part of a robustness assessment.

8. *In considering the uncertainty in earth conductivity data, Figure 6 is for the Cascade Sierra region but the subsequent analysis is for a Florida region. Is it even important to mention that initial investigations used the Florida region*

Backhaus/Rivera: No, it is not. We did the same analysis for all the regions. The text is a holdover from a previous draft when a second figure showing the analysis for Florida was inserted. That figure was removed, but the text inexplicably remained. The text has been removed.

9. *In undertaking the earth conductivity uncertainty analysis, it would be interesting to do Monte-Carlo with values chosen uniformly over their ranges to explore the uncertainty space more fully.*

Backhaus/Rivera: This was, in fact, how we originally approached the problem. The difficulty was in the selection of a probability distribution function to drive the simulation for each conductivity layer. We started with a log-normal distribution, but soon abandoned that for just a flat distribution between the bounds because we could not justify the use of a log-normal. Ultimately, we decided to look simply at worst and best case scenarios for both ease of reporting because that the conductivity data was in such a state that no distribution could be justified.

10. *The analysis builds on the one-dimensional plane wave model. How good is that model?*

Backhaus/Rivera: As with all models, it depends on how well the physical reality you are applying the model to matches the underlying model assumptions. For the particular case of the plane wave model, the underlying assumption is spatial homogeneity in the lateral directions (North-South and East-West). It is likely this assumption works well enough in certain regions (e.g., the middle of a continental plate) and less well in others (the edge of a continental plate).