Polynomial Fitting and Analysis of Geomagnetic Disturbance Impacts

Rhett Guthrie, Katherine Davis
Department of Electrical and Computer Engineering
Texas A&M University
College Station, TX, USA
rhett.guthrie@tamu.edu, katedavis@tamu.edu

2023 IEEE Texas Power and Energy Conference (TPEC) | 978-1-6654-9071-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/TPEC56611.2023.10078446

***Abstract*—Geomagnetic disturbances (GMDs) are a threat to power systems with the potential to cause large impacts on grids around the world through the creation of geomagnetically induced currents (GICs). Modeling GMDs with detailed power system impact analysis is crucial, as GICs cause transformers to overheat and saturate, causing reactive power losses and lower voltages. The scale of impact is related to the severity of the storm. While the ability to systematically study the impacts of GICs on power systems has greatly improved over the past decade, it remains a challenge to predict these storms and to study and prepare for impacts of future storms. Hence, the goal of this paper is to create a mathematical model of GMD storms to predict how stronger storms will impact the grid in the future. To obtain this model, polynomial regression is applied to magnetic field data to obtain an equation. The equation models the general shape of GMDs, and it allows for alterations based on the location, time, and strength of each GMD. Analyzing these features helps to understand the effect they have on GMDs and to generate better models. The models in this work generate the general shape of past GMD storms, however there are variations by location, adding challenges for analysis.**

***Index Terms*—Geomagnetic disturbances; Geomagnetically induced current; Polynomial fitting; Latitude; Longitude**

I. Introduction

GEOMAGNETIC disturbances (GMDs) are caused by solar eruptions that send magnetic particles hurtling toward Earth. A GMD is characterized by variations in the earth's magnetic field causing electric fields at its surface which produce low-frequency quasi-dc currents running along transmission lines and through transformer windings [1]. The flow of these geomagnetically-induced currents (GICs) through the system is related to available paths to ground, such as grounded transformers [2]. One way to simulate GICs is to use Maxwell's equations to find the changing properties of Earth's electric field, based on magnetic field data. To analyze the magnetic field data, a method of modeling is needed that captures the behavior of the magnetic field during a GMD storm, which poses various challenges, such as the magnetic field shifting every year, location dependence, and magnetic field direction differences. [3]. One approach is to use mathematical modeling to obtain an equation, with time as the input and magnetic field value as the output, that can be used to find electric field changes through Faraday's law (equation 1). Within the mathematical modeling approach for storm and impact analysis, various methods are of interest, and this work introduces an approach for one of these methods, *polynomial fitting*. Polynomial fitting works best when there is a curvilinear relation between two variables, and for GMDs, there is reason to believe that such a relationship based on time exists [4], [5]. The major appeals of polynomial fitting include its simplicity and speed, such that it could model various GMD storms in a timely manner. The benefit of polynomial fitting is generating equations for the different directions (X, Y, and Z) and different days of each storm at multiple locations around the United States. After creating an equation through polynomial fitting, the coefficients can be changed to represent features of GMD storms, such as location (latitude, longitude, and height), for scaling particular coefficients.

The authors would like to acknowledge the support of the U.S. Department of Energy, Office of Cybersecurity, Energy Security, and Emergency Response (DOE-CESER) and the Pacific Northwest National Laboratory (PNNL) for work conducted under Master Agreement 401273, Task Order 610876.

This paper has three objectives. First, the paper motivates and presents the rationale for modeling GMD storms through polynomial fitting and details the approach. Second, is an analysis on different days of the storm, looking at each magnetic field direction. The third and final objective is to see how location and time play a role in GMD impact through polynomial fitting.

The paper is structured as follows: Section II introduces an overview of GMD storms, polynomial fitting, and location importance for GMDs The case studies and developed scenarios are presented in Section III. The analysis and results of polynomial fitting for GMDs are presented in Section IV. Lastly, conclusions and future work are discussed in Section V.

II. Background

A. Geomagnetic Disturbances

Geomagnetic disturbances are caused by coronal mass ejections (CMEs), which are solar eruptions that propagate away from the sun until reaching Earth. When CMEs reach Earth, it causes a disturbance in the Earth's magnetic field, and through Faraday's Law, where E and B are electric and magnetic fields, respectively, with ∇ being the curl operator, it can be

978-1-6654-9071-9/23/$31.00 ©2023 IEEE

shown that a change in a magnetic field will induce a change in an electric field:

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{1}$$

Thus, the disturbance of Earth's magnetic field leads to a change in Earth's electric field, which induces currents in conductors like transmission lines. These currents are called geomagnetically induced currents (GICs), which are low-frequency, quasi-dc waveforms that can affect electric grid components like transformers, protective relays, and transmission lines [6].

These effects become disruptive to power grid operations and can damage or destroy equipment. The strength of the GICs depends on the strength of the GMD storm, G1 (minor storm) to G5 (extreme storm) [7]. Large scale GMD events are infrequent but can cause substantial impact to the power grid. Hence, due to the infrequent nature of these GMDs, there is a lack of data to construct models, but with improved modeling for past disturbances, protective measures can be created for future GMD storms on the power grid.

B. Polynomial Fitting

Polynomial fitting describes a data-driven mathematical modeling approach that applies a polynomial regression to a data set, resulting in an equation that fits that particular data set of the form below, where f(t) is a function dependent on t or time, and a_n are coefficients.

$$f(t) = a_0 + a_1 \cdot t + a_2 \cdot t^2 + ... + a_n \cdot t^n \tag{2}$$

A concern with polynomial regression is overfitting, and to avoid this each degree of polynomial was checked to find the lowest value polynomial that fit the data best, capturing the underlying waveform, not the noise, and this is further discussed in IV-A. The benefits of polynomial regression are that data sets can be sampled quickly, in a matter of seconds to minutes, generate an equation capturing the underlying waveform of given data sets and that the generated equation's coefficients can be altered for individual problems. Hence, polynomial equations are quick, effective, and flexible models for fitting data sets [8].

C. Locational Importance

The polynomial regression approach is also motivated by the need to better understand, how GMD impacts relate to geographical location. In GMD storms, the effects are stronger near Earth's magnetic poles, which can be seen from changes in the geoelectric and magnetic fields [9]. While the effects of a storm are not as strong near the equator, GICs can still occur under a strong GMD. For example, Southern Africa is close to the equator, but during the 2003 Halloween GMD event (one of the strongest storms in recent history), GICs were experienced [10]. Locational analysis of past GMD storms will help predict these situations by determining where strong GMD storms will affect.

III. Data Analysis and Modeling

A. Data Analysis

Large-scale GMD impact analysis requires a wide-area view to support comparing multiple locations at once. This allows for comparison of the magnetic field changes at these different locations, to answer the question, *how can the data help quantitatively compare how storm impact behaves in different locations around Earth*? Hence, this work applies data from magnetometers across the United States,shown in Fig. 1, which use Coordinated Universal Time (UTC), from INTERMAGNET. [11]. However, not all locations have data for all storms. From these locations, the magnetic field data can be plotted over the course of a GMD.

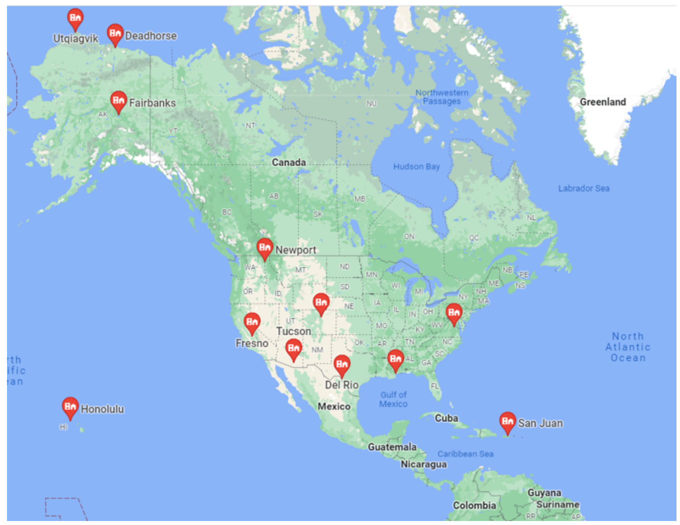

Fig. 1: Magnetometers locations for the analysis (image from Google Maps, data from [11])

The GMD events analyzed in this paper are two G3 storms and one G4 storm. The two G3 storms are from Nov. 01-07, 2021 and Sept. 06-10, 2017, respectively. The G4 storm occurred November 06-14, 2004. The GMD storms are separated into individual days and by magnetic field directions X, Y, and Z, to better analyze the patterns. The X, Y, and Z directions are north/south, east/west, and in/out of Earth, respectively.

B. Case Study Scenarios

The experiments in this work involve applying a polynomial regression to GMD storm data and obtaining a fit for the magnetic field waveforms. Using polynomial fitting, an equation will be obtained similar to the form of equation 2. The dependent function $f(t)$ will be the magnetic field value, B, taken from the magnetometers, and the a_i are coefficients determined when running the polyfit function, with t time as the independent/input variable. To note t^n is time to the power of n, so there is still only 1 independent variable in the equation. The benefit of such an equation is the ability to modify the coefficients to fit features of the storm as discovered. As such, the equation could be easily scaled, by multiplying the entire equation by a single value, or individual coefficients by a single value, to model more intense storms, and can be used to find the derivative with respect to time for calculating the electric field, via Faraday's law 1. Polynomial regression is quick for a large range of degree polynomials, the higher the degree and more data, the longer it will take to process, but unless the degree is greater than or equal to 30, with hundreds of thousands of data points, the time should be less than two minutes to complete the regression.

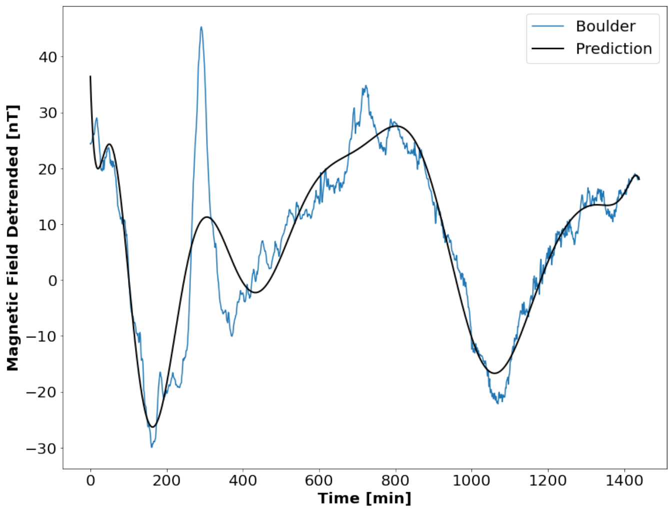

Fig. 2: Polynomial Fitted of Boulder for G3 storm Nov. 02, 2021 in the X direction

Fig. 2 shows the polynomial fit of the detrended magnetic field data in Boulder in the X-direction on Nov. 02, 2021 for the G3 storm. The reason for detrending the data is to equalize each location, as the initial magnetic value for each location is different. The polynomial model fits Boulder's magnetic waveform well; however, as can be seen, it does not quite model the large spike seen at 300 minutes. Fig. 2 shows how the polynomial works, per magnetometer. The same process is applied to all magnetometer locations. For each location, direction, and day, a polynomial fit is made; these are averaged to obtain the underlying waveform of a GMD. Then, all waveforms are cross-analyzed to see if a particular day, direction, or location share features or if each storm is individualistic.

IV. Results

A. Polynomial Regression Fitting

After calculating the polynomial curve fit for each location, mentioned in the last section, the average is taken and plotted with the magnetic field data for each magnetometer location. Computing the average of the data allows for an equation that can be altered for each location. As an example, the equation for Fig. 3 is shown below. The equation is lengthy, as the polynomial degree is set to 14, but this complexity is justified, since it created the best fit to the magnetometer data. The code was run in a Jupyter Notebook using the polyfit function [4]. When testing different degrees for the polynomial regression, any degree lower than 10, missed the increase around 250 minutes, while any degree higher than 17 resulted in minimal improvement. Increasing the degree from 10 to 17 does improve the shape, but benefits decrease as the polynomial number increases. When increasing the degree from 13 to 14, a noticeable improvement occurs, whereas the improvement seen in going from 14 to 15 is minimal. Similarly, the other waveforms also had the best fit when using a 14th-degree polynomial. The 14th degree polynomial works best as the underlying waveform, but when wanting to model a smaller portion of the waveform, a smaller degree is wanted. Having a 14 degree polynomial on a 100-300 minutes interval would cause overfitting, so a degree between 2-4 is preferable for the larger surges during a GMD. Testing revealed that a degree over 4, started to capture more noise, then the waveform for the smaller intervals, hence overfitting. The smaller intervals will create a piece-wise polynomial equation that will fit the maximum and minimums better than the underlying waveform. The piece-wise polynomial model will be discussed later in this section.

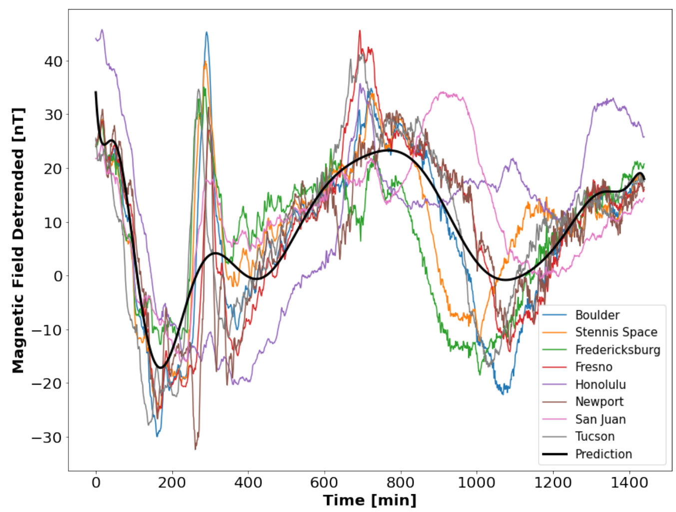

Fig. 3: Average polynomial fit of the G3 storm on November 02, 2021 in the X direction detrended

B. Magnetic Field Direction Differences

The y-axes in Figs. 3, 4, and 5 shows X, Y, and Z directions of the detrended magnetic field, respectively. Looking at these figures, key differences can be seen. Usually, the X direction is the main concern, as it is the primary source of GICs. However, these figures show that Y is also notable, as the peak value is more than double that of the X direction. Since GICs are caused by *changes* in the magnetic field inducing an electric field, the greater the change, the stronger the GICs. Hence, while the X direction has more variation in the waveform, the Y direction has larger values at certain times. The Z direction is not as important, as its maximum are less than X, and the majority of the waveform is centered around 0. However, this is not always the case, as sometimes the maximum of Z can be greater than the maximum of

TABLE I: THE COEFFICIENTS OF THE POLYNOMIAL EQUATIONS FOR EACH FIG.

Storm	Nov21 G3	Nov21 G3	Nov21 G3	Nov21 G3	Nov04 G4	Nov21 G3	Nov21 G3
Day	Day 2	Day 2	Day 2	Day 2	Day 6	Day 2	Day 2 [250-400]
Direction	Boulder_X	X	Y	Z	X	Bou_scaled_X	Piecewise_X
t^{14}	4.13E-36	2.09E-36	1.94E-36	7.15E-37	-7.84E-38	2.51E-36	0
t^{13}	-4.25E-32	-2.15E-32	-1.96E-32	-7.17E-33	1.01E-33	-2.58E-32	0
t^{12}	1.96E-28	9.88E-29	8.80E-29	3.20E-29	-5.45E-30	1.19E-28	0
t^{11}	-5.31E-25	-2.69E-25	-2.32E-25	-8.33E-26	1.68E-26	-3.22E-25	0
t^{10}	9.44E-22	4.79E-22	4.00E-22	1.40E-22	-3.30E-23	5.74E-22	0
t^{9}	-1.15E-18	-5.88E-19	-4.70E-19	-1.60E-19	4.38E-20	-7.05E-19	0
t^{8}	9.91E-16	5.07E-16	3.85E-16	1.26E-16	-4.06E-17	6.08E-16	0
t^{7}	-5.99E-13	-3.08E-13	-2.21E-13	-6.76E-14	2.65E-14	-3.70E-13	0
t^{6}	2.52E-10	1.30E-10	8.79E-11	2.41E-11	-1.22E-11	1.56E-10	0
t^{5}	-7.14E-08	-3.68E-08	-2.35E-08	-5.40E-09	3.95E-09	-4.42E-08	0
t^{4}	1.29E-05	6.55E-06	4.08E-06	6.72E-07	-8.73E-07	7.86E-06	-8.86E-08
t^{3}	-0.00134	-0.000651	-0.000427	-3.36E-05	0.000125	-0.000782	0.000213
t^{2}	0.0672	0.029	0.0246	-0.000311	-0.0104	0.0348	-0.152
t^{1}	-1.41	-0.566	-0.643	0.0503	0.413	-0.679	43.3
t^{0}	32	31.5	4.84	-3.26	-17.7	37.8	-4.28E+03

X; this is rare, and typically only occurs on more northern magnetometers, based on this analysis.

The directions of the magnetic data are important for analysis, but similarly important is a time-based analysis to determine timings for sudden increases or decreases. As an example, Fig. 3 has surges around 300 min (5 AM), 700 min, and roughly 1000 min. Similarly, Y has surges around 300 min, 700 min, 950 min, and 1100 min (6:20 PM). Finally, Z has surges around 300 min, 700 min, and around 1000-1200, showing that all three directions have spikes around similar times. While the more drastic changes occur at a similar time, there seems to be a small delay, where X changes, then 10-20 minutes later Y changes, and then another 10 or so minutes, and Z changes. Hence, it seems that in the timings of a GMD storm, X might be a good indicator, as the other directions soon follow up with their own changes in the magnetic field. These same time variations have been seen in the 3 GMD storms analyzed for this paper.

Lastly, for each magnetic field direction, the rise and fall of the waveforms are analyzed. In Figs. 3, 4, and 5, starting with the surges around 300 min, all directions have a positive spike. At 700 min, X has a positive spike, while Y and Z have negative spikes. For the last major time step, 1100 min, all directions have negative spikes. These effects are difficult to model when focusing on only one direction, ideally each direction would have its own model, with slight differences, that coincide with the features for each direction.

C. Locational GMD Effects

Location plays a major role in the waveforms and effects of GMDs. This can be seen in all the figures with waveforms from multiple locations, as well as in some figures from [3]. Location of a GMD functions such that the closer a location is to Earth's magnetic poles, the stronger the effects of GICs.

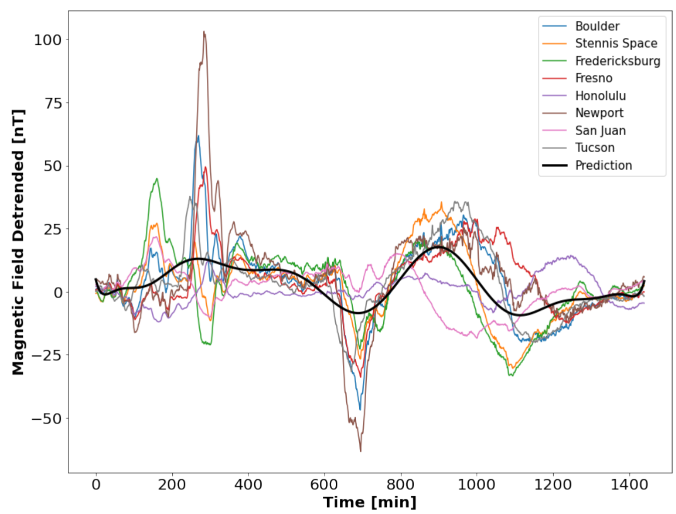

Fig. 4: Average polynomial fit of the G3 storm on November 02, 2021 in the Y direction

Hence, NERC's Benchmark Report on GMDs has a scaling factor based on latitude, increasing in value at higher latitudes [9]. Similarly, the magnetic field fluctuates more in locations closer to the poles. This can be seen in Fig. 6, as it shows the locations separately, prior to detrending. The physical location for the magnetometers is seen in Fig. 1. As Alaska is closest to the North Pole, the magnetic field suffers more magnetic disturbances, shown in Fig. 6 for a G4 storm, with Barrow, AK (orange) and Fairbanks, AK (red) suffering disturbances from 2500-7000 min.

While these drastic changes are relatively easy to observe without a computational model, it is more nuanced to analyze the waveforms for smaller surges. For example, two locations are affected later than most of the locations, and to a lesser

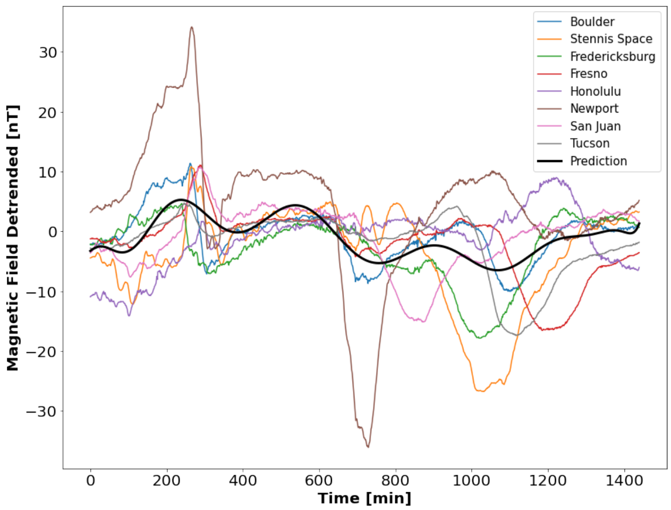

Fig. 5: Average polynomial fit of the G3 storm on November 02, 2021 in the Z direction

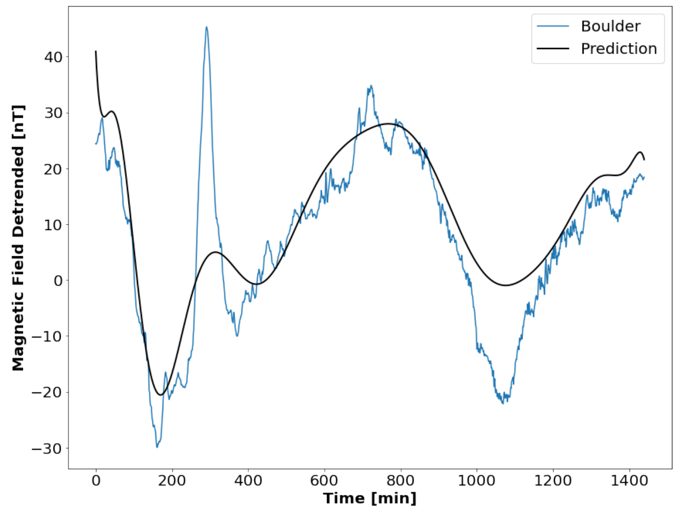

Fig. 7: Boulder for the G3 storm on November 02 2021 in the X direction with a scaled polynomial fit

extent: Honolulu (Hawaii) and San Juan (Puerto Rico). These locations are closest to the Equator for these magnetometers. For this reason, their waveforms do not show the same drastic changes as the other locations; the waveforms are closer to zero (due to detrended data) and delayed. This delayed effect is seen in Fig. 3, as both Honolulu and San Juan have a larger jump than the other locations toward the end of day two. The locations closer to the equator do not have as quick of changes in the magnetic field, thus causing a weaker electric field to be induced and hence weaker GICs. While these locations are less susceptible to GICs, that does not mean severe impacts cannot happen, as seen in southern Africa for strong GMD storms [10]. Hence, the results of this study motivate a quantitative data-driven approach to ensure that more subtle, currently hidden or unknown, relationships are captured in models for GMD analysis.

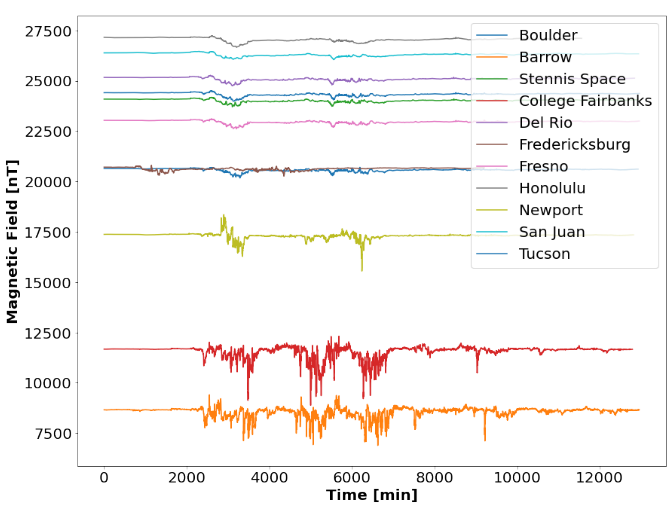

Fig. 6: Magnetic field waveforms for G4 storm on November 06-14, 2004 in the X direction

In summary, to better predict GMD impacts in places deemed less susceptible, a way to alter the average waveform is needed. A good way to alter the waveform is to scale up (multiply by a value greater than one) or down (multiply by a value less than one) the average polynomial waveform based on location. This work looks at scaling the entire polyfit with a single value, but future work will scale individual coefficients based on different features of GMD storms. Fig. 7 shows an example of scaling the average polynomial fit for the magnetic data during day 2 of the November 2021 storm in the X direction. This can be directly compared with Fig. 2.

Comparing the two figures shows that the scaling, while not as precise, fits the underlying curves of the waveform, however the downsides are that the local minimum and maximums are not modeled as accurate, seen around 1100 minutes. The pros of scaling are that the waveforms still keep their general shape, and that the waveforms can be altered according to each location. Whereas the cons of scaling, are that the entire waveform is affected, and the local maximums and minimums are still not modeled as accurately as wanted. To improve on the modeling of minimum and maximums, the current idea is to model each max or min by themselves and treat the storm as a collection of functions, based on time. For example, looking at Fig. 8, there are two separate polynomial functions, one for the underlying waveform, and one for the maximum around 300 minutes. The main polynomial function is the same as Fig. 3, and shows how creating the piece-wise polyfit improves the modeling of GMD storms.

Generating consistent piece-wise functions for the waveforms, involves finding timings of maximums and minimums for GMD storms. The main issue with determining when these times occur is that the beginning of GMD storms differ for each individual storm, whereas the endings of each storm are similar. In addition to the timings of maxs and mins, time lag needs to be taken into consideration, where time lag is referring to a small lag between magnetometer locations. This

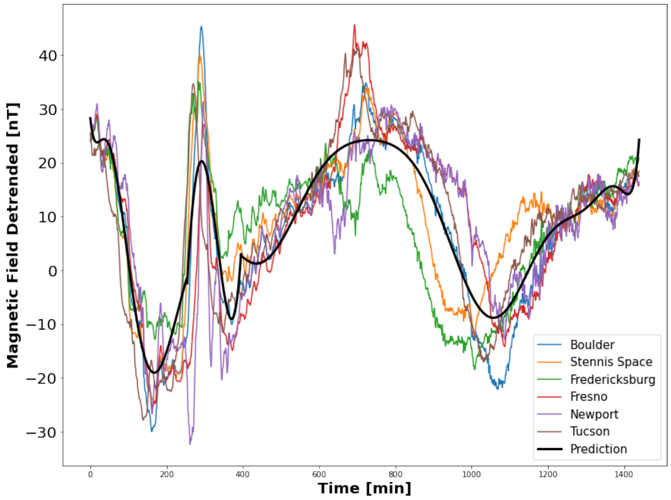

Fig. 8: Average piece-wise polynomial fit of the G3 storm on November 02, 2021 in the X direction

was mentioned before when looking at Honolulu and San Juan, and how the waveforms had a small delay compared to other locations. Timing is another key aspect that is important to keep in mind, that correlates with location.

V. Summary

A. Conclusion

This paper analyzes historical data from distributed magnetometer data in North America for three historical GMD storms. The results have shown that the polyfit works well in obtaining the general shape of the storm, and generating an equation that is able to determine the derivative of the magnetic field to obtain electric field data through Faraday's Law, Eq. 1. While the polyfit works great for the general shape of a GMD storm, to model the maximums and minimums with polyfit more accurately, a piece-wise function is needed. A challenge in analyzing GMD storms are directional effects, where X has been shown to have more fluctuations in magnetic field data, while Y has less fluctuations with higher maximums, and Z shares aspects to a lesser extent. Lastly is the importance of location, where the closer a place is to the magnetic poles the stronger the effects of GMD storms, and increased disturbances [9]. The different strengths of disturbances in the magnetic field create challenges when scaling the polynomial fit, for temporal and spatial features. Overall, the polyfit is helpful in obtaining more features related to GMD storms, and reveals aspects of GMDs to look into for future methods.

B. Future Works

As mentioned in section IV, location is important and plays a vital role in GMD and GIC analysis. The future work is to do a study on the topography of locations, correlation analysis between distance and magnetic field data between magnetometer locations, and looking into ground conductivity as an additional feature. Through such analyses, it would be made clear how exactly location and topography play a role in GMD storms, and determining places most susceptible to GICs.

On the thought of modeling though, there are many more mathematical modeling techniques to implement that could yield positive results, such as a sinusoid that changes in frequency, and peak values to match the magnetic field waveforms. Aside from mathematical modeling, there is machine learning, and physics-based modeling. If no single method creates a highly correlated curve, a combination of multiple methods might yield the best results. Modeling the storms is good, but unless measures are taken to protect the grid, the models wouldn't matter as much for power systems. There is a plan to look into ways to reduce GICs, harmonics, and damage to grid equipment.

All of these methods and ideas are to help gather data to better understand and analyze GICs and GMD impacts to protect the power grid. Understanding 'how' and 'why' something happens is key in creating plans to protect the grid from future events. The 'why' is already understood, as the sun emits coronal mass ejections which disturb Earth's magnetic field, inducing an electric field, creating a quasi-dc current. The 'how' it happens is the difficult part, as there are many factors that play a role in GMD events, and while some of these factors have been discussed, there always seem to be more features that appear. The goal is to find the most impactful features of GMD storms, to understand the 'how', and create ways to protect the power grid.

References

[1] T. J. Overbye, K. S. Shetye, T. R. Hutchins, Q. Qiu, and J. D. Weber, "Power grid sensitivity analysis of geomagnetically induced currents," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4821–4828, 2013.

[2] J. L. Wert, P. Dehghanian, J. Snodgrass, and T. J. Overbye, "The effects of correctly modeling generator step-up transformer status in geomagnetic disturbance studies," *arXiv preprint arXiv:2208.11232*, 2022.

[3] J. J. Love, "Magnetic monitoring of earth and space," *Physics Today*, vol. 61, no. 2, p. 31, 2008.

[4] "Numpy.polyfit." [Online]. Available: https://numpy.org/doc/stable/reference/generated/numpy.polyfit.html

[5] Y. Feldstein and G. Starkov, "Dynamics of auroral belt and polar geomagnetic disturbances," *Planetary and Space Science*, vol. 15, no. 2, pp. 209–229, 1967. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0032063367901900

[6] P. Dehghanian and T. J. Overbye, "Temperature-triggered failure hazard mitigation of transformers subject to geomagnetic disturbances," in *2021 IEEE Texas Power and Energy Conference (TPEC)*, 2021, pp. 1–6.

[7] "NOAA space weather scales." [Online]. Available: https://www.swpc.noaa.gov/noaa-scales-explanation

[8] E. Ostertagová, "Modelling using polynomial regression," *Procedia Engineering*, vol. 48, pp. 500–506, 2012.

[9] S. D. Team, "Benchmark geomagnetic disturbance event," 2016.

[10] C. Gaunt and G. Coetzee, "Transformer failures in regions incorrectly considered to have low gic-risk," in *2007 IEEE Lausanne power tech*. IEEE, 2007, pp. 807–812.

[11] I. contributors, "International real-time magnetic observatory network." [Online]. Available: https://intermagnet.github.io/