The influence of geomagnetic storms on magnetotelluric data analysis

Hao Chen (✉ chenhaomagnetic@gmail.com)

Department of Earth Resources Engineering, Graduate School of Engineering, Kyushu University, Fukuoka 819-0395, Japan

Hideki Mizunaga

Department of Earth Resources Engineering, Faculty of Engineering, Kyushu University, Fukuoka 819-0395, Japan

Toshiaki Tanaka

Department of Earth Resources Engineering, Faculty of Engineering, Kyushu University, Fukuoka 819-0395, Japan

Research Article

Keywords: geomagnetic storm, impedance tensor, magnetotelluric

Posted Date: January 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-144983/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The influence of geomagnetic storms on magnetotelluric data analysis

Hao Chen¹⁾, Hideki Mizunaga²⁾ and Toshiaki Tanaka²⁾

1) Department of Earth Resources Engineering, Graduate School of Engineering, Kyushu University, Fukuoka 819-0395, Japan

2) Department of Earth Resources Engineering, Faculty of Engineering, Kyushu University, Fukuoka 819-0395, Japan

Hao Chen: chenhaomagnetic@gmail.com

Hideki Mizunaga: mizunaga@mine.kyushu-u.ac.jp

Toshiaki Tanaka: tanaka@mine.kyushu-u.ac.jp

Abstract: The geomagnetic storm is a temporary disturbance of the earth's magnetosphere caused by a solar wind shock wave interacts with the earth's magnetic field. It is rarely acquired in the practical magnetotelluric (MT) survey. The rare MT researcher pays attention to the influence of geomagnetic storms on the MT data. MT data include the natural electromagnetic signals and artificial noises (instrumental, humanmade, and so on). Therefore, not all the time series contain usable information about the electrical conductivity distribution at depth, particularly when the signal-to-noise ratio is low. However, the signal-to-noise ratio will increase when there is a geomagnetic storm. In this paper, we focus on research the influences of the geomagnetic storm on MT data. Three case studies were demonstrated to show the positive effect of the geomagnetic storm on MT data. As a result, we could obtain reliable MT impedances at the noisy site using the geomagnetic storm data. It is difficult to get a reliable impedance tensor under electromagnetic environments contaminated by continuous noise. Therefore, predicting the geomagnetic storm by the space weather forecast before acquiring the MT data is effective. Utilizing the MT data during a geomagnetic storm may get a reliable result at the site contaminated by the continuous noise.

25 Keyword: geomagnetic storm, impedance tensor, magnetotelluric

26

27 **1 INTRODUCTION**

28 The magnetotelluric method is a passive electromagnetic geophysical method for inferring the
29 earth's subsurface electrical conductivity from measurements of natural geomagnetic and
30 geoelectric field variation on the earth's surface. It was proposed by Cagniard (1953) and
31 Tikhonov (1950) first. The natural MT sources from the earth's magnetosphere and ionosphere
32 or global lighting are far enough from the observation site. Therefore, we can assume the
33 electromagnetic signals as plane waves. Many works have researched the earth's
34 electromagnetic environment (Constable, 2016; Constable and Constable, 2004; McPherron,
35 2005). The general speaking that the lower frequencies signal (< 1 Hz) is originated from the
36 interaction of the solar wind with the earth's magnetosphere and ionosphere. The high-
37 frequency signal (>1 Hz) is originated from the worldwide thunderstorm activity. Constable
38 (2016) reviewed the electromagnetic sources in high frequency (>1 Hz). McPherron (2005)
39 reviewed the ultra-low frequency (ULF) band electromagnetic source (< 1 Hz). Garcia (2002)
40 also used the MT data to research the sources in the high-frequency domain. The geomagnetic
41 storm is a temporary disturbance of the earth's magnetosphere caused by a solar wind shock
42 wave interacts with the earth's magnetic field. In this paper, we will research the influence of
43 the geomagnetic storm on the MT data

44 The first step in MT data processing is to estimate the frequency domain's impedance
45 tensor from the measured time-series data. The field data include natural electromagnetic
46 signals and noises. We can subdivide the noises and signals into three parts. We can rewrite
47 the fields **E** and **H** as follow:

48

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

49
$$\mathbf{E} = \mathbf{E}^{MT} + \mathbf{E}^{CN} + \mathbf{E}^{UN},$$

50
$$\mathbf{H} = \mathbf{H}^{MT} + \mathbf{H}^{CN} + \mathbf{H}^{UN}, \tag{1}$$

51 where the CN denotes the correlated noise that the noise is coherent between electric
52 and magnetic fields, the UN denotes uncorrelated noise, and the MT denotes electromagnetic
53 signals (Oettinger et al., 2001).

54 Szarka (1988) and Junge (1996) summarized the active and passive noise sources observed
55 in MT measurements. Not all MT time-series include usable information about the electrical
56 conductivity distribution at depth, particularly in low signal-to-noise ratio. It can occur when
57 the natural signal level is comparable to or below the instrument noise level or in the presence
58 of some types of cultural noise. The low signal-to-noise ratio data can be regarded as noisy
59 data. Robust procedures only can get reliable results from a reasonable proportion of the noisy
60 data, typically not more than 40-50% (Smirnov, 2003).

61 A geomagnetic storm starts when enhanced energy comes from the solar wind into the
62 magnetosphere. During the geomagnetic storm, the variation of the MT signal increases
63 significantly. Sometimes the amplitude of MT signals in the storm period can be 100 times that
64 of the non-storm period. The noise can be neglected at that condition, and the noisy data
65 segment converts to high signal-to-noise ratio data but depending on the strength of the
66 geomagnetic storm and the noise. In this paper, a statistical analysis of the geomagnetic storm
67 is performed at first in section 2. Section 3 introduces the impedance estimator, Bounded
68 Influence Remote Reference Processing (BIRRP), which was used to calculate the impedance
69 tensor. Section 4 shows three case studies and an analysis of the MT data influenced by the
70 geomagnetic storm. In section 5, the practical meaning of this research is discussed.
71 Finally, we suggest that it is better to perform MT measurement in Spring and Fall from the
72 signal-to-noise ratio. It is not suitable to carry out the MT survey around 2021 because the

probability of a geomagnetic storm is the lowest. When there are continuous noises, it is not easy to get a reliable result from the current method. Therefore, we should try to acquire the MT data during the geomagnetic storm. Though the geomagnetic storm doesn't occur frequently, we could predict the geomagnetic storm using the space weather forecast and acquire the MT data during the geomagnetic storm period. Using storm period data may bring a reliable result from the contaminated data with continuous noises.

79

80 2 GEOMAGNETIC STORM

81 The geomagnetic storm is a temporary disturbance of the earth's magnetosphere caused by a solar wind shock wave interacts with the earth's magnetic field. Geomagnetic storms start when enhanced energy transfer from the interplanetary magnetic field (IMF) of the solar wind into the magnetosphere leads to the magnetospheric ring current's intensification. A magnetic storm is seen as a rapid drop in the earth's magnetic field strength at the earth's surface. Fig.1 shows the X (E-W) component of geomagnetic data during a storm day and a quiet day from Kakioka(KAK) station in Japan. In 1973, the KAK Magnetic Observatory was designated as one of four facilities worldwide to calculate the Dst index representing the strength of the equatorial ring current encircling the earth. The blue line denotes the non-storm day's data, and the red line denotes the storm day. The left figure shows the data in the time domain, and the right figure shows the data in the frequency domain. The intensity on storm day can be almost twice-order stronger than that on the non-storm day.

93

Fig.1 The geomagnetic intensities along the E-W direction during a storm-day and a quiet-day. The blue line denotes the data in the non-storm day. The red line denotes the storm day. The left is a profile in the time domain, and the right is a profile in the frequency domain.

The Disturbance Storm Time (Dst) index is a negative index of geomagnetic activity used to estimates the averaged change of the horizontal component of the earth's magnetic field based on measurements from a few magnetometer stations. It is derived from hourly scalings of low-latitude horizontal magnetic variation and expressed in nanoteslas. When the Dst index is less than -50 nT, it is categorized as a geomagnetic storm. When the Dst index is less than -100 nT, it is categorized as a strong geomagnetic storm. In this section, we analyzed the geomagnetic storm event statistically by the Dst index.

Fig.2 shows the distribution of the Dst index from 1957 to 2020; the orange line denotes the boundary of the geomagnetic storm ($Dst \leq -50$ nT), and the light blue line denotes the boundary of the strong geomagnetic storm ($Dst \leq -100$ nT). It shows that geomagnetic storms did not appear frequently. The probability of a strong storm is less than 1%.

Fig.3 shows the duration of each strong geomagnetic storm event by the hour. The upper figure shows the number of strong geomagnetic storm events versus the strong storm event's

length by the hour. The horizontal axis denotes the length of the strong storm event, and the y-axis denotes the number. The lower figure shows the cumulative distribution of the upper figure. It shows that about 46% of strong geomagnetic events lasted more than 4 hours, and 8% of the strong geomagnetic event lasted more than one day. The longest strong geomagnetic event lasted 55 hours. There are 688 strong geomagnetic storms from 1957 to 2020; one year had about ten strong geomagnetic events, and about five events lasted more than 4 hours on average.

Fig.4 shows the monthly count of strong geomagnetic storms. One hour was recorded as one count in this figure. The high probability of a strong geomagnetic storm occurred in Spring and Fall. Fig.5 shows the yearly count of geomagnetic storms that occurred each year. Fig.6 shows the FFT result of the yearly count of storms from 1957 to 2020. There is a 10.7-year peak, which corresponds to the 11-year solar cycle. This section concludes that the geomagnetic storm has a seasonal and 11-year solar cycle and causes significant EM field variations observed on the earth's surface.

Fig.2 The distribution of strong storms based on the Dst index between 1957-2020, the orange line denotes Dst (≤ -50 nT), and light blue line denotes Dst (≤ -100 nT).

Fig.3 The duration of each strong geomagnetic storm event. The upper figure shows the number of each strong geomagnetic storm event in a different storm event length. The lower figure shows the cumulative distribution of the upper figure.

Fig.4 The monthly count of strong geomagnetic storms based on the Dst index.

Fig.5 The yearly count of geomagnetic storms based on the Dst index from 1957 to 2020.

Fig.6 The calculated periods after Fourier analysis using the yearly count of geomagnetic storms from 1957 to 2020.

3 METHOD

The first step in MT data processing is to estimate the impedances in the frequency domain from the measured time-series data. All of the MT data interpretation is based on the result of the impedance tensor.

In the MT theory, the magnetic field (**H**) and the electric field (**E**) has a linear relationship in

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

the frequency domain. The impedance tensor at a specific frequency can be calculated in the frequency domain as follow:

$$\begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix} = \begin{pmatrix} Z_{xx}(\omega) & Z_{xy}(\omega) \\ Z_{yx}(\omega) & Z_{yy}(\omega) \end{pmatrix} \begin{pmatrix} H_x(\omega) \\ H_y(\omega) \end{pmatrix}. \quad (2)$$

This section introduces the impedance estimators (BIRRP) used in section 4.

3.1 BIRRP

Bounded Influence Remote Reference Processing, BIRRP (Chave and Thomson, 2004) is a typical conventional robust estimator to calculate the impedance tensor based on windowed FFT.

The first estimator (least-square estimator; Sims et al., 1971) is based on the least-squares theory, and the solution of impedance tensor at a specific frequency is given by:

$$\mathbf{Z} = (\mathbf{H}^\dagger \mathbf{H})^{-1} (\mathbf{H}^\dagger \mathbf{E}), \quad (3)$$

where $\mathbf{E}$ and $\mathbf{H}$ are the horizontal electric and magnetic field components at a specific frequency, and $\mathbf{Z}$ is the MT impedance tensor. The superscript $\dagger$ denotes the complex conjugate transpose of the matrix. The M-estimator gives a weight $\mathbf{w}$ to the outlier based on the residual between the output (electric field) of the least-square estimator and the observation data. And the solution of M-estimator at a specific frequency is given by:

$$\mathbf{Z} = (\mathbf{H}^\dagger \mathbf{wH})^{-1} (\mathbf{H}^\dagger \mathbf{wE}). \quad (4)$$

M-estimator can reduce the influence of unusual data (outliers) in the electric field but are not sensitive to exceptional input (magnetic field) data, which are termed leverage points. The hat matrix is widely used to detect unusual input data. The hat matrix is N by N matrix and defined as follows:

$$\mathbf{H}_{hat} = \mathbf{H}(\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger. \quad (5)$$

Chave (2003) suggested that the hat matrix's diagonal element, which is more than several

times $2/N$, is problematic. N denotes the number of data. The bounded influence (BI) estimator combines the robust M-estimator with Leverage weighting ($\mathbf{v}$) based on the statistics of the hat matrix diagonal element (Chave et al., 2004). And the solution of BI-estimator at a specific frequency is given by:

$$\mathbf{Z}=(\mathbf{H}^{\dagger} \mathbf{w} \mathbf{v} \mathbf{H})^{-1}(\mathbf{H}^{\dagger} \mathbf{w} \mathbf{v} \mathbf{E}). \quad (6)$$

The detail of the M-estimator and BI estimator was described by Chave (2012).

The remote reference processing can improve the estimator's performance using the cross-spectral instead of the auto-spectral when performing the regression based on the least-squares estimator (Gamble et al., 1979). And the solution of remote reference based on BI-estimator at a specific frequency is given by:

$$\mathbf{Z}=(\mathbf{H}_r^{\dagger} \mathbf{w} \mathbf{v} \mathbf{H})^{-1}(\mathbf{H}_r^{\dagger} \mathbf{w} \mathbf{v} \mathbf{E}), \quad (7)$$

where $\mathbf{H}_r$ denotes the remote reference magnetic field. BIRRP is an abbreviation for the Bounded Influence Remote Reference Processing method.

The error bar of the BIRRP code is calculated by jackknife at a specified frequency. The jackknife is a resampling method used to estimate the bias of a large population. It involves a leave-one-out strategy to estimate the error bar for the data set.

4 CASE STUDIES

This section shows three case studies to research the influence of geomagnetic storms on the MT data. Fig.7 shows the map of site locations in the three case studies (Sawauchi, USArray, KAP03). The left map shows the detailed site location used in USArray, and the right map shows the detailed survey line for KAP03. All of the case studies include geomagnetic storm data, the detailed information on those data described in each case study.

Fig.7 The location map of three sites in the case studies (KAP03, USArray, Sawauchi). The left map shows the detailed site location used in USArray, and the right map shows the detailed survey line for KAP03.

4.1 Case Study 1: Sawauchi, Japan

In the first case study, the broadband frequency 5-component magnetotelluric time-series data observed by the Phoenix Geophysics system was used. The data was observed from Aug. 20 to Aug. 29, 2018, at Sawauchi station, Japan. The geomagnetic storm occurred on Aug. 26. The MT time-series data are stored in three files. Two files store the high- and middle- frequency bands (2,400 and 150 Hz); the other file stored the low-frequency data (15 Hz) continuously. The high-frequency band (2,400 Hz) sampled 1 second at the interval of 4 minutes from the beginning of the minute, the middle-frequency band (150 Hz) sampled 16 seconds at the interval of 4 minutes from the beginning of the minute.

At first, we analyzed the spectrum variation along with the Dst index. To obtain robust spectral information from these data sets, we first applied a set of Slepian tapers and then used a fast Fourier transform algorithm to the time-series (Garcia and Jones, 2002). Fig.8 shows the

time-frequency distribution against the Dst index and the Hx component time-series data. The sampling rate is 15 Hz, and the upper figure shows the spectrum variation from Aug. 20 to Aug. 29. The color denotes the value of $10 \cdot \log_{10}(\text{amp.})$, and the "amp" denotes the spectrum's amplitude. The lower figure shows the Hx component time-series along with the Dst index. This figure shows that the amplitude increase correlates with the geomagnetic storm.

Fig.8 The time-frequency distribution against the Dst index variation and the sampling rate is 15 Hz. The upper figure shows the time-frequency distribution from Aug. 20 to Aug. 29. The color denotes the value of $10 \cdot \log_{10}(\text{amp.})$. The lower figure shows the time variations of the Hx component and the Dst index.

Fig.9 The magnetotelluric sounding curves during both the storm day (Oct. 26) and non-storm day (Oct. 23). The blue markers indicate impedance tensor on the non-storm day; the yellow markers indicate impedance tensor on the storm day.

Fig.10 The result of the TE mode impedance curve calculated by each day data at the period of 75-second. The horizontal axis denotes the date. The upper figures show the apparent resistivity, and the lower figures show the impedance phase.

Next, we computed the impedance tensor on each day. Fig.9 shows the typical magnetotelluric sounding curves calculated during both the storm day (Oct. 26) and non-storm day (Oct. 23). The non-storm day data generated the blue curve; the storm day data generated the yellow curve. The result calculated using the storm day data in the TE mode was much stable

than that of the non-storm day during the periods from 100 to 1,000 seconds. And TM mode is almost the same.

Fig.10 shows the result of the TE mode impedance curve calculated using each day data at the period of 75-second. The horizontal axis denotes the date. The upper figures show the apparent resistivities, and the lower figures show the impedance phases. Usually, the result does not change with time. The result of Fig.10 is almost the same, which means there was no strong coherent noise in the data site. And it is considered the closest to the true solution.

Fig.11 shows the variation of the TE mode impedance curve calculated by each hour of time-series data at the period of 75-second. We use one-hour time-series data to calculate one result. The result shows that the signal to noise ratio is varied hour by hour. The impedance curve's result becomes stable and close to the true model correlates with the geomagnetic storm event. This result shows that the signal to noise ratio increase during the storm period and has a positive influence on the calculation of MT impedance tensor.

Fig.11 The time variation of the impedance and phase curves, which was calculated using each hour time-series data at the period of 75-second. The horizontal axis denotes the date. One result was calculated using one-hour data.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

259

260 Fig.12 shows the time-frequency distribution against the Dst index. The sampling rate is 150

261 Hz, and the content is the same as Fig.8. There were no obvious changes that were correlated

262 with the storms in this figure. But distinct peaks appeared around 7.83, 14.3, 20.8 and 27.3 Hz.

263 These frequencies correspond to the frequencies of Schumann resonances (SR). SR is a set of

264 spectrum peaks in the extremely low frequency (ELF) of the earth's electromagnetic field

265 spectrum. Lightning discharges generate such global electromagnetic resonances in the cavity

266 formed by the earth's surface and the ionosphere.

267 Fig.13 shows the amplitude variation in periods 16, 8, 4 and 1 s along with the Dst index. The

268 upper figure shows the amplitude variation from Aug. 20 to Aug. 29. And the lower figure shows

269 the time variation of the Hx component along with the Dst index. The amplitude increased in the

270 16, 8, 4 seconds correlated with the geomagnetic storm. In the variation of 1 second, there was

271 no increase correlated with the storm. Fig.14 shows the amplitude variation in frequencies 2, 4,

272 8 and 16 Hz along with the Dst index, and the content is the same as Fig.13. There is no

273 increasing correlation with the geomagnetic storm.

274 Fig.15 shows the time-frequency distribution against the Dst index. The sampling rate is 2,400

275 Hz, and the content is the same as Fig.8. There is also no increasing correlation with the

276 geomagnetic storm. In this figure, the high-frequency data have daily changes. It increased in

277 the local night time and decreased in the local day time. In general, the high-frequency signal (>1

278 Hz) is originated from worldwide thunderstorm activities. These signals are the most active part

279 of these frequency bands, known as ELF. During the daytime, the atmospheric conductivity is

280 greater than the nighttime because sunlight ionizes the atmosphere. The increased

281 conductivities result in a stronger attenuation of the high-frequency signal (Garcia and Jones,

282 2002). So the variations of high-frequency signal increase during the nighttime. From Fig. 12 to

Fig.15, we concluded that the geomagnetic storm did not influence the high-frequency (>1 Hz) impedances.

Fig.12 The time-frequency distribution against the Dst index. The sampling rate is 150 Hz. The content is the same as Fig.8.

Fig.13 The amplitude variation in periods 16, 8, 4 and 1 second against the Dst index. The upper figure shows the amplitude variation from Aug. 20 to Aug. 29. The lowest figure shows the time variation of the Hx component along with the Dst index variation.

Fig. 14 The amplitude variation in frequency 2, 4, 8, and 16 Hz against the Dst index. The content is the same as Fig.13.

Fig.15 The time-frequency distribution against the Dst index. The sampling rate is 2,400 Hz. The content is the same as Fig.8.

4.2 Case Study 2: USArray, USA

In the second case study, long-period 5-component magnetotelluric time-series data at two sites (ALW48 and TNU48) was used. The data sets were recorded with a 1-second sampling rate for around two weeks in 2015 from the USArray project. The geomagnetic storm occurred between June 22 and June 24.

This study compared four results calculated from site TNU48, using ALW48 as remote reference data. Table 1 shows the name of each result. All the results were calculated using the BIRRP program. The Quiet was calculated from a time-series from June 20 to June 22. The QuietRR is derived from a time-series from June 20 to June 22 and using ALW48 as the remote site. The storm was calculated from a time-series from June 22 to June 24. The StormRR was calculated from a time-series from June 22 to June 24 and using ALW48 as the remote site.

Table 1 The classification of results. The Quiet was calculated from the non-storm data. The QuietRR was calculated from the non-storm data and using ALW48 as the remote site. The storm was calculated from the storm data. The StormRR was calculated from the storm data and using ALW48 as the remote site.

	2015.06.20-06.22	06.22-06.24
TNU48	Quiet	Storm
TNU48- ALW48	QuietRR	StormRR

All the results are shown in Fig.16. The apparent resistivity of Quiet in the period from 8 to 30 seconds is severe down biased. And the long period of Quiet and QuietRR is scattered. The result calculated from the storm data is much stable compared with the result calculating by the non-storm day data. After compared all results, the StormRR result is the closest to a true structure. The result calculated by a single site from storm data is very close to the true model. From the case study, we can also see the positive influence of geomagnetic storms on MT impedance tensor calculation.

Fig.16 The magnetotelluric sounding curves which were calculated during both storm day and the non-storm day. The black curve denotes Quiet's result; the blue curve denotes the QuietRR's result; the red curve denotes the storm's result; the yellow curve denotes the StormRR's result.

4.3 Case Study 3: KAP03, South Africa

In the third case study, long-period 5-component magnetotelluric time-series data observed at the Kaapvaal 2003 (KAP03) was used. Data were recorded with a 5-second sampling rate for almost a month at each site using GSC LIMS systems in 2003 as a part of the SAMTEX project. The 26 long-period sites distributed in a NE-SW profile are shown in the lower right corner of Fig.7. Data for sites in the middle of the profile (KAP127-KAP145) were heavily contaminated by DC signals from the DC train line running between Kimberley and Johannesburg.

Fig.17 Time-series of magnetotelluric data whose sampling rate was 5 seconds at site 130. The red vertical lines show the data gaps, and the black lines show the 5-component MT data. The blue line shows the Dst index of magnetic data. The unit of the electric field is mV/km, and the magnetic field in nT.

Fig.17 shows the time-series MT data at site 130. The red vertical lines indicate the data gaps, and the black lines indicate the 5-component MT time-series data. The sampling rate is 5 seconds. The unit of the electric field is mV/km, and the magnetic field in nT. The blue line shows the Dst (disturbance storm time) index. In this data set, there were two geomagnetic storm events in the observation periods. Both of the storms lasted about two days. The first storm was from Oct. 29 to Oct. 31, 2003; the second storm was from Nov. 20 to Nov. 22, 2003. We used the different period time-series data of the KAP03 to analyze the geomagnetic storm's influence on the impedance tensor calculation. The result calculated by BIRRP at the noisy site 130 is shown first, and then the result calculated at a quiet site 160 and another noisy site 133 is shown.

Fig. 18 shows the magnetotelluric sounding curves during storm day and the non-storm day at site 130. The black curves were calculated using the non-storm day data from 00:00:00 on Oct. 22 to 00:00:00 on Oct. 25. The red curves were calculated using the first storm period data from

00:00:00 on Oct. 29 to 00:00:00 on Oct. 31. The blue curves were calculated using the second geomagnetic storm data from 06:00:00 on Nov. 20 to 06:00:00 on Nov. 21. The upper figures show the apparent resistivity. The lower figures show the impedance phase.

Using the non-storm data, we failed to recover a reliable result. The phase of the YX and XY components was close to 180° or 0°. That is the phenomenon of artificial noise (Zonge and Hughes, 1987); 180° or 0° would correspond to a dipole electric source, which could be caused by the train line. And the long-period of the XY component was highly perturbed.

In the XY component, we think the period below 200-second is improved and gets a reliable result using the storm data. The results calculated by the first and second storm data almost keep the same and change reasonable. But for the long period results, the result is biased. Coming back to Fig.17, the time-series data at site 130, we find the sine signal dominates the Ex. In general, the natural MT signal usually shows as a random waveform. The continuous noise was too strong to get a reliable result in the longer period data, even if we used the geomagnetic storm signal. On the other hand, the YX component is improved using the geomagnetic storm data. The MT sounding curve changes reasonably.

Fig.18 The magnetotelluric sounding curves calculated during both storm day and the non-storm day at site 130. The non-storm data calculated the black curve. The first storm data calculated the red curve; the second storm data calculated the blue curve; the upper figures show the apparent resistivity; the lower figures show the impedance phase.

Fig.19 shows the magnetotelluric sounding curves calculated by the different period data at site 130. The non-storm data from Oct. 21 to Oct. 31 is used to calculate the black result, including the storm data. The data from Nov. 1 to Nov. 4 is used to calculate the blue result. The green result is calculated by the data from Nov. 1 to Nov. 4, using the data from site 163 as the remote reference data. The storm data between Oct. 29 to Oct. 31 is used to calculate the red result. Compared to the result, the red (storm) and green (remote reference) result close to the true model.

Moreover, the result calculated using storm data is more reliable. The data set from 21 to 31 failed despite including the storm data. The proportion of high signal-to-noise ratio data was approximately 20%. In contrast, robust estimators need a reasonable noisy data ratio (typically 40-50% or less) to yield reliable results.

Fig.19 Magnetotelluric sounding curves calculated by the different period time series at site 130.

Both Fig.20 and 21 show the magnetotelluric sounding curves calculated both storm day and the non-storm day at site 163 and 133, respectively. The non-storm data was used to calculate the black curve; The storm data was used to calculate the red curve. Fig.20 shows the result calculated using the quiet data at site 163. Comparing the result, no matter there is a geomagnetic storm or not, the result is excellent. Fig.21 shows the result calculated using the noisy data at site 133. The XY phase calculated by non-storm day data is close to 0°, and the apparent resistivity increases as a line in the log scale. The impedance curves were improved by using the geomagnetic storm data, and the result became more reasonable than the curves using non-storm data.

Fig.20 The magnetotelluric sounding curves calculated during both storm day and the non-storm day at site 163. The storm data was used to calculate the red curve; the non-storm was used to calculate the black curve.

Fig.21 The magnetotelluric sounding curves calculated during both storm day and the non-storm day at site 133. The colors are the same as Fig.20.

5 DISCUSSION

MT field data include natural signal sources and noise sources (e.g., instrumental, humanmade, and so on). Not all time segments contain information on the depth resistivity distribution. Before MT data processing, noisy data should be considered in the impedance estimation. Both the amplitude of the signal and noise vary with time. We can consider a time segment as noise data if the signal-to-noise ratio is low.

In contrast, we can consider a time segment as noise-free data. By the method based on windowed FFT, we divide the whole time series into the segment, and one segment was used to calculate one data set when we estimate the impedance tensor in the frequency domain. Robust estimators need a reasonable noise ratio (typically 40-50% or less) to yield reliable results (Smirnov, 2003). When there is a geomagnetic storm, the high signal-to-noise ratio data will increase and improve the calculation precision dramatically.

The basic technique to calculate impedance is the least-squares estimator. In the presence of intermittent contaminated data, the result can be improved by the following method.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 429 1) Robust procedure: M-estimator will fail to reconstruct the reliable response function when it
- 430 contains more than 30% noise. The repeated median (RM) algorithm is no more than 50%
- 431 (Smirnov, 2003; Chave,1989; Chave,2003; Chave and Thomson, 2004).
- 432 2) Remote reference technique: Using the cross-spectral instead of the auto-spectral when
- 433 performing the regression based on the least-squares estimator (Gamble et al., 1979).
- 434 3) Multiple stations processing (Egbert, 1997).
- 435 4) Signal noise separation in the frequency domain (Oettinger et al., 2001).
- 436 5) Noise identification and modifying (Kappler, 2012; Li et al., 2018; Wang et al., 2017).
- 437 6) Using a selection scheme before impedance tensor estimation (Weckmann et al., 2005).
- 438 Those methods are possible to improve the result in the presence of intermittent
- 439 contaminated data. Along with urban constructions, artificial disturbances to the
- 440 electromagnetic observations are becoming more and more serious. The observation
- 441 occasionally contains continuous noise. That is difficult to get a reliable result from the current
- 442 technique. As an alternative way of the situation, it is possible to get a reliable result using the
- 443 time-series data between geomagnetic storms. Sometimes, the storm period variation can be
- 444 100 times greater than in the non-storm period data. In that condition, the noise can be
- 445 neglected. Though the geomagnetic storm doesn't occur frequently, it will happen every year. It
- 446 is effective if the MT method is measured during a geomagnetic storm. It is possible to predict
- 447 the geomagnetic storm by the space weather forecast information. The Space Weather
- 448 Prediction Center (SWPC; see the website in references) provides information about space
- 449 weather in the coming three days. Utilizing geomagnetic storm data may give us a reliable result
- 450 despite the site contaminated by continuous noise.
- 451 We should sample at least two points in one period by the sampling theory. But to get the
- 452 accurate complex coefficient from the time series. We suggest that it is better to sample 4

points in one period, and one continuous segment should contain at least four times the period. For 1,000 s, we need a 4,000 seconds time-series segment to get an accurate spectrum. The overlay rate is set to be 50% to keep each data's independence and get more sample data. By the continuous 4-hour time-series data, we may get about eight samples to do the impedance estimation in the frequency domain by FFT. That means if there is continuous 4-hour geomagnetic storm data. We may get a relatively reliable impedance tensor until 1,000 seconds. It just depends on the length of the geomagnetic storm. The longer the geomagnetic storm is, the more stable result we can get. By the result of statistical analysis of the geomagnetic storm, one year had about ten strong geomagnetic events, and about five events lasted more than 4 hours on average. That is practical, meaningful for MT exploration.

6 CONCLUSIONS

First, we carried out the statistical analysis of the geomagnetic storm. The geomagnetic storm has seasonal and 11-year solar-cycle fluctuations and causes significant signal amplitude variations. Each year, the geomagnetic storm probability is higher in Spring and Fall, especially April and October, than in other seasons. The geomagnetic storm probability will be the lowest around 2021, judging from the 11-year cycle. The probability of a strong-storm is less than 1% during all periods. However, the probability of a week-storm is greater than it of a strong-storm. On average, strong-storms happen about ten times in one year. 46% of strong-storm events lasted over 4 hours, and 8% of the strong-storm events lasted over a day. That means five events lasted longer than 4 hours each year on average. That is very meaningful for the calculation of MT impedance tensor.

Next, this paper introduced three case studies to show the positive influence of geomagnetic storms on MT impedance tensor calculation. In the first case study, we used the broadband-

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

frequency MT data observed using the Phoenix Geophysics system at Sawauchi station. We analyzed the TE mode impedance curve variation using each hour time-series data at the period of 75-second. The result of the geomagnetic storm was much stable than the other non-storm period. We also analyzed the time-frequency variation against the Dst index. We confirmed that the solar wind interaction with the earth's magnetosphere originated from a low-frequency signal (< 1 Hz). Therefore, the geomagnetic storm significantly influences the low-frequency signal; the low-frequency signal amplitude increases with the geomagnetic storm. There are seldom influences on the high-frequency data.

In the second case study, we used long-period 5-component MT time-series data from USArray at two sites (ALW48 and TNU48). We compared four results, using the TNU48 as a local site and ALW48 as the remote reference site between the storm and non-storm period data. The period from 8 to 30 seconds was down-biased significantly by the no-storm data. Besides, the result was improved effectively by using storm data. And the result became more stable, especially in a long period.

In the third case study, we used the impedance estimator BIRRP to analyze the MT data. From the result calculated from site 163. When the data is quiet, no matter if there is a geomagnetic storm or not, we can recover the true model. While from the result calculated from site 130 and 133, we failed to calculate the true model from the non-storm data. Due to the geomagnetic storm, the signal amplitude increased during the storm period. We could get a reliable result from the YX component. However, we failed to get a reliable result from the XY component because the Ex field noise was too strong. Reliable results depend on the signal-to-noise ratio of time-series data.

499
500

DECLARATION**Availability of data and materials**

The magnetic time-series data observed at kak station is downloaded from the INTERMAGNET (International Real-time Magnetic Observatory Network). The SAMTEX team and USArray team provided the long period time-series data to investigate. Nittetsu Mining Consultants Co., Ltd. provided the broadband frequency MT time-series data observed at Sawauchi, Japan. The WDC for Geomagnetism, Kyoto, provided the Dst index data to do the statistical analysis of the geomagnetic storm. Alan Chave provided the BIRRP code.

Competing interests

We know of no conflicts of interest associated with this publication. We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

Funding

'Not applicable.'

Authors' contributions

Hao Chen processed the time series data, created the result and wrote the paper. Hao Chen contributes about 60%. Hideki Mizunaga reviewed the paper and contributed about 30%; Toshiaki Tanaka contributed about 10% to this work.

Acknowledgments

1
2
3 524 We thank the INTERMAGNET (International Real-time Magnetic Observatory Network) for
4
5 525 providing the magnetic time-series data observed at the KAK station. We thank all SAMTEX and
6
7 526 USArray team members for providing the long period time-series data to investigate. We also
8
9 527 thank Nittetsu Mining Consultants Co., Ltd. for providing the broadband frequency MT time-
10
11 528 series data observed at Sawauchi, Japan. We thank the WDC for Geomagnetism, Kyoto, for
12
13 529 providing the Dst index data to do the statistical analysis of the geomagnetic storm. Finally, we
14
15 530 express special thanks to Maik Neukirch for his constructive comments to improve the
16
17 531 manuscript.
18
19

20 532

21
22 533 **REFERENCES**
23

24 534 Berdichevsky, M.N., 1973, Magnetotelluric sounding with applications to mathematical filters,
25
26 535 Phys. Earth, **3**, 76–92.
27
28 536 Cagniard, L., 1953, The basic theory of the magneto-telluric method of geophysical prospecting.
29
30 537 *Geophysics*, **18**, 605–635.
31
32 538 Constable, C., 2016, earth's electromagnetic environment. *Surv, Geophysics*, **37**, 27–45.
33
34 539 Constable, C. G. and Constable, S.C., 2004, Satellite magnetic field measurements: applications
35
36 540 in studying the deep earth. *State Planet Front. Chall. Geophysics*, **19**, 147.
37
38 541 Chave, A. D. and Thomson, D. J., 1989, Some comments on magnetotelluric response function
39
40 542 estimation, *J.Geophys.Res.*, **94**, 14215–14225.
41
42 543 Chave, A. D. and Thomson, D. J., 2003, A bounded influence regression estimator based on the
43
44 544 statistics of the hat matrix, *J. R. Stat. Soc., Ser. C*, **52**, 307–322.
45
46 545 Chave, A. D. and Thomson, D. J., 2004, Bounded influence estimation of magnetotelluric
47
48 546 response functions, *Geophys. J. Int*, **157**, 988–1006.
49
50 547 Chave, A. D. and Jones, A. G. eds., 2012, The magnetotelluric method: Theory and practice.
51
52 548 Cambridge University Press.
53
54
55
56
57
58
59
60

- 549 Egbert, G. D. and Booker, J. R., 1986, Robust estimation of geomagnetic transfer functions,
550 *Geophys. J. R. Astr. Soc.*, **87**, 173-194.
- 551 Egbert, G. D., 1997, Robust multiple-station magnetotelluric data processing, *Geophys. J. Int.*,
552 **130**, 475–496.
- 553 Gamble, T. D., Goubau, W. M. and Clarke, J., 1979, Magnetotellurics with a remote reference,
554 *Geophysics*, **44**, 53–68.
- 555 Garcia, X. and Jones, A. G., 2002, Atmospheric sources for audio-magnetotelluric (AMT)
556 sounding, *Geophysics*, **67**, 448-458.
- 557 Kappler, K. N., 2012, A data variance technique for automated despiking of magnetotelluric data
558 with a remote reference, *Geophys. Prospect.*, **60**, 179–191.
- 559 Li, J., Zhang, X., Gong, J., Tang, J., Ren, Z., Li, G., Deng, Y., Cai, J., 2018, Signal-noise identification
560 of magnetotelluric signals using fractal-entropy and clustering algorithm for targeted de-noising.
561 *Fractals* **26**, 1840011.
- 562 McPherron, R.L., 2005, Magnetic pulsations: their sources and relation to solar wind and
563 geomagnetic activity, *Surv. Geophysics*, **26**, 545–592.
- 564 Oettinger, G., Haak, V., Larsen, J. C., 2001, Noise reduction in magnetotelluric time-series with a
565 new signal–noise separation method and its application to a field experiment in the Saxonian
566 Granulite Massif, *Geophys. J. Int*, **146**, 659–669.
- 567 Sims, W. E., Bostick, F. X., and Smith, H. W., 1971, The estimation of magnetotelluric impedance
568 tensor elements from measured data, *Geophysics*, **36**, 938–942.
- 569 Smirnov, M. Y., 2003, Magnetotelluric data processing with a robust statistical procedure having
570 a high breakdown point, *Geophys. J. Int.*, **152**, 1–7.
- 571 SWPC (The Space Weather Prediction Center),
572 <https://www.swpc.noaa.gov/content/wmo/geomagnetic-activity>.

1
2
3 573 Tikhonov, A. N., 1950, On determining electrical characteristics of the deep layers of the earth's
4
5 574 crust, in Doklady, Citeseer, 295–297.
6
7 575 Weckmann, U., Magunia, A. and Ritter, O., 2005, Effective noise separation for magnetotelluric
8
9 576 single site data processing using a frequency domain selection scheme, *Geophys. J. Int.*, **161**,
10
11 577 635–652.
12
13
14 578 Wang, H., Campanyà, J., Cheng, J., Zhu, G., Wei, W., Jin, S., and Ye, G., 2017, Synthesis of natural
15
16 579 electric and magnetic Timeseries using Interstation transfer functions and time-series from a
17
18 580 Neighboring site (STIN): Applications for processing MT data, *J. Geophys. Res. Solid Earth*, **122**,
19
20 581 5835–5851.
21
22
23 582 Zonge, K. L. and Hughes, L. J., 1987, Controlled source audio-frequency magnetotellurics, in
24
25 583 Electromagnetic Methods in Applied Geophysics. Applications (ed. Nabighian, M. N.), SEG, 713–
26
27 584 809.
28
29
30 585
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figures

Figure 1

The geomagnetic intensities along the E-W direction during a storm-day and a quiet-day. The blue line denotes the data in the non-storm day. The red line denotes the storm day. The left is a profile in the time domain, and the right is a profile in the frequency domain.

Figure 2

The distribution of strong storms based on the Dst index between 1957-2020, the orange line denotes Dst (≤ -50 nT), and light blue line denotes Dst (≤ -100 nT).

Figure 3

The duration of each strong geomagnetic storm event. The upper figure shows the number of each strong geomagnetic storm event in a different storm event length. The lower figure shows the cumulative distribution of the upper figure.

Figure 4

The monthly count of strong geomagnetic storms based on the Dst index.

Figure 5

The yearly count of geomagnetic storms based on the Dst index from 1957 to2020.

Figure 6

The calculated periods after Fourier analysis using the yearly count of geomagnetic storms from 1957 to 2020.

Figure 7

The location map of three sites in the case studies (KAP03, USArray, Sawauchi). The left map shows the detailed site location used in USArray, and the right map shows the detailed survey line for KAP03. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Figure 8

The time-frequency distribution against the Dst index variation and the sampling rate is 15 Hz. The upper figure shows the time-frequency distribution from Aug. 20 to Aug. 29. The color denotes the value of $10 \cdot \log_{10}(\text{amp.})$. The lower figure shows the time variations of the Hx component and the Dst index.

Figure 9

The magnetotelluric sounding curves during both the storm day (Oct. 26) and non-storm day (Oct. 23). The blue markers indicate impedance tensor on the non-storm day; the yellow markers indicate impedance tensor on the storm day.

Figure 10

The result of the TE mode impedance curve calculated by each day data at the period of 75-second. The horizontal axis denotes the date. The upper figures show the apparent resistivity, and the lower figures show the impedance phase.

Figure 11

The time variation of the impedance and phase curves, which was calculated using each hour time-series data at the period of 75-second. The horizontal axis denotes the date. One result was calculated using one-hour data.

Figure 12

The time-frequency distribution against the Dst index. The sampling rate is 150 Hz. The content is the same as Fig.8.

Figure 13

The amplitude variation in periods 16, 8, 4 and 1 second against the Dst index. The upper figure shows the amplitude variation from Aug. 20 to Aug. 29. The lowest figure shows the time variation of the Hx component along with the Dst index variation.

Figure 14

The amplitude variation in frequency 2, 4, 8, and 16 Hz against the Dst index. The content is the same as Fig.13.

Figure 15

The time-frequency distribution against the Dst index. The sampling rate is 2,400 Hz. The content is the same as Fig.8.

Figure 16

The magnetotelluric sounding curves which were calculated during both storm day and the non-storm day. The black curve denotes Quiet's result; the blue curve denotes the QuietRR's result; the red curve denotes the storm's result; the yellow curve denotes the StormRR's result.

Figure 17

Time-series of magnetotelluric data whose sampling rate was 5 seconds at site 130. The red vertical lines show the data gaps, and the black lines show the 5-component MT data. The blue line shows the Dst index of magnetic data. The unit of the electric field is mV/km, and the magnetic field in nT.

Figure 18

The magnetotelluric sounding curves calculated during both storm day and the non-storm day at site 130. The non-storm data calculated the black curve. The first storm data calculated the red curve; the second storm data calculated the blue curve; the upper figures show the apparent resistivity; the lower figures show the impedance phase.

Figure 19

Magnetotelluric sounding curves calculated by the different period time series at site 130.

Figure 20

The magnetotelluric sounding curves calculated during both storm day and the non-storm day at site 163. The storm data was used to calculate the red curve; the non-storm was used to calculate the black curve.

Figure 21

The magnetotelluric sounding curves calculated during both storm day and the non-storm day at site 133. The colors are the same as Fig.20.