

# Theoretical Consideration on the Application of Continuous Wavelet Transform to Time-Series Processing for Magnetotelluric Data

Hiroki Ogawa (✉ [hiroki.ogawa@asagi.waseda.jp](mailto:hiroki.ogawa@asagi.waseda.jp))

Nihon Genshiryoku Kenkyu Kaihatsu Kiko Tono Chikagaku Center <https://orcid.org/0000-0001-7848-551X>

Yuki Hama

Nittetsu Mining Co., Ltd

Koichi Asamori

Japan Atomic Energy Agency

Takumi Ueda

Waseda University Faculty of Science and Engineering: Waseda Daigaku Riko Gakujutsuin

---

## Technical report

**Keywords:** magnetotelluric, time series, continuous wavelet transform, short-time Fourier transform, signal processing

**Posted Date:** September 8th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-72251/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

1 **Theoretical consideration on the application of continuous wavelet transform to**  
2 **time-series processing for magnetotelluric data**

3

4 • Hiroki OGAWA,

5 1 Department of Earth Sciences, Resources and Environmental Engineering, Graduate

6 School of Creative Science and Engineering, Waseda University, 3-4-1, Ookubo,

7 Shinjuku-ku, Tokyo, 169-8555, Japan

8 2 Nittetsu Mining Co., Ltd., Yusen Bldg., 2-3-2, Marunouchi, Chiyoda-ku, Tokyo, 100-

9 8377, Japan

10 3 Tono Geoscience Center, Japan Atomic Energy Agency (JAEA), 959-31, Jorinji,

11 Izumi-cho, Toki-shi, Gifu, 509-5102, Japan

12 [hiroki.ogawa@asagi.waseda.jp](mailto:hiroki.ogawa@asagi.waseda.jp) (corresponding author)

13 • Yuki HAMA,

14 2 Nittetsu Mining Co., Ltd., Yusen Bldg., 2-3-2, Marunouchi, Chiyoda-ku, Tokyo, 100-

15 8377, Japan

16 [hama@nittetsukou.co.jp](mailto:hama@nittetsukou.co.jp)

17     • Koichi ASAMORI,  
18     3 Tono Geoscience Center, Japan Atomic Energy Agency (JAEA), 959-31, Jorinji,  
19             Izumi-cho, Toki-shi, Gifu, 509-5102, Japan  
20     4 Nuclear Waste Management Organization of Japan (NUMO), Mita NN Bldg., 4-1-23,  
21             Shiba, Minato-ku, Tokyo, 108-0014, Japan  
22     kasamori@numo.or.jp  
23     • Takumi UEDA,  
24     1 Department of Earth Sciences, Resources and Environmental Engineering, Graduate  
25             School of Creative Science and Engineering, Waseda University, 3-4-1, Ookubo,  
26             Shinjuku-ku, Tokyo, 169-8555, Japan  
27     takumi.ueda@waseda.jp  
28

29 **Abstract**

30 In the magnetotelluric (MT) method, the responses of the natural electromagnetic fields  
31 are evaluated by transforming time-series data into spectral data and calculating the  
32 apparent resistivity and phase. The continuous wavelet transform (CWT) can be an  
33 alternative to the short-time Fourier transform, and the applicability of CWT to MT data  
34 has been reported. There are, however, few cases of considering the effect of numerical  
35 errors derived from spectral transform on MT data processing. In general, it is desirable  
36 to adopt a window function narrow in the time domain for higher-frequency  
37 components and one in the frequency domain for lower-frequency components. In  
38 conducting the short-time Fourier transform, because the size of the window function is  
39 fixed unless the time-series data are decimated, there might be difference between the  
40 calculated MT responses and the true ones due to the numerical errors. Meanwhile,  
41 CWT can strike a balance between the resolution of the time and frequency domains by  
42 magnifying or reducing the wavelet, according to the value of frequency. Although the  
43 types of wavelet functions and their parameters influence the resolution of time and  
44 frequency, those calculation settings of CWT are often determined empirically. In this

45 study, focusing on the frequency band between 0.001 Hz and 10 Hz, we demonstrated  
46 the superiority of utilizing CWT in MT data processing and determined its proper  
47 calculation settings in terms of restraining the numerical errors caused by the spectral  
48 transform of time-series data. The results obtained with the short-time Fourier transform  
49 accompanied with gradual decimation of the time-series data, called cascade  
50 decimation, were compared with those of CWT. The shape of the wavelet was changed  
51 by using different types of wavelet functions or their parameters, and the respective  
52 results of data processing were compared. Through these experiments, this study  
53 indicates that CWT with the complex Morlet function with its wavelet parameter  $k$  set  
54 to  $6 \leq k < 10$  will be effective in restraining the numerical errors caused by the spectral  
55 transform.

56

57 **Keywords**

58 magnetotelluric, time series, continuous wavelet transform, short-time Fourier  
59 transform, signal processing

60

## 61 **Introduction**

62 The magnetotelluric (MT) method is an effective geophysical tool for investigating the  
63 resistivity structure of the subsurface (e.g., Chave and Thomson, 2004). It is usual to  
64 evaluate the responses of the natural electromagnetic fields by transforming observed  
65 time-series data into spectral data and obtaining the apparent resistivity and phase  
66 profiles (Vozoff, 1990). The short-time Fourier transform (STFT) has been widely used  
67 to divide the time-series data into several small segments and calculate the spectral data  
68 in individual segments (e.g., Chave and Thomson, 1987). As the natural MT signal can  
69 be assumed to be non-stationary (Escalas et al., 2013), the time-varying frequency  
70 components are evaluated by means of STFT. Meanwhile, the continuous wavelet  
71 transform (CWT) can be an alternative method of time-frequency analysis for non-  
72 stationary signals. The applicability of CWT to MT data has been reported since the  
73 theory of CWT was devised. Zhang and Paulson (1997) focused on the frequency band  
74 (approximately 10 Hz – 10 kHz) where lightning activity is dominant and were the first  
75 to use CWT for audiomagnetotellurics (AMT) data processing. They proposed a  
76 technique that applied CWT to calculating the spectra in each segment and selected

77 those showing a high signal-to-noise ratio. They reported the effectiveness of their  
78 proposed technique in obtaining AMT responses. Garcia and Jones (2008) utilized  
79 CWT to precisely detect AMT responses in the band of 1 kHz through 5 kHz, called the  
80 AMT dead band, where the natural signal significantly attenuates. They proposed a  
81 method that contained both CWT and robust estimation algorithm. Larnier et al. (2016)  
82 processed time-series data by applying different types of wavelet functions to the higher  
83 band (above about 7 Hz) and the lower band (below 0.1 Hz) respectively. They showed  
84 improvement in the quality of the spectral data by comparing the results of their  
85 proposed method with those of the BIRRP algorithm (Chave and Thomson, 2004).

86 There are, however, few cases in which the effect of numerical errors derived from the  
87 spectral transform on MT data processing has been examined. It is desirable to adopt a  
88 window function narrow in the time domain for higher-frequency components that have  
89 short periods so that the electromagnetic signal can be observed to oscillate fractionally.  
90 Whereas, a window function narrow in the frequency domain should be adopted for  
91 lower-frequency components that oscillate smoothly, which enables us to accurately  
92 differentiate the frequency values. However, owing to the uncertainty principle, the

93 resolution of time and frequency cannot be enhanced simultaneously (Cohen, 1995).  
94 Thus, there might be difference between the calculated MT responses and the true  
95 responses due to the numerical errors caused by transforming time-domain data into  
96 frequency-domain data. When using STFT, processing is conducted with a window of  
97 constant size unless the time-series data are decimated (Farge, 1992). CWT is intended  
98 to strike a balance between the resolution of time and frequency by magnifying or  
99 reducing its wavelet, according to the value of frequency (Farge, 1992). In addition, the  
100 types of the wavelet functions and their parameters are often determined empirically.  
101 The shape of the wavelet determined by such calculation settings influences the time  
102 and frequency resolution (De Moortel et al., 2004). In this study, we proposed the  
103 superiority of CWT in MT data processing and its proper calculation settings in terms of  
104 restraining the numerical errors caused by the spectral transform of time-series data.  
105 This study focused on the frequency band approximately between 0.001 Hz and 10 Hz,  
106 which corresponds to the ultra-low frequency (ULF) band originating from pulsations.

107

108 **Theoretical estimation**

109 Let us consider a Gaussian function  $g(t)$  that vanishes fast at positive and negative  
110 infinity as follows:

$$111 \quad g(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}} \quad (1)$$

112 The STFT coefficient of a time-series function  $f(t)$ , denoted by  $G(\tau, \omega)$ , is defined as  
113 the convolution of  $f(t)$  with  $g(t)$  (Ishikawa and Arai, 2014):

$$114 \quad G(\tau, \omega) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\tau)^2}{2\sigma^2}} e^{-i\omega t} dt. \quad (2)$$

115 The width of  $g(t)$  is determined by the standard deviation  $\sigma$ , which is independent of  
116 the frequency value  $\omega$ . In conducting STFT, therefore, the size of the window function  
117 is constantly fixed regardless of the frequency value, and the resolution of time  $\Delta t$  and  
118 that of frequency  $\Delta\omega$  should be invariant.

119 Among the methods of the wavelet transform which balances the resolution of the time  
120 and frequency domains, CWT is generally superior in detecting time variation of  
121 individual frequency components (Rhif et al., 2019). Thus, CWT is highly compatible  
122 with the non-stationary MT signal. A function  $\psi(t)$  called “wavelet”, which  
123 corresponds to the window function  $g(t)$  in Equation (1), is introduced in Equation  
124 (3),

125 
$$\psi(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}} e^{ikt}, \quad (3)$$

126 where  $k$  controls the number of oscillations present in the wavelet and defines the  
 127 shape of the wavelet being used. Though  $\psi(t)$  is independent of the frequency, the  
 128 function derived by translating  $\psi(t)$  along the time axis or dilating  $\psi(t)$  can be  
 129 expressed as follows:

130 
$$\psi_{b,s} = s^{-1/2} \psi\left(\frac{t-b}{s}\right), \quad (4)$$

131 where  $s$  is a dilation parameter, which is inversely proportional to the frequency value,  
 132 and  $b$  denotes the central position of the wavelet along the time axis. As the value of  $s$   
 133 increases, the wavelet is magnified in the time domain. To normalize the wavelet  
 134 function at each frequency, the translated and dilated function is divided by the square  
 135 root of  $s$  (Torrence and Compo, 1998). Using these equations, the CWT coefficient is  
 136 expressed as:

137 
$$Wf(b,s) = \int_{-\infty}^{\infty} f(t) s^{-1/2} \psi^*\left(\frac{t-b}{s}\right) dt$$

138 
$$= s^{-1/2} \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{-(t-b)^2}{2\sigma^2 s^2}} e^{-ik\left(\frac{t-b}{s}\right)} dt.$$

139 (5)

140 Note that  $*$  denotes a complex conjugate. Both  $\Delta t$  and  $\Delta \omega$  change depending on the

141 frequency value because the size of the wavelet is variable. Figure 1 shows the  
 142 conceptual diagrams of time-series processing using STFT and CWT.

143 The remote reference method (Gamble et al., 1979a) is often used to reduce noise  
 144 included in the magnetotelluric impedance tensors by calculating the cross-correlation  
 145 between the electromagnetic components at a local site and those at a reference site.

146 Suppose that the electric field or the magnetic field at each site can be classified into the  
 147 natural signal  $FS$  and the noise  $FN$  and be represented by  $F_i = FS_i + FN_i$  ( $i=1$ : the  
 148 local site  $i=2$ : the reference site). Considering the numerical errors caused by the  
 149 transformation from the time domain to the frequency domain, each electromagnetic  
 150 component can be expressed as follows:

$$151 \quad F_i = (FS_{iT} + FS_{iE}) + (FN_{iT} + FN_{iE}), \quad (6)$$

152 where subscript T denotes the true value and subscript E denotes the numerical error.

153 The cross-power spectrum between the electromagnetic components at the local site and  
 154 at the reference site is represented as:

$$155 \quad \langle F_1 F_2^* \rangle = \langle [(FS_{1T} + FS_{1E}) + (FN_{1T} + FN_{1E})] \cdot [(FS_{2T}^* + FS_{2E}^*) + (FN_{2T}^* + FN_{2E}^*)] \rangle$$

$$156 \quad = \langle [FS_{1T} \cdot FS_{2T}^* + \underline{FS_{1T} \cdot FS_{2E}^* + FS_{1E} \cdot FS_{2T}^* + FS_{1E} \cdot FS_{2E}^*}] + \dots \dots \rangle. \quad (7)$$

157 In the case where the numerical errors (underlined terms with subscript E in Equation  
158 (7)) included in FS are not negligible or where they have a high correlation with the true  
159 value of the natural signal, the calculated MT responses might be different from the true  
160 responses that reflect the correct information on the subsurface structure. In the  
161 frequency domain, error  $\delta\omega$  occurred by obtaining the frequency components is given  
162 uniquely once the width of the time window is determined as  $T$ . Lower-frequency  
163 components are subject to the effect of  $\delta\omega$ , and it becomes difficult to distinguish the  
164 value of  $\omega$  correctly. In the time domain, meanwhile, the width of the time window  
165 should be smaller as the frequency being analyzed increases; otherwise, the shorter-  
166 period signal that oscillates fractionally might be difficult to detect accurately.

167 Based on the above, we conducted several numerical experiments on MT data  
168 processing, considering the effect of the numerical errors caused by the spectral  
169 transform on the results such as the apparent resistivity and the phase.

170

### 171 **Data processing using CWT for ULF signal**

172 Based on Equation (5), we applied CWT to processing of MT time series to derive the

173 spectra in the ULF band. Equation (5) can be written in the frequency domain as

174 follows:

$$175 \quad Wf(b, s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \Psi_{b,s}^*(\omega) e^{ib\omega} d\omega. \quad (8)$$

176 In practice, the convolution of  $f(t)$  with  $\psi(t)$  was accelerated by discretizing

177 Equation (5) and utilizing both the fast Fourier transform (FFT) and the inverse fast

178 Fourier transform (IFFT), following the technique of Torrence and Compo (1998). Note

179 that  $F(\omega)$  and  $\Psi(\omega)$  represent  $f(t)$  and  $\psi(t)$  in the frequency domain,

180 respectively.

181 MT data at site 001 and site 002 (Figure 2), both of which were in Kirishima City,

182 Kagoshima Prefecture, were acquired in January 2015. As each site was located in a

183 forest zone where there was hardly any electromagnetic noise source, their sufficiently

184 high signal-to-noise ratios enabled us to estimate the true MT responses easily. The

185 remote reference method was applied to calculation of the spectra using horizontal

186 magnetic components acquired at a reference site. The reference site where

187 geomagnetic data were constantly observed was located in Nishiwaga Town, Iwate

188 Prefecture, approximately 1250 km away from Kirishima City. All measurements were

189 conducted with an MTU system built by Phoenix Geophysics Ltd. Two horizontal  
190 electric components ( $E_x$ ,  $E_y$ ) were measured using Pb-PbCl<sub>2</sub> non-polarizing electrodes,  
191 and two horizontal magnetic components ( $H_x$ ,  $H_y$ ) were measured using MTC  
192 induction coil sensors. The specifications of the observations at three MT sites used in  
193 this chapter are presented in Table 1. Figure 3 illustrates the examples of the time series  
194 of the horizontal electromagnetic components at both site 001 and site 002.

195

## 196 **Comparison with STFT**

197 In conducting CWT, a wavelet that corresponds to a window function can  
198 continuously expand its width in the time domain as the frequency being analyzed  
199 decreases.  $T$  is the double of the time span from a time when the wavelet indicates a  
200 maximum value until it becomes 0 by attenuating in the positive (or negative) direction.  
201 Once the convolution of the wavelet with the time series within the range of  $T$  was  
202 computed, the wavelet was translated along the time axis. These operations were  
203 repeated throughout the observation period to derive the wavelet coefficients. The  
204 wavelet function used in this section was the complex Morlet function (Torrence and

205 Compo, 1998), which is defined as follows:

$$206 \quad \psi(t) = \pi^{-1/4} e^{-t^2/2} e^{ikt} \quad (9)$$

207 We compared the results of CWT with those of STFT accompanied by the gradual  
208 decimation of the time-series data, called cascade decimation (Wight and Bostick,  
209 1980). It is a mathematical technique used to obtain the components of wideband  
210 frequencies by gradually reducing the value of the sampling frequency, which leads to  
211 the decimation of the time-series data. Desired frequencies were classified into several  
212 octaves in advance. The frequency components were calculated from the convolution of  
213 the window function with the time series within the range that differs according to the  
214 octaves. To obtain lower-frequency components, the sampling frequency was reduced  
215 by half, and the above operations were repeated. A simple method of dividing the time-  
216 series data into several segments and computing FFT within each segment needs to  
217 increase the width of each segment in order to obtain lower-frequency components.  
218 However, this causes the number of segments in all observation times to decrease,  
219 making it difficult to obtain a sufficient number of stacks for calculating MT impedance  
220 tensors (Miura and Takasugi, 1991). To overcome this drawback, cascade decimation

221 has been introduced to computations of STFT. This technique is effective in wideband  
222 analyses because it enables us to obtain components in the lower band and computes  
223 faster than the simple STFT method (Miura and Takasugi, 1991). Appendix 1 illustrates  
224 the decimation level used in this section and the corresponding values of the frequencies  
225 obtained. As all the time-series data were discretized as well as the method of CWT,  
226 FFT was computed within the range of the convolution  $L$ . Let  $n$  be the sample number  
227 along the time axis, and the Hamming window function (Neska, 2006) is expressed as:

$$228 \quad Hw(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{L-1}\right). \quad (10)$$

229 MT data processing using CWT was compared with that of STFT with cascade  
230 decimation by calculating the magnetotelluric impedance tensors at each site under the  
231 calculation settings illustrated in Table 2. Figure 4 shows the comparison between the  
232 apparent resistivity and phase profiles using CWT and STFT. High-quality MT  
233 responses were also calculated by multi-day stacking and editing with the software  
234 “SSMT2000” and “MTEditor” made by Phoenix Geophysics Ltd., which helped to  
235 estimate the shapes of the true curves derived from the natural signal. As a result, the  
236 method using CWT yielded higher-quality MT responses that were close to the multi-

237 day responses throughout all the ULF band, in comparison with the method using  
238 STFT. Regarding site 001, the responses obtained with STFT contain more scatter  
239 below 0.1 Hz than those obtained with CWT. At site 002, the results of CWT show MT  
240 responses that smoothly vary with frequency in the bands below 0.01 Hz and between  
241 0.1 Hz and 1.0 Hz where energy of the natural signal attenuates. As the sampling  
242 frequency of 15 Hz was common between CWT and STFT at frequencies more than 0.3  
243 Hz, it was a type of window function that differed between the two methods. The effect  
244 of the numerical errors including  $FS_{1E}$  and  $FS_{2E}$  in Equation (7) is almost the same,  
245 considering that there is little difference in the quality of MT responses between the two  
246 methods. However, below 0.3 Hz, the time-series data were downsampled in the case  
247 where STFT were used, which resulted in more significant difference in the apparent  
248 resistivity and phase profiles between CWT and STFT. These results indicate that STFT  
249 increases the effect of the numerical errors by downsampling the time-series data below  
250 a certain frequency to obtain the components of wideband frequencies. In contrast, it is  
251 suggested that CWT, which dilates the wavelet without decimation of the time-series  
252 data, can restrain the numerical errors more easily in the time and frequency domains

253 throughout the ULF band.

254

## 255 **Proper calculation settings**

256 Basis function

257 When a wavelet vanishes quickly in positive and negative directions, i.e., when a

258 wavelet attenuates sufficiently and its time mean value is zero, there can be the

259 invertibility, and thus the time series can be recovered from the wavelet coefficients.

260 This characteristic is generally called “the admissibility condition” (Farge, 1992).

261 However, the Morlet function satisfies this condition for only  $k \geq 6$ . Meanwhile, the

262 complex Paul function (De Moortel et al., 2004), in Equation (11), can be considered a

263 complex wavelet that is generally utilized as well as the Morlet function.

264 
$$\psi(t) = \frac{2^k i^k k!}{\sqrt{\pi(2k)!}} (1 - it)^{-(k+1)}. \quad (11)$$

265 As shown in Figure 5, the Paul function, whose width along the time axis is narrower

266 than that of the Morlet function, satisfies the admissibility condition for any value of  $k$

267 (De Moortel et al., 2004). Larnier et al.(2016) reported that CWT with a wavelet with a

268 high time resolution, such as the Paul function, can accurately detect the pulsive and

269 transient signal, at frequencies above 7 Hz, where lightning activity is dominant. Thus,  
270 we investigated the effect of using the Paul function on the quality of MT responses  
271 obtained in the ULF band.

272 In Figure 6, we plotted the apparent resistivity and phase profiles of both site 001 and  
273 site 002 calculated using the Morlet function with  $k = 6$  and the Paul function with  
274  $k = 6$  and  $k = 4$ . In addition, “standard curves” were calculated by applying a  
275 smoothing spline (Garcia, 2014) with its parameter set to 5 to the multi-day responses  
276 shown in the previous section. The results obtained by using the Paul function display  
277 substantially smaller error bars than those obtained by using the Morlet function. Phase  
278 values calculated from the Paul function are almost equal to those of the “standard  
279 curves” and the Morlet function. Meanwhile, the shapes of the apparent resistivity  
280 profiles obtained with the Paul function appear to be distorted so greatly that they differ  
281 from any other profile. However, it is difficult to evaluate the quality of such distorted  
282 apparent resistivity profiles because the true values derived from the natural signal are  
283 unknown in actual observations.

284 Additionally, we attempted to recover the time-series data from the wavelet

285 coefficients of each frequency by means of the inverse continuous wavelet transform  
 286 (ICWT). In contrast to CWT in Equation (5), ICWT is defined as follows (Meyers et al.,  
 287 1993):

$$288 \quad \hat{f}(t) = C^{-1} \int_{0+}^{\infty} \int_{-\infty}^{\infty} s^{-2} Wf(b, s) \left( s^{-\frac{1}{2}} \psi \left( \frac{t-b}{s} \right) \right) db ds, \quad (12)$$

289 with

$$290 \quad C^{-1} = \int_{0+}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega, \quad (13)$$

291 where  $\hat{\cdot}$  denotes an estimated value. Note that  $C$  is independent of frequency and is a  
 292 constant for each wavelet function. In practice, Equations (12) and (13) were discretized  
 293 to reconstruct the time-series data quickly and approximately. Using the sum of the real  
 294 part of the wavelet coefficients  $\text{R}\{Wf(b, s)\}$  over frequencies below the Nyquist  
 295 frequency (Torrence and Compo, 1998), we computed ICWT as follows:

$$296 \quad \hat{f}(t) = \frac{dj dt^{1/2}}{C_{\delta} \psi(0)} \sum_{j=0}^J \frac{\text{R}\{Wf(b, s_j)\}}{s_j^{1/2}}, \quad (14)$$

297 where  $dt$  is a time spacing that equals the inverse of the sampling frequency, and  $dj$   
 298 determines the spacing of the scale (i.e., the inverse of the frequency) being considered.  
 299 Here, we chose a value of  $dj = 0.25$ . A constant  $C_{\delta}$  was derived from the technique of  
 300 the delta function (Torrence and Compo, 1998). The root mean square error (RMSE)

301 between the time series  $\hat{f}(t)$  reconstructed by ICWT and the original  $f(t)$  is defined

302 as follows:

$$303 \quad \text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (f(t_i) - \hat{f}(t_i))^2}{N}}, \quad (15)$$

304 Figure 7 illustrates an example of the comparison of the original time-series data to  
305 those reconstructed by ICWT using the Morlet function and the Paul function. Table 3  
306 lists the RMSE values generated by the processing of the 15-hour magnetic data at sites  
307 001 and 002. The results obtained using the Morlet function corresponded well with the  
308 original time-series data, whereas those obtained using the Paul function generated  
309 significant errors at many time points.

310

311 Wavelet parameter

312 As shown in Figure 8, a wavelet becomes narrower in the frequency axis as its wavelet  
313 parameter  $k$  becomes larger (De Moortel et al., 2004). A better frequency resolution  
314 enables us to distinguish frequency components more finely. We have examined the  
315 range of the value of  $k$  which should be appropriate for MT data processing for the  
316 ULF band.

317 The time-series data observed at sites 001 and 002 were additionally processed using  
318 CWT with the Morlet function. Figure 9 illustrates the comparison of the apparent  
319 resistivity and phase profiles by increasing the value of  $k$ . Note that  $k$  ranges from 6  
320 to 16. If  $k \geq 10$ , each profile gradually deviates from the “standard curves” below 0.3  
321 Hz. Consequently, there appears to be slightly more scatter and larger error bars  
322 included in each response. To confirm these findings quantitatively, we defined the  
323 deviation of the apparent resistivity and phase profiles from the “standard curves” as the  
324 root mean square deviation (RMSD):

$$325 \quad \text{RMSD} = \sqrt{\frac{1}{2J} \sum_{j=1}^J \left\{ \left( \log_{10} \frac{\rho_j}{\rho_{SDj}} \right)^2 + \left( \frac{(\phi_j - \phi_{SDj})}{65.96} \right)^2 \right\}}, \quad (16)$$

326 where  $J$  denotes the number of data points in the frequency domain,  $\log_{10} \rho$  is the  
327 logarithm of the apparent resistivity, and  $\phi$  is the phase. Subscript SD is added to  
328 each data point of the “standard curves”. Note that 65.96 in Equation (16) is a  
329 coefficient that relates the apparent resistivity and phase, the basis of which is described  
330 in Appendix 2. The change in RMSD with  $k$  is shown in Figure 10, which illustrates  
331 that RMSD increases significantly when  $k$  at site 001 exceeds 12, or when  $k$  at site  
332 002 exceeds 8.

333

## 334 Discussion

335 As a result of applying CWT using the Paul function to the time-series data, the shapes  
336 of the calculated apparent resistivity profiles were significantly distorted such that they  
337 differed from those of the “standard curves”. The time-series data reconstructed by  
338 ICWT using the Paul function contained large errors; i.e., large RMSE. Based on the  
339 results of both the apparent resistivity profiles given by CWT and the time series  
340 recovered by ICWT, numerical experiments demonstrated that MT data processing  
341 using the Paul function can be so unreliable that it can lead to incorrect MT responses  
342 that deviate from the true values in the ULF band. Though the natural signal caused by  
343 lightning activity at frequencies more than 10 Hz is often transient (Zhang and Paulson,  
344 1997), the signal originating from pulsations in the ULF band assumes gentle time  
345 variations defined as  $e^{i\omega t}$  (Fowler et al.,1967). Therefore, the results of this numerical  
346 experiment suggest that the wavelet including a sine wave component, such as the  
347 Morlet function generates more reliable solutions. Regarding the wavelet parameter  $k$ ,  
348 the quality of the calculated MT responses began to deteriorate when  $k$  became

349 approximately 10 or more. Scatter found in the apparent resistivity and phase profiles  
350 reflects the effect of the numerical errors derived from the spectral transform. This  
351 indicates that a value of  $k \geq 10$  makes the wavelet so wide in the time domain that its  
352 time resolution becomes significantly low.

353 Here we also discuss the error bars, namely, the standard error  $\sigma/\sqrt{N_{\text{stk}}}$ , where  $\sigma$  is  
354 the standard deviation and  $N_{\text{stk}}$  is the number of stacks. Large error bars are attributed  
355 to both (i) a decrease in  $N_{\text{stk}}$  caused by an increase in the width of the wavelet in the  
356 time domain and (ii) an increase in the effect of the numerical errors derived from the  
357 spectral transform. As for (i), because the Paul wavelet is narrower than the Morlet  
358 wavelet in the time domain, the time span within which we compute the convolution of  
359 the Paul wavelet with the time-series data is shorter. This should lead to an increase in  
360  $N_{\text{stk}}$  and, simultaneously, a decrease in the error bars. Furthermore, a wavelet becomes  
361 larger in the time domain with the use of a larger value of  $k$ , which should result in a  
362 decrease in  $N_{\text{stk}}$  and an increase in the error bars. Regarding (ii), the observed electric  
363 fields  $E$  and electric fields estimated from observed magnetic fields  $H$  are necessary  
364 to derive the variance (i.e. the square of the standard deviation) of each spectral data



381 numerical errors derived from the spectral transform. However, proper calculation  
382 settings will make  $E$ ,  $H$  and  $Z_R$  close to their respective true values, which will  
383 consequently prevent the error bar from becoming excessively large.

384

### 385 **Conclusion**

386 In this study, we performed several numerical experiments on MT data processing  
387 focusing on numerical errors caused by the spectral transform of time-series data. We  
388 evaluated MT responses at frequencies of approximately 0.001 Hz through 10 Hz,  
389 which corresponds to the ultra-low frequency (ULF) band originating from pulsations.  
390 The first experiment supports that CWT yields higher-quality MT responses that can be  
391 used to estimate the true responses of the natural signal than STFT accompanied by  
392 cascade decimation. In addition, this study proposes the use of the complex Morlet  
393 function with its wavelet parameter  $k$  set to approximately  $6 \leq k < 10$  as the proper  
394 calculation settings of CWT. These settings were confirmed to balance between the  
395 resolution of the time and frequency domains well and therefore effectively restrain the  
396 numerical errors that may be caused by the spectral transform. However, there is still

397 room for improvement in the technique of MT data processing using CWT to reduce  
398 computational steps more effectively and sufficiently restrain the effect of strong  
399 artificial noises.

400

#### 401 **Abbreviations**

402 MT: magnetotelluric; STFT: short-time Fourier transform; CWT: continuous wavelet  
403 transform; AMT: audio magnetotelluric; ULF: ultra-low frequency; FFT: fast Fourier  
404 transform; IFFT: inverse fast Fourier transform; ICWT: inverse continuous wavelet  
405 transform; RMSE: root mean square error; RMSD: root mean square deviation.

406

#### 407 **Authors' contributions**

408 The manuscript was written by HO and reviewed by all authors. All authors read and  
409 approved the final manuscript.

410

#### 411 **Competing interests**

412 The authors declare that they have no competing interests.

413

414 **Funding**

415 Not applicable.

416

417 **Availability of data and materials**

418 Not applicable.

419

420 **Acknowledgements**

421 Not applicable.

422

423 **Author details**

424 1 Department of Earth Sciences, Resources and Environmental Engineering, Graduate

425 School of Creative Science and Engineering, Waseda University, 3-4-1, Ookubo,

426 Shinjuku-ku, Tokyo, 169-8555, Japan.

427 2 Nittetsu Mining Co., Ltd., Yusen Bldg., 2-3-2, Marunouchi, Chiyoda-ku, Tokyo, 100-

428 8377, Japan.

429 3 Tono Geoscience Center, Japan Atomic Energy Agency, 959-31, Jorinji, Izumi-cho,

430 Toki-shi, Gifu, 509-5102, Japan.

431 4 Nuclear Waste Management Organization of Japan (NUMO), Mita NN Bldg., 4-1-23,

432 Shiba, Minato-ku, Tokyo, 108-0014, Japan.

433

434 **Authors' information**

435 This work was performed when HO was on loan to JAEA. KA is currently being sent

436 on loan to NUMO.

437

438 **Appendix 1**

439 The settings utilized in conducting STFT accompanied by cascade decimation are

440 described in this Appendix. Table 1 shows the relationship between the width of the

441 window function in the time domain  $L$  and the values of frequencies obtained by  $L$ .

442

443 **Appendix 2**

444 Here we derive the coefficient 65.96, which relates the apparent resistivity and the

445 phase. Equations (1) and (2) represent the method of converting the variance of the  
446 impedance tensor described by Gamble et al. (1979b) to the variance of the logarithm of  
447 the apparent resistivity and that of the phase, respectively.

$$448 \quad \text{Var}(\log_{10} \rho) \cong 0.3772 \times \text{Var}(Z)/|Z|^2, \quad (1)$$

$$449 \quad \text{Var}(\phi) \cong 0.5 \times \text{Var}(Z)/|Z|^2, \quad (2)$$

450 where  $\text{Var}()$  denotes the variance and  $Z$  denotes the impedance tensor. Based on  
451 these equations, Negi et al. (2010) correlated the standard error (i.e. the square of the  
452 variance) of the logarithm of the apparent resistivity and that of the phase as follows:

$$453 \quad \text{Std}(\phi) / \text{Std}(\log_{10} \rho) \cong 65.96 \quad (3)$$

454 Thus, we can examine the deviations of both apparent resistivity and the phase in the  
455 same dimension by adopting a value of 65.96 in Equation (3).

456

457 **References**

458 Chave AD, Thomson DJ (2004) Bounded influence magnetotelluric response function  
459 estimation. *Geophys J Int* 157(3):988-1006.

460 Chave AD, Thomson DJ, Ander ME (1987) On the robust estimation of power spectra,  
461 coherencies, and transfer functions. *J Geophys Res* 92(B1):633-648.

462 Cohen L (1995) *Time-Frequency Analysis*. Prentice-Hall, Englewood Cliffs, New  
463 Jersey.

464 De Moortel I, Munday SA, Hood, AW (2004) Wavelet analysis: The effect of varying  
465 basic wavelet parameters. *Sol Phys* 222(2):203-228.

466 Escalas M, Queralt P, Ledo J, Marcuello A (2013) Polarisation analysis of  
467 magnetotelluric time series using a wavelet-based scheme: A method for detection  
468 and characterisation of cultural noise sources. *Phys Earth Planet Inter* 218:31-50.

469 Farge M (1992) Wavelet transforms and their applications to turbulence. *Annu Rev*  
470 *Fluid Mech* 24:395-458.

471 Fowler RA, Kotick BJ, Elliott RD (1967) Polarization analysis of natural and artificially  
472 induced geomagnetic micropulsation. *J Geophys Res* 72(11):2871-2883.

473 Gamble TD, Goubau WM, Clarke J (1979a) Magnetotellurics with a remote reference.  
474 Geophysics 44(1):53-68.

475 Gamble TD, Goubau WM, Clarke J (1979b) Error analysis for remote reference  
476 magnetotellurics. Geophysics 44(5):959-968.

477 Garcia D, 2014, Matlab File Exchange: SMOOTH1Q  
478 <[mathworks.com/matlabcentral/fileexchange/37878-quick---easy-smoothing](http://mathworks.com/matlabcentral/fileexchange/37878-quick---easy-smoothing)>.

479 Garcia X, Jones AG (2008) Robust processing of magnetotelluric data in the AMT dead  
480 band using the continuous wavelet transform. Geophysics 73(6):F223-F234.

481 Ishikawa T, Arai S (2014) Variable resolution Gabor transform using the discrete  
482 convolution of Gaussian in the frequency domain. Proceedings of the 2014  
483 International Symposium on Nonlinear Theory and its Applications (NOLTA2014),  
484 120-123, Luzern, Switzerland, 14-18 September 2014.

485 Larnier H, Sailhac P, Chambodut A (2016) New application of wavelets in  
486 magnetotelluric data processing: reducing impedance bias. Earth, Planets Space  
487 68:70. doi: 10.1186/s40623-016-0446-9.

488 Meyers SD, Kelly BG, O'Brien JJ (1993) An introduction to wavelet analysis in

489 oceanography and meteorology: With Application to the dispersion of Yanai  
490 waves. *Mon Weather Rev* 121(10):2858-2866.

491 Miura Y, Takasugi S (1991) Comparison of cascade decimation techniques with FFT  
492 techniques in MT data processing. Proceedings of the 84th SEGJ Conference,  
493 Japan, SEGJ, 201-206, 1991 (in Japanese).

494 Negi T, Umeda K, Matsuo K, Asamori K, Yokoi K, Ohara H (2010) An efficient and  
495 effective stacking method for MT spectrum data; —Effectiveness against coherent  
496 noise—. *Geophys Explor* 63(5):395-408 (in Japanese).

497 Neska A (2006) Remote reference versus signal-noise separation: A least-square based  
498 comparison between magnetotelluric processing techniques. PhD thesis, Institut für  
499 Geologische Wissenschaften, Freie Universität Berlin.

500 Rhif M, Abbas AB, Farah IR, Martinez B, Sang Y (2019) Wavelet transform  
501 application for/in non-stationary time-series analysis: A review. *Appl Sci* 9(7). doi:  
502 10.3390/app9071345.

503 Stodt JA (1983) Processing of conventional and remote reference magnetotelluric data.  
504 PhD thesis, University of Uta.

505 Torrence C, Compo GP (1998) A practical guide to wavelet analysis. Bull Am Meteorol  
506 Soc 79(1):61-78.

507 Vozoff K (1990) Magnetotellurics: Principles and practice. Proceedings of the Indian  
508 Academy of Sciences- Earth and Planetary Sciences, the Indian Academy of  
509 Sciences, 99(4):441-471.

510 Wight D, Bostick F (1980) Cascade decimation: A technique for real time estimation of  
511 power spectra. Proceedings of IEEE International Conference on Acoustic Speech  
512 and Signal Processing, IEEE, 626-629, Denver, USA, 1980.

513 Zhang Y, Paulson KV (1997) Enhancement of signal-to-noise ratio innatural-source  
514 transient MT data with wavelet transform. Pure Appl Geophys 149(2):405-419.

515

516

517

518

519

520

521 **Figures**

522

523 **Figure 1** (a) A conceptual diagram of time-series processing using short-time Fourier  
524 transform (STFT). Convolution of time-series data with window functions (top), and the  
525 resolution of time and frequency (bottom) are illustrated. (b) A conceptual diagram of  
526 time-series processing using continuous wavelet transform (CWT). Convolution of  
527 time-series data with wavelets (top), and the resolution of time and frequency (bottom)  
528 are illustrated.

529 **Figure 2** (a) An index map. (b) A location map for site 001 and site 002.

530 **Figure 3** Examples of the electromagnetic time series at each MT site. They display  
531 13 hours of the total time series. Note that the system responses included in the raw data  
532 were corrected.

533 **Figure 4** Comparison of MT responses obtained with CWT to those obtained with  
534 STFT. Results of multi-day stacking and editing are also shown. Navy inverted triangles  
535 and red circles represent the spectral data of XY and YX modes, respectively. The  
536 impedance error bars were calculated using the technique of Gamble et al. (1979b) and  
537 Stodt (1983).  
538

539 **Figure 5** Real parts of the Morlet wavelet and the Paul wavelet.

540

541 **Figure 6** Comparison of MT responses by altering the function. The “standard  
542 curves” are also shown. Navy inverted triangles and cyan lines represent the spectral  
543 data of XY modes, whereas red circles and magenta dotted lines represent those of YX  
544 modes. The impedance error bars were calculated using the technique of Gamble et al.  
545 (1979b) and Stodt (1983).

546

547 **Figure 7** An example of comparison of the Hx at site 001 to those reconstructed by  
548 ICWT. Note that the raw magnetic data prior to correction of the system response were  
549 used as the original time-series data.  
550

551 **Figure 8** The Morlet wavelet for different values of wavelet parameter  $k$  in the  
552 frequency domain.  
553

554 **Figure 9** Comparison of MT responses by increasing the value of wavelet parameter  
555 *k*. The “standard curves” are also shown. Navy inverted triangles and cyan lines  
556 represent the spectral data of XY modes, whereas red circles and magenta dotted lines  
557 represent those of YX modes. The impedance error bars were calculated using the  
558 technique of Gamble et al. (1979b) and Stodt (1983).

559 **Figure 10** RMSD between the “standard curves” and the calculated MT responses for

560 each wavelet parameter  $k$ .

561

562 **Table 1** Outline of specification of the observation at three MT sites.

563

564 **Table 2** Outline of specification of the time-series processing.

565

566

567 **Table 3** Comparison of RMSE between the original time-series data and those

568 reconstructed by ICWT.

569

570 **Appendix**

571 **Table 1** Decimation level of STFT and the corresponding values of frequencies

572 obtained with cascade decimation.

573

# Figures

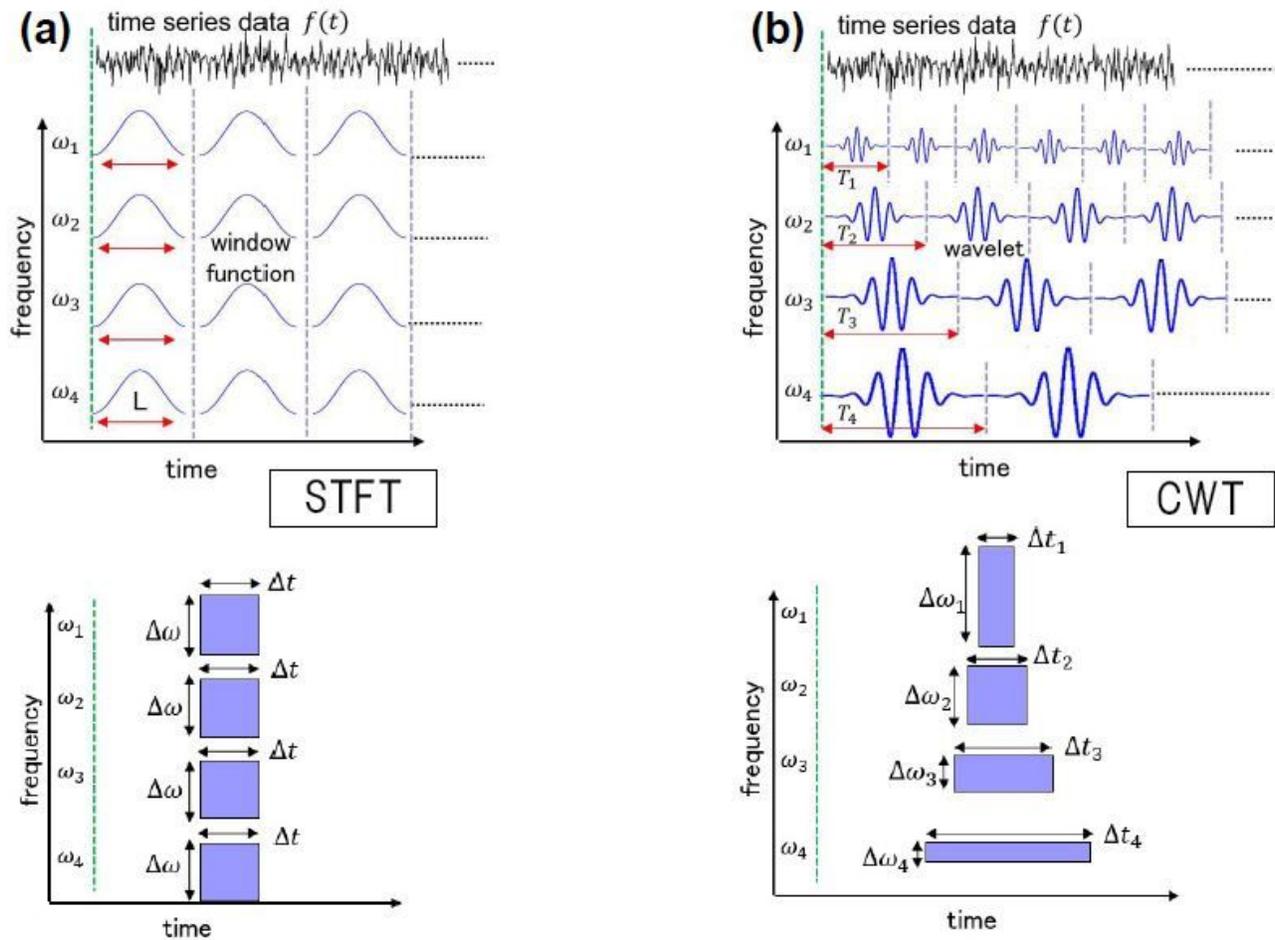


Figure 1

(a) A conceptual diagram of time-series processing using short-time Fourier transform (STFT). Convolution of time-series data with window functions (top), and the resolution of time and frequency (bottom) are illustrated. (b) A conceptual diagram of time-series processing using continuous wavelet transform (CWT). Convolution of time-series data with wavelets (top), and the resolution of time and frequency (bottom) are illustrated.

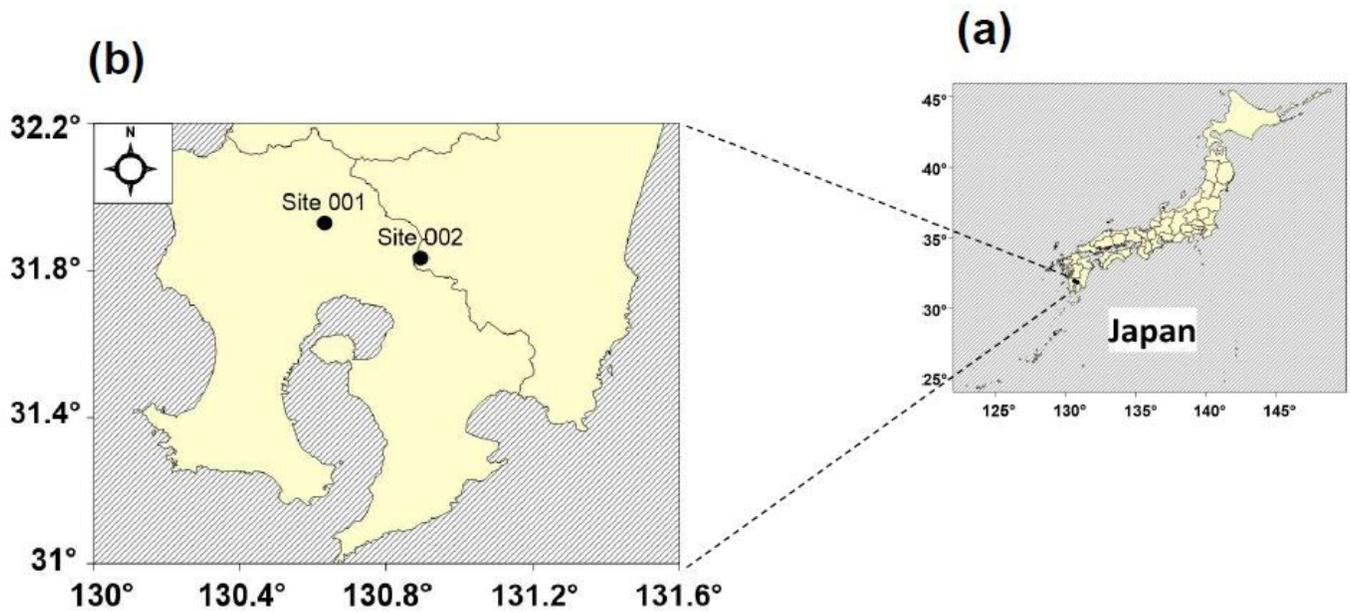


Figure 2

(a) An index map. (b) A location map for site 001 and site 002.

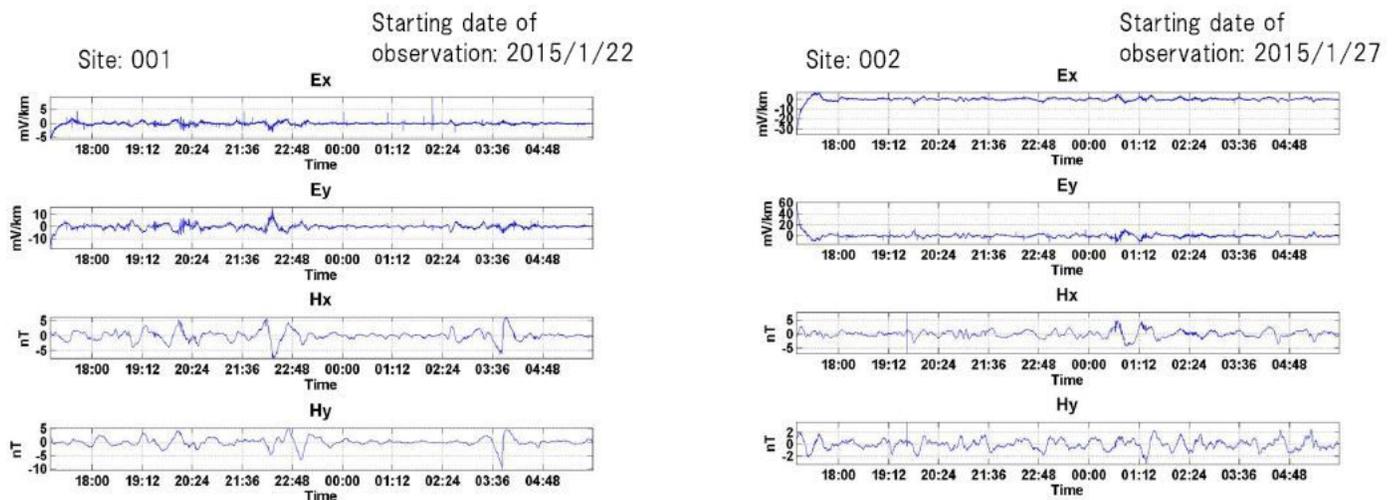


Figure 3

Examples of the electromagnetic time series at each MT site. They display 13 hours of the total time series. Note that the system responses included in the raw data were corrected.

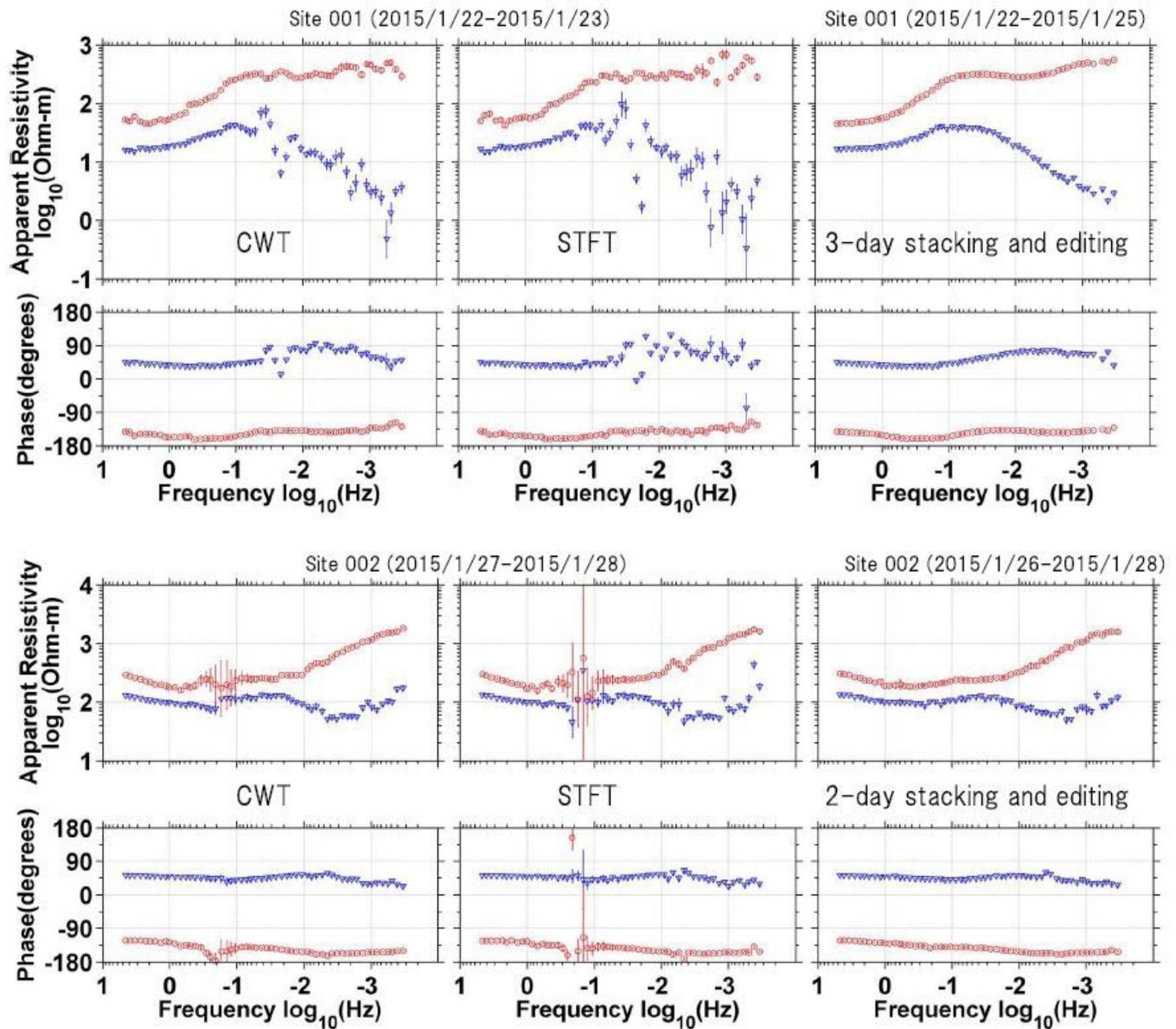


Figure 4

Comparison of MT responses obtained with CWT to those obtained with STFT. Results of multi-day stacking and editing are also shown. Navy inverted triangles and red circles represent the spectral data of XY and YX modes, respectively. The impedance error bars were calculated using the technique of Gamble et al. (1979b) and Stodt (1983).

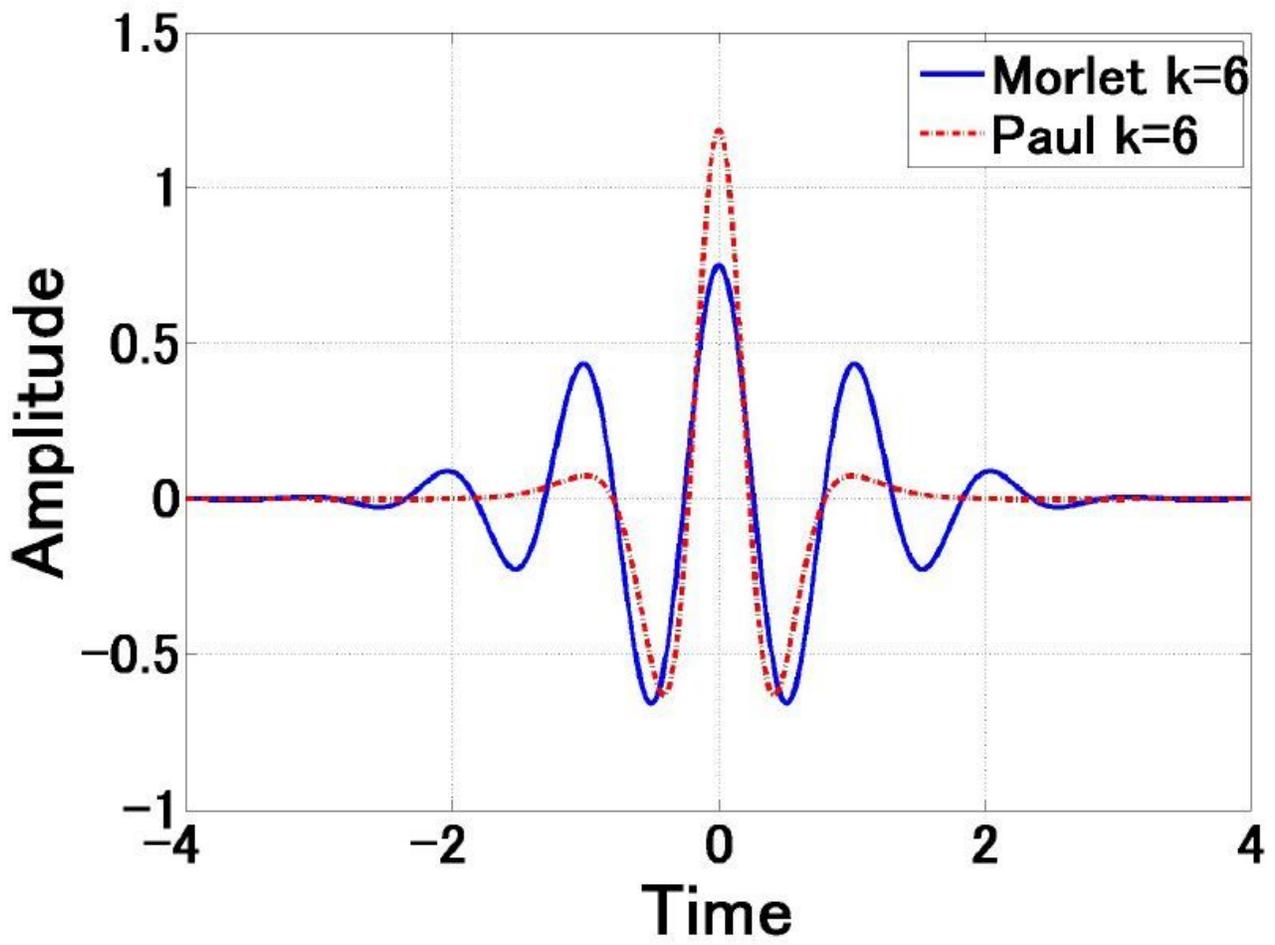
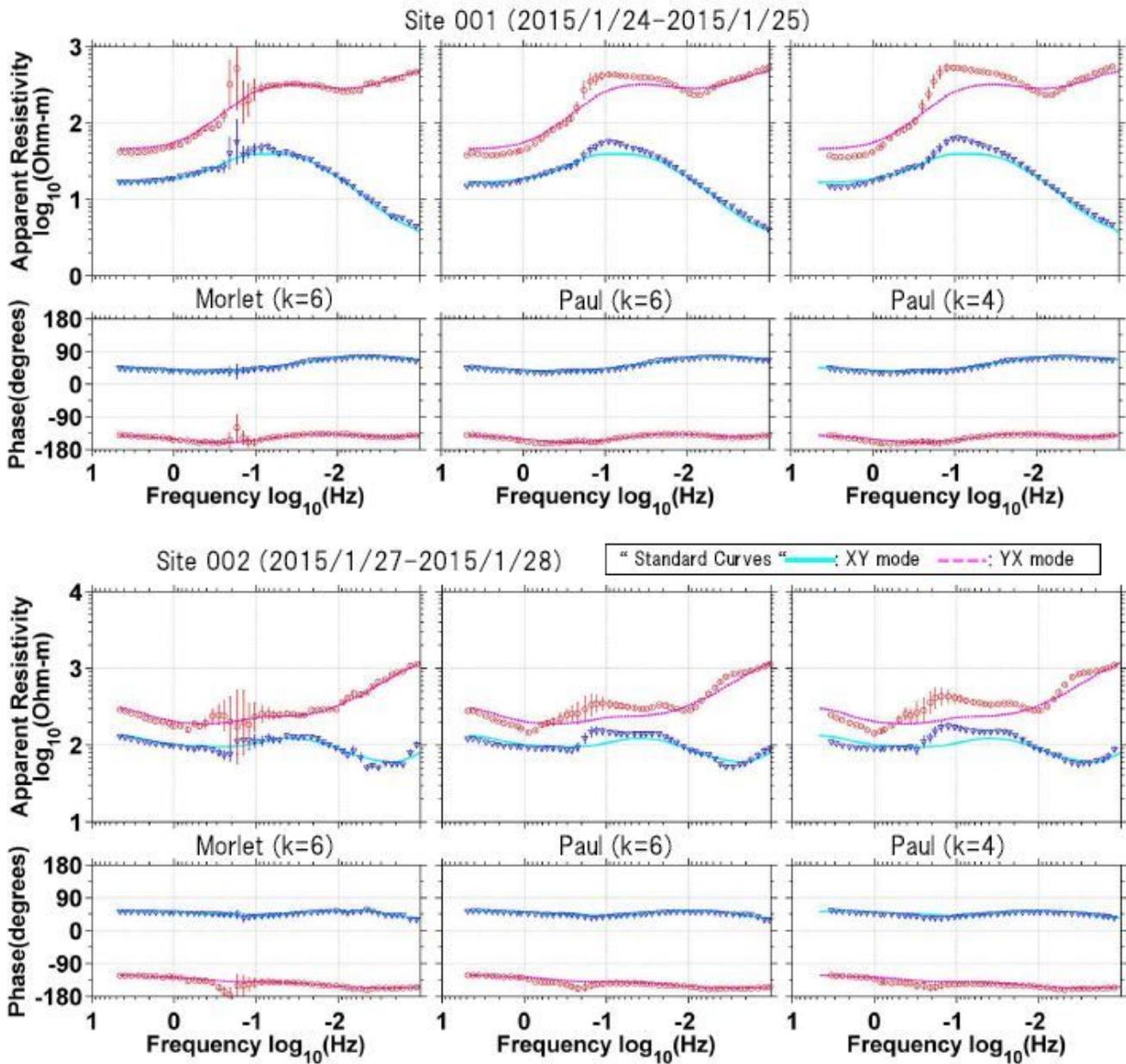


Figure 5

Real parts of the Morlet wavelet and the Paul wavelet.



**Figure 6**

Comparison of MT responses by altering the function. The “standard curves” are also shown. Navy inverted triangles and cyan lines represent the spectral data of XY modes, whereas red circles and magenta dotted lines represent those of YX modes. The impedance error bars were calculated using the technique of Gamble et al. (1979b) and Stodt (1983).

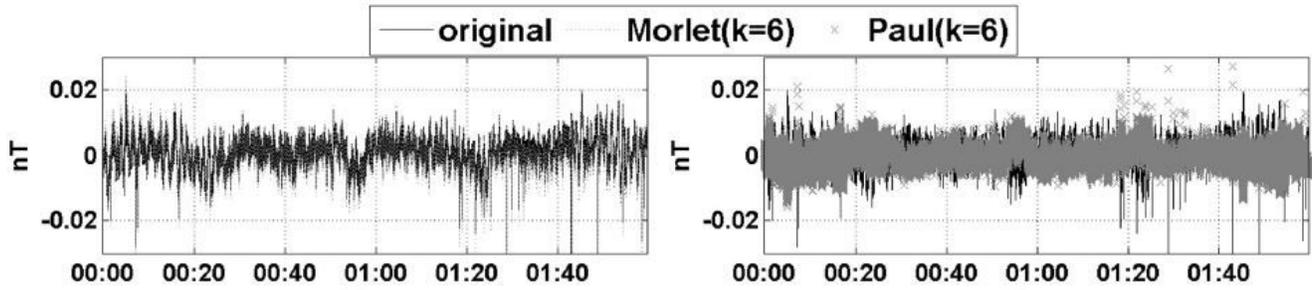


Figure 7

An example of comparison of the Hx at site 001 to those reconstructed by ICWT. Note that the raw magnetic data prior to correction of the system response were used as the original time-series data.

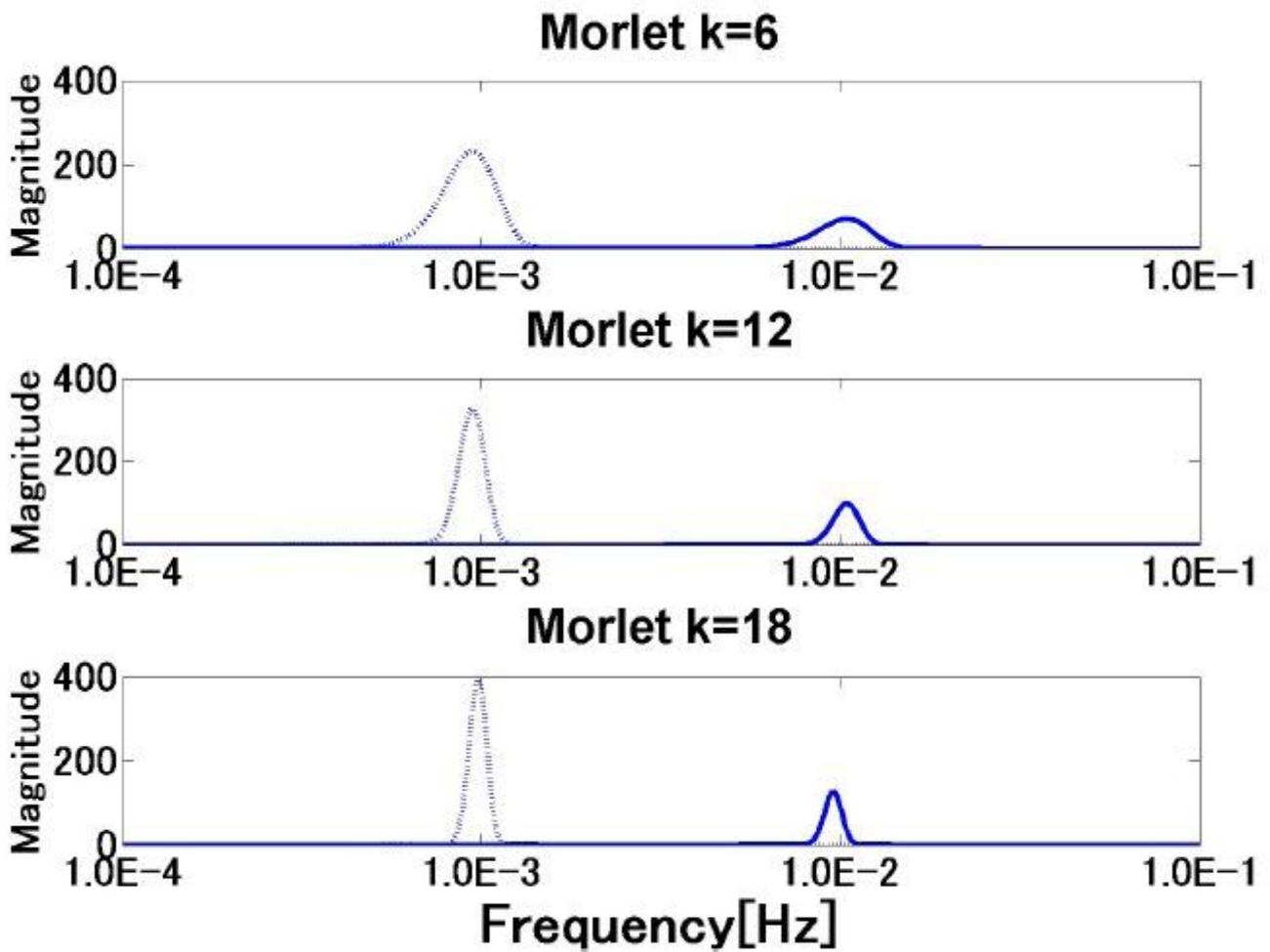


Figure 8

The Morlet wavelet for different values of wavelet parameter k in the frequency domain.

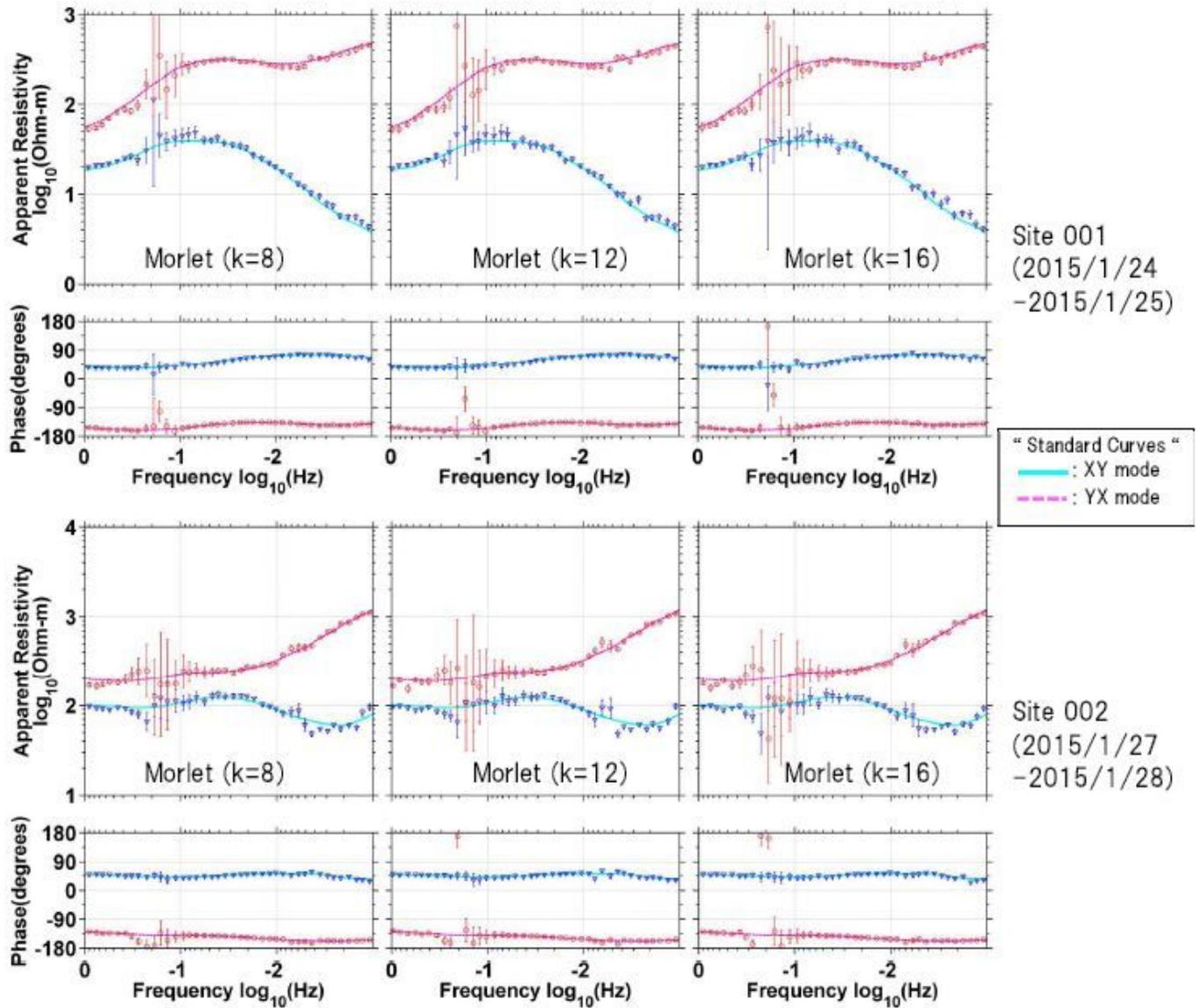


Figure 9

Comparison of MT responses by increasing the value of wavelet parameter k. The "standard curves" are also shown. Navy inverted triangles and cyan lines represent the spectral data of XY modes, whereas red circles and magenta dotted lines represent those of YX modes. The impedance error bars were calculated using the technique of Gamble et al. (1979b) and Stodt (1983).

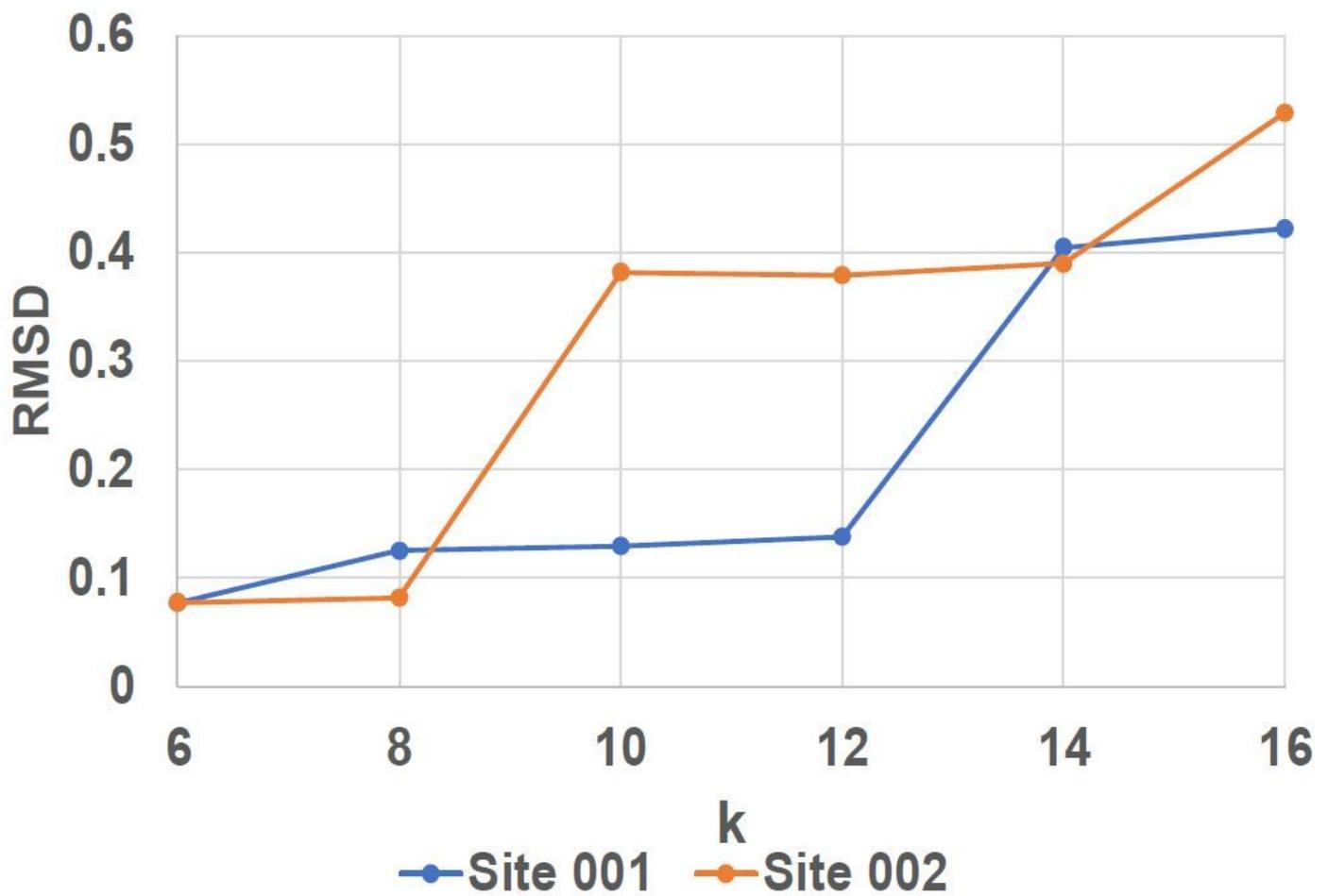


Figure 10

RMSD between the “standard curves” and the calculated MT responses for each wavelet parameter  $k$ .

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AppendixTable1Ogawaetal.2020.pdf](#)
- [TablesOgawaetal.2020.pdf](#)
- [GraphicalAbstractOgawaetal.2020.pdf](#)