Development of Reliable and Stable QL1-NMF Algorithm for Analyzing Environmental ELF Magnetic Signals

Motoaki Mouri, *Member, IEEE,* Ichi Takumi, *Member, IEEE,* and Hiroshi Yasukawa, *Member, IEEE*

***Abstract*—We previously developed two NMF algorithms (QL1-NMF1 and QL1-NMF2) using the quasi-L1 norm for analyzing environmental ELF magnetic field measurements. When the data included many outliers, the QL1-NMF algorithms returned better results than other BSS algorithms using the L1 norm. However, the derivative of the cost function in QL1-NMF1 was not based on a monotonically increasing function. This problem decreased the validity of the algorithm. QL1-NMF2 had serious problems with stability though it was based on a monotonically increasing derivative. The method therefore required an improvement of validity and stability. In the work described in this paper, we introduced new update functions that were based on a monotonically increasing derivative. Computer simulation results and real data results confirmed the new algorithm worked more stability than the previous one. Moreover, we showed the new algorithm was fast and accurate.**

***Index Terms*—magnetic field measurement, earthquake, BSS, outlier, L1 norm**

I. Introduction

Environmental electromagnetic (EM) waves have many possible uses in the geoscience field. It is known that EM waves are radiated in geodynamic processes related to the preparation and post-seismic stage of earthquakes (EQs) [1]. Anomalous radiation has been reported to be a pre-seismic EQs [2]-[5]. This fact means the possibility of detecting crustal activities or predicting EQs using environmental EM measurements. The major traditional EQ prediction technique is based on finding active faults and EQ cycles from trench surveys. Strain evaluation from GPS data is also used to improve accuracy[6]. These techniques estimate the possibility of an EQ occurring within the next several years or tens of years. In industry, this time-scale is not useful because they will need at least monthly, daily if possible, EQ possibility estimates. In the case of EM waves, it is said that crustal activity-related anomalies may appear several days or even weeks before an EQ occurrence [7][8]. There is significant worth if we can detect EM anomaly-related crustal activities within this short time-scale.

The main stream of this research field is using ultra low frequency (ULF, less than 1 Hz) EM waves[7][8]. ULF EM waves from deep ground levels can reach the surface directly. However, data from ULF measurements can only be obtained for 2-4 hours after midnight due to strong radiation factors, e.g. solar signals and artificial noises, occurring during other times of the day. On the other hand, we have been measur-

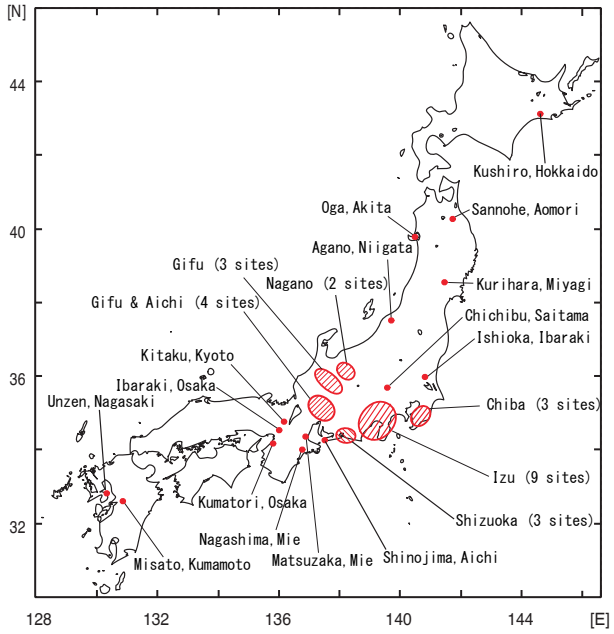


Fig. 1. Arrangement of observation sites

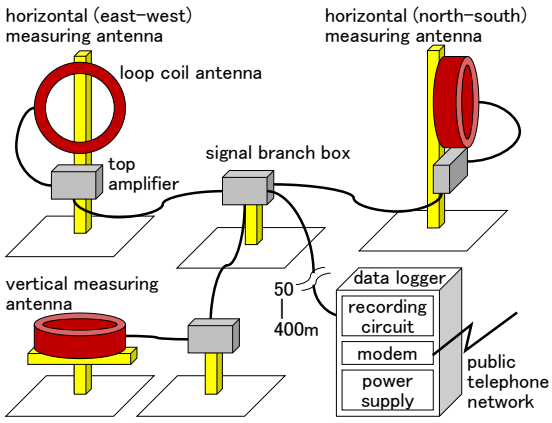


Fig. 2. Antennas measuring electromagnetic fields

M. Mouri is with Faculty of Business Administration, Aichi University, Aichi, Japan Email: mouri@aichi-u.ac.jp

I. Takumi is with Dept. of Computer Science and Engineering, Nagoya Institute of Technology, Aichi, Japan.

H. Yasukawa are with Grad. School of Information Science and Technology, Aichi Prefectural University, Aichi, Japan.

Manuscript received September 15, 2017.

Manuscript revised January 14, 2018.

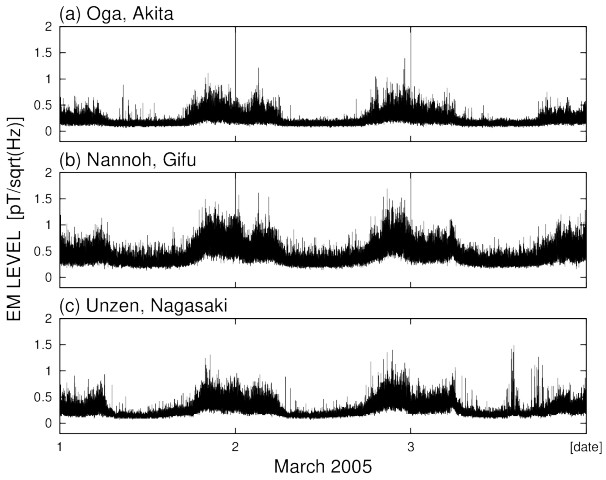
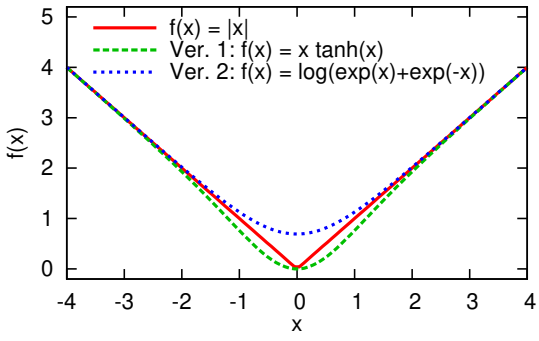
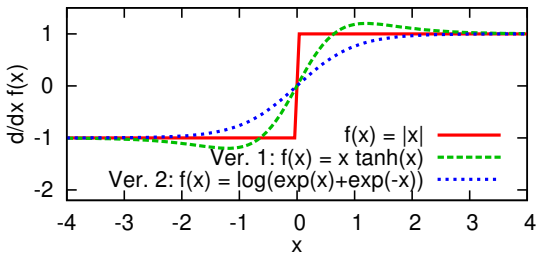


Fig. 3. Typical signals observed over 3 days



(a) Quasi-absolute functions



(b) Derivative functions

Fig. 4. Quasi-absolute functions and their derivatives in previous methods

ing extremely low frequency (ELF) band magnetic fields of 223 Hz all over Japan (Fig. 1) since 1985 [9]-[11]. Each observation site has three axial loop antennas of east-west, north-south and vertical orientations (Fig. 2). The sensors are tuned to measure 223 Hz with notch filters (for 50/60 Hz and their harmonics) and a narrow bandwidth (1 Hz) filter. The observation devices record the average absolute value of the EM field every 6-second periods, which is an equivalent 14,400 samples per day per direction. An example of our typical observed signals is shown in Fig. 3. The vertical axes indicate the level of magnetic flux density $[\mathrm{pT}/\sqrt{\mathrm{Hz}}]$, and the horizontal axis indicates the period of time (date). All of the observed signals have daily changes with higher levels at night and lower levels during the day time though they are recorded separately. We consider these changes are caused by background signals with large energy levels that are far from Japan. The ELF measurements enable us to use day-time data to detect anomalous signals because the background signal levels are lower during the day.

The ELF measurements are mixtures of signals associated with crustal activities, solar activities, thunderclouds, human activities, and other phenomena. We need to separate the signals for each factor or extract signals depending on a specific factor. It is therefore important to develop a blind source separation (BSS) method that is appropriate for analyzing the ELF magnetic signals. We have concluded that nonnegative matrix factorization (NMF) [12][13] is suitable for analyzing our data because of its mathematical model [14]. In previous research, we developed two NMF algorithms, hereafter referred to as QL1-NMF1 [15] and QL1-NMF2 [16], on the basis of minimizing the quasi-L1 norm of an error matrix [15]-[17]. The quasi-L1 norm is a kind of linear norm that is defined by a quasi-absolute function. Figure 4(a) shows the quasi-absolute functions used in previous methods. The red solid line corresponds to the true-absolute function, and the green dashed line and the blue dotted line correspond to the quasi-absolute functions of QL1-NMF1 and QL1-NMF2, respectively. Both values come close to that of the absolute function at a distance from zero. The reason why we need to use low-order norm, like these quasi-L1 norms, in cost function is a measure against outliers. Outliers should not be ignored because the amount of data we obtain sometimes becomes large and often includes important information.

We found that the QL1-NMF algorithms worked well when the data included many outliers, and performed better than other BSS algorithms using L1 norm. However, the derivative of the quasi-absolute function in QL1-NMF1 does not monotonically increase. Figure 4(b) shows the derivatives of the quasi-absolute functions in previous methods. The distortion parts of the function, which are seen at about $0.7 \leq |x| \leq 3$, corresponding to QL1-NMF1 may change the search direction of a solution into an undesirable one. The derivative function corresponding to QL1-NMF2 monotonically increases. However, the algorithm is not stable; the solutions it provides sometimes diverge to infinity or not-a-number. Its stability needs to be controlled by small and decreasing step-size parameters. Nevertheless, when the solutions diverged, we restarted the algorithm with new initial matrices and smaller step-size parameters. To improve the reliability of the result, a stable algorithm based on a distortion-free function is required. Such an algorithm may be useful for other application cases.

The aim of this paper is to develop a new, reliable, and stable NMF algorithm on the basis of minimizing the quasi-L1 norm. The partial derivative of its cost function should be distortion-free. This paper is organized as follows. Section II presents the development of the update functions for the new QL1-NMF algorithm. Section III presents numerical comparison results of the proposed method with other algorithms using computer simulations. Section IV describes the application of our proposed method to ELF magnetic signals and presents the comparison results with other algorithms' results. Finally,

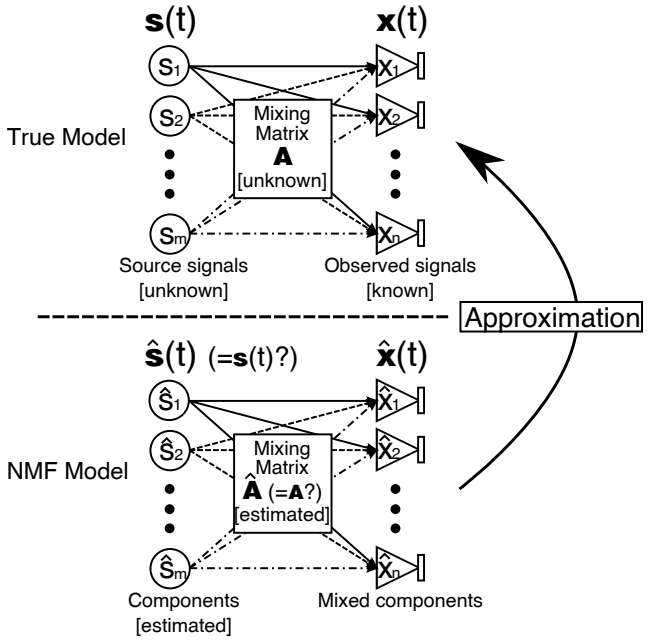
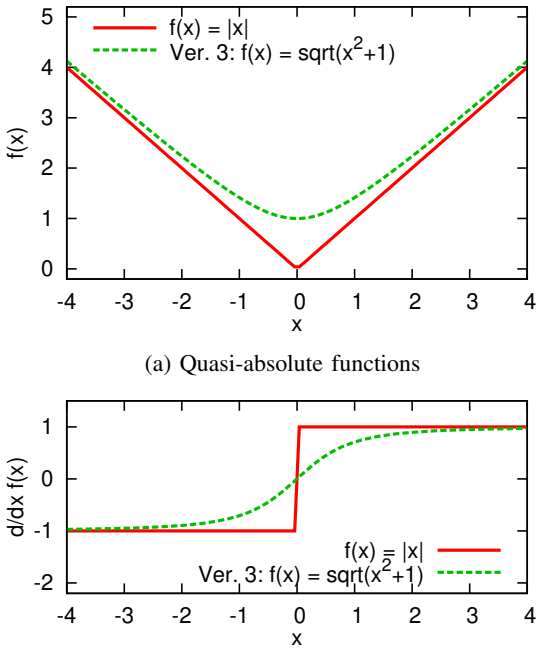


Fig. 5. NMF model



(a) Quasi-absolute functions

(b) Derivative functions

Fig. 6. Quasi-absolute function and its derivative in proposed method

the conclusion and future works are presented in Section V.

II. Nonnegative Matrix Factorization on the Basis of Minimizing Quasi-L1 Norm

A. Outline of NMF

The NMF algorithm model approximately factorizes a given nonnegative matrix under nonnegativity constraints. With this model the $n \times T$ matrix $\boldsymbol{X}$, which has only nonnegative values, is approximated by NMF as

$$\boldsymbol{X} \approx \boldsymbol{AS} \quad \boldsymbol{X}, \boldsymbol{A}, \boldsymbol{S} \geq 0 \tag{1}$$

where $\boldsymbol{A}$ is an $n \times r$ mixing matrix and $\boldsymbol{S}$ is an $r \times T$ component matrix. Both $\boldsymbol{A}$ and $\boldsymbol{S}$ have only nonnegative values. The rank of factorization, r, should be chosen as $r < \min(n, T)$ [12]. Eq. (1) can be written column by column as $\boldsymbol{x}(t) \approx \boldsymbol{A}\boldsymbol{s}(t)$, where $\boldsymbol{x}(t)$ and $\boldsymbol{s}(t)$ correspond to the tth columns in $\boldsymbol{X}$ and $\boldsymbol{S}$. This model is an approximation of a linear mixture signal model (Fig. 5).

B. Previous QL1-NMF algorithms

NMF searches $\boldsymbol{A}$ and $\boldsymbol{S}$ by using iterative updates based on an arbitrary cost function. To robustly analyze data including outliers, we developed algorithms named QL1-NMF1 [15] and QL1-NMF2 [16] based on the cost functions using quasi-L1 norm we defined. The quasi-L1 norms for QL1-NMF1 and QL1-NMF2 are follows:

$$D_1(\boldsymbol{X}||\boldsymbol{AS}) \equiv \sum_{i,k} \mathrm{qabs}_1(E_{ik}) = \sum_{i,k} \equiv E_{ik} \tanh(E_{ik}) \tag{2}$$

$$\begin{aligned} D_2(\boldsymbol{X}||\boldsymbol{AS}) &\equiv \sum_{i,k} \mathrm{qabs}_2(E_{ik}) \\ &= \sum_{i,k} \log(\exp(E_{ik}) + \exp(-E_{ik})) \end{aligned} \tag{3}$$

where E_{ik} is the value of the error matrix $\boldsymbol{E} = \boldsymbol{X} - \boldsymbol{AS}$ whose index is (i, k). The shapes of $\mathrm{qabs}_1(\mathrm{x})$ and $\mathrm{qabs}_2(\mathrm{x})$ are shown in Fig. 4(a) and their derivatives are shown in Fig. 4(b). The first problem are that the derivative function in QL1-NMF1 does not monotonically increase. This problem was solved by developing QL1-NMF2, however, the second problem, the stability weakness, became more critical.

From cost function of eq. (2), the nonnegative update functions for QL1-NMF1 are follows:

$$A_{ij} \leftarrow (1-\beta)A_{ij} + \frac{\sum_k S_{jk} \frac{\exp(2E_{ik})+4X_{ik}}{4\cosh^2(E_{ik})}}{\sum_k S_{jk} \frac{\exp(-2E_{ik})+4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})}} \beta A_{ij} \tag{4}$$

$$S_{jk} \leftarrow (1-\beta)S_{jk} + \frac{\sum_i A_{ij} \frac{\exp(2E_{ik})+4X_{ik}}{4\cosh^2(E_{ik})}}{\sum_i A_{ij} \frac{\exp(-2E_{ik})+4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})}} \beta S_{jk} \tag{5}$$

where β is an adjustment parameter that behaves like a step-size parameter ($0 < \beta \leq 1$). The reason why these functions have stability weakness might be the divergence speed of $\cosh^2(\cdot)$.

From cost function of eq. (3), the nonnegative update functions for QL1-NMF2 are follows:

$$A_{ij} \leftarrow (1-\beta)A_{ij} + \frac{\sum_k S_{jk} \frac{\exp(E_{ik})}{\exp(E_{ik})+\exp(-E_{ik})}}{\sum_k S_{jk} \frac{\exp(-E_{ik})}{\exp(E_{ik})+\exp(-E_{ik})}} \beta A_{ij} \tag{6}$$

$$S_{jk} \leftarrow (1-\beta)S_{jk} + \frac{\sum_i A_{ij} \frac{\exp(E_{ik})}{\exp(E_{ik})+\exp(-E_{ik})}}{\sum_i A_{ij} \frac{\exp(-E_{ik})}{\exp(E_{ik})+\exp(-E_{ik})}} \beta S_{jk} \tag{7}$$

The reason why these functions have critical stability weakness might be that the denominators in second terms sometimes become extremely small values.

C. New QL1-NMF algorithm

On the basis of consideration in previous studies, we propose the new quasi-absolute function as follows:

$$\mathrm{qabs}_3(x) \equiv \sqrt{x^2 + 1/\alpha^2} \tag{8}$$

where α is an approximation parameter. The larger α is, the more the curve fits to the absolute function. When $\alpha^2 = 2.0$, eq. (8) is equivalent to the $l_2 - l_1$ norm in [12]. Mostly we set $\alpha^2 = 1.0$. Figure 6(a) shows the graph of $\mathrm{qabs}_3(x)$. Then, the cost function using quasi-L1 norm becomes

$$D_3(\boldsymbol{X}||\boldsymbol{AS}) \equiv \sum_{i,k} \mathrm{qabs}_3(E_{ik}) = \sum_{i,k} \sqrt{E_{ik}^2 + 1/\alpha^2} \tag{9}$$

where E_{ik} is the value of the error matrix $\boldsymbol{E} = \boldsymbol{X} - \boldsymbol{AS}$ whose index is (i, k).

The derivative function of eq. (8) is as follow.

$$\frac{d}{dx}\mathrm{qabs}_3(x) = \frac{x}{\sqrt{x^2 + 1/\alpha^2}} \tag{10}$$

Figure 6(b) shows that the value of this function monotonically increases. From this equation, the gradient descent update function based on eq. (9) for an element A_{ij} in matrix $\boldsymbol{A}$ becomes as follows.

$$A_{ij} \leftarrow A_{ij} + \eta \sum_k S_{jk} \frac{E_{ik}}{\sqrt{E_{ik}^2 + 1/\alpha^2}} \tag{11}$$

where η is a step-size parameter. However, this update function does not ensure the nonnegativity of the solution because E_{ik} often become a negative value. We have extended η by referring the literatures [15]-[17] [19] to:

$$\eta_{Aij} \equiv \beta A_{ij} \left\{ \sum_k S_{jk} \frac{[\boldsymbol{AS}]_{ik}}{\sqrt{E_{ik}^2 + 1/\alpha^2}} \right\}^{-1} \tag{12}$$

where β is an adjustment parameter that behaves like a step-size parameter ($0 < \beta \leq 1$). Previous algorithms have the same parameters, and we mostly set $\beta = 0.2$ and gradually decreased them in iterations to improve the algorithm's stability and convergence. When the results were diverged to infinity, the initial βs decreased to 90 % when the algorithm was restarted.

The nonnegative update function for an element A_{ij} is as follows.

$$A_{ij} \leftarrow (1-\beta)A_{ij} + \frac{\sum_k S_{jk} \frac{X_{ik}}{\sqrt{E_{ik}^2+1/\alpha^2}}}{\sum_k S_{jk} \frac{[\boldsymbol{AS}]_{ik}}{\sqrt{E_{ik}^2+1/\alpha^2}}} \beta A_{ij} \tag{13}$$

Similarly, the nonnegative update function for an element S_{jk} is as follows.

$$S_{jk} \leftarrow (1-\beta)S_{jk} + \frac{\sum_i A_{ij} \frac{X_{ik}}{\sqrt{E_{ik}^2+1/\alpha^2}}}{\sum_i A_{ij} \frac{[\boldsymbol{AS}]_{ik}}{\sqrt{E_{ik}^2+1/\alpha^2}}} \beta S_{jk} \tag{14}$$

Additionally, we standardize matrices by each iteration as follows:

$$S_{jk} \leftarrow \max_j(\boldsymbol{A}) \cdot S_{jk}, \qquad A_{ij} \leftarrow \frac{A_{ij}}{\max_j(\boldsymbol{A})}. \tag{15}$$

where $\max_j(\cdot)$ is a function that returns the maximum value of the jth column in the matrix.

The algorithm judges the solution is converged when the following conditions are satisfied, and terminates updating.

$$\frac{\sum_{ij} \left(A_{ij}^{\mathrm{new}} - A_{ij}^{\mathrm{old}}\right)^2}{\sum_{ij} \left(A_{ij}^{\mathrm{new}}\right)^2} + \frac{\sum_{jk} \left(S_{jk}^{\mathrm{new}} - S_{jk}^{\mathrm{old}}\right)^2}{\sum_{jk} \left(S_{jk}^{\mathrm{new}}\right)^2} < \varepsilon \tag{16}$$

where suffixes $\cdot^{\mathrm{old}}$ and $\cdot^{\mathrm{new}}$ are defined as identifiers of before and after updates, respectively. The ε should be small positive value. We empirically set $\varepsilon = 10^{-7}$ based on the balance between processing time and performance. Moreover, we set the upper limit of the number of updates to 1000 because algorithm should stop in finite number of updates, and we empirically judge 1000 is enough number to converge. We also consider the estimation was failed if the algorithm needs more updates. These convergence conditions have also been used in QL1-NMF1 and QL1-NMF2 [17]. Hereafter we call this new algorithm QL1-NMF3.

This algorithm does not provide unique solutions, and greatly depends on initial values, because the degree of freedom (or space for the solution) is too large. The local solutions are located eccentrically, and they provide undesired results. However, this is common problem of most NMF algorithms. We often apply singular value decomposition (SVD) initialization to matrices $\boldsymbol{A}$ and $\boldsymbol{S}$ (see Section 1.3.3 in [12]) because it is suitable for our data. We do not discuss this problem in this paper.

III. Computer Simulation

A. Basic Research

First, we generated four source signals of $\boldsymbol{s}(t)$ similar to those shown in Fig. 7. In the figure, the horizontal axis indicates the sampling index and the vertical axis indicates amplitude. One source signal, $s_1(t)$, is a large common signal observed at all virtual observation sites. We assumed it to be background signal in the ELF band environmental magnetic signals. The signals $s_2(t)$ and $s_3(t)$ are common signals observed at several virtual observation sites. We assumed them to be electromagnetic waves from thunderclouds and artifacts. The signal $s_4(t)$ is outliers observed at only one virtual observation site. In the case of ELF data, it is impossible to know the number of sourse signals. However, the number of common source signals and large signals might be not so many because it need huge energy when the EM wave propagates a distance of hundreds km. Therefore, we usually set the rank r, the assumed number of source signals, for the number of observation sites n based on following equation:

$$r = \mathrm{round}(0.2n) + 1 \tag{17}$$

where $\mathrm{round}(\cdot)$ is round off function to the nearest integer.

We then generated a mixing matrix. The values corresponding to background signal $s_1(t)$ are almost uniform, while those in the second and third column vary. In the fourth column, only one value is non-zero. An example of a generated mixing matrix is shown in Fig. 8. The two horizontal axes indicate the index of the components and the index of the virtual

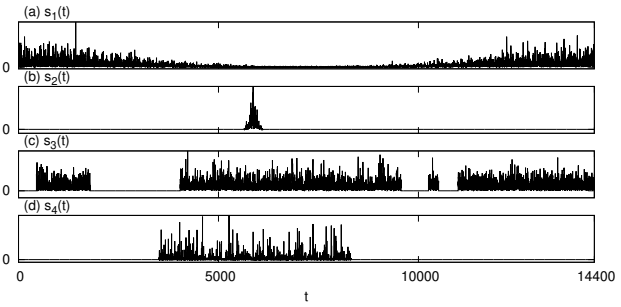


Fig. 7. An example of generated source signals

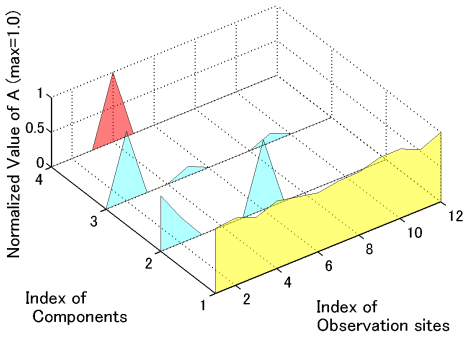


Fig. 8. An example of generated mixing matrix

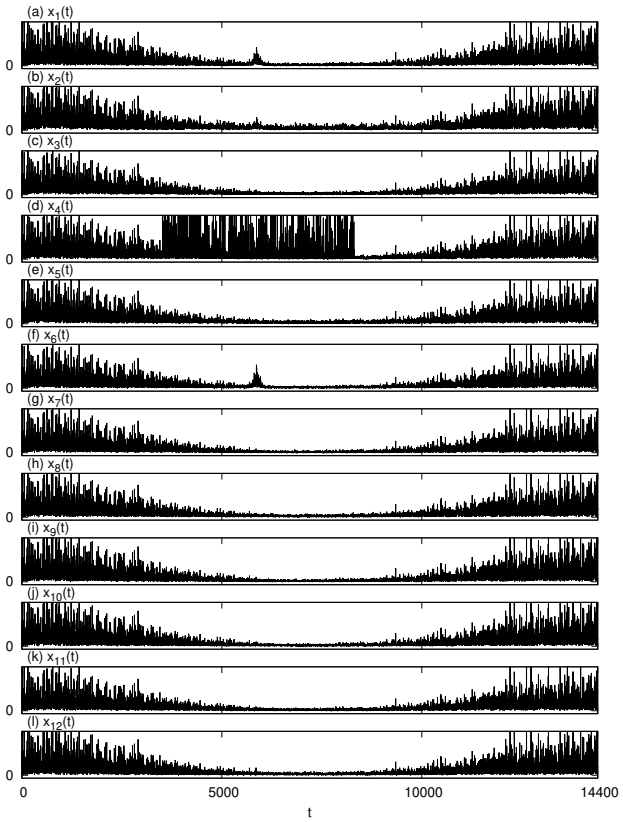


Fig. 9. An example of mixed signals including nonnegative noise

observation sites, respectively, and the vertical axis indicates values in matrix $\boldsymbol{A}$. We made 12 observed signals $\boldsymbol{x}(t)$ by a mixture signal of

$$\boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{l}(t) \tag{18}$$

where $\boldsymbol{l}(t)$ are small unique signals that distribute absolute Gaussian and its amplitude. We assumed them to be not just noise but also including local specific signals. However, NMF algorithm will deal $\boldsymbol{l}(t)$ as just noise because it is impossible to separate them with the blind condition. An example of mixed signals is shown in Fig. 9. The average of signal-to-noise ratio (SNR) of these signals, between $\boldsymbol{A}\boldsymbol{s}(t)$ and $\boldsymbol{l}(t)$, is 23.47 dB. We applied NMF to these signals with $r = 4$. Though this r was different with eq. (17), it was the true number.

Finally, we evaluated the accuracy of the estimated components using a criterion defined as

$$C_j \equiv \frac{100}{n} \sum_{i=1}^{n} \frac{\sum_t \left(A_{ij} s_j(t) - \hat{A}_{ij} \hat{s}_j(t) \right)^2}{\sum_t x_i^2(t)} \tag{19}$$

where $\hat{A}_{ij}$ is an estimated A_{ij}, and $\hat{s}_j(t)$ is an estimated $s_j(t)$. Smaller values of C_j produce more accurate solutions. We also evaluated the stability of the methods by comparing the number of restarts.

We processed 100 trials using the same procedure, generating new values of $\boldsymbol{s}(t)$, $\boldsymbol{A}$ and $\boldsymbol{l}(t)$ in each trial. Each of average SNR between $\boldsymbol{A}\boldsymbol{s}(t)$ and $\boldsymbol{l}(t)$ was around 23 dB. The specifications of the computer used for the processing is shown in Table I. Special calculation functions such as Parallel Toolbox are not used when calculating in this experiment.

TABLE I
COMPUTER SPECIFICATIONS

CPU	Intel Core i5 750 2.66 [GHz] (4 cores)
RAM	DDR3 4 [GB]
OS	Lubuntu 16.04.1 LTS (64 [bit])
Software	MATLAB 2017a (64 [bit])

TABLE II
AVERAGES OF ACCURACY CRITERIA FOR 100 TRIALS

Method	C_1	C_2	C_3	C_4
ISRA	7.70	6.21	7.63	0.28
PRMF	0.81	2.97	1.12	3.71
VSMF	7.19	8.12	3.68	0.10
BPGD	1.90	1.83	3.28	1.75
QL1-NMF1 ($\beta \leq 0.2$)	0.38	2.07	1.99	0.08
QL1-NMF1 ($\beta = 1.0$)	0.38	2.92	1.97	0.05
QL1-NMF2 ($\beta \leq 0.2$)	0.38	3.76	1.98	0.06
QL1-NMF2 ($\beta = 1.0$)	145.97	34.12	105.19	5.39
QL1-NMF3 ($\beta \leq 0.2$)	0.39	2.07	1.99	0.06
QL1-NMF3 ($\beta = 1.0$)	0.35	2.89	1.98	0.04

The components corresponding to those in Fig. 7 estimated by QL1-NMF3 are shown in Fig. 10. The mixing matrix corresponding to that in Fig. 8 estimated by QL1-NMF3 is shown in Fig. 11. Though the components are of different orders, they are well estimated.

Table II shows the averaged values of C_j for 100 trials. The method ISRA [19] is a basic NMF algorithm that use the L2 norm. The methods PRMF [20] and VSMF [21] are

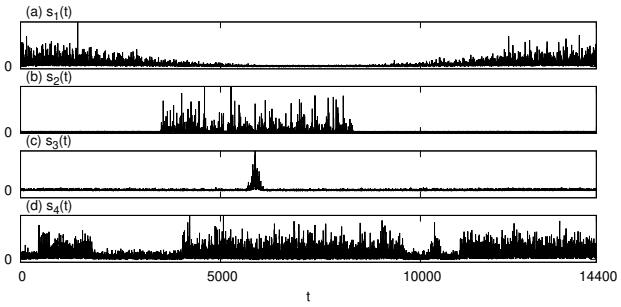


Fig. 10. Components corresponding to those in Fig. 7 estimated by QL1-NMF3

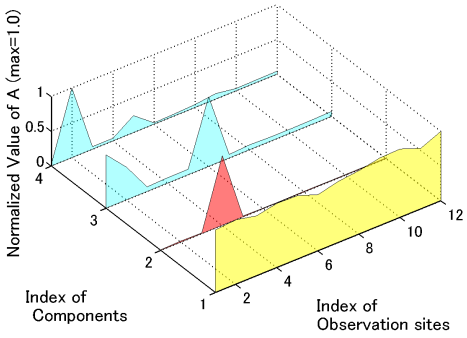


Fig. 11. Mixing matrix corresponding to that in Fig. 8 estimated by QL1-NMF3

TABLE III
RESTART STATISTICS FOR 100 TRIALS

Method	Trials with restarts	Average number of restarts
QL1-NMF1 ($\beta \leq 0.2$)	0	—-
QL1-NMF1 ($\beta = 1.0$)	100	8.10
QL1-NMF2 ($\beta \leq 0.2$)	4	2.00
QL1-NMF2 ($\beta = 1.0$)	100	9.21
QL1-NMF3 ($\beta \leq 0.2$)	0	—-
QL1-NMF3 ($\beta = 1.0$)	0	—-

other NMF algorithms that use L1 norm. The method BPGD [22] is a NMF algorithm with outlier consideration. $\beta \leq 0.2$ indicates that the method sets $\beta = 0.2$ and decreases them in iterations. $\beta = 1.0$ indicates that the method sets $\beta = 1.0$ and does not decreases them in iterations. In both cases, the βs are decreased when the algorithm restarts. From this table, the criteria for the QL1-NMF3 are small and almost the same with those for the QL1-NMF1. There are no significant differences in the C_js of QL1-NMF1 and QL1-NMF3 by two-sample t-test ($p < 0.05$). In the case of QL1-NMF2 ($\beta = 1.0$), all C_j are large because the solutions did not converge enough. However, C_2 of BPGD and C_3 of PRMF and are smaller than the others. The reasons for these difference are not known yet. C_4 of BPGD is large. This might be because BPGD estimates signals with outliers separately, the estimated sourse signals do not include outliers. The results of VSMF are almost the same degree with ISRA.

We maintained restart statistics to compare the stability of the methods. The results are shown in Table III. In the case of using QL1-NMF1 ($\beta = 1.0$) or QL1-NMF2 ($\beta = 1.0$), restarts occurred in all trials, and the average number of restarts was over 8. This means the solutions were diverged to infinity when βs were large when using the previous method. In the case of using QL1-NMF2, its stability was low because it needed to restart when $\beta \leq 0.2$. However, in the case of using QL1-NMF3, it did not need to restart. We can conclude the stability of QL1-NMF3 has a significant improvement over previous methods.

TABLE IV
AVERAGE PROCESSING TIMES FOR 100 TRIALS

Method	Time [sec]
ISRA	7.17
PRMF	22.50
VSMF	311.62
BPGD	27.34
QL1-NMF1 ($\beta \leq 0.2$)	3.16
QL1-NMF1 ($\beta = 1.0$)	7.75
QL1-NMF2 ($\beta \leq 0.2$)	28.58
QL1-NMF2 ($\beta = 1.0$)	32.05
QL1-NMF3 ($\beta \leq 0.2$)	2.30
QL1-NMF3 ($\beta = 1.0$)	12.11

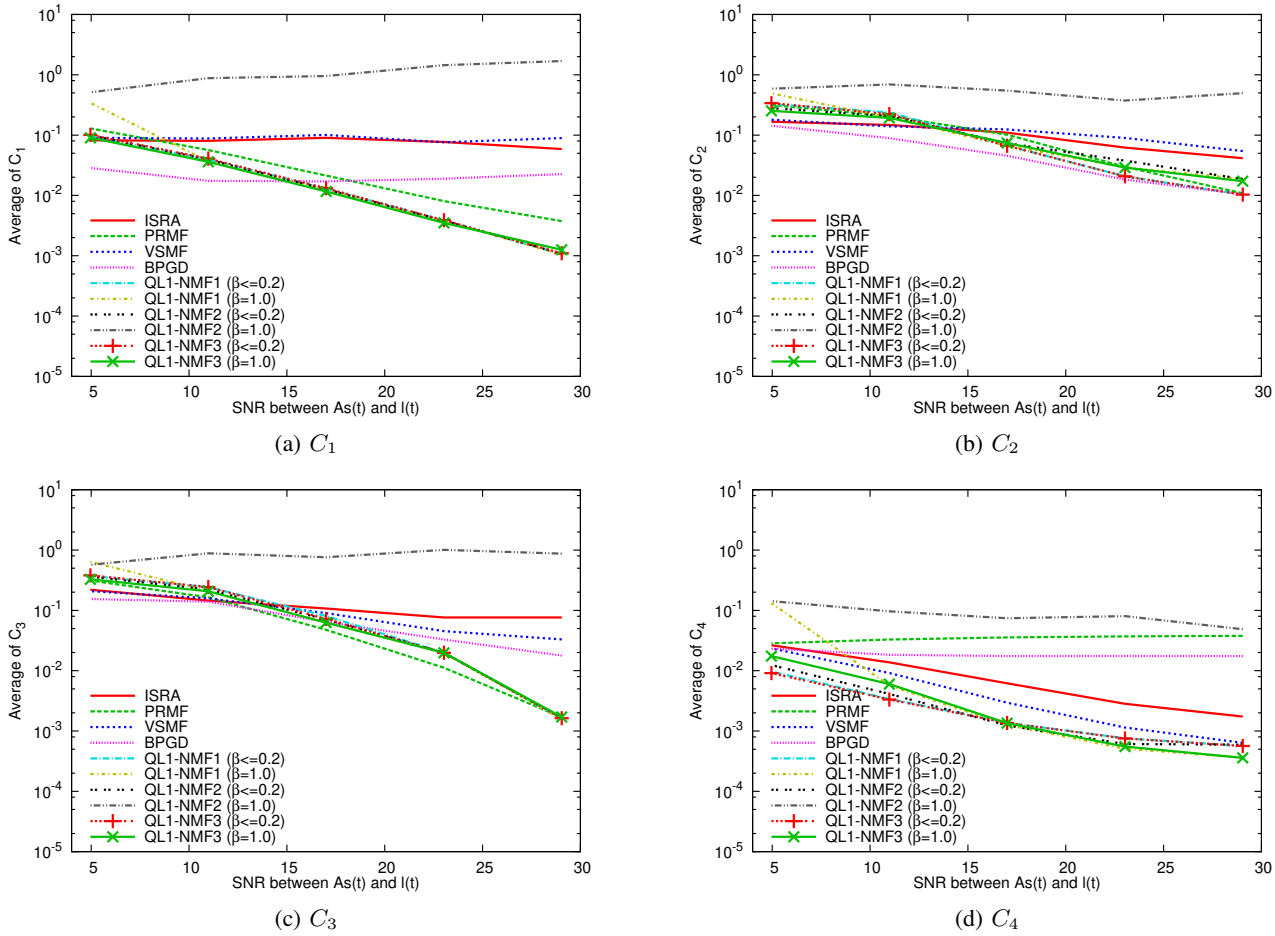
We also compared the processing speed using the cputime function in MATLAB. Table IV shows the average processing time for 100 trials. These data include time wasted by restarts. Comparing with other algorithms, QL1-NMF1 and QL1-NMF3 find solution rapidly. The reason why the processing time of QL1-NMF1 is short because it requires smaller number of updates for convergence. The reason why the processing time of QL1-NMF3 is short because it does not include costly calculations, e.g. exponent function, in update functions. The case of $\beta = 1.0$, convergence speed is not so fast. This fact suggests that we can treat β as a step-size parameter, not for control stability like previous methods.

B. Research Using Data of Various SNR

Usually, as shown in Fig. 3, the energy of a local specific signal including noise in our ELF band environmental magnetic data is not so larger than the background signal. However, we researched by computer simulation using data of various SNR for check the capacity of algorithm. The base of procedure was the same as basic research in Section III-A. We changed the amplitude of local specific signals $\boldsymbol{l}(t)$ 0.5 to 8 times that in the basic research.

Figure 12 shows average C_js corresponding to various SNR. (a)-(d) in the figure correspond to C_1 to C_4. The horizontal axis indicates average SNR between $\boldsymbol{As}(t)$ and $\boldsymbol{l}(t)$ for 100 trials. The vartical axis indicates the value C_j for 100 traials. The results for our proposed algorithms, QL1-NMF3, are plotted with points "+" and "×".

In the case of low SNR (less than 15 dB), BPGD works well especially estimating $s_1(t)$ of background signal. The reason why C_4 of BPGD is large might be that BPGD estimates sourse signals which do not include outliers. In the case of middle and high SNR (greater than 15 dB), QL1-NMFs, excepting QL1-NMF2 ($\beta = 1.0$), work well. If the SNR of



(a) C_1 (b) C_2

(c) C_3 (d) C_4

Fig. 12. Average C_js corresponding to various SNR between $\boldsymbol{As}(t)$ and $\boldsymbol{l}(t)$

the ELF data falls within that range, an accurate estimation result can be expected.

IV. Applying Algorithms to Environmental Magnetic Measurements

A. In the case of before Fukuoka earthquake

We experimented with our algorithms by applying them to ELF band environmental magnetic data. The target was 31 sets of data for each day from March 1st to 31st, 2005. Note that the 2005 Fukuoka EQ (Mj 7.0) occurred on March 20th.

Figure 13 shows our observed ELF signals (4 of 29 sites) on March 17, 2005. The values are the square sum of measurements observed at east-west and north-south in order to convert from amplitude to energy. Each vertical axis indicates the EM energy $[(\mathrm{pT})^2/\mathrm{Hz}]$, and each horizontal axis indicates the period of time (hour). (a)-(d) in the figure correspond to measurements taken in Akita, Gifu, Osaka, and Nagasaki, with each having common changes caused by a background signal. An anomalous signal was observed at Nagasaki on March 17th. This site is 112 km from the epicenter of the Fukuoka EQ.

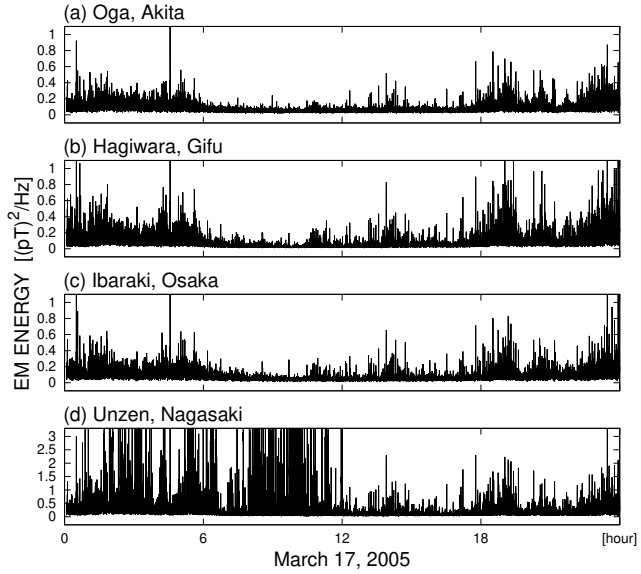


Fig. 13. Observed signals (4 of 29 sites)

Figures 14 and 15 show the source signals and mixing matrix estimated by QL1-NMF3 ($\beta = 1.0$). In Fig. 14, each vertical axis indicates the amplitude of the signals and each horizontal axis indicates the period of time (hour). In Fig. 15, the two horizontal axes indicate the index of the estimated components and that of the observation sites, and the vertical

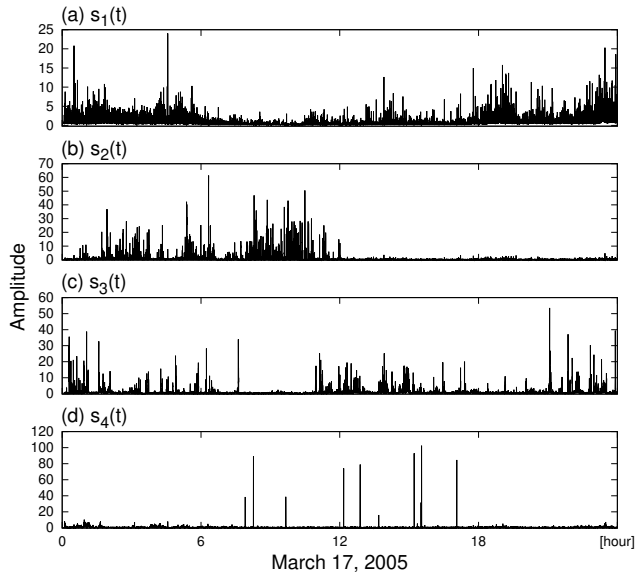


Fig. 14. Source signals estimated by QL1-NMF3 (4 of 7)

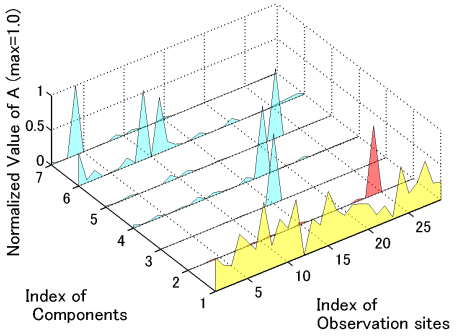


Fig. 15. Mixing matrix estimated by QL1-NMF3

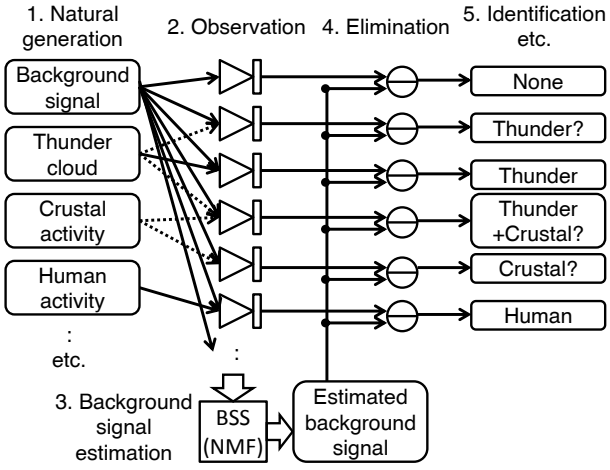


Fig. 16. Flow of background signal elimination for ELF data

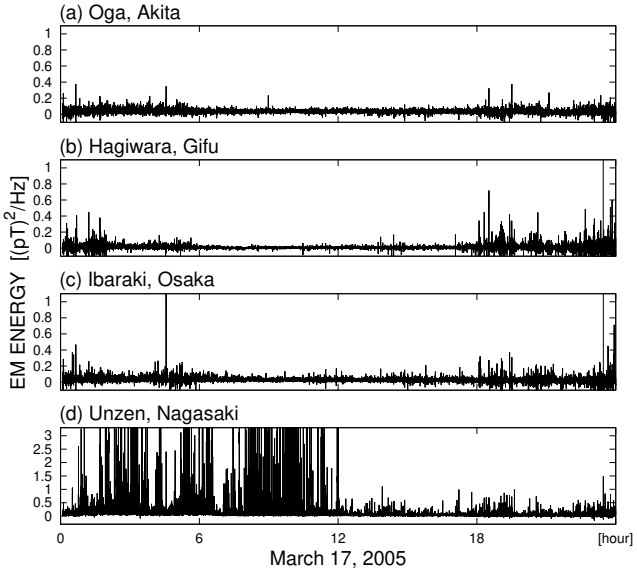


Fig. 17. Local signals estimated by QL1-NMF3 (4 of 29 sites)

axis indicates values in the estimated matrix $\boldsymbol{A}$. The index numbers of Akita, Gifu, Osaka, and Nagasaki are 2, 5, 26, and 24, respectively. The signal $s_1(t)$ was similar to the common changes. From the estimated 1st mixing weight A_1, $s_1(t)$ was observed at most sites. Therefore it was the background signal. The signal $s_2(t)$ was similar to the anomalous signal observed at Nagasaki. From the estimated 2nd mixing weight A_2, $s_2(t)$ was virtually observed only at Nagasaki (24). There is a possibility that $s_2(t)$ is an EQ-related source signal.

It is impossible to calculate the SNR directly from our ELF signals because the true source signal of each factor is necessary to calculate the SNR. Therefore, we evaluate effectiveness using the global information criteria GIC [14] we previously proposed. Here we describe the procedure to calculate GIC.

1) Estimate the source components from observed signals using BSS algorithm.
2) Identify the background signal component from the estimated components.
3) Estimate local signals by subtracting the global signal from the observed signals.
4) Calculate the averaged mutual information between each pair of local signals.

The procedure (1)-(3) is called background (or global) signal elimination and its flow is shown in Fig. 16. When the common background signal was successfully estimated and eliminated from the observed signals, the mutual information between local signals became quite small. Therefore, the smaller the GIC, the better the background signal estimation.

Figure 17 shows the estimated local ELF signals corresponding to those shown in Fig. 13. The common daily change is well eliminated from all observation sites. It suggests that the background signal and its mixing weights were well estimated.

The calculated GICs and processing times are shown in Table. V. "Mar. 17" refers to the results on March 17, 2005. "Average" refers to the averaged results in March, 2005. The GICs of QL1-NMF3 ($\beta \leq 0.2$) were a little larger than QL1-NMF1($\beta \leq 0.2$), however, there were no significant differences in the two-sample t-test ($p < 0.05$). It was the

TABLE V
*GIC*S AND PROCESSING TIMES FOR 31 ELF DATA

Method	*GIC*		Time [sec]	
	Mar. 17	Average	Mar. 17	Average
Observed signal	0.2238	0.2787	—	—
ISRA	0.1029	0.1329	11.86	9.87
PRMF	0.1026	0.1189	45.81	43.82
VSMF	0.1037	0.1379	283.04	313.96
BPGD	0.1093	0.1405	5.88	21.89
QL1-NMF1 ($\beta \leq 0.2$)	0.0904	0.1097	33.48	44.10
QL1-NMF1 ($\beta = 1.0$)	0.0887	0.1193	50.17	76.48
QL1-NMF2 ($\beta \leq 0.2$)	0.0888	0.1098	96.22	89.44
QL1-NMF2 ($\beta = 1.0$)	0.1092	0.1282	183.02	167.30
QL1-NMF3 ($\beta \leq 0.2$)	0.0942	0.1129	19.41	29.26
QL1-NMF3 ($\beta = 1.0$)	0.0881	0.1166	13.52	50.87

TABLE VI
RESTART STATISTICS FOR 31 ELF DATA

Method	Trials with restarts	Average number of restarts
QL1-NMF1 ($\beta \leq 0.2$)	4	3.25
QL1-NMF1 ($\beta = 1.0$)	31	8.61
QL1-NMF2 ($\beta \leq 0.2$)	3	3.33
QL1-NMF2 ($\beta = 1.0$)	31	10.19
QL1-NMF3 ($\beta \leq 0.2$)	0	—
QL1-NMF3 ($\beta = 1.0$)	0	—

same in the cases of $\beta = 1.0$. The processing time of QL1-NMF3 ($\beta \leq 0.2$) was virtually the smallest because ISRA's GIC was comparatively large.

The restart statistics are shown in Table VI. In the case of using QL1-NMF1 ($\beta = 1.0$) or QL1-NMF2 ($\beta = 1.0$), restarts occurred in all trials, and the average number of restarts was over 8. Additionally, QL1-NMF1 ($\beta \leq 0.2$) and QL1-NMF2 ($\beta \leq 0.2$) sometimes needed restart. On the other hand, QL1-NMF3 was quite stable because it did not need any restarts.

B. In the case of before Iwate and Miyagi earthquakes

The target was June 1st, 2008. A couple of large EQs occurred at Iwate and Miyagi on June 14, 2008. Their magnitudes are Mj 7.2 and Mj 5.7. Their epicenters were inland and their depth are less than 10 km.

Figure 18 shows observed ELF signals (4 of 30 sites) on June 1st, 2008. The values are the square sum of measurements observed at east-west and north-south in order to convert from amplitude to energy. The axes are the same with 13. (a)-(d) in the figure correspond to measurements taken in Miyagi, Gifu, Osaka, and Kumamoto. An anomalous signal was observed at Miyagi on June 1st. This site is about 40 km from the epicenters of the EQs.

Figures 19 and 20 show the source signals and mixing matrix estimated by QL1-NMF3 ($\beta = 1.0$). The axes are the same as Fig.14 and 15. The index numbers of Miyagi, Gifu, Osaka, and Kumamoto are 15, 3, 19, and 13, respectively. The signal $s_3(t)$ was similar to the common changes. From the estimated 3rd mixing weight A_3, $s_3(t)$ was observed at most sites. Therefore it was the background signal. The anomalous

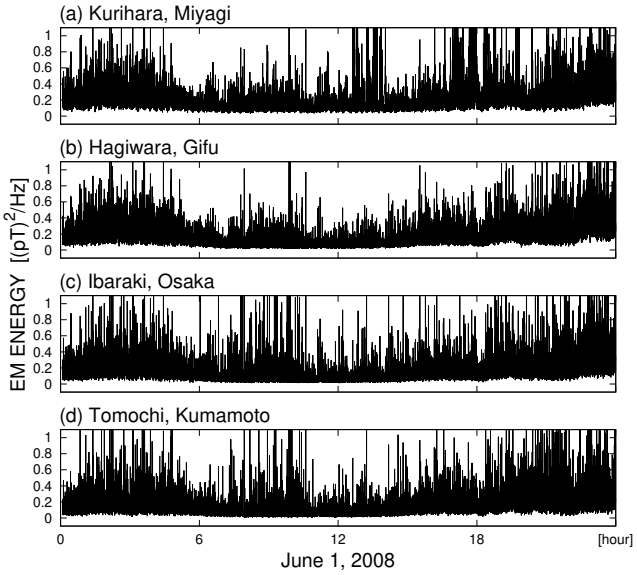


Fig. 18. Observed signals (4 of 30 sites)

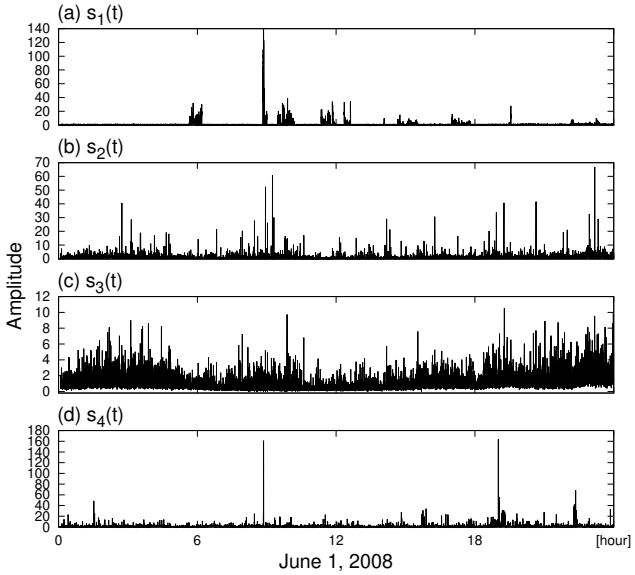


Fig. 19. Source signals estimated by QL1-NMF3 (4 of 7)

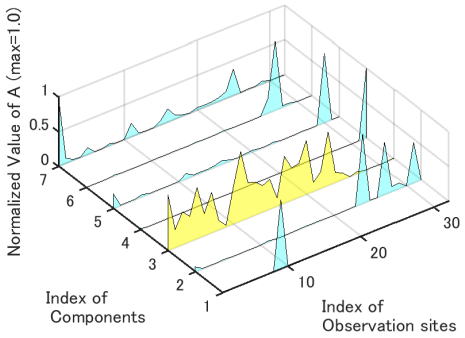


Fig. 20. Mixing matrix estimated by QL1-NMF3

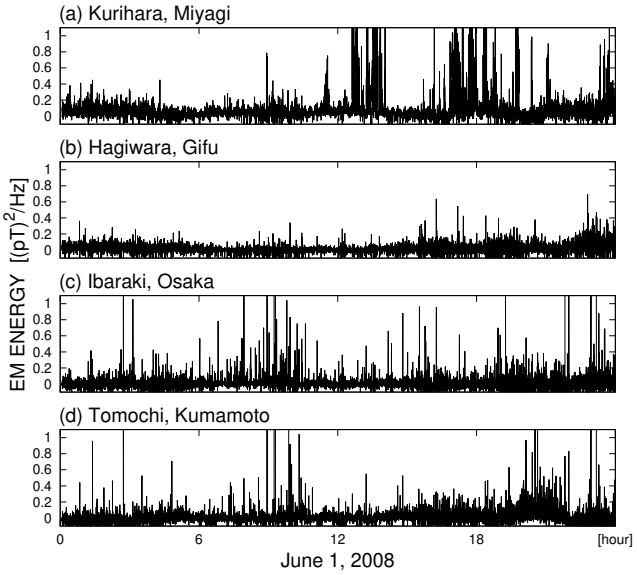


Fig. 21. Local signals estimated by QL1-NMF3 (4 of 30 sites)

TABLE VII
GICS AND PROCESSING TIMES

Method	GIC	Time [sec]
Observed signal	0.2406	—
ISRA	0.0957	10.54
PRMF	0.0937	42.51
VSMF	0.0958	291.76
BPGD	0.1081	14.08
QL1-NMF1 ($\beta \leq 0.2$)	0.0956	48.51
QL1-NMF1 ($\beta = 1.0$)	0.0871	96.76
QL1-NMF2 ($\beta \leq 0.2$)	0.0914	106.93
QL1-NMF2 ($\beta = 1.0$)	0.0937	147.86
QL1-NMF3 ($\beta \leq 0.2$)	0.0982	22.23
QL1-NMF3 ($\beta = 1.0$)	0.0856	58.39

signal observed at Miyagi was not estimated as one of the source signals.

Figure 21 shows the estimated local ELF signals corresponding to those shown in Fig. 18. The common daily change is well eliminated from all observation sites. It suggests that the background signal and its mixing weights were well estimated. On the other hand, Miyagi's large signals were remained. These signas has possibility of EQ pre-seismic precursor. The cause of outliers in other observation sites are unknown. It requires more research using estimated source signals and estimated local spesific signals.

The calculated GICs and processing times are shown in Table. VII. The GIC of QL1-NMF3 ($\beta = 1.0$) becomes the smallest though the one of QL1-NMF3 ($\beta \leq 0.2$) is large. It might suggest that the parameter β is unnecessary for QL1-NMF3 because small β makes achievement of the termination condition early even if the convergence of solution is not enough. The processing times for QL1-NMF3 are shorter than previous QL1-NMFs.

V. Conclusion

The aim of this paper is to develop a new, reliable, and stable NMF algorithm on the basis of minimizing the quasi-L1 norm. The main target of appling this algorithm is ELF band environmental magnetic data observed in Japan. Previous algorithm called QL1-NMF1 [15] had problem that the cost function is not based on a monotonically increasing derivative. It had also problem that the solution sometimes diverges. Another previous algorithm called QL1-NMF2 was not sistorted, however, it had critical weakness about stability. Our developed new NMF algorithm, QL1-NMF3, whose cost function is distortion-free unlike in the case of QL1-NMF1.

We compared the performance of new algorithm with previous algorithms and other NMF algorithms of ISRA [19], PRMF [20], VSMF [21] and BPGD [22] using computer simulations. In comparing the accuracy criteria C_j (Table II), the accuracy of solutions estimated by QL1-NMF3 were at the best as the same level as QL1-NMF1. In comparing the restart statistics (Table III), QL1-NMF3 is enough stable that not require restarts. In comparing the processing time (Table IV), QL1-NMF3 find solution rapidly as the same level as QL1-NMF1. We also researched by using data of various SNR for check the capacity of algorithm. From the Fig. 12, BPGD works well in condition of low SNR especially estimating $s_1(t)$ of background signal. In the case of middle and high SNR, QL1-NMF3 works well.

We applyed NMF algorithms to observed ELF data and compared their results. QL1-NMF3 could estimate anomalous outliers, which has possibility of being pre-seismic precursor of Fukuoka's EQ on March 20 in 2005, as one of the source signals. In comparing the accuracy criteria GIC (Table V), the accuracy of background signals estimated by QL1-NMF3 were well. There was no significant difference between the result of QL1-NMF1 and QL1-NMF3. The processing speed is also improved because QL1-NMF3 does not include costly calculations in update functions.

The results of computer simulations and using ELF data demonstrated that QL1-NMF3 did not need the restart technique and a decrease in the adjustment parameter β in order to stabilize. We can conclude the stability of QL1-NMF3 has a significant improvement over previous algorithms.

We applyed algorithms to ELF data preliminary Iwate an Miyagi EQs on June 1 in 2008. QL1-NMF3 could not estimate anomalous signal as one of source signas, however, anomalous signals observed at Miyagi were remained by background signal elimination. It suggests that the background signal and its mixing weights were well estimated by QL1-NMF3.

VI. Future Work

Subjects for future work will include researching characteristics of QL1-NMF3, e.g. how to optimize the parameter α and β. For that purpose, it is important to study the theoretical verification of the algorithm. Studying measures against low SNR data is also important in order not to limit the application or the target data. We will apply it to other kinds of signals. We want to apply it to multispectral satellite images if possible,

because many NMF target them. In that case, comparison with other BSS algorithms is necessary.

Moreover, in order to know EQ pre-seismic precursor, additionaly analyzing data estimated by QL1-NMF3 is important. Correspondence with other natural phenomena is also important because EM waves have many possible uses in the geoscience field.

ACKNOWLEDGMENT

This work was supported by The Nitto Foundation and Scientific Research (A) 26249060 from the JSPS

REFERENCES

[1] V. A. Gavrilov, E. V. Poltavtseva, A. V. Desherevsky, Yu. Yu. Buss, Yu. V. Morozova, "Geological environmental monitoring based on synchronous borehole geoacoustic and electromagnetic measurements: Use of natural electromagnetic radiation," *Seismic Instruments*, Springer, Vol. 52, Issue 3, pp. 266-277, 2016.

[2] M. B. Gokhberg, V. A. Morgunov, T. Yoshino and I. Tomizawa, "Experimental measurements of EM emissions possibly related to earthquakes in Japan," *Journal of Geophysical Research*, vol. 87, pp. 7824-7829, 1982.

[3] M. Hayakawa and Y. Fujinawa, "EM phenomena related to earthquake prediction," *Terra Scientific (TERAPUB)*, 1994.

[4] K. Maeda and N. Tokimasa, "Decametric radiation at the time of the Hyogo-ken Nanbu earthquake near Kobe in 1995," *Journal of Geophysical Research Letters*, vol. 23, pp. 2433-2436, 1996.

[5] E. Petraki, D. Nikolopoulos, C. Nomicos, J. Stonham, D. Cantzos, P. Yannakopoulos and S. Kottou, "Electromagnetic Pre-earthquake Precursors: Mechanisms, Data and Models-A Review," *Journal of Earth Science & Climatic Change*, vol. 6, issue 1, 205, 2015.

[6] C. Hashimoto, A. Noda, T. Sagiya and M. Matsu'ura, "Interplate seismogenic zones along the Kuril-Japan trench inferred from GPS data inversion," *Nature Geoscience*, vol. 2, pp. 141-144, 2009.

[7] K. Hattori, P. Han, C. Yoshino, F. Febriani, H. Yamaguchi, C-H. Chen, "Investigation of ULF Seismo-Magnetic Phenomena in Kanto, Japan During 2000-2010: Case Studies and Statistical Studies," *Surveys in Geophysics*, vol. 34, pp. 293-316, 2013.

[8] P. Han, K. Hattori, J. Zhuang, C-H. Chen, Y-J Liu and S. Yoshida, "Evaluation of ULF seismo-magnetic phenomena in Kakioka, Japan by using Molchan's error diagram," *Geophysical Journal International*, vol. 208, Issue 1, pp. 482-490, 2017.

[9] M. Hata, I. Takumi, S. Yabashi: "A Model of Earthquake Seen by Electromagnetic Observation – Gaseous Emission from the Earth as Main Source of Pre-Seismic Electromagnetic Precursor and Trigger of Followed Earthquake," *Proceedings of European Geophysical Society*, C131, 1998.

[10] M. Hata, I. Takumi and H. Yasukawa: "Electromagnetic-Wave Radiation due to Diastrophism of Magma Dike Growth in Izu-Miyake Volcanic Eruptions in Japan in 2000," *Natural Hazards and Earth System Sciences*, Vol. 1, No. 1, pp. 43-51, 2001.

[11] M. Hata, I. Takumi, K. Ohta, J. Izutsu, T. Fujii, T. Sato, S. Yahashi, N. Watanabe, "Development of ELF Band Receiver of Detecting Extreme Low Frequency Magnetic Flux Variation Due to Earthquakes," *Journal of Atmospheric Electricity*, Vol. 30, No. 1, pp. 37-52, 2010 (in Japanese).

[12] A. Cichocki, R. Zdunek, A.-H. Phan and S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Chichester: Wiley, 2009.

[13] A. Cichocki, R. Zdunek: NMFLAB toolboxes, [http://www.bsp.brain.riken.jp/ICALAB/nmflab.html], 2006.

[14] M. Mouri, A. Funase, I. Takumi, A. Cichocki, H. Yasukawa and M. Hata, "Global Signal Elimination from Environmental Electromagnetic Signals by Nonnegative Matrix Factorization," *Journal of Signal Processing*, Vol. 14, No. 6, pp. 415-425, 2010.

[15] M. Mouri, A. Funase, A. Cichocki, I. Takumi, H. Yasukawa, Analysis of Environmental Electromagnetic Signal Using Nonnegative Matrix Factorization Minimizing Quasi-l1 Norm, *2011 IEEE International Geoscience and Remote Sensing Symposium Proceedings*, pp. 1846-1849, 2011.

[16] M. Mouri, I. Takumi, H. Yasukawa, A. Cichocki: "Revising Algorithm for Nonnegative Matrix Factorization Based on Minimizing Quasi-L1 Norm," *Proceedings 2014 IEEE Asia Pacific Conference on Circuits and Systems*, pp. 767-770, 2014.

[17] M. Mouri, I. Takumi, H. Yasukawa, A. Cichocki, "Usefulness of Quasi-L1 Norm-Based Nonnegative Matrix Factorization Algorithm to Estimate Background Signal Using Environmental Electromagnetic Field Measurements at ELF Band," *IEEJ Transactions on Fundamentals and Materials*, Vol. 136, No. 5, pp. 241-251, 2016 (in Japanese).

[18] M. Mouri, I. Takumi, H. Yasukawa, "Implementation of Valid and Stable Algorithm of QL1-NMF for Analyzing Environmental ELF Magnetic Signals", *Proceedings 2017 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5763-5766, 2017.

[19] D. D. Lee and H. S. Seung, "Algorithms for Non-negative Matrix Factorization, *Advances in Neural Information Processing 13. MIT Press*, pp. 556-562, 2001.

[20] N. Wang, T. Yao, J. Wang, D. Yeung, "A Probabilistic Approach to Robust Matrix. Factorization," *Computer Vision - ECCV 2012*, LNCS, Vol. 7578, pp. 126-139, 2012.

[21] Y. Li and A. Ngom, "The non-negative matrix factorization toolbox for biological data mining," *Source Code for Biology and Medicine 2013*, pp. 8-10, 2013.

[22] R. Zhao, V. Y. F. Tan, "Online Nonnegative Matrix Factorization With Outliers," *IEEE Transactions on Signal Processing*, vol. 65, issue 3, pp. 555-570, 2016.

Motoaki Mouri received his B.Eng., M.Eng. and Dr.Eng. degrees in Computer Science and Engineering from Nagoya Institute of Technology, Nagoya Japan, in 2005, 2007 and 2011, respectively. Since April 2011, he has been with Aichi University, where he is now a Assosiate Professor in the Faculty of Business Administration. His current research interests include digital signal processing and blind signal separation. He is a member of IEICE and IEEE.

Ichi Takumi received his B.Eng., M.Eng. and Dr.Eng degrees in electronic engineering from Nagoya Institute of Technology, Nagoya Japan, in 1982, 1984 and 1994, respectively. Since December 1985, he has been with Nagoya Institute of Technology, where he is currently a Professor with the Department of Computer Science. His current research interests include digital signal processing and digital communications. He is a member of IEEJ, IEICE, SICE and IEEE.

Hiroshi Yasukawa received his B.Eng., M.Eng. and Ph.D. degrees in electrical and electronics engineering from Shizuoka University, Hamamatsu Japan, in 1970, 1972, and 1993, respectively. He worked on the research and development of analog and digital communication systems in the Laboratories of Nippon Telegraph and Telephone corporation (NTT) from 1972 to 1998. He was a Professor of Aichi Prefectural University (APU) from 1998 to 2013, and now is a Professor Emeritus of APU. His research interests include digital signal processing, sensor information systems, communication systems and information networks. He is a member of IEEE, IEICE and EURASIP.