RADIO SCIENCE

Volume 38 Number 6 November-December 2003

Published by
American Geophysical Union
Cosponsored by
International Union of Radio Sciences

RADIO SCIENCE Volume 38, Number 6 November–December 2003

Contents

(continued on inside back cover)

(continued from outside back cover)

RADIO SCIENCE, VOL. 38, NO. 6, 1093, doi:10.1029/2002RS002687, 2003

Use of wavelet analysis for detection of seismogenic ULF emissions

L. Alperovich

Raymond and Beverly Sackler Faculty of Exact Sciences, Department of Geophysics and Planetary Sciences, Tel Aviv University, Tel Aviv, Israel

V. Zheludev

School of Computer Sciences, Tel Aviv University, Tel Aviv, Israel

M. Hayakawa

Department of Electronic Engineering, University of Electro-Communications, Chofu, Tokyo, Japan

Received 25 March 2002; revised 12 August 2003; accepted 28 August 2003; published 4 November 2003.

[1] Wavelet analysis is applied to high-resolution magnetic ULF data in a seismoactive region to determine whether there is evidence of ULF electromagnetic emissions that precede or accompany earthquakes. We have developed an algorithm that is specially adapted to the single-station wavelet detection of geomagnetic events. For this purpose we have constructed wavelet-based magnetic signatures of certain earthquakes. Namely, we have used the distribution of energies among blocks consisting of coefficients of wavelet packet transforms. Our computer experiments have shown that common features preceding two strong earthquakes appear in geomagnetic fields recorded close (around 20 km) to the epicenter of the earthquakes. The anomalies occupy a wide range of periods (from 10 s to 250 s). Evidence was also found on the presence of short period seismogenic pulses associated with a strong earthquake (M = 5.6) on March 26, 1997, in Kyushu, Japan. A comparison of extracted geomagnetic variations at two observatories located in its epicentral zone has indicated that seismogenic geomagnetic disturbances occurred 6 to 7 hours prior to the earthquake. *INDEX TERMS:* 7218 Seismology: Lithosphere and upper mantle; 6934 Radio Science: Ionospheric propagation (2487); 6974 Radio Science: Signal processing

Citation: Alperovich, L., V. Zheludev, and M. Hayakawa, Use of wavelet analysis for detection of seismogenic ULF emissions, *Radio Sci.*, *38*(6), 1093, doi:10.1029/2002RS002687, 2003.

1. Introduction

[2] A great deal of evidence of electromagnetic phenomena associated with earthquakes has been accumulated in recent years [see *Hayakawa and Molchanov*, 2002, and references therein]. It is thought that electromagnetic emissions appear before and after an earthquake in a wide frequency range from DC to VHF [*Hayakawa and Molchanov*, 2002]. In particular, significant progress has been achieved in ULF (ultralow frequency, frequencies less than ~1 Hz) emissions during the last decade. First of all, convincing results were reported on precursory ULF emissions for two famous large earthquakes (Spitak and Loma Prieta) [*Fraser-Smith et al.*, 1990; *Bernardi et al.*, 1991; *Molchanov et al.*, 1992; *Kopytenko et al.*, 1993; *Merzer and Klemperer*, 1997]. These earthquakes are extremely large because their magnitudes are M = 6.9 for Spitak and M = 7.1 for Loma Prieta. *Molchanov et al.* [1992] compared the ULF magnetic field characteristics for these earthquakes and found that substantial ULF emissions started a few days before both earthquakes. An additional significant similarity is that the ULF emissions occurred in the same frequency range of 0.01–0.1 Hz. The magnetic intensity is likely to be different for these earthquakes because of the different epicentral distance (about 5 nT recorded at 7 km from the epicenter for Loma Prieta earthquake, which is extremely intense).

[3] Subsequent extensive studies of other earthquakes have supported those characteristics for the two famous

0048-6604/03/2002RS002687$11.00

earthquakes. *Hayakawa et al.* [1996] have studied seismogenic ULF emissions for an earthquake in a Guam with a magnitude ~8.0. Not only by using simple amplitude information, but also by proposing a new signal processing technique (polarization analysis), they have succeeded in finding significant and convincing evidence of precursory ULF emissions. Polarization analysis means that they measured the ratio of vertical to horizontal magnetic field components to distinguish seismogenic ULF signals from other ULF noises. A few points were again confirmed for this earthquake: (1) similar temporal behavior (the first maximum one to two weeks before the quake, then a quiet period, and followed by a sharp increase a few days before the quake), and (2) similar frequency ranges (0.01 ~ 0.1 Hz). This polarization analysis has been found to be very effective in detecting earthquake precursory ULF emissions [*Kopytenko et al.*, 2001; *Hattori et al.*, 2002]. Later *Hayakawa et al.* [1999, 2000] and *Smirnova et al.* [2001] proposed the use of fractal analysis of the ULF data. They suggested that this fractal analysis on the basis of a self-organized criticality concept would be useful for detecting seismogenic ULF emissions.

[4] In addition to the analysis of a single-observatory measurement, there have been successful attempts to use network observations [*Hayakawa*, 2001]. For example, *Gotoh et al.* [2002] applied the principal component analysis to the three-stationed ULF data obtained on the Izu Peninsula to search for precursory ULF emissions associated with the Izu Island earthquake swarm. Also, *Ismaguilov et al.* [2001] have suggested a new direction-finding technique to locate the ULF noise source for the same event on the basis of a gradient magnetometer system.

[5] A few possible mechanisms have been proposed to describe the generation of seismogenic ULF emissions. The first possibility is the so-called electrokinetic effect connected with the generation of streaming potential due to water diffusion through the inhomogenously stressed rock medium [*Fitterman*, 1979]. Another proposed possibility [*Molchanov and Hayakawa*, 1995, 1998] is a mechanism based on stochastic microcurrent activity due to a microfracturing process.

[6] Despite the progress achieved during the last decade in searching for the ULF signature of earthquakes [*Hayakawa and Molchanov*, 2002], there are still significant unanswered questions associated primarily with weak imperceptible signals. One of the key questions is how to construct a formal procedure to distinguish seismogenic ULF signals from the other factors such as magnetic oscillations and man-made noise. In this sense we need to develop an improved technique to detect ULF seismogenic emissions in addition to the few signal processing techniques mentioned above. In this paper we propose a new technique to complement these previous methods. Our proposed method is to construct ULF magnetic "portraits" before and after an earthquake and to investigate whether there is an extinction between them.

Table 1. List of Earthquakes

	Month	Day	Time, UT	Latitude	Longitude	Depth	*M*	*Mw*
1997	01	11	05:50	31.63	131.54	52	5.0	5.5
1997	01	17	15:53	28.81	129.95	33	5.9	6.3
1997	03	26	08:32	31.92	130.43	10	5.6	6.1
1997	04	02	19:33	31.82	130.09	10	5.1	5.5
1997	04	02	21:47	33.36	132.22	45	4.9	5.1
1997	04	05	04:25	31.95	130.43	33	4.7	5.1
1997	05	13	05:38	31.82	130.28	33	5.6	6.1

2. Outline of the Algorithm

[7] We undertook an examination of half-year high-resolution three-component magnetic recordings (1-s sampling rate) taken at the geomagnetic observatory at Kagoshima (geographic coordinates; 31.50N, 130.70E) [*Yumoto et al.*, 1992] to determine whether there were any distinguishable magnetic variations that might have had an earthquake origin. Table 1 summarizes the dates (year, month, time), geographic coordinates, depth (km) and the values of surface-wave magnitude (M) and momentum (Mw) for 7 earthquakes ($Mw \geq 5.0$) that occurred within a radius of 400 km. On January 17, March 26, and May 13 three strong earthquakes took place. We used the data obtained by the World Data Center A for Seismology (National Earthquake Information Center).

[8] We assume that magnetic variations generated by earthquakes are somehow different from conventional oscillations caused by extraterrestrial sources. We also believe that there are properties common to all seismogenic signals that were recorded during the analyzed interval. First, these signals are quasi-periodic in the sense that dominating frequencies exist in each signal. However, these frequencies may vary with the location of the seismic source. For the close sources, these variations are confined to narrow frequency bands.

[9] Therefore, we think that the distribution of the energy (or some energy-like parameters) of signals belonging to a class over different areas of the frequency domain may provide a reliable characteristic signature for this class. In this paper, we focus on the separation of two classes of geomagnetic signals, namely signals of space and tectonic origins.

[10] We develop a special mathematical procedure based on the wavelet technique and suggest an approach based on the identification of wavelet signatures of the observed field during 'quiet' and seismic active periods preceding an earthquake. Our technique is based on a

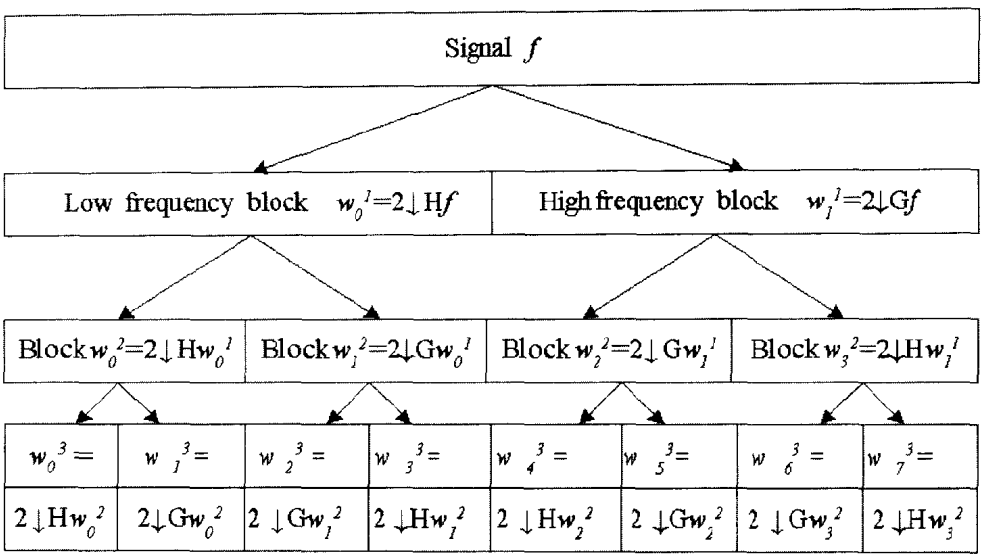

Figure 1. Diagram of the wavelet packet transform up to third level.

generic algorithm for identification of quasi-periodic signals [*Averbuch et al.*, 2001]. Modifications of this algorithm were successfully applied to the classification and detection of moving vehicles and airborne targets [*Averbuch et al.*, 2000], and in medical diagnostics.

2.1. Wavelet Analysis

[11] Wavelet transforms, which have recently become widespread, have been described comprehensively in the literature [see, e.g., *Daubechies*, 1992; *Mallat*, 1999]. Therefore, we provide in Appendix A only the relevant facts that are necessary to understand the construction of the algorithm.

[12] The basic assumption justifying an application of wavelet analysis is that the essential structure of an analyzed signal does not consist of a large number of various finite-length waveforms. The best way to reveal this structure is to represent a signal by a set of basic elements contained in the waveforms coherent with the signal (see Figure 1). Large coefficients are attributed to a few basic waveforms for the structures of signals coherent with the basis. On the other hand, we expect small coefficients for the noise and structures incoherent with all basic waveforms.

[13] Wavelet packet analysis is a highly relevant tool for adaptive search for valuable frequency bands of a signal or class of signals. The wavelet packet transform of a signal produces a set of correlation coefficients of the signal with a multitude of finite-length waveforms whose spectra yield a variety of different partitions of the frequency domain.

[14] An additional way to adapt the tool to the problem stems from a variety of available classes of wavelet packet transforms. A crucial factor for the correct choice of a class is a proper tradeoff between the time-domain localization, the frequency-domain resolution, and the shape of waveforms inherent to this class. After a series of experiments, we chose as a working tool the wavelet packets generated by the Spline filters of order 8 [*Mallat*, 1999]. In Figure 2 we show the waveforms corresponding to the third level of the wavelet packet decomposition generated by these filters. Figure 3 dis-

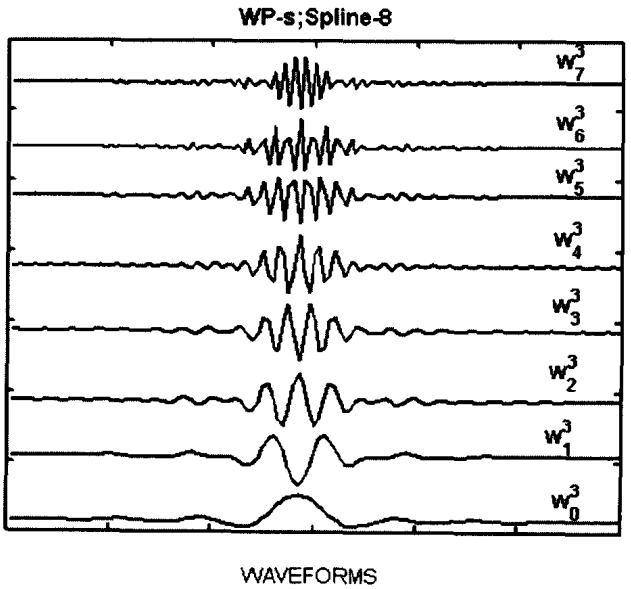

Figure 2. Wavelet packet waveforms after three levels of decomposition generated by Spline filters of 8th order.

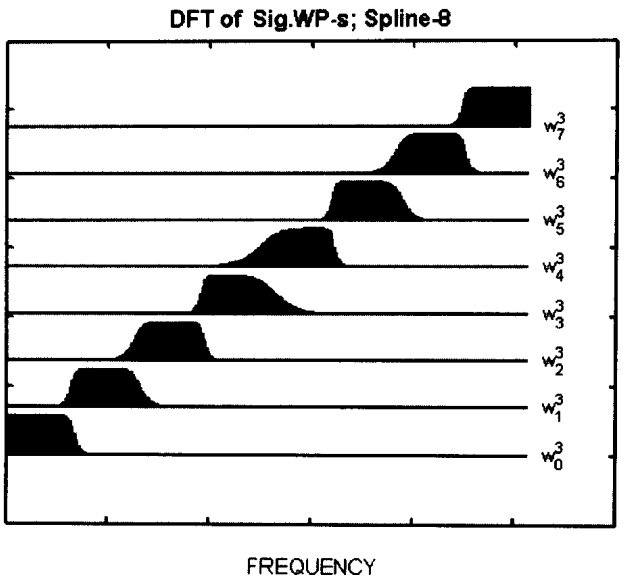

Figure 3. Fourier spectra of wavelet packet waveforms after three levels of decomposition generated by Spline filters of the 8th order.

plays their Fourier spectra. They combine good time localization with the refined split of the frequency domain. We achieve additional time localization by imposing a comparatively short window on each input signal followed by a shift of this window along the signal so that adjacent sections overlap to some extent.

2.2. Formulation of the Approach

[15] The basic assumption is that the general signature of the geomagnetic field in the given region can be obtained as a combination of energies inherent in a small set of the most essential blocks of the wavelet packet decompositions of the recorded signals. We assume a recognizable disturbance of this configuration before and during an earthquake event.

[16] Two intrinsically interesting problems based on geomagnetic information are the problems of classification of geomagnetic signals emitted by the preparation processes of an earthquake and the detection of the presence of a quake-caused signal via analysis of its wavelet signature against the existing database. A crucial factor in having a successful classification is to construct signatures built from characteristic features that enable us to discriminate between the recorded classes. In this paper we only address the detection problem. However, this problem can be treated as a two-class classification problem.

[17] Multiscale wavelet analysis provides a promising methodology for extraction of characteristic features of classes of signals. In the learning phase, we select from a set of signals with known membership a few blocks of coefficients of the wavelet packet transform (WPT) that efficiently discriminate between the given classes of signals and regard the energies within these blocks as the characteristic features of classes of signals. We use these features to train the classifiers used to determine membership of a given signal in the predetermined class. We use the conventional classifiers: Linear Discriminant Analysis (LDA) [*Fisher*, 1936; *Saito and Coifman*, 1996] and Classification and Regression Trees (CART) [*Breiman et al.*, 1993]. While LDA is a common knowledge classifier, the CART method has recently been introduced. We outline its basics in Appendix B. Once we want to classify an unknown signal, we apply the WPT to the signal, and then calculate the energies in the previously selected blocks of coefficients. Finally, we submit the extracted features to one of the above-mentioned classifiers. The classifier, being appropriately trained beforehand, decides which class this signal belongs to.

2.3. Algorithm

[18] The algorithm is centered on two basic issues: (1) selection of the discriminant blocks of the wavelet packet coefficients and (2) discrimination of the signals. We use the WPT based on Spline 8 filters. These transforms reduce the overlapping among the frequency bands associated with different decomposition blocks while retaining suitable time-domain localization.

[19] We treat our problem as a two-class classification problem where one class comprises signals recorded during 'quiet' time intervals. We attribute signals recorded during earthquakes to the second class. Initially, we gather as many recordings as possible for each class. We prepare, from each selected recording that belongs to a certain class, a number of overlapping slices, shifted with respect to each other. These groups of slices form the training set for the search of discriminant blocks.

[20] Each slice is subjected to the WPT up to a level L (vocabulary of notations is given in the Notation section). Typically we choose $L = 8$. The energies of each block of coefficients are calculated in accordance with the chosen measure. As a result we obtain a distribution of the 'energies' of the chosen slice over various frequency bands of widths from $N_F/2$ to $N_F/2^L$, where N_F is the Nyquist frequency. In our case,

$$\begin{aligned}\text{Nyquist frequency} &= N_F \\ &= 0.5 \text{ of the sample frequency} \\ &= f_{\max} = 0.5 \text{ Hz}.\end{aligned}$$

We recall that, at the level j, we have 2^j blocks of coefficients. Respectively, the whole frequency range is divided at this level into $S = 2^j$ subintervals, and the

coefficients of the k^{th} block of j^{th} level correspond approximately to the following frequency band

$$\frac{N_F(k-1)}{S} \sim \frac{N_F k}{S}.$$

After L levels of the WPT, we have altogether $R = 2^{L+1} - 2$ blocks associated with different frequency bands.

[21] The energies of all R blocks of a slice number ν are gathered into the energy vector $E_\nu{}^l$ of length R. The energy vectors along the training set of the class are averaged as follows:

$$E^l = \frac{1}{G^l} \sum_{\nu=1}^{G^l} E_\nu^l,$$

where G^l is the number of slices in the whole set of training signals belonging to the class C^l. The average energy map E^l indicates how the distribution of the 'energies' among various blocks of the decomposition and frequency bands, respectively, takes place within the whole class.

2.4. Evaluation of the Discriminant Power of Decomposition Blocks and Selection of Discriminating Blocks

[22] The average energy map E^l yields some sort of characterization for the class C^l, but it is highly redundant, and, therefore, insignificant information is mixed with significant information. We select the most discriminating blocks to gain a more concise and meaningful representation of the class.

[23] One possible way to do this is to, first, note that for a two-class problem, the difference between two maps provides some insight into the matter. The differences for most blocks are nearly zero. It means that, unlike a few blocks with large values in their differences, they are of no use for discrimination. Therefore, the term-wise difference (absolute values) of the energy maps serves as the discriminant power map for the decomposition blocks: $DP(1, 2) = |E^1 - E^2|$.

[24] Now we are in a position to select a few discriminant blocks that form a type of signatures for the classes. This is not possible immediately because we are in a situation where the frequency bands of the blocks overlap. For example, the blocks w_2^3 and w_3^3 of the third level (see Appendix A) together occupy the same band as the block w_1^2 of the second level, which is considered their "parent." If the latter has a strong discriminant power, then probably at least one of the "children" blocks has the same. To avoid this frequency overlap, we apply a procedure similar to the Best Basis Selection Algorithm [*Coifman and Wickerhauser*, 1992].

[25] The idea is to compare the discriminant power of each pair of the "children" with the power of their parent. When the discriminant power of the parent exceeds the sum of the children powers, the children blocks are discarded and vice versa. As a result, we obtain some nonoverlapping sets of blocks that map the whole frequency domain of our signals, which are referred to as the "most discriminating basis." Typically, this set contains a relatively large number of blocks, especially if the number L of levels of the decomposition is large. Therefore, we select a few blocks with the highest discriminant factors. Moreover, if we are interested in certain frequency bands, we can select the corresponding blocks.

[26] We performed the wavelet packet analysis presenting 27 days with rather strong earthquakes (M > 4.0) (class C^1) within a radius of 500 km from the observatory (depth 0–50 km) and 27 quiet days (class C^2). As a result of the operations described above, we found 12 decomposition blocks such that the distribution of energies among them characterizes the classes to be distinguished.

2.5. Preparation of the Reference Set

[27] We have chosen a number of recordings that belong to the classes to be distinguished. We prepare, from the recordings related to the class C^l, a number (let it be Q^l) of overlapping slices of length n, each shifted with respect to each other by s samples. All the Q^l slices are gathered into a $Q^l \times n$ matrix. Each row of this matrix is operated by the wavelet packet transform up to level L. In the decomposed slice we calculate the "energies" of the t blocks that are selected beforehand. In doing so, we obtain the vector of length t, which we regard as a representative of the chosen slice. These vectors form the $Q^l \times t$ reference matrix R^l associated with the class C^l. We do the same for both classes. These two matrices R^l form reference sets, which are used for the construction of the classification tree (CART; see Appendix B), and as pattern sets for LDA.

[28] After the construction of the classification tree and having pattern sets for LDA, we are in a position to classify test signals. To do so, we must preprocess these signals.

2.6. Preparation of the Test Set

[29] Suppose we are given a signal f whose membership in a certain class has to be established. We form, from the signal f, a number (let it be Q) of overlapping slices of length n, each shifted with respect to each other by s samples. All the Q slices are gathered into a $Q \times n$ matrix, each row of which is operated by the wavelet packet transform up to level L. In the decomposed slice we calculate the "energies" of the t blocks that are selected beforehand. The produced vectors form the $Q \times t$ test matrix T associated with the signal f.

2.7. Making the Decision

[30] Once the test matrix T is ready, we present each row T_i of the matrix to two classifiers: (1) LDA

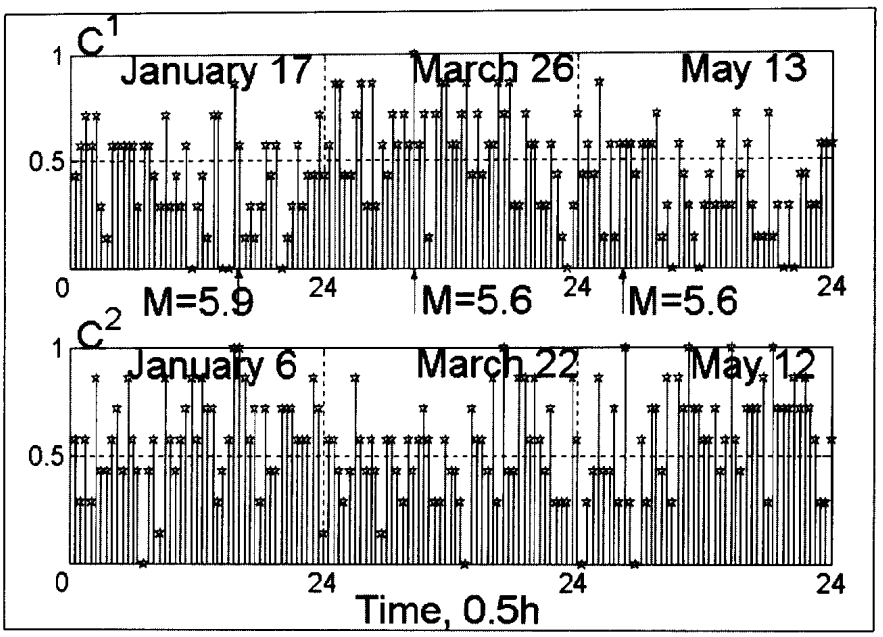

Figure 4. Results of training and classification. The following parameters were used: Spline8 wavelets, CART classifiers, and every 0.5-hour designs. The signals were decomposed up to the 8th level. As training signals, we took recordings for 27 days ('disturbed' days, class C^1) corresponding to the days with earthquakes within a radius of 500 km for 0.5 year and 27 'quiet' days of class C^2. The signals of the C^1 class were January 17 (M = 5.9, 15:53UT), March 26 (M = 5.6, 08:31UT), and May 13, 1997 (M = 5.6, 05:38UT). The distances from the observation point to epicenters and the depths of seismic sources are shown in the top panel. For signals of the C^2 class, we chose 27 days from the 0.5-year recordings. We used all blocks. For decisions, we used the Classification and Regression TREE (CART) classifiers, bottom panel. The upper picture in the panel corresponds to the C^1 class, and the bottom picture to the C^2 class.

calculates a sort of distance of the vector from the pattern sets associated with the classes (C^l, $l = 1, 2$) and attributes it to the class whose distance is the least. (2) CART uses the tree that constructed on the basis of the pattern sets. Once a vector is presented to the tree, it is assigned to one of the subsets of the input set. This determines the most probable membership of the vector.

[31] Then we count the numbers of vectors T_i attributed to each class and make the decision in favor of the class C^l which gets the majority of the vectors. The robustness of the decision is checked by the percentage of the vectors T_i attributed to the class C^l.

3. Analysis of Results

[32] We have conducted a series of experiments on discriminating the signals generated by earthquakes and the background. We processed the signals using the scheme explained above. The recording is processed with sliding overlapped windows of size $n = 1024$. The window is shifted along the signal with a step of $s = 128$ samples. Each window is processed by the WPT up to the 8th level. As a result we have selected various sets of discriminant blocks. To improve the comprehension of the training phase, we have used not only the component recordings (H, D, Z), but also their combinations such as the variations of horizontal orientation angle and inclination. During the training phase we have employed differences between the data for a chosen quake day and the data for a day in the same month with the same Kp index to diminish the influence of the magnetic activity and daylight conditions.

[33] The top picture in Figure 4 illustrates the classification rate for the variations of the Z-component. Signals of class C^1 ('disturbed' days) are from the events, January 17, M = 5.9; March 26, M = 5.6; May 13, M = 5.6. For the selection of discriminant blocks we used geomagnetic recordings around the times of 27 earthquakes of magnitude M > 4.0 appearing at distances up to 500 km from the Kagoshima observatory.

[34] We have used Spline8 WPT up to the 8th level. The CART classifier made decisions on every 0.5-hour fragment of a signal. Each star in the picture corresponds to a single 0.5-hour interval of class C^l. Its height, h^l,

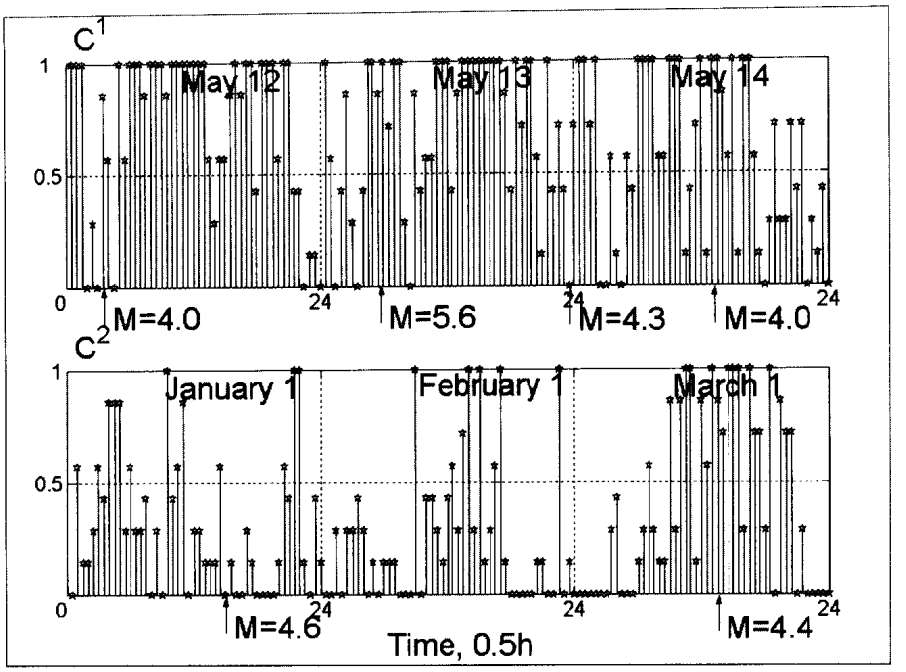

Figure 5. The following parameters were used for training and classification: Spline8 wavelets, LDA classifiers, and every 1-hour design. The signals were decomposed up to the 8th level. Recordings of March 25–27 ('disturbed' days, class C^1) and March 9–11 ('quiet' days, class C^2) are training signals. Signals of the C^1 class submitted to classification were May 12–14, 1997 (M = 5.6, May 13, 05:38UT). Signals of the C^2 class were January 1, February 1 and March 1; 1:2, 4:6 blocks were used for the results presented in the left pictures. The upper picture in each panel corresponds to the C^1 class, the bottom picture to the C^2 class.

may range from 0 to 1 and reflects the probability of an incorrect addressing of the fragment to the class C^l. So, if $h^l = 0$, then the signal S^l is completely classified. If $0 < h < 0.5$ then the probability of a correct answer prevails over the probability of a wrong answer, and the signal is classified. If $h^l = 0.5$, the signal is nonclassified, the closer h^l gets to 0, the more reliable the answer is. Finally, the signal is determined to be misclassified if $0.5 < h^l < 1$.

[35] As seen in the upper panel of Figure 4, within the time interval 0.5–1.5 hours before and 2 hours after the January 17 earthquake almost all signals from classes C^1 and C^2 are classified correctly. Beyond the January quake the classification for March and May seems to fail.

[36] The results of next experiment are shown in Figure 5. The signals are decomposed using the Spline8 WPT up to the 8th level. We have employed as the training signals magnetograms recorded on March 25–27, 1997 (class C^1) and on March 9–11 (class C^2). We have submitted the following signals to classification: the signals of the C^1 class are May 12–14, 1997 (an earthquake with M = 5.6 on May 13). Its epicenter is located almost at the same point as the epicenter of the March 26 earthquake. We chose January 1, February 1 and March 1 for the signals of the C^2 class.

[37] For the decisions, we used both classifiers: Classification and Regression Tree (CART) and Linear Discriminant Analysis (LDA). Figure 6 presents the result of LDA classification. The upper panel corresponds to the disturbed period and the bottom to the quiet period. We can see that the classification rate for the signals of class C^2 is generally less than 0.5; that is, both classifiers (though the result by CART is not shown) correctly classify quiet magnetograms. At the same time a majority of the C^1 signals are misclassified. One can see from the figure that the classification procedure reliably separates the observed magnetic field into two classes only for the two isolated recordings, the one made on May 12, one day before the earthquake, and the made four hours before the earthquake. There the substantial fraction of signals is classified well (LDA classifier).

[38] Common features are found to appear in the geomagnetic field preceding both the March 26 and May 13 earthquakes. The anomalies occupy a wide range of periods (from 10s to 250s). Narrowing of the interval and excluding low-frequency bands degrade classifica-

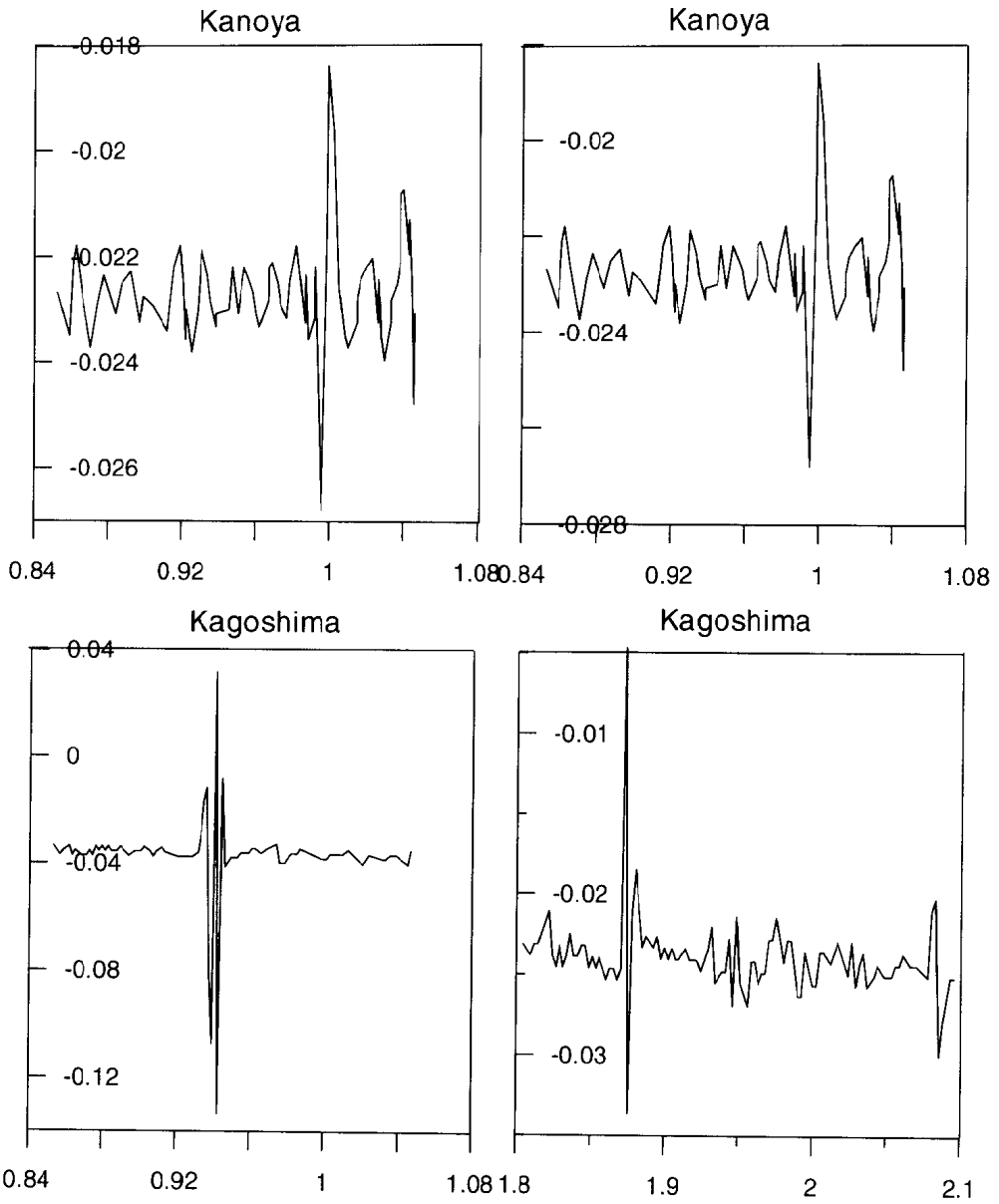

Figure 6. Two sequential pulses observed simultaneously at the Kagoshima and Kanoya observatories. The time delay between the first pulses fixed by both observatories is 3.2 min; that for the second pulses is 6.8 min. The distance between the two observation points is about 25 km, and so the horizontal velocities are ≈8 km/min and ≈4 km/min.

tion. Numerical experiments indicate that some irregularity disturbed the energy signature of geomagnetic variations recorded close (around 20 km) to the epicenter of the earthquake.

[39] As the next stage, we constructed a training set for C^1 class, including new time intervals corresponding to the moments of earthquakes taken from different distances. The classification results did not change up to a radius of 300 km from the Kagoshima observatory when we included six 3-day intervals containing five additional earthquakes. A further increase in radius led to degradation of the classification.

4. Discussion and Conclusions

[40] There are several significant results from this work.

[41] 1. We have constructed and applied a wavelet approach to the problem of searching for geomagnetic precursors of an earthquake. The main idea is to create a

wavelet signature in the ULF field associated with an earthquake and for 'quiet' intervals not accompanied by an earthquake.

[42] 2. Our computer experiments have shown that common features preceding two strong earthquakes appear in the geomagnetic fields recorded close (around 20 km) to the epicenters of the earthquakes. The anomalies occupy a wide range of periods (from 10s to 250s).

[43] The localization of the geomagnetic signature we have discovered does not directly indicate the locality of generated signals because a signal can sometimes spread away from the source over a large distance. During this propagation, the energetic portrait, a distinctive relationship between the energies within different levels and blocks can be lost, and an emitted signal behaves as though it forgot everything about its source.

[44] Three pieces of evidence indicate that the discovered ULF variations are neither magnetosphere nor industrial-induced.

[45] 1. We have studied the sensitivity of the method used to change the training sets of both classes. We have included new time intervals corresponding to the moments of earthquakes taken from different distances. Classification results did not change up to a radius of 300 km from the Kagoshima observatory when we included six 3-day intervals containing five additional earthquakes. A further increase in the radius leads to degradation of the classification.

[46] 2. To verify our results, we have performed the wavelet analysis of records at Kagoshima and Kanoya located almost in the epicentre zone and spaced 20 km apart. We have seen two successive pulses of the same shape with a 3.2 min and a 6.8 min time delay, respectively, 6–7 hours prior to the March 26 earthquake (Figure 6). The pulses manifest themselves predominantly in the H-component. The distance between the two observation points is about 25 km, hence the horizontal wave velocities of the pulses are 8 km/min and 4 km/min. Figure 6 shows how the intensity of the pulses is strongly dependent on the distance. Taking into account the relative location of the observation points and the epicentre of the quake, one can see that the wave propagated from the epicentre via Kagoshima to Kanoya with strong longitudinal magnetic component. It follows that the attenuation rate α is 0.1 km^{-1}. Leaving aside the question of the generation of such impulses [*Gershenzon and Gokhberg*, 1994; *Molchanov and Hayakawa*, 1995, 1998], in principle, the availability or lack of magnetic pulses can be explained not only by the remoteness of sources from the ground observer, but also by the ground conductivity and propagation conditions.

[47] Deep magnetotelluric soundings performed in the active ('hot') regions revealed a layer of conductivity $\sim 0.1 S/m \approx 10^9 s^{-1}$ at a depth ~30 km [*Vanyan*, 1997]. If a source of the ULF seismogenic radiation is located in that layer, then the wave propagation in such a waveguide is defined not only by the thickness of the waveguide but also by damping on the walls. The damping α of the electromagnetic wave with a strict transverse electric field (wave of magnetic type) in a plane waveguide of thickness b is $\alpha = \omega\zeta'/ck_z b$) [*Landau and Lifshitz*, 1984]. Here k_z is the wave number along the waveguide axis, c is the speed of light, and ζ' is the surface impedance of the surrounding media ($\zeta' = (1 - i)\sqrt{\omega/8\pi\sigma}$) ($\sigma$ is the specific conductivity (s^{-1})). The group velocity u_g is given by $u_g = \partial\omega/\partial k_z = (c^2 k_z)/\omega$. Hence, the propagation velocity along the waveguide axis for the wave H_{10} can be estimated as $u_g = c/(2\alpha b\sqrt{\sigma T})$. Let $\sigma = 10^9 s^{-1} T \approx 100s$, $b = 30$ km, then we have $u_g \approx 10$ km/min.

[48] Direct comparisons of recordings preceding another earthquake with the epicentre at the same place (Kyushu, M = 5.6, May 13, 1997) have not demonstrated any such simultaneous visible anomalies. Although it occurred in the same place, the reason why the pulses were not observed during the second earthquake may be due to their distinct depths. The hypocentre of the March 26 quake was at a depth of 10 km, and the May 13 at 33 km (see Table 1). Due to the proximity to the ground surface, the signal of the March quake should be of high intensity.

[49] Thus, for the May 13 quake, the source is located in a highly conductive layer surrounded by a medium of low conductivity. While, the situation of the March 26 earthquake is a waveguide of low conductivity with highly conductive walls. An electromagnetic wave can propagate here as a 'diffusive' wave of low velocity and high damping.

[50] 3. Confirmation that the discovered impulses are not man-made but natural was obtained by thorough analysis of records for half the year of 1997 at both the Kagoshima and Kanoya observatories. We have tried to find impulses as discussed above appearing simultaneously at these sites to exclude the artificial interference field source. These impulses are unique in this sense.

[51] Although we do not claim to have solved the problem of finding ULF electromagnetic precursors, we feel that the developed method is rather strong and universal. In this paper, our approach was based on single-point measurement and two-class recognition, namely 'with' and 'without' an earthquake. This is not a limitation imposed by the method, and there is no reason why the data could not be classified as a multiclass datum, which corresponds, for example, to various magnetic activities. The method is also not limited to one or multicomponent recognition.

[52] We need further work to develop the approach of exploring vector features and multipoint observations [*Alperovich and Zheludev*, 1998] of the geomagnetic variations of two separate groups, namely: (1) variations

caused by magnetospheric processes and (2) variations produced by the preparation processes of an earthquake.

Appendix A

[53] The result of application of the wavelet transform to a signal f of length $N = 2^J$ is a set of correlation coefficients of the signal with scaled and shifted versions of two basic waveforms—'father' and 'mother' wavelets. The transform is implemented by step-wise subband filtering of the signal by a conjugate pair of low (H) and high (G) pass filters followed by downsampling. In the first decomposition step, the filters are applied to the signal f and, after downsampling, the result has two blocks of coefficients $w_0^1 = 2 \downarrow Hf$ and $w_1^1 = 2 \downarrow Gf$ of the first scale, each of a size $n/2$. (The symbol $2 \downarrow x$ means downsampling of an array x resulting from the removal of all terms with odd indices: $y = 2 \downarrow x \iff y_k = x_{2k}$.)

[54] These blocks consist of correlation coefficients for signals with 2-sample shifts of the low frequency 'father' wavelet and the high frequency 'mother' wavelet, respectively. The block w_0^1 contains the coefficients necessary for reconstruction of the low-frequency component of the signal. Similarly, the high-frequency component can be reconstructed from the block w_1^1. In this sense, each decomposition block is linked to a certain half of the frequency domain of the signal.

[55] While the block w_1^1 is stored, it is subjected to the same decomposition procedure to generate the second level (scale) blocks w_0^2 and w_1^2, each of a size $n/4$. These blocks consist of the correlation coefficients with 4-sample shifts of the two-times dilated versions of the 'father' and 'mother' wavelets. Their spectra share the low-frequency band previously occupied by the spectrum of the original father wavelet. Then w_0^2 is decomposed in the same manner and the procedure is repeated m times. Finally, the signal is transformed into a set of blocks $f \rightarrow \{w_0^m, w_1^m, w_1^{m-1}, w_1^{m-2}, \ldots, w_1^2, w_1^1\}$ up to the m decomposition level.

[56] This transform is orthogonal. One block is remained at each scale (level) except for the last one. Each block is related to a single waveform. Thus the total number of waveforms involved in the transform is $m + 1$. Their spectra cover the whole frequency domain and split it in a logarithmic manner. Each decomposition block is linked to a certain frequency band, and since the transform is orthogonal, the l_2 norm of the block coefficients is equal to the l_2 norm of the component of the signal f whose spectrum occupies this band.

[57] Through the application of wavelet packet transform many more waveforms, namely, 2^j waveforms at the j^{th} decomposition level are involved in the transformation of the signal. The difference between the wavelet packet and wavelet transforms begins in the second step of the decomposition. Now both blocks w_0^1 and w_1^1 are stored at the first level, and at the same time both are processed by a pair of filters, H and G, which generate four blocks w_0^2, w_1^2, w_2^2, and w_3^2 in the second level. These are the correlation coefficients of the signal with 4-sample shifts of the four waveforms whose spectra split the frequency domain into four parts. All of these blocks are stored in the second level and transformed into eight blocks in the third level, etc. The involved waveforms are well localized in time and frequency domains. Their spectra form a refined partition of the frequency domain (into 2^j parts in the j^{th} scale). Correspondingly, each block of the wavelet packet transform represents a certain frequency band.

[58] The flow of the wavelet packet transform is given in Figure 1. The partition of the frequency domain corresponds approximately to the location of blocks in the diagram. We may argue that the wavelet packet transform bridges the gap between time-domain and frequency-domain representations of a signal. As we advance to a coarser level (scale) we see a better frequency resolution at the expense of time domain resolution and vice versa. In principle, the transform of a signal of a length $n = 2^J$ can be implemented up to the J^{th} decomposition level. At that level there exist n different waveforms, which are close to the sine and cosine waves with multiple frequencies.

[59] There is a duality in the nature of the wavelet coefficients of a certain block. On the one hand, they indicate the presence of the corresponding waveform in the signal and measure its contribution. On the other hand, they evaluate the contents of the signal inside the related frequency band.

[60] There are many wavelet packet libraries. They differ from each other by their generating filters H and G, by the shape of the basic waveforms, and by the frequency content. In Figure 2 we describe the wavelet packets corresponding to the third level of decomposition generated by the so-called Spline filters of the 8th order. While the waveforms are not localized as well as some other wavelets in time domain, they produce satisfactory splitting of the frequency domain.

Appendix B: Outline of the CART Algorithm

[61] A comprehensive exposition of the CART scheme can be found in *Breiman et al.* [1993]. For simplicity we consider a two-class classification problem.

B1. Building the Tree

[62] The space X of input patterns from the reference set consists of two reference matrices V^l ($l = 1, 2$) of sizes $\mu_l \times n$, respectively. We assume that $\mu_1 = \mu_2$. The i^{th}

row of the matrix V^l, is a vector $V^l(i, j)$ of length n, representing the signal s_i^l, which belongs to the class C^l. In our case, n is equal to the number of discriminant blocks. All row vectors should be normalized as follows

$$0 \le V^l(i, j) \le 1, \qquad \sum_{j=1}^{n} V^l(i, j) = 1.$$

The tree-structured classifier to be constructed has to divide our space X into J disjoint subspace (terminal nodes):

$$X = \bigcup_{\nu=1}^{J} X_\nu^t. \tag{B1}$$

Each subspace X_ν^t must be "pure" in the sense that the percentage of vectors from one of the atrices V^l must prevail as the percentage of vectors from the other matrix. (In the original space both are 50%.) The construction of the binary tree starts by a split of X into two descendant subspaces:

$$X = X_1 = X_2 \bigcup X_3, \qquad X_2 \bigcap X_3 = \emptyset.$$

Here $\cup$, $\cap$, and $\otimes$ are, respectively, a sign of joining, a common part, and an empty intersection.

[63] To do this, the CART chooses a split variable y_j and a split value z_j in a way to achieve minimal possible "impurity" of the subspaces X_2 and X_3. The split rule for the space X_1 is as follows: If a vector $y = (y_1, \ldots, y_n)$ satisfies the condition $y_j \le z_j$, then it is directed to the subspace X_2, otherwise it is directed to the subspace X_3. In addition, we divide the subspace X_2 in a similar manner:

$$X_2 = X_4 \bigcup X_5, \qquad X_4 \bigcap X_5 = \emptyset.$$

[64] The subsequent split variable y_k and split value z_k are selected so that the data in each of the descendant subspaces are purer than the data in the parent subspace. Then one of the subspaces X_4 or X_5 can be further divided recursively until we reach the so-called terminal subspace X_1^t, which is not spitted further.

[65] The decisions whether a subspace is classified as a terminal subspace depends on the predetermined minimal "impurity" and the minimal size of the subspace. The terminal subspace X_1^t is assigned to the class C^l with the probability

$$p_1^l = \frac{m_1^l}{m_1},$$

where m_1^l is the number of points in the node X_1^t belonging to class C^l, and m_1 is the total number of points in the subspace X_1^t. As the termination is reached

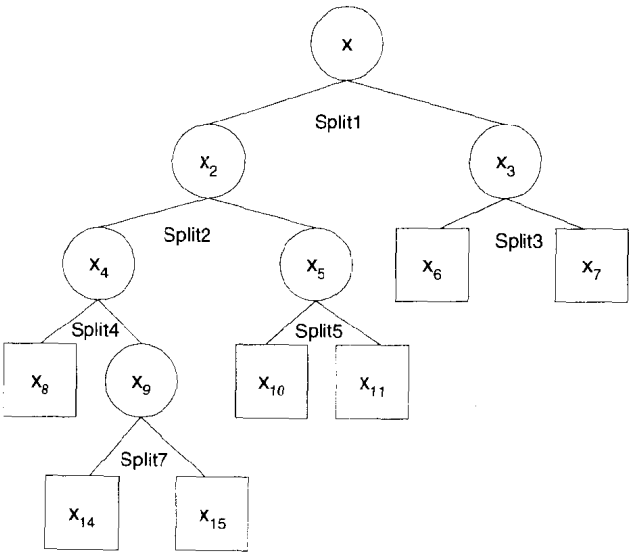

Figure B1. Example of the classification and regression tree for a two-class problem. The terminal nodes are indicated by rectangular boxes and are designated by a class label; circles indicate the nonterminal nodes.

in subspace X_1^t, we return to subspace X_3 still unsplit. Similarly, we reach the next terminal subspace X_2^t. We do the same with one of the yet nonsplit subspaces and finally come to the tree (B1). In the terminology of graph theory, the space X is called the root node; the nonterminal and terminal subspaces are the nonterminal and terminal nodes. This process is illustrated in Figure B1. The terminal nodes are marked as rectangles.

B2. Classification

[66] A vector $x = (x_1, \ldots, x_n)$ is submitted to the tree. In the first step its coordinate x_j is checked. If the inequality $x_j \le z_j$ holds then x is directed to the node X_3; otherwise it is directed to the node X_3. Finally, by checking subsequent split variables, the vector is forwarded to a terminal node X_r^t, which and labeled as class C^l with the probability p_r^l.

Notation

j index of a level.
k index of a block at jth level.
K number of shifted slices in the signal f.
l class ($l = 1, 2$).
L number of levels (or index of the last one).
n length of the training slice ($n = 1024$).
t number of chosen blocks which characterize the signal.
w block; thus w_k^j - kth block on the level j.

[67] **Acknowledgments.** The geomagnetic data of the Kagoshima observatory were kindly provided by K. Yumoto and the members of the 210^0 MM team. We are grateful to them for their contribution to the successful operation of the station. One of the authors (M.H.) expresses his gratitude to the Mitsubishi Foundation for its support.

References

Alperovich, L., and V. Zheludev, Wavelet transform as a tool for detection of geomagnetic precursors of earthquakes, *J. Phys. Chem. Earth*, *23*(9–10), 965–967, 1998.

Averbuch, A. Z., V. A. Zheludev, and I. Kozlov, Wavelet based algorithm for acoustic detection of moving ground and airborne targets, in *Wavelet Applications in Signal and Image Processing IX*, edited by A. Aldroubi, A. F. Laine, and M. A. Unser, *Proc. SPIE*, *4119*, 2000.

Averbuch, A. Z., I. Kozlov, and V. Zheludev, Wavelet packet based algorithm for identification of quasi-periodic signals, in *Wavelet Applications in Signal and Image Processing IX*, edited by A. Aldroubi, A. F. Laine, and M. A. Unser, *Proc. SPIE*, *4478*, 353–360, 2001.

Bernardi, A., A. C. Fraser-Smith, P. R. McGill, and O. G. Villard Jr., ULF magnetic field measurements near the epicentre of the Ms 7.1 Loma Prieta earthquake, *Phys. Earth Planet. Inter.*, *68*, 45–63, 1991.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Chapman and Hall, New York, 1993.

Coifman, R. R., and M. V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. Inf. Theory*, *38*, 713–719, 1992.

Daubechies, I., Ten lectures on wavelets, Soc. for Ind. and Appl. Math., Philadelphia, Pa., 1992.

Fisher, R. A., The use of multiple measurements in taxonomic problems, *Ann. Eugenics*, *7*, 179–188, 1936.

Fitterman, D. V., Theory of electrokinetic-magnetic anomalies in a faulted half-space, *J. Geophys. Res.*, *84*, 6031–6040, 1979.

Fraser-Smith, A. C., P. R. McGill, R. A. Helliwell, and O. G. Villard Jr., Ultralow-frequency magnetic field measurements near the epicenter of the Ms 7.1 Loma Prieta earthquake, *Geophy. Res. Lett.*, *17*, 1465–1468, 1990.

Gershenzon, N. I., and M. B. Gokhberg, On the origin of anomalous ultralow-frequency geomagnetic disturbances prior to Loma Prieta, California, earthquake, *Phys. Solid Earth*, *30*, 112–118, 1994.

Gotoh, K., Y. Akinaga, M. Hayakawa, and K. Hattori, Principal component analysis of ULF geomagnetic data for Izu Islands earthquakes in July 2000, *J. Atmos. Electr.*, *22*(1), 1–12, 2002.

Hattori, K., Y. Akinaga, M. Hayakawa, K. Yumoto, T. Nagao, and S. Uyeda, ULF magnetic anomaly preceding the 1997 Kagoshima earthquakes, in *Seismo Electromagnetics: Lithosphere-Atmosphere-Ionosphere Coupling*, edited by M. Hayakawa and O. A. Molchanov, pp. 19–28, Terra Sci., Tokyo, 2002.

Hayakawa, M., NASDA's earthquake remote sensing frontier research: Seismo-electromagnetic phenomena in the lithosphere, atmosphere and ionosphere, final report, 228 pp., Univ. of Electro-Commun., Tokyo, March 2001.

Hayakawa, M., and O. A. Molchanov (Eds.), *Seismo Electromagnetics: Lithosphere-Atmosphere-Ionosphere Coupling*, 477 pp., Terra Sci., Tokyo, 2002.

Hayakawa, M., R. Kawate, O. A. Molchanov, and K. Yumoto, Results of ultralow-frequency magnetic field measurements during the Guam earthquake of 8 August 1993, *Geophys. Res. Lett.*, *23*, 241–244, 1996.

Hayakawa, M., T. Itoh, and N. Smirnova, Fractal analysis of ULF geomagnetic data associated with the Guam earthquake on August 8, 1993, *Geophys. Res. Lett.*, *26*(18), 2797–2800, 1999.

Hayakawa, M., T. Itoh, K. Hattori, and K. Yumoto, ULF electromagnetic precursors for an earthquake at Biak, Indonesia, on February 17, 1996, *Geophys. Res. Lett.*, *27*(10), 1531–1534, 2000.

Ismaguilov, V. S., Y. A. Kopytenko, K. Hattori, P. M. Voronov, O. A. Molchanov, and M. Hayakawa, ULF magnetic emission connected with under sea bottom earthquakes, *Nat. Hazards Earth Syst. Sci.*, *1*, 23–31, 2001.

Kopytenko, Y. A., T. G. Matiashvili, P. M. Voronov, E. A. Kopytenko, and O. A. Molchanov, Detection of ultralow-frequency emissions connected with the Spitak earthquake and its aftershock activity, based on geomagnetic pulsations data at Dusheti and Vardzia observatories, *Phys. Earth Planet. Inter.*, *77*, 85–95, 1993.

Kopytenko, Y. A., V. Ismaguilov, M. Hayakawa, N. Smirnova, V. Troyan, and T. Peterson, Investigation of the ULF electromagnetic phenomena related to earthquakes: Contemporary achievements and the perspectives, *Ann. Geofis.*, *44*(2), 325–334, 2001.

Landau, L. D., and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed., edited by E. M. Lifshitz and L. P. Pitaevskii, 460 pp., Pergamon, New York, 1984.

Mallat, S., *A Wavelet Tour of Signal Processing*, Academic, San Diego, Calif., 1999.

Merzer, M., and S. L. Klemperer, Modelling low-frequency magnetic field precursors to the Loma-Prieta earthquake with a precursory increase in fault-zone conductivity, *Pure Appl. Geophys.*, *150*, 217–232, 1997.

Molchanov, O. A., and M. Hayakawa, Generation of ULF electromagnetic emissions by microfracturing, *Geophys. Res. Lett.*, *22*, 3091–3094, 1995.

Molchanov, O. A., and M. Hayakawa, On the generation mechanism of ULF seismogenic electromagnetic emissions, *Phys. Earth Planet. Inter.*, *105*, 201–210, 1998.

Molchanov, O. A., Y. A. Kopytenko, P. M. Voronov, E. A. Kopytenko, T. G. Matiashvili, A. C. Fraser-Smith, and A. Bernardi, Results of ULF magnetic field measurements near the epicenters of the Spitak (*Ms* = 6.9) and Loma Prieta (Ms = 7.1) earthquakes: Comparative analysis, *Geophys. Res. Lett.*, *19*, 1495–1498, 1992.

Saito, N., and R. R. Coifman, Improved local discriminant bases using probability density estimation, *Proc. Am. Stat. Assoc.*, *5*, 312–321, 1996.

Smirnova, N., M. Hayakawa, K. Gotoh, and D. Volobuev, Scaling characteristics of ULF geomagnetic field at the Guam seismoactive area and their dynamics in relation to the earthquake, *Nat. Hazards Earth Syst. Sci.*, *1*, 119–126, 2001.

Vanyan, L. L., *Electromagnetic Soundings* (in Russian), 218 pp., Nauchnyi MIR, Moscow, 1997.

Yumoto, K., et al., Globally coordinated magnetic observations along 210^0 magnetic meridian during STEP period: 1. Preliminary results of low-latitude Pc3's, *J. Geomagn. Geoelectr.*, *44*, 261–276, 1992.

L. Alperovich, Raymond and Beverly Sackler Faculty of Exact Sciences, Department of Geophysics and Planetary Sciences, Tel Aviv University, 69978 Tel Aviv, Israel.

M. Hayakawa, Department of Electronic Engineering, University of Electro-Communications, Chofu, 182-2525 Tokyo, Japan. (hayakawa@whistler.ee.uec.ac.jp)

V. Zheludev, School of Computer Sciences, Tel Aviv University, 69978 Tel Aviv, Israel.