

Application of a wavelet technique for the detection of earthquake signatures in the geomagnetic field

L. Alperovich¹, V. Zheludev², and M. Hayakawa³

¹Dept. of Geophysics and Planetary Sciences, Ramat-Aviv, 69978 Tel-Aviv University, Israel

²School of Computer Science, Tel-Aviv University, 69978 Tel Aviv, Israel

³Dept. of Electronic Engineering, The University of Electro-Communications, Chofu, Tokyo, Japan

Received: 23 May 2001 – Revised: 17 October 2001 – Accepted: 24 October 2001

Abstract. We developed an algorithm especially adapted to *single-station* wavelet detection of geomagnetic events, which precede or accompany the earthquakes. The detection problem in this situation is complicated by a great variability of earthquakes and accompanied phenomena, which aggravates finding characteristic features of the events. Therefore we chose to search for the characteristic features of both “disturbed” intervals (containing earthquakes) and “quiet” recordings. In this paper we propose an algorithm for solving the problem of detecting the presence of signals produced by an earthquake via analysis of its signature against the existing database of magnetic signals. To achieve this purpose, we construct the magnetic signature of certain earthquakes using the distribution of the energies among blocks, which consist of wavelet packet coefficients.

1 Introduction

It is expected that local geomagnetic disturbances caused by under the surface seismic and tectonic activity contribute substantially to the total geomagnetic field, including magnetospheric and anthropogenic fields. Through the interaction of the flux of interporous conductive liquid with the field of microcracks, these disturbances can be generated by such a flux (Gershenzon et al., 1993; Molchanov 1995). To achieve a better understanding of the sources, properties and behaviour of the local perturbations, it is important, therefore, to learn to recognize these disturbances. Alperovich and Zheludev (1998) suggested a method based on the detection of the differences in the spatial distributions of magnetospheric and tectonogenic perturbations. In addition to conventional wavelet analysis, wavelet packet transforms of simultaneous recordings of geomagnetic signals from a number of observatories were also employed. The coefficients obtained were subjected to either thresholding or mutual comparison in or-

der to reveal the events we were looking for. The procedure of comparison of simultaneous wavelet coefficients allowed for the detection of two types of perturbations spreading from the epicentral zone. The technique of stationary wavelet transform has been used for the best time-domain localization of events.

It was discovered that a quake produces geomagnetic disturbances of two types. The first type consists of the low-frequency oscillations with periods ranging within several hours. The second type consists of relatively high-frequency oscillations with periods of about 10 min. The low-frequency oscillations are propagating along the ground surface with the velocity of sound. The high-frequency ones are radiating outward from the seismoactive region at a rate of about 5 km/s. The aim of this paper is to contribute to the research of geomagnetic perturbations caused by a subsurface source using single-point, long-term geomagnetic observations in a seismic active region. We supply a new approach to the problem of the search of earthquake precursors based on the creation of a wavelet signature of the observed field for quiet periods and the period preceding an earthquake. As original data, we used 3-component recordings of a geomagnetic observatory Kagoshima (58.5° N, 130.7° E) within a 6-month period with a 1 s sampling rate (Yumoto et al., 1992). Within this period two significant earthquakes ($M = 6.2$, 26 March 1997/ 08:31 GMT, 32.0° N, 130.3° E; $M = 6.1$, 13 May 1997/ 05:38 GMT, 31.9° N, 130.3° E) occurred. We used data from the World Data Center A for Seismology (National Earthquake Information Center (<http://neic.usgs.gov/epic>)).

2 General approach

We tried to develop an algorithm which could distinguish between a magnetic field recorded just before an earthquake and the field recorded within a quiet time interval. Obviously, it is a classification problem. The basic assumption is that the general signature of the geomagnetic field in the given region could be obtained as a combination of energies

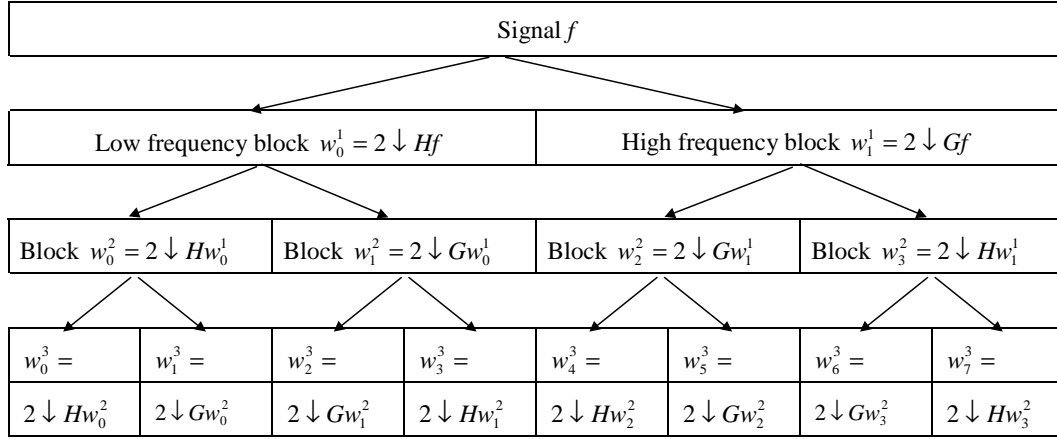


Fig. 1. Flow of the wavelet packet transform. The partition of the frequency domain corresponds approximately to the location of the blocks in the diagram. There are seven levels of the subdivision scheme for the frequency band on different levels L and block numbers N_{block} .

inherent in a small set of the most essential blocks of the wavelet packet decompositions of the recorded signals (see Appendix A). We assume a remarkable disturbance of this configuration before and during the event of earthquake.

A crucial factor in having a successful classification is to construct signatures built from characteristic features that enable one to discriminate among classes. Multiscale wavelet analysis of recorded signals provides a promising methodology for this purpose. Recently, several wavelet-based techniques for feature extraction were developed. We mention the Local Discriminant Bases (LDB) algorithm (Saito and Coifman, 1995), Discriminant Pursuit (Buckheit and Donoho, 1995) and the Matching Pursuit method (Mallat, 1998) for the construction of the wavelet packet bases that separate classes of signals. The LDB algorithm was applied to classification of some geological phenomena (Saito and Coifman, 1997).

In the final phase of the process, in order to identify the earthquake signatures of the predetermined classes of signals, we used conventional classifiers such as Linear Discriminant Analysis (LDA) (Saito and Coifman, 1995) and Classification and Regression Trees (CART) (Breiman, 1993). For training purposes, we use a set of signals with known membership. From this set, we select a few blocks which discriminate efficiently between the given classes of signals. Then, we apply the wavelet packet transform on the signal to be classified. We use as its characteristic features the wavelet packet coefficients normalized by the energy contained in the selected blocks. Finally, we submit the extracted features to one of the classical classifiers, who is appropriately trained beforehand, and decides which class this signal belongs to.

The *algorithm* is centered on two basic issues: (1) selection of the discriminant blocks of the wavelet packet coefficients; (2) discrimination among the signals.

2.1 Selection of discriminant blocks

We used Spline and Coiflet wavelet transforms. These filters reduce the overlapping among frequency bands associated with different decomposition blocks. Initially, we gathered as many recordings as possible for each class, which have to be separated (including to an earthquake and “quiet” interval). Then, we prepared from each selected recording, which belongs to a certain class, a number of overlapping slices of length $n = 2^J$ samples each, shifted with respect to each other. These groups of slices form the training set for the search of discriminant blocks.

We use the l_2 or l_1 norms to measure the energy in the block. Then, the wavelet packet transform is applied up to scale m on each slice of length n from a given class C^l , $l = 1, 2$. This procedure produces mn coefficients arranged into $2^{m+1} - 1$ associated with different frequency bands (see Fig. 1).

The slice is decomposed. The “energies” of each block are calculated in accordance to the chosen measure. As a result we obtain, to some extent, the distribution of the “energies” of the slice over various frequency bands of widths from $N_F/2$ to N_F/m , where N_F is the Nyquist frequency. In our case,

$$N_{Nyquist} = 0.5_{\text{sample frequency}} = f_{\max} = 0.5 \text{ Hz.}$$

The whole frequency range is divided by 2^L sub-intervals (blocks). So the maximal frequency f_{block} of the block of number N_{block} is

$$f_{block} = \frac{N_{block} \cdot f_{\max}}{2^N}.$$

Figure 1 demonstrates a subdivision scheme of the whole frequency range for $L = 7$ to allow us to go to a finer resolution of the wavelet packet transform. The energy is presented by

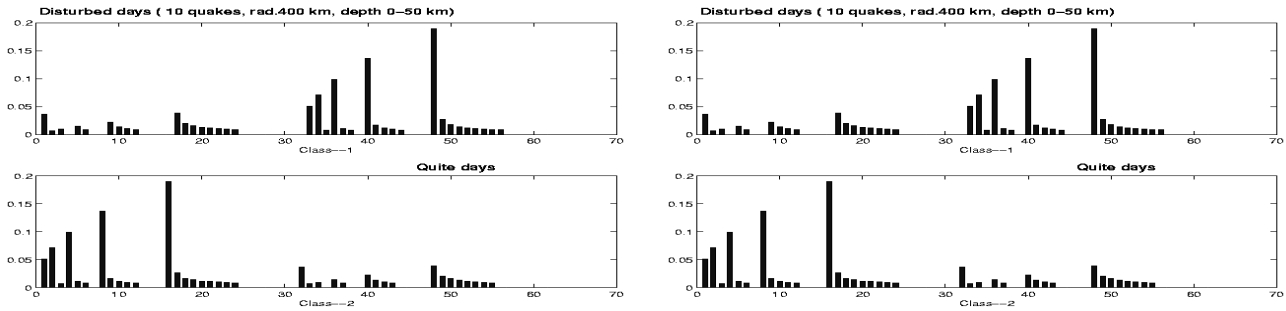


Fig. 2. Energy map for 5 decomposition levels of a two-class problem (left picture) and the difference in class C^1 (10 earthquakes, radius 400 km, depth 0–50 km) and class C^2 (seismic “quiet” days) maps (right picture). The length of a slice is $n = 1024$ samples.

an energy vector E_i^l of length $2^{m+1} - 1$. The energy vectors along the training set of the class are averaged:

$$E^l = \frac{1}{M} \sum_{i=1}^M E_i^l.$$

The average energy map E^l indicates how the distribution of the “energies” among the various blocks of the decomposition and frequency bands, respectively, is taking place within the whole class C^l . The left picture in Fig. 2 displays a typical energy map for 5 decomposition levels of a two-class problem. The heights of the bars indicate the normalized energy of each of the 63 decomposition blocks.

2.2 Discriminant power and selection of discriminating blocks

The average energy map yields some sort of characterization for the chosen class, but it is highly redundant. To gain a more concise and meaningful representation of the class, we select the most discriminating blocks. One possible way to do so is the following. First, note that for a two-class problem, the difference between two maps provides some insight into the problem (Fig. 2, right). The differences for most blocks are near zero. It means that they are of no use for discrimination, unlike a few blocks with large values in their differences. Therefore, the term-wise difference (absolute values) of the energy maps serves as the discriminant power map for the decomposition blocks $DP(1, 2) = |E^1 - E^2|$. As a result of the operations described above, we discover a relatively small set of decomposition blocks, such that the distribution of energies amongst them characterizes the classes to be distinguished.

2.3 Preparation of the reference set

Initially, we chose a number of recordings that belong to the classes to be distinguished, from which we form the reference set. These recordings are sliced similarly to that which was used for the preparation of the training set. For a certain class, we form a number of overlapping slices each of length n form from each selected recording belonging to the class. These are shifted with respect to each other by s samples.

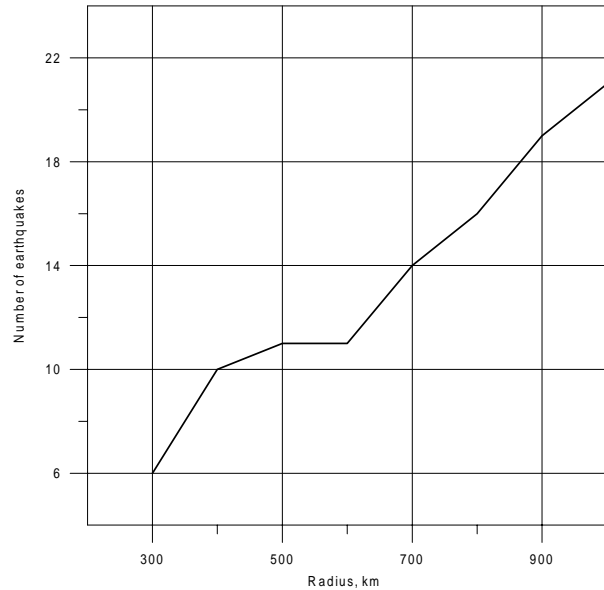


Fig. 3. Dependency of numbers of earthquakes of magnitude range of $M = 5.07.0$ on distance from Kagoshima for January–May, 1997.

Then, we apply the wavelet packet transform up to level m on each row of this matrix. After the decomposition, we calculate only the “energies” of the blocks that were selected before. We do the same for both classes. The reference sets are used for two purposes: (1) as pattern sets for LDA, (2) as training sets for building CART.

The construction of the classification tree is done by a binary split of the space of input patterns so that once a appears in the subspace X_k , its membership could be predicted with a reasonable reliability. The basic idea with the split is that the data in each descendant subset is more “pure” than the data in the parent subset.

After the construction of the classification tree, with pattern sets for LDA, we are in a position to classify test signals.

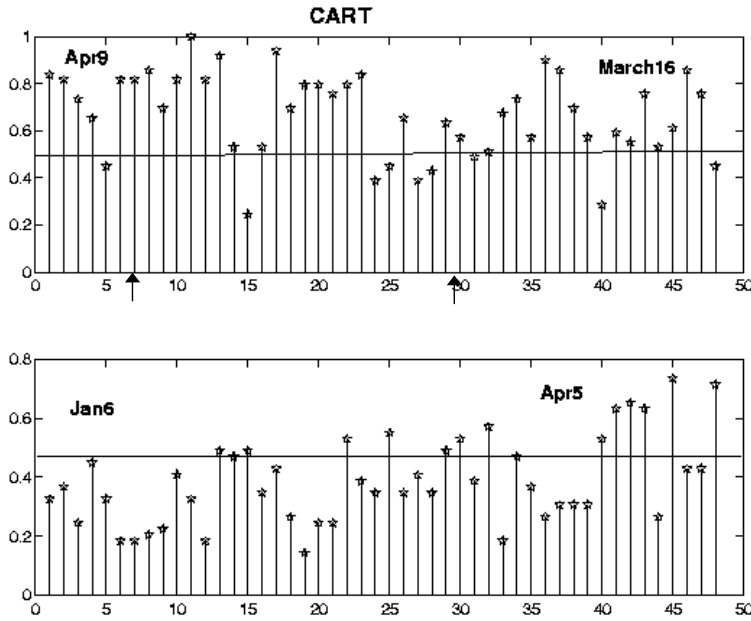


Fig. 4. The top picture illustrates the classification rate for the signals of class C^1 (16 March, 05:51, $M = 5.7$, $\varphi = 35$, $\lambda = 137$; 9 April, 07:02, $M = 5.4$, $\varphi = 26$, $\lambda = 128$) and the bottom picture illustrates the signals of C^2 class. Each (*) in the picture corresponds to a 1 h interval. Vertical lines are the ratio h of the number of vectors attributed incorrectly to the C^l class ($l = 1, 2$) to the total number of the vectors associated with the signals of the l -th class. So, if $0 < h^l < 0.5$, then the number of correct answers prevails over the number of wrong ones, and the signal is well classified. If $h^l = 0$, then the signal is classified completely. If $0.5 < h^l < 1$, then the signal is misclassified.

3 Results

We conducted two series of experiments: (1) classification of signals emitted by earthquakes; and (2) discrimination between signals emitted by earthquakes and the background. We tested various families of wavelet packets and various norms for the feature extraction and various combinations of features presented to the LDA and CART classifiers. The best results were achieved with wavelet packets based on Coiflet 5 filters with 10 vanishing moments and splines of the order 8. LDA classifiers in most experiments outperformed CART.

We processed the signals using the scheme explained above. For the selection of the discriminant blocks we used geomagnetic recordings around the time of 21 earthquakes appearing at distances up to 1000 km from Kagoshima (Yumoto et al., 1992). Figure 3 shows the dependency of the number of earthquakes with magnitudes of 5–8 and depth ranges of 0–100 km in distance. The recording is processed under sliding overlapped windows of size $n = 1024$. The window was shifted along the signal with a step of $s = 128$ samples. Each window was processed by the wavelet packet transform up to the 6th level. The best results were achieved when we used spline 8 wavelet packets and either l_1 or l_2 norms as the “energy” measure for the blocks. As a result, we selected various sets of discriminant blocks.

The top picture of Fig. 4 illustrates the classification rate for the signals of class C^1 (“disturbed” days: 16 March, 05:51 GMT, $M = 5.7$, $\varphi = 35$, $\lambda = 137$; 9 April, 07:02 GMT, $M = 5.4$, $\varphi = 26$, $\lambda = 128$) and the bottom picture illustrates the signals of C^2 class: 6 January (numbers 0–24) and 5 April (24–48). Each (*) in the picture corresponds to a single 1 h interval of class C^1 . Its height, h^1 ,

may range from 0 to 1 and is

$$h^1 = \frac{K^l}{K_c^l}$$

where K^l is the total number of $s W(i, :)$ associated with the signal and K_c^l is the number of the $s W(i, :)$ attributed correctly 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 number of correct answers prevails over the number of wrong ones, and the signal is classified. The closer h^l gets to 0, the more reliable the answer. If $h^l = 0.5$, then the number of correct answers is equal to the number of wrong ones, and the signal is non-classified. And, finally, if $0.5 < h^l < 1$, then the signal is misclassified.

In Fig. 5, we present the results of training and classification. The following parameters were used: Spline 6 wavelets; CART and LDA classifiers. The signals were decomposed up to the 7th level the chosen wavelet was close to rectangular. We took the magnetograms of 25–27 March (“disturbed” days, class C^1) and 9–11 March (“quiet” days, class C^2) as the training signals. We submitted to be classified new disturbed and quiet days using the same training signals as before. The signals of the C^1 class were 12–14 May 1997 ($M = 5.6$, 13 May, 05:38 GMT). The epicenter was almost at the same point as the 26 March earthquakes. For the signals of the C^2 class, we choose 1 January, 1 February and 1 March. In the results presented in the left pictures of Fig. 5, we used 1:2, 4:6 blocks and in the right picture, all blocks. For a decision, we used the Classification and Regression TREE (CART) classifiers in the top panel and the Linear Discriminant Analysis (LDA) in the bottom panel. The upper picture in each panel corresponds to the disturbed period and the bottom picture corresponds to the quiet period.

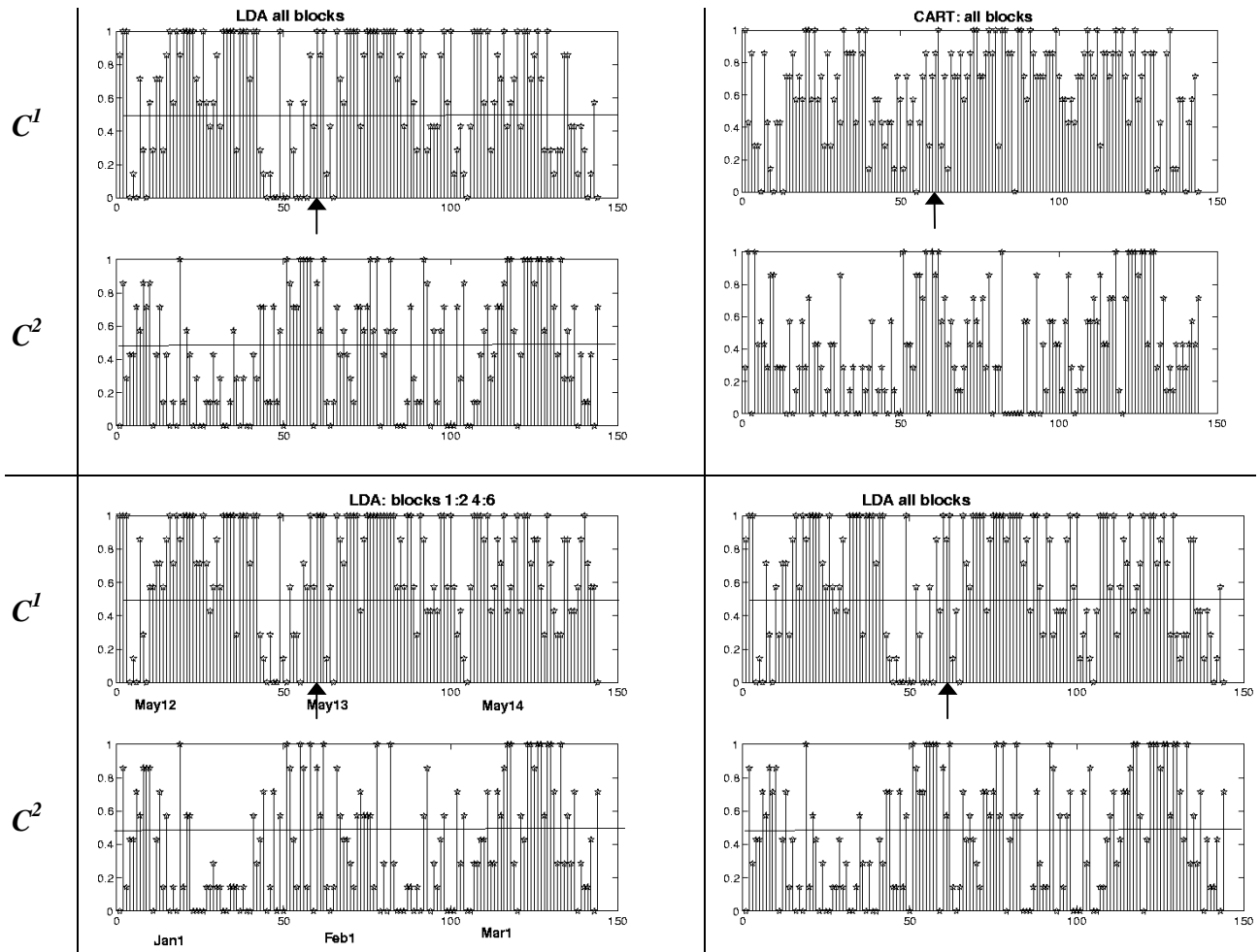


Fig. 5. The results of training and classification. The following parameters were used: Spline 6 wavelets; CART and LDA classifiers. The signals were decomposed up to the 7th level. As the training signals, we took recordings on 25–27 March (“disturbed” days, class C^1) and 9–11 March (“quiet” days, class C^2). We submitted to be classified new “disturbed” and “quiet” days using the training signals as before. The signals of the C^1 class were 12–14 May 1997 ($M = 5.6$, 13 May, 05:38). The epicenter was almost at the same point as the 26 March earthquakes. For the signals of the C^2 class, we chose 1 January, 1 February and 1 March. In the results presented in the left pictures, we used 1:2, 4:6 blocks and in the right picture, all blocks. For the decision we used the Classification and Regression TREE (CART) classifiers in the top panel and the Linear Discriminant Analysis (LDA) in the bottom panel. The upper picture in each panel corresponds to the C^1 class and bottom picture corresponds to the C^2 class.

We can see that the classification rate for the signals of class C^2 are less than 0.5, in general, i.e. both classifiers correctly classify the quiet magnetograms. At the same time the majority of the C^1 signals are misclassified. The classifiers cannot separate them from the C^2 signals, except for the 3–4 hour period before the $M = 5.6$ 26 March earthquake. Here the substantial fraction signals are well classified (LDA classifier). A performed numerical experiment demonstrates that indeed there is something irregular in the energy signature of the geomagnetic variations recorded closely (around 20 km) to the earthquake epicenter.

One can see that the classification procedure separates the observed magnetic field into two classes with a level of confidence just on the isolated parts of recordings, namely before 1 day, and 4 hours before the earthquake. Common features

appear in the geomagnetic field preceding both the 26 March and 13 May earthquakes. The anomalies occupy a wide period range (from 10 s to 250 s). Narrowing of the interval and excluding the low-frequency band degrades the classification (see, for comparison, the right panel of Fig. 5 where we used all blocks and the left panel where rather high frequency blocks were used).

On the next stage, we constructed a training tree for the C^1 class, including new time intervals corresponding to moments of earthquakes taken from different levels of remoteness. The results of the classification did not change until the radius up to 300 km from Kagoshima when we included six 3-day intervals containing five additional earthquakes. The following increase of the radius leads to a degradation of the classification.

Discovering the localization of the geomagnetic signature does not mean directly locality of the generated signals, since a signal can spread away from the source on large distances. In doing so, the energetic portrait which is a distinctive relationship between the energies on different levels and blocks, can be lost, and an emitted signal forgets about the source. For example, Alperovich and Zheludev (1998) applied the wavelet-based approach to the US geomagnetic network data and extracted the wave components by aligning the arrival time of each observation point. Oscillations within the range from 10 min to a few hours, 5 hours and 2 days prior to the occurrence of the strong Loma Prieta earthquake, have been revealed. The main feature of the discovered variations is spreading northward from the epicentral zone. Direct comparison of the results from Alperovich and Zheludev (1998) with field geomagnetic observations close to the epicenter (see Hayakawa et al., 1996; Merzer and Klemperer, 1997; Ismaguilov et al., 2001; and references herein) confirm our conclusion regarding the time and characteristic time-scale of the precursors.

Appendix A Wavelet packet transforms

The result of the application of the m -level wavelet packet transform to a signal f of length $n = 2^J$ is a set of mn correlation coefficients of the signal with shifted versions of $2^{m+1} - 2$ basic waveforms: the wavelet packets. The transform is implemented through iterated application of 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, w_0^1 and w_1^1 , of the first scale, each of size $n/2$. These blocks consist of the correlation coefficients of the signal with 2-sample shifts of the low-frequency father wavelet and high-frequency mother wavelet related to the filters (H) and (G), respectively. The block w_0^1 contains the coefficients necessary for the reconstruction of the low-frequency component of the signal. Due to the orthogonality of the filters, the energy (l_2 norm) of the block w_0^1 is equal to that of the component W_0^1 . Similarly, the high-frequency component W_1^1 can be reconstructed from the block w_1^1 . In that sense, each decomposition block is linked to a certain half of the frequency domain of the signal. The signal f is the sum:

$$f = W_0^1 + W_1^1.$$

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 $w_0^2, w_1^2, w_2^2, w_3^2$ in the second level. These are the correlation coefficients of the signal with 4-sample shifts of the four new 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^m parts at the level m). Correspondingly, each block of coefficients of the wavelet packet transform describes the component of the signal f related to a certain frequency band. The l_2 norm of this block is equal to the norm of the component.

There are many wavelet packet libraries. They differ from each other by their generating filters H and G , the shape of the basic waveforms and the frequency content. It was important for our investigation to have refined frequency resolution. Therefore, we have chosen the wavelet packets derived from the splines of the 8th order.

Acknowledgements. We are grateful to Dr. K. Yumoto and the members of the 210° magnetic meridian team who contributed to the successful operation of the Kagoshima station (Yumoto et al., 1992).

References

- Alperovich, L. and Zheludev, V.: Wavelet transform as a tool for detection of geomagnetic precursors of earthquakes, *J. Phys. Chem. Earth*, 23, 965–967, 1998.
- Averbuch, A. Z., Hulata, E., Zheludev, V. A., and Kozlov, I.: A wavelet packet algorithm for classification and detection of moving vehicles, *Multidimensional Systems and Signal Processing*, 12, 9–31, 2001.
- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J.: *Classification and Regression Trees*, Chapman and Hall, Inc., New York, 1993.
- Buckheit, J. and Donoho, D.: Improved linear discrimination using time-frequency dictionaries, *Proc. SPIE*, 2569, 540–551, 1995.
- Coifman, R. R., Meyer, Y., and Wickerhauser, M. V.: Adapted waveform analysis, wavelet-packets, and applications, In *Proceedings of ICIAM'91*, SIAM Press, Philadelphia, 41–50, 1992.
- Eom, K. B.: Analysis of acoustic signatures from moving vehicles using time-varying autoregressive models, *Multidimensional Systems and Signal Processing*, 10, 357–378, 1999.
- Fisher, R. A.: The use of multiple measurements in taxonomic problems, *Ann. Eugenics*, 7, 179–188, 1936.
- Gershenson, N. I., Gohberg, M. B., Karakin, A. V., Petviashvili, N. V., and Rykunov, A. L.: Modelling the connection between earthquake preparation process and crustal electromagnetic emission, *Phys. Earth. Planet. Inter.*, 77, 85–95, 1993.
- Hayakawa, M., Kawate, R., Molchanov, O. A., and Yumoto, K.: Results of ultra-low-frequency magnetic field measurements during the Guam earthquake of 8 August 1993, *Geoph. Res. Lett.*, 23, 241–244, 1996.
- Ismaguilov, V. S., Kopytenko, Yu. A., Hattori, K., Voronov, P. M., Molchanov, O. A., and Hayakawa, M.: ULF magnetic emissions connected with under sea bottom earthquakes, GRA3, 8957, *Natural Hazards, Geophys. Res. Abst.*, 3, 2001.
- Mallat, S. and Zhang, Z.: Matching pursuit with time-frequency dictionaries, *IEEE Trans. Sign. Proc.*, 41, 3397–3415, 1993.
- Mallat, S.: *A wavelet tour of signal processing*, Acad. Press, 1998.
- Merzer, M. and Klemperer, S. L.: Modeling low-frequency magnetic field precursors to the Loma Prieta earthquake with a precursory increase in fault-zone conductivity, *Pure and Appl. Geophys.*, 150, 217–248, 1997.

- Molchanov, O. A. and Hayakawa, M.: Generation of ULF electromagnetic emissions by microfracturing, *Geoph. Res. Lett.*, 22, 3091–3094, 1995.
- Saito, N. and Coifman, R. R.: Local discriminant bases and their application, *J. Mathematical Imaging and Vision*, 5, 337–358, 1995.
- Saito, N. and Coifman, R. R.: Improved local discriminant bases using probability density estimation, *Proc. Am. Statist. Assoc., Statistical Computing Section*, 5, 312–321, 1996.
- Saito, N. and Coifman, R. R.: Extraction of geological information from magnetic well-logging waveforms using time-frequency wavelets, *Geophysics*, 62, 1921–1930, 1997.
- Yumoto, K., Tanaka, Y., Oguti, T., Shiokawa, K., Yoshimura, U., Isono, A., Fraser, B. J., Menk, F. W., Lynn, J. W., Seto, M., and 210° MM magnetic observation group: Globally coordinated magnetic observations along 210° magnetic meridian during STEP period-1, Preliminary results of low-latitude Pc3's, *J. Geomagn. Geoelectr.*, 44, 261–276, 1992.