RUSSIAN JOURNAL OF EARTH SCIENCES, VOL. 11, ES2006, doi:10.2205/2009ES000387, 2009

PROCEEDINGS OF THE INTERNATIONAL CONFERENCE
Electronic Geophysical Year: State of the Art and Results
3–6 June 2009 • Pereslavl-Zalessky, Russia

Detection of hardware failures at INTERMAGNET observatories: application of artificial intelligence techniques to geomagnetic records study

A. A. Soloviev,[1] Sh. R. Bogoutdinov,[1] S. M. Agayan,[1] A. D. Gvishiani,[1] and E. Kihn[2]

Received 14 November 2009; accepted 14 November 2009; published 17 December 2009.

INTERMAGNET global network consists of 114 observatories around the world, which monitor the Earth's magnetic field in real time. Data from geomagnetic observatories represent time series. In spite of high quality of measuring instruments all of them are subjected to external impact, which affects the quality of records. The present work is devoted to detection of "Spike", "Baseline jump" and "Baseline drift" failures on records. A search is implemented in the framework of a new geoinformatics approach entitled Discrete Mathematical Analysis (DMA) developed at the Geophysical Center of RAS. It is based on fuzzy logic methods and intended for study of multidimensional data sets and time series. Failures on records are treated as anomalies of particular shapes: e.g., jump/spike is an anomaly on a record leading/not leading to its baseline shift. Preliminarily anomalies on records are detected by FCARS algorithm (Fuzzy Comparison Algorithm for Recognition of Signals). A further search of spikes, baseline jumps and baseline drifts among them is done by additional testing, which uses DMA technique. *KEYWORDS: time series; pattern recognition; geomagnetic observations.*

Citation: Soloviev, A. A., Sh. R. Bogoutdinov, S. M. Agayan, A. D. Gvishiani, and E. Kihn (2009), Detection of hardware failures at INTERMAGNET observatories: application of artificial intelligence techniques to geomagnetic records study, *Russ. J. Earth. Sci., 11,* ES2006, doi:10.2205/2009ES000387.

INTERMAGNET global network consists of 114 observatories around the world, which monitor the Earth's magnetic field in real time. Data from geomagnetic observatories represent time series. Each observatory is equipped with instrumentation needed for digital registration of three components and module of geomagnetic field. Data are transmitted to INTERMAGNET data centers in France, Japan, USA, Canada and United Kingdom with a minimal delay [*Kerridge,* 2001].

In spite of high quality standard of measuring instruments all of them are subjected to external impact, which affects the quality of records. Figure 1 illustrates examples of the most common failures, which have to be eliminated prior to further use of obtained observations.

The present work is devoted to recognition of "Spike", "Baseline jump" and "Baseline drift" failures on records. Such failures are characterized by certain morphology; therefore the present work is a continuation of morphologic analysis of time series developed by the authors [*Gvishiani*, 2008]. We interpret magnetogram as a record $y = \{y_k = y(kh)\}$ given on a segment (registration period) T of discrete semi-axis $R_h^+ = \{kh, h > 0,\ k = 1, 2, ...\}$.

Spike. Informal logic which underlies such failure recognition can be formulated in the following way: "Spike is a vertically significant and horizontally insignificant (i.e. singular) disturbance on a record lying on its one side".

Since a spike is vertically significant it should contain points with very large derivatives (Figure 2). We search them adaptively using fuzzy comparisons [*Gvishiani*, 2008; *Kolmogorov*, 1981] and collect into small compact groups. The point is that a real spike often differs from an ideal one, which consists of 3 points (beginning, peak and ending). Nevertheless, as a rule its length is ≤ 10. We set the corresponding Δ, collect points with large derivatives (red) into small compact groups basing on Δ-coherence (i.e. distant from each other not more than by Δ) and fill possible

[1]Institution of the Russian Academy of Sciences Geophysical Center, Moscow, Russia
[2]National Geophysical Data Center, Boulder, Colorado, USA

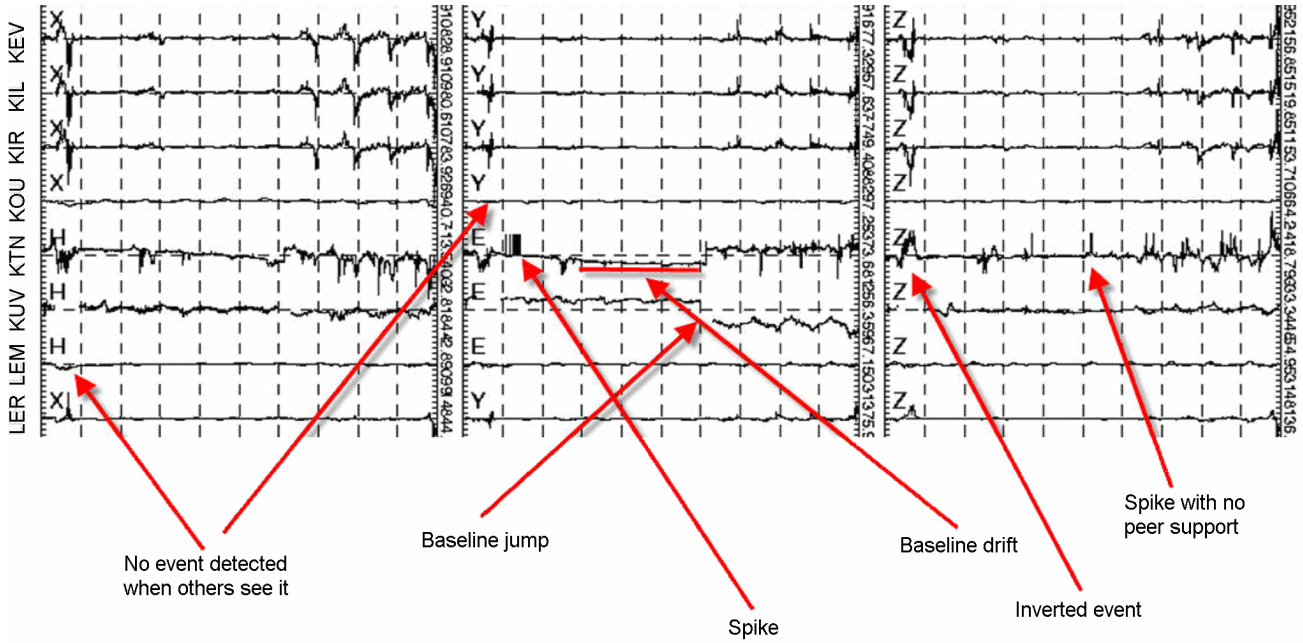

Figure 1. Examples of magnetometer hardware failures reflected on records.

gaps inside them with other points (green). As a result we get segments of green and red points with red ends.

We proceed with spike recognition among such segments basing on the second test for homogeneity.

S – set of nodes with large derivatives on a record y, Δ – coherence parameter. $S = \vee_{\alpha=1}^{m(\Delta)} S_\alpha$ – decomposition into components of Δ-coherence.

Test: Component S_α contains a spike if

$$(y(t) - l_{S_\alpha}(t))(y(t) - r_{S_\alpha}(t)) \underset{\text{fuzzy}}{\geq} 0 \quad \text{on } S_\alpha$$

where

$$l_{S_\alpha}(t) = \text{regr}^1(y|_{[\min S_\alpha - \Delta, \min S_\alpha - 1]})$$

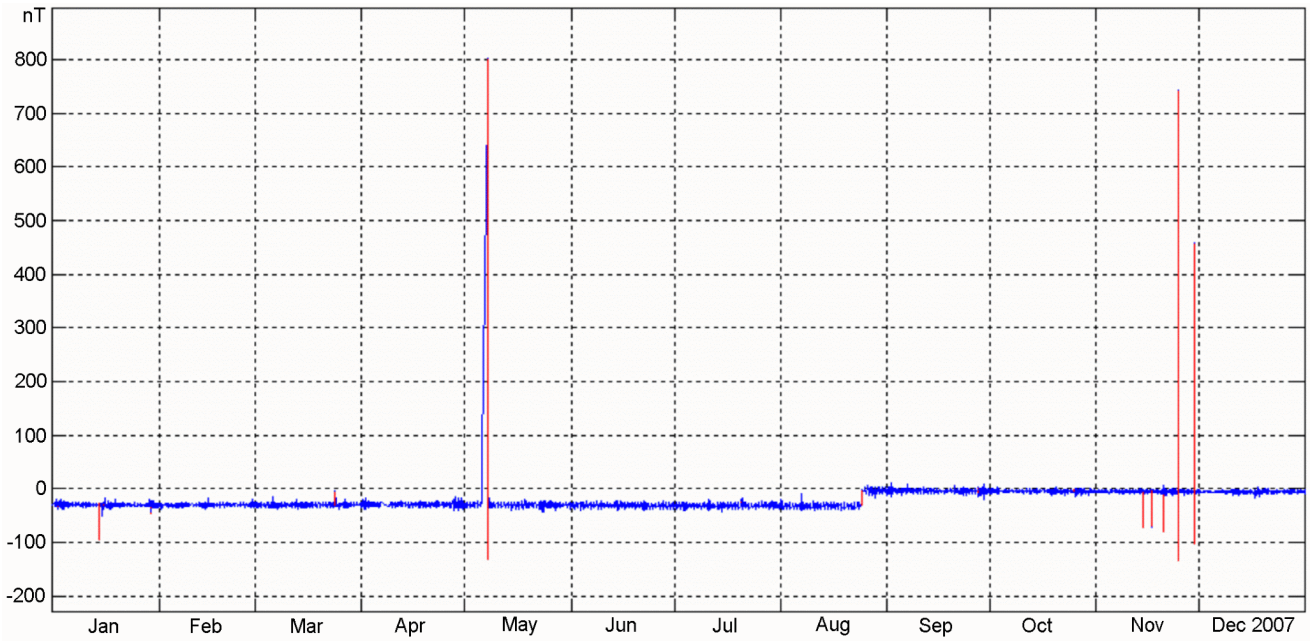

Figure 2. 29 anomalies with large derivatives are recognized.

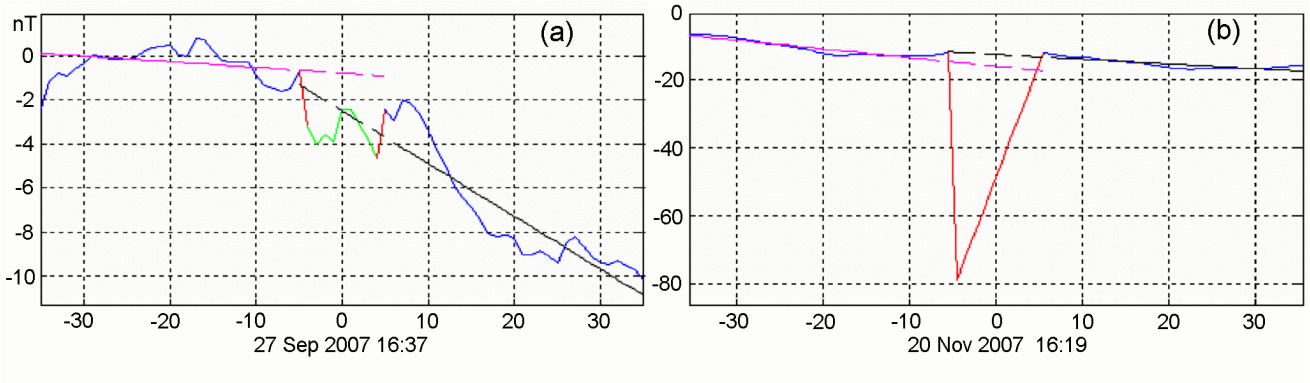

Figure 3. Anomaly on the right has passed the test Homogeneity test verifies same sign definiteness of differences.

$$r_{S_\alpha}(t) = \mathrm{regr}^1(y|_{[\max S_\alpha+1, \max S_\alpha+\Delta]})$$

Any red-green segment should neighbor from both sides with relatively calm segments (blue) with a length not less than Δ. According to them we build special linear regressions l_s and r_s, each of which we continue to initial red-green segment and consider as variants of background for it (Figure 3).

Homogeneity test verifies same sign definiteness of differences $y(t) - l_s(t)$ and $y(t) - r_s(t)$ on a current testing segment. Sign definiteness is understood in a special soft, fuzzy way. Segments, which have passed the test, are considered as spike carriers. As a result, 7 spikes were recognized on a magnetogram given on Figure 2.

Baseline jump. The following informal logic underlies recognition: "Jump is a disturbance on a record leading to its baseline shift". Since jump is considered to be an anomaly first we recognize all anomalies on a record using FCARS [*Gvishiani*, 2008] algorithm (Figure 4).

The first step is local – recognition of anomalies, which contain jumps. Recognition is implemented using fuzzy bounds of a numeric set introduced by us: fuzzy inf and sup. We get corridors of anomaly from the left and from the right, which should be conformed with each other, i.e. lie at significantly different levels.

$$A = y|_{[a,b]}$$

is anomaly recognized by FCARS, $t \in [a, b]$,

$$l \inf y(t), \quad l \sup y(t)$$

are fuzzy bounds of anomaly A on $[a, t]$,

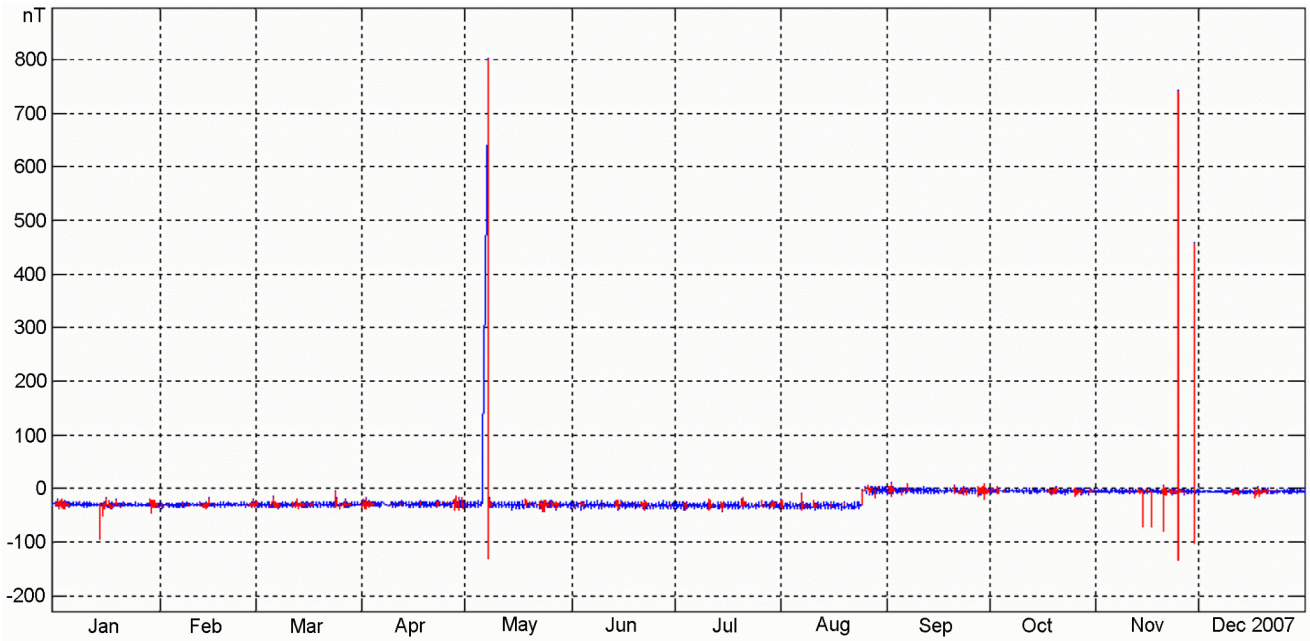

Figure 4. 352 anomalies have been recognized by FCARS.

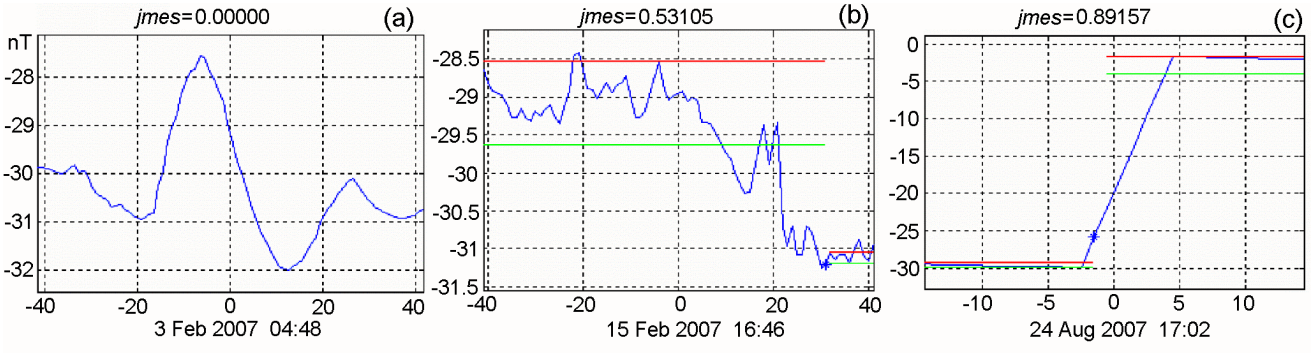

Figure 5. Anomalies on the second and on the third pictures have passed the test.

$$r \inf y(t), \quad r \sup y(t)$$

are fuzzy bounds of anomaly A on $[t, b]$.

Local test (Figure 5): anomaly $A = y|_{[a,b]}$ contains jump if $\exists\, t \in [a, b]$:

$$(l \inf y(t) \le l \sup y(t) < r \inf y(t) \le r \sup y(t)) \wedge$$

$$\left(jmes(A) = \frac{(r \inf y(t) - l \sup y(t))}{(r \sup y(t) - l \inf y(t))} \ge 0.5 \right)$$

(jump up), or

$$(r \inf y(t) \le r \sup y(t) < l \inf y(t) \le l \sup y(t)) \wedge$$

$$\left(jmes(A) = \frac{(l \inf y(t) - r \sup y(t))}{(l \sup y(t) - r \inf y(t))} \ge 0.5 \right)$$

(jump down), where $jmes(A)$ is a local measure of jumpiness.

If an anomaly has passed the local test and contains a jump it is tested for stability in a more global scale in the same way using fuzzy bounds. Λ – parameter of global observation; t – point of local jump $A = y|_{[a,b]}$.

$l \inf y(t, \Lambda)$, $l \sup y(t, \Lambda)$ – fuzzy bounds of record y on $[a - \Lambda, t]$.

$r \inf y(t, \Lambda)$, $r \sup y(t, \Lambda)$ – fuzzy bounds of record y on $[t, b + \Lambda]$.

Global test (Figure 6): local jump of anomaly A at point t is global if

$$\begin{cases} l \inf y(t, \Lambda) \le l \sup y(t, \Lambda) < r \inf y(t, \Lambda) \le \\ r \sup y(t, \Lambda), \\ Jmes(A) = \dfrac{(r \inf y(t, \Lambda) - l \sup y(t, \Lambda))}{(r \sup y(t, \Lambda) - l \inf y(t, \Lambda))} \ge 0.5 \end{cases}$$

or

$$\begin{cases} r \inf y(t, \Lambda) \le r \sup y(t, \Lambda) < l \inf y(t, \Lambda) \le \\ l \sup y(t, \Lambda), \\ Jmes(A) = \dfrac{(l \inf y(t, \Lambda) - r \sup y(t, \Lambda))}{(l \sup y(t, \Lambda) - r \inf y(t, \Lambda))} \ge 0.5 \end{cases}$$

where $Jmes(A)$ is a global measure of jumpiness.

As a result of application of two algorithms given above 7 spikes and 2 jumps were recognized on preliminary data magnetogram recorded at BOU observatory (Boulder, Colorado, USA) in the time period of 01.01.2007–31.12.2007, which corresponds completely with results of visual data control.

Baseline drift. Suppose that at time t algorithm FCARS [*Gvishiani*, 2008] has recognized an anomaly $A_t = y|_{[c(t),d(t)]}$, $d(t) \le t$ on a record $y|_{T(t)}$, $T(t) = \{\bar{t} \in T : \bar{t} \le t\}$.

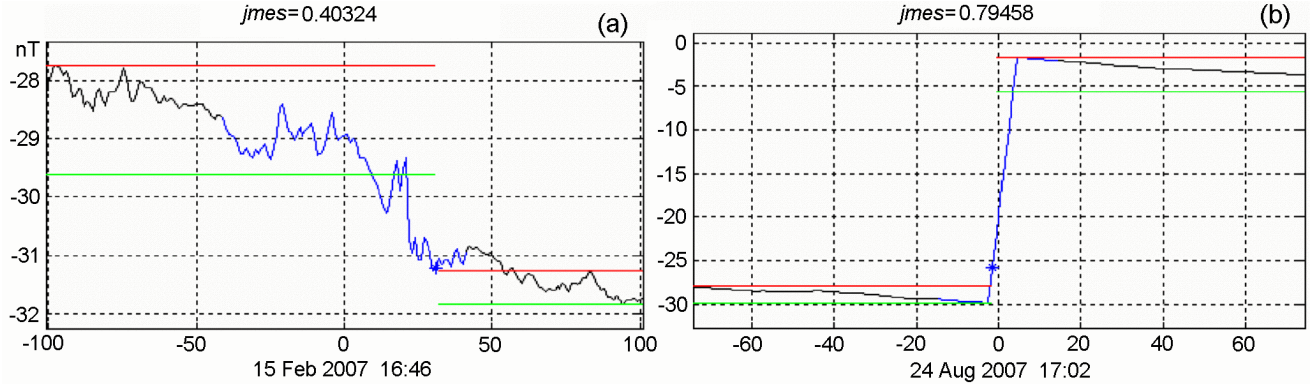

Figure 6. Anomaly on the right has passed the test.

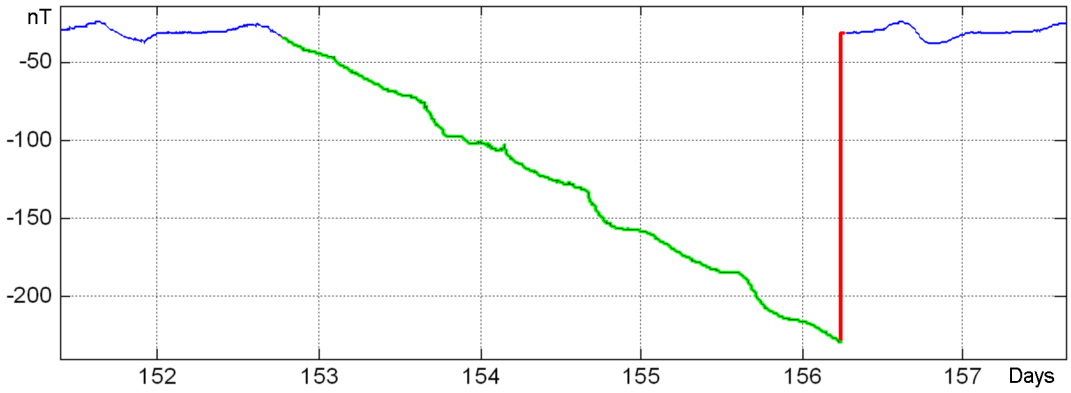

Figure 7. Example of drift recognition.

If anomaly A_t is "fresh" enough: $t - d(t) < \Delta$ then we wait for a time moment $d(t)+\Delta$ to determine whether A_t contains a jump $\bar{J}(A_t) = [a(t), b(t)]$. Suppose that it does. But jump is accompanied with a record baseline shift. A question of the beginning (origin) of such shift arises.

We search it with a loop over $\bar{t}$ to leftward from the beginning of jump $c(t)$. Thus, $\bar{t} < c(t)$ is proposed as the beginning of a shift ended with a jump $y|_{[a(t),b(t)]}$. Further using FCARS algorithm we study docking between fragments $y|_{[\bar{t}-\Delta,\bar{t}]}$ and $A_t = y|_{(d(t),d(t)+\Delta]}$. Docking takes place (point $\bar{t}$ is "good") if rectification $F_{\tilde{y}}(\bar{t}|\Delta)$ of union $\tilde{y} = y|_{[\bar{t}-\Delta,\bar{t}]\cup(d(t),d(t)+\Delta]}$ at docking position $[\bar{t}, d(t) + h]$ is vertically background:

$$F_{\tilde{y}}(\bar{t}|\Delta) < \alpha_s$$

where α_s is a strong level of vertical anomality in FCARS algorithm. Point $t^* + h$ where t^* is the closest to $c(t)$ "good" point is considered as the beginning of a shift. Therefore, a shift itself represents a record fragment $y|_{[t^*+h,a(t))}$ (Figure 7).

In total, 30 INTERMAGNET magnetograms were processed; 100% of spikes, jumps and drifts recognized visually were recognized by the algorithms given above as well. All of the magnetograms processed represent preliminary data records. The list of the records processed with registration time periods indicated is given below:

AIA observatory:
1. F component, 01/01/2007–31/12/2007
2. X component, 01/01/2007–31/12/2007
3. Y component, 01/01/2007–31/12/2007
4. Z component, 01/01/2007–31/12/2007

BOU observatory:
5. F component, 01/01/2007–31/12/2007
6. X component, 01/01/2007–31/12/2007
7. Y component, 01/01/2007–31/12/2007
8. Y component, 22/04/2007–18/05/2007
9. Y component, 14/08/2007–09/09/2007
10. Z component, 01/01/2007–31/12/2007

GUA observatory:
11. F component, 01/01/2007–31/12/2007
12. X component, 01/01/2007–31/12/2007
13. Y component, 01/01/2007–31/12/2007
14. Z component, 01/01/2007–31/12/2007

NVS observatory:
15. F component, 01/01/2007–31/12/2007
16. X component, 01/01/2007–31/12/2007
17. Y component, 01/01/2007–31/12/2007
18. Z component, 01/01/2007–31/12/2007

PPT observatory:
19. F component, 01/01/2007–31/12/2007
20. X component, 01/01/2007–31/12/2007
21. Y component, 01/01/2007–31/12/2007
22. Z component, 01/01/2007–31/12/2007

SOD observatory:
23. F component, 01/01/2007–31/12/2007
24. X component, 01/01/2007–31/12/2007
25. Y component, 01/01/2007–31/12/2007
26. Z component, 01/01/2007–31/12/2007

TAM observatory:
27. F component, 01/01/2007–31/12/2007
28. X component, 01/01/2007–31/12/2007
29. Y component, 01/01/2007–31/12/2007
30. Z component, 01/01/2007–31/12/2007

The records contain a few number of baseline drifts therefore additionally 13 synthetic examples containing drifts of different shapes were generated. All of them were recognized by the corresponding algorithm.

Acknowledgment. The results presented in this paper rely on data collected at magnetic observatories. We thank the national institutes that support them and INTERMAGNET for promoting high standards of magnetic observatory practice. (http://www.intermagnet.org)

References

Gvishiani, A. D., S. M. Agayan, Sh. R. Bogoutdinov (2008), Fuzzy recognition of anomalies in time series, *Dokl. Earth Sci., 421*(5), 838–842.

Kerridge, D. (2001), Intermagnet: worldwide near-real-time geomagnetic observatory data, *Proceedings of the Workshop on Space Weather,* BGS, Edinburgh, UK.

Kolmogorov, A. N., S. V. Fomin (1981), *Elements of Function Theory and Functional Analysis, 5-th ed.* (in Russian), Nauka, Moscow.

S. M. Agayan, Sh. R. Bogoutdinov, A. D. Gvishiani, and A. A. Soloviev, Institution of Russian Academy of Sciences Geophysical Center, 3 Molodezhnaya Str., 119296 Moscow, Russia. (a.soloviev@gcras.ru)

E. Kihn, National Geophysical Data Center (NOAA's NGDC), Boulder, Colorado, USA