ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ ПРОБЛЕМИ І МЕТОДИ СИСТЕМНОГО АНАЛІЗУ

УДК 519.68

ALGORITHM FLARS AND RECOGNITION OF TIME SERIES ANOMALIES

A.D. GVISHIANI, S.M. AGAYAN, SH.R. BOGOUTDINOV, S.A. TIKHOTSKY,
J. HINDERER, J. BONNIN, M. DIAMENT

As a rule, algorithms of recognition of time series anomalies are based on time frequency or statistical analysis . This article is devoted to detailed formal description of new fuzzy set based algorithm FLARS (Fuzzy Logic Algorithm for Recognition of Signals). It recognizes time series anomalies by means «smooth» modelling (in fuzzy mathematics sense) of interpreter's logic, which searches for anomalies at the record.

INTRODUCTION

The present article extends the fuzzy set theory approach [11] to recognition of high-frequency anomalies on time series. It may be considered as an extension of the series «Mathematical methods of geoinformatics» [1–5]. This article is devoted to detailed formal description of new fuzzy set based algorithm FLARS (Fuzzy Logic Algorithm for Recognition of Signals). Applications of FLARS to wide spectrum of geophysical data time series along with the application of DRAS algorithm of the same family give successful results [10].

As a rule, algorithms of recognition of time series anomalies are based on time frequency or statistical analysis [6–8, 12]. An exclusion is constituted by the algorithm of «probabilistic» modelling of interpreter's logic [9], which was developed by B. Naimark already in 1966. In a certain sense, our article extends the ideas [9].

The algorithm FLARS may be considered as a results of «smooth» modelling (in fuzzy mathematics sense) of interpreter's logic, which searches for anomalies at the record. Moreover, FLARS is really based on fuzzy mathematic principals. We proceed from the following understanding of interpreter's logic. At first, interpreter glances at the record and estimates activity of its sufficiently small fragments by positive numbers. At the same time, he puts some numeric marks to the centres of the fragments. In this way, from initial record interpreter necessarily proceeds to a non-negative function. It is naturally to call this function by rectification of the initial time series. Indeed, greater values of this function (Figs. 1*a*, 1*b*) correspond to more anomalous points (centres of fragments). In the other words searching the anomalies on the record can be found by searching the

© *A.D. Gvishiani, S.M. Agayan, Sh.R. Bogoutdinov, S.A. Tikhotsky, J. Hinderer, J. Bonnin, M. Diament,* 2004

uplifts on its rectification (Fig. 1b). Thus, FLARS works on two levels. The local one is rectification of the record, and the global one is search of the uplifts on the rectification.

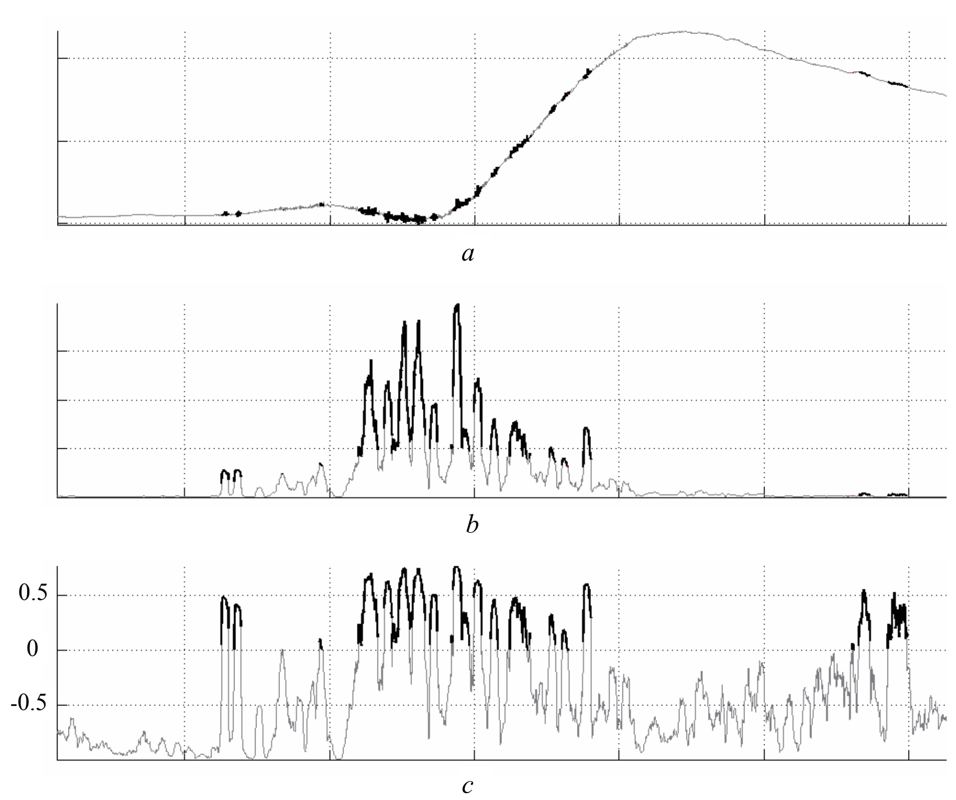

Fig. 1. The example of FLARS working: *a* — the initial record «*y*»; *b* — the rectification $\chi_y(k)$; *c* — the measure of anomality $\mu(k)$

Let us proceed with exact statement.

FLARS LOCAL LEVEL. RECORD RECTIFICATION

Activity at a record (time series) is a multivalued notion, which can change from record to record and within a record. The notion of activity doesn't have a formal mathematical definition like notions as «set», «element», «function», *etc.* Thus we interpret «activity of a record» as intuitively clear notion. For its adequate modelling we introduce strictly defined «rectifications» opened for replenishments. Rectifying functionals serve as basis of the described below FLARS algorithm.

Let us proceed with exact statement.

We consider a discrete semiaxis $\mathbf{R}_h^+ = \{kh, h > 0\ ,\ k = 1, 2, 3, \ldots\}$ and a record (discrete function, time series) $y = \{y_k = y(kh),\ k = 1, 2, 3, \ldots\}$, which is defined on a «coherent» subset (registration period) $Y \subset \mathbf{R}_h^+$. At the semiaxis $\mathbf{R}_h^+$ we introduce the parameter of a local observation $\Delta > 0$, which is divisible by «*h*», $\Delta \in \mathbf{R}_h^+$. The segment of the time series «*y*» with centre at a point «*kh*»

$$\Delta^k y = \left\{ y_{k-\Delta/h}, \ldots, y_k, \ldots, y_{k+\Delta/h} \right\} \in \mathbf{R}^{\frac{2\Delta}{h}+1}$$

we call by a fragment of local observation.

Definition 1. A non-negative mapping $\chi_y : \Im \to \mathbf{R}^+$ defined on the set of fragments $\Im = \{\Delta^k y\} \subset \mathbf{R}^{\frac{2\Delta}{h}+1}$ we call by a functional of the time series «*y*».

In the set $\{\chi_y\}$ there is a subset $\Phi_y(k)$ of the functionals, which transfer anomalous fragments $\{\Delta^k y\}$ into «uplifted» points on the graph of the corresponding curve $\{\chi_y\}$.

Definition 2. We call any function $\Phi_y(k)$ by rectification of the initial time series «*y*». Correspondingly, we call functional Φ_y by a rectifying functional of the given record «*y*».

In FLARS we use the following functionals of time series, which in many important cases occur to be the rectifying functionals:

1. Length of the fragment

$$L\left(\Delta^k y\right) = \sum_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}-1} \left| y_{j+1} - y_j \right| .$$

2. Energy of the fragment

$$E\left(\Delta^k y\right) = \sum_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} \left(y_j - \bar{y}_k \right)^2, \text{ where } \bar{y}_k = \frac{h}{2\Delta + h} \sum_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} y_j .$$

3. Difference of the fragment from its regression of order n

$$R_n(\Delta^k y) = \sum_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} [y_j - \mathrm{Regr}^n_{\Delta^k y}(jk)]^2 ,$$

here as usual $\mathrm{Regr}^n_{\Delta^k y}$ is an optimal mean squares approximation of order *n* of the fragment $\Delta^k y$. If $n = 0$ we get the previous functional «energy of the fragment» (see 2.):

$$\mathrm{Regr}^0_{\Delta^k y} = \frac{h}{2\Delta + h} \sum_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} y_j = \bar{y}_k ,$$

$$R_0(\Delta^k y) = \sum_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} \left(y_j - \mathrm{Regr}^0_{\Delta^k y}(jh) \right)^2 = \sum_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} \left(y_j - \bar{y}_k \right)^2 = E(\Delta^k y) .$$

4. Oscillation of the fragment

$$O(\Delta^k y) = \max_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} y_j - \min_{j=k-\frac{\Delta}{h}}^{k+\frac{\Delta}{h}} y_j \,.$$

Illustration of rectification in the case of natural electric potential in the region of La Furnaise volcano (Reunion island, France) is given in Fig. $1, b$.

FLARS GLOBAL LEVEL

Having the results of the rectification procedure (local level operations) FLARS starts to search uplifts on the corresponding graph. As it is shown at the Fig.1b the relief of the rectification may be enough complicated. Therefore, generally speaking analysis of only vertical level of the curve $\Phi_y(k)$ is insufficient to define of the uplifts. Anomalies may not possess constant high intensity. Often, the anomalies they are heterogeneous. There is a majority of active fragments, divided by short non-anomalous segments. In this case the corresponding fragments of rectification occur to be oscillating uplifts. Analysis of rectifications $\Phi_y(k)$ shows that its domain of definition Y consists of three types of points: anomalous, which belonging to uplifts, background (calm), located far enough from the uplifts, and potentially anomalous (not calm), which have intermediate character. The latter formally not belong to uplifts, but locate close enough to them to experience their influence.

FLARS makes such division of Y into above three subsets in two steps. At the one the algorithm constructs an alternating-sign measure of extremality $\mu(k) \in [-1,1]$, which characterises Φ_y in a point «*kh*». Possibility of growth of $\mu(k)$ from –1 to +1 corresponds to growth of «maximality of Φ_y in the point «kh» and in this way to growth of anomality of the initial record in that point. Apparently, $\mu(k) = 1$ corresponds to maximum anomality and when $\mu(k) = -1$ — to maximum calmness in the point «*kh*». Being based on $\mu(k)$, FLARS chooses in Y the segment *A*, that consists of certainly anomalous points. On the next step the complement $\overline{A} = Y \setminus A$ of non-anomalous points, is divided by FLARS into the set of calm (background) points S' and the set A' of not calm (potentially anomalous) points. One-sided background measures $\mathsf{L}_\alpha \Phi_y(k)$ and $\mathsf{R}_\alpha \Phi_y(k)$, which characterise the rate of calmness of the rectification Φ_y from left and right hand sides of the point «*kh*», play key role in the division $\overline{A} = S' \bigcup A'$. It should be noticed here that FLARS principally differs from DRAS [11], which is based on the same principle of rectification. Indeed, FLARS at first recognizes the fragments of a certain anomality which form the set A. On the next stage fragments of a potential anomality A' are added to the set A. On the contrary, DRAS at the first stage recognises the set $A \bigcup A'$ of potential anomalies and at the second stage establishes certainly anomalous zones $A \subset A \bigcup A'$.

Let us proceed with exact statements.

First we introduce the FLARS free parameter $\Lambda >> \Delta$, $\Lambda \in \mathbf{R}_h^+$, which we call by parameter of a global observation,. The following set we call as a fragment of a global observation

$$\Lambda^k \Phi_y = \left\{ \Phi_y (k - \frac{\Lambda}{h}), \ldots, \Phi_y (k), \ldots, \Phi_y (k + \frac{\Lambda}{h}) \right\} \in \mathbf{R}^{\frac{2\Lambda}{h}+1}. \quad (1)$$

We denote by $\aleph$ the set of fragments (1).

In FLARS we will use the following weight function on the segment $[kh - \Lambda, kh + \Lambda]$

$$\delta_k (\bar{k}) = \frac{\Lambda + h - h\left|\bar{k} - k\right|}{\Lambda + h}. \quad (2)$$

It is easy to see that the graph of this weight function (2) looks like an isosceles triangle with the base $2 \cdot \Lambda$ and the height 1.

The function δ_k represents another model of global observation of the fragment $\Lambda^k \Phi_y$. In this model we, in a certain sense, «look at the fragment» giving bigger weight to the points $\bar{k}h \in \Lambda(k)$, which are located closer to the centre «*kh*».

We symbolize

$$\Lambda^+ (k) = \{k^+ h \in [kh - \Lambda, kh + \Lambda] : \Phi_y (k^+) \geq \Phi_y (k)\},$$

$$\Lambda^- (k) = \{k^- h \in [kh - \Lambda, kh + \Lambda] : \Phi_y (k^-) < \Phi_y (k)\}.$$

The following sum will be an «argument» for minimality (backgroundness) of the point «*kh*»

$$\sigma^+ (\Lambda, k) = \sum \Big(\Phi_y (k^+) - \Phi_y (k) \Big) \delta_k (k^+) : k^+ h \in \Lambda^+ (k). \quad (3)$$

At the same time the following sum will be an «argument» for maximality (anomality) of the point «*kh*»

$$\sigma^- (\Lambda, k) = \sum \Big(\Phi_y (k) - \Phi_y (k^-) \Big) \delta_k (k^-) : k^- h \in \Lambda^- (k). \quad (4)$$

The measure $\mu(k)$ is a result of the comparison of the «arguments» (3) and (4)

$$\mu(k) = \mu(\sigma^- < \sigma^+) = \frac{\sigma^- (\Lambda, k) - \sigma^+ (\Lambda, k)}{\max (\sigma^+ (\Lambda, k), \sigma^- (\Lambda, k))} \in [-1, 1].$$

The measure $\mu(k)$ gives «more delicate assessment» of difference between positive numbers than just a simple difference $|x - y|$. Indeed, the difference in age between five-year-old and ten-year-old children is more considerable than the same difference between 70-year-old and 75-year-old people. Apparently, it is reflected in the values of the measure μ: ($\mu(5 < 10) = \frac{1}{2}$ and $\mu(70 < 75) = \frac{1}{15}$); while $10 - 5 = 5 = 75 - 70$.

Decision on the presence of an anomaly is taken in FLARS by analysis of the vertical level.

Definition 3. Let $\alpha \in [-1,1]$. A point $kh \in Y$ is α-anomalous, if $\mu(k) > \alpha$.

We denote by A_α the whole set of the anomalous points.

The set A_α, unfortunately, does not give a complete picture of the anomality of the record «*y*». Indeed, as it is seen on the Fig.1, c, there can be narrow zones between the fragments of anomality located closely with each other. These zones should not be considered as calm zones since they are too small. To take it into account we transform the initial measure $\mu(k)$ and construct above-mentioned one-sided background measures $\mathsf{L}_\alpha \Phi_y(k)$ and $\mathsf{R}_\alpha \Phi_y(k)$. For that we need «the helping function» $\psi_\alpha(x)$ (Fig. 2).

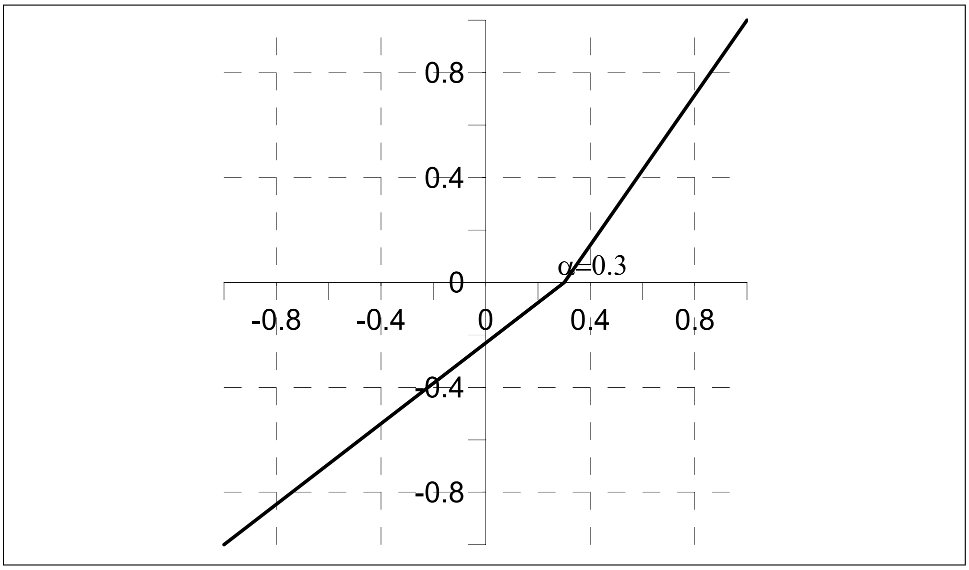

Fig. 2. The helping function $\psi_\alpha(x)$

$$\psi_\alpha(x) = \begin{cases} \dfrac{x-\alpha}{1-\alpha}, & \alpha \le x \le 1 \\ \dfrac{x-\alpha}{1+\alpha}, & -1 \le x \le \alpha \end{cases}, \text{ where } -1 \le \alpha \le 1.$$

It's easy to see that $\mu_\alpha(k) = \psi_\alpha(\mu(k))$ possesses the following properties:

$$\begin{aligned} &\mu_\alpha(k) > 0 \Leftrightarrow \mu(k) > \alpha, \\ &\mu_\alpha(k) = 0 \Leftrightarrow \mu(k) = \alpha, \\ &\mu_\alpha(k) < 0 \Leftrightarrow \mu(k) < \alpha. \end{aligned} \tag{5}$$

From (5) we can simply state.

Statement 1. If $\mu_\alpha(k) > 0$, the point $kh \in Y$ is α-anomalous.

Let us introduce left and right background measures. We assign Θ — the parameter of intermediate observation: $\Delta < \Theta \le \Lambda$. Thus,

$$\mathsf{L}_\alpha \Phi_y(k) = \frac{\sum \mu_\alpha(\bar{k}h)\delta_k(\bar{k})}{\sum \delta_k(\bar{k})}, \quad k \in [k - \Theta/h, 0], \tag{6}$$

$$\mathsf{R}_\alpha \Phi_y(k) = \frac{\sum \mu_\alpha(\bar{k}h)\delta_k(\bar{k})}{\sum \delta_k(\bar{k})}, \quad k \in [0, k + \Theta/h]. \tag{7}$$

The parameter Θ in FLARS plays similar role to the parameter of a global observation in the algorithm DRAS [11]. It is necessary to emphasise that formulas (6) and (7), which define left and right background measures in our case, differ from functions $L_\alpha \Phi_y(k)$ and $R_\alpha \Phi_y(k)$, which define background measures for the algorithm DRAS [11], although both are introduced for the same purpose.

Definition 4. We call the point $kh \in \bar{A}$ α-not calm (potentially anomalous), if $\max\left(\mathsf{L}_\alpha \Phi_y(k), \mathsf{R}_\alpha \Phi_y(k)\right) > 0$.

Definition 5. We call the point $kh \in \bar{A}$ is α-calm (background), if $\max\left(\mathsf{L}_\alpha \Phi_y(k), \mathsf{R}_\alpha \Phi_y(k)\right) \le 0$.

Informaly the constructed measures can be interpreted as follows. The less are values of $\mathsf{L}_\alpha \Phi_y(k)$ and $\mathsf{R}_\alpha \Phi_y(k)$ — the calmer (α-calmer) is the record «y» in the point «kh» (correspondently from the left and from the right). Contrariwise, the greater are values of the membership function the «more anomalous» (more α-anomalous) is the initial record in the given point. Thus, $\mathsf{L}_\alpha \Phi_y(k)$ and $\mathsf{R}_\alpha \Phi_y(k)$ may be interpreted as alternating-sign membership functions of the fuzzy set of not calm points $kh \in Y$.

We denote by S' the set of the background points, and A' the set of potentially anomalous points. Thus, the searched division of the set Y is constructed:

$$Y = A \vee A' \vee S'. \tag{8}$$

FLARS FORMAL ALGORITHMICAL MODEL

FLARS has the following free parameters to be chosen by user:

a) $\Delta \in \mathbf{R}_h^+$ — parameter of local observation;

b) $\Lambda \in \mathbf{R}_h^+$, $\Lambda >> \Delta$ — parameter of global observation;

c) $\alpha \in [-1,1]$ — vertical level of extremality;

d) $\Theta \in \mathbf{R}_h^+$, $\Delta < \Theta \le \Lambda$ — parameter of intermediate observation.

In more «flexible» implementation of the algorithm additional parameter $-1 \le \beta \le 1$ is introduced. It is horizontal background level. By substituting the condition $\max\left(\mathsf{L}_\alpha \Phi_y(k), \mathsf{R}_\alpha \Phi_y(k)\right) > 0$ by the condition $\max\left(\mathsf{L}_\alpha \Phi_y(k), \mathsf{R}_\alpha \Phi_y(k)\right) > \beta$ in the definition 2 and the in equation $\max\left(\mathsf{L}_\alpha \Phi_y(k), \mathsf{R}_\alpha \Phi_y(k)\right) \le 0$ by the condition $\max\left(\mathsf{L}_\alpha \Phi_y(k), \mathsf{R}_\alpha \Phi_y(k)\right) \le \beta$ in definition 3 for $\beta \in [-1,1]$ we obtain another, «more flexible» version on FLARS. In other words, upon proper «tuning» of the algorithm it pays attention only to those outputs above level «α», which are massive enough by time: they will be surrounded by an aureole of not calm points. Such «tuning» is achieved with the help of a free parameter $\beta \in [-1,1]$ and conditions $\mathsf{L}_\alpha \Phi_y(k) \le \beta$, $\mathsf{R}_\alpha \Phi_y(k)) \le \beta$.

For each «β» a corresponding dichotomy of non-anomalous points of $\overline{A}$ arises:

$$S' = [kh] \in \overline{A} : \max \{\mathsf{L}_\alpha \Phi_y(k),\ \mathsf{R}_\alpha \Phi_y(k) \le \beta\},$$

$$S' = [kh] \in \overline{A} : \max \{\mathsf{L}_\alpha \Phi_y(k),\ \mathsf{R}_\alpha \Phi_y(k) > \beta\}.$$

It is natural to call parameter «β» a horizontal level of a background, since a massiveness of anomaly by time is its horizontal extension. The bigger is «β» the bigger segment of a record is related to a background and the smaller aureole will correspond to each anomalous fragment.

Block-scheme of the algorithm FLARS is given in Fig.3.

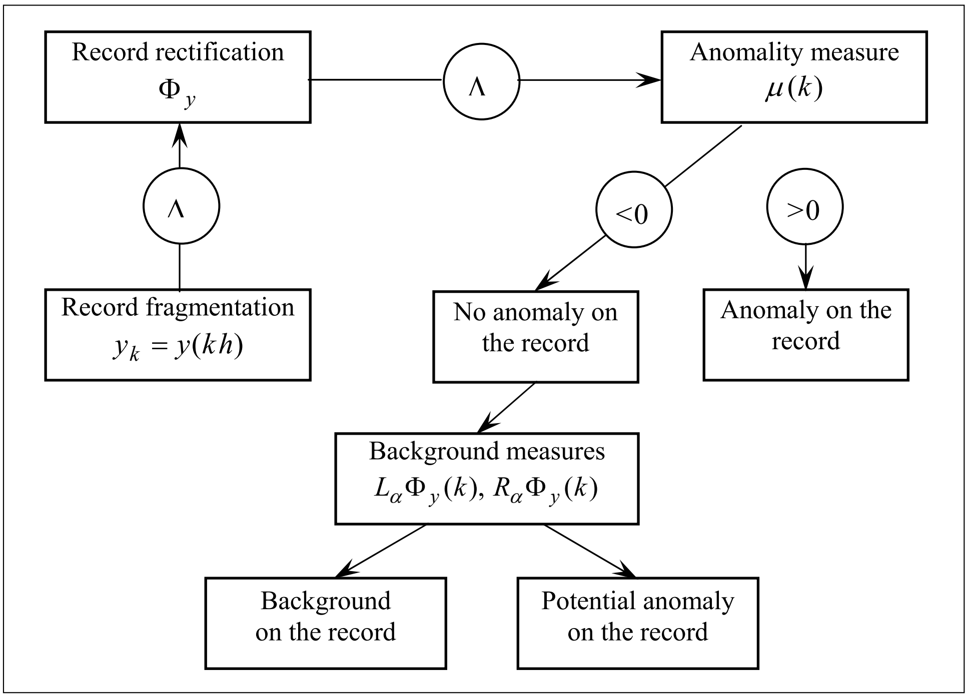

Fig. 3. The block-scheme of the algorithm FLARS

THE APPLICATION OF FLARS FOR THE SUPERCONDUCTING GRAVIMETER DATA PREPROCESSING

The network of the Superconducting Gravimeters (SG) was started in the course of Global Geodynamic Project at the 1^{st} of July 1997. Now it consists of about 20 instruments deployed all around the globe that continuously measure the time variations of gravity with the accuracy reaching the 10^{-9} m/s^2 and sampling interval less or equal to 1 minute. Thus it is clear that the total amount of SG data is now really huge and the processing and analysis of this data possess a real scientific and technical challenge.

The main purpose of the GGP is to reveal the information about the Earth's inner structure and dynamics by interpreting the SG data in terms of Earth's tides as well as non-tidal gravity variations that are related to the Earth axis precession, core nutations and core modes. Some processes that could manifest themselves in very weak periodical gravity variations were theoretically predicted, thus the problem arises of proving or rejecting these hypotheses on the basis of comprehensive analysis of SG data.

The scientific approaches that are used to reveal the periodical variations are generally based on the harmonic (or Fourier) analysis. A process of revealing the weak peaks in the Fourier power spectra could be very unstable with respect to noise in the data. On the other hand it is clear that SG time series always contain a significant noise associated with a wide spectra of sources. Even the small rain at the site of the SG deployment would produce a substantial signal of tens of microgals. Other noise sources are the earthquakes, atmospheric pressure variations, soil humidity variations, snow cover and technical sources of external (autos, trains and nearby factories) as well of internal (electrical power instabilities, manipulations with instrument, etc.) origin. It is important that all these sources produce the noise of a wide variability of features such as amplitudes, frequencies and specific shapes that may include spikes, gaps, offsets and others. Thus it is difficult to remove all these noise signals on the base of convenient noise reduction approaches such as frequency filtering or cutting by the signal amplitude or amplitude of signal time derivative. But FLARS algorithm by its fuzzy logic nature provides a powerful tool for the solution of this problem.

We have successfully applied the FLARS algorithm for the processing of data recorded by the SG deployed in Strasbourg, France, in the period from April to December 1997 (Fig. 4). These records were processed previously by different authors [13] and the detailed comparison of different approaches could be performed. It has been demonstrated, in particular, that algorithm FLARS is able to find almost all noise events on SG records.

Рис. 4. An example of noise (interval marked black) detection in SG data. Note the stable detection of different noise patterns: first four signals are the seismic waves distant earthquakes, the fifth disturbance of unknown origin and the sixth is the offset associated with the instrument operation

Depending on choice of free parameters it could act in a more «strong» or «soft» rejection approaches used by different authors. The present version of the FLARS algorithm software adopt the «TSoft» data format which is the standard for the SG community.

REFERENCES

1. *Gvishiani A.D., Agayan S.M.*, Bogoutdinov Sh.R. Mathematical geoinformatics methods. I. On a new approach to clusterisation // Kibernetika i sistemny analiz. — 2002. — № 2. — P. 104–122.
2. *Gvishiani A.D., Diament M., Mikhailov V.O., Galdeano A., Agayan S.M., Bogoutdinov Sh.R., Graeva E.M.* Artificial Intelligence Algorithms for Magnetic Anomaly Clustering // Physics of the Solid Earth. — 2002. — № 7. — P. 13–38.
3. *Mikhailov V.O., Galdeano A., Diament M., Gvishiani A., Agayan S., Bogoutdinov Sh., Graeva E., Sailhac P.* Application of artificial intelligence for Euler solution clustering // Geophysics. — 2003. — **68**, № 1. P. 168–180.
4. *Gvishiani A., Dubois J.O.* Dynamic Systems and Dynamic Classification Problems in Geophysical Applications // Springer, 1998. — 256 p.
5. *Gvishiani A., Dubois J.O.* Artificial Intelligence and Dynamic Systems for Geophysical Applications // Springer, 2002. — 347 p.
6. *Kedrov O.K., Polikarpova L.A., Steblov G.M.* Detection Algorithm for Weak Short-Period Seismic Signals, Based on the Frequency–Time Analysis of Three-Component Records in the Real Time Mode // Izvestiya, Physics of the Solid Earth. — 1998. — № 8. — P. 30–45.
7. *Magotra N., Ahmed N., Chael E.* Seismic event detection and source location using single-station (three-component) data // Bull. Seism. Soc. Am. — 1987. — **77**. — P. 958–971.
8. *Plesinger A., Hellweg M., Seidl D.* Interactive high-resolution polarization analysis of broadband seismogram // J. Geophys. — 1986. — **59**. — P. 129–139.
9. *Naymark B.M.* Algoritmy dlya obnaruzheniya seysmicheskogo signala na fone mikroseym // Vychislitelnaya seysmologiya. — 1966. — № 1. — P. 5–9.
10. *Zlotnicki J., Mikhailov V., Le Mouël J., Gvishiani A., Yvetot P., Agayan S., Fanquet F., Bogoutdinov Sh., Donnadieu G.* Ulf magnetic and electric signals related to volcanic activity: la Fournaise case (Reunion island) // III International Workshop on Magnetic, Electric and Electro Magnetic Methods in Seismology and Volcanology (MEEMSV-2002). — 2002. — 127 p.
11. *Gvishiani A.D., Agayan S.M., Bogoutdinov Sh.R., Ledenev A.V., Zlotnicki J., Bonnin J.* Mathematical geoinformatics methods. II. The fuzzy logic algorithms for recognition of anomalies on time series // Kibernetika i sistemny analiz. — 2003. — № 4. — P. 103–111.
12. *Levin B.R.* Teoreticheskie osnovy statisticheskoy radiotekhniki. — Moscow: Sov. radio. — 1968. — 504 p.
13. *Hinderer J., Rosat S., Crossley D., Amalvict M., Boy J.P. & Gegout P.* (2002) Influence of different processing methods on the retrieval of gravity signals from GGP data // Bulletin d'Information des Marées Terrestres (BIM). — **135**. — P. 10653–10668.

Received 30.04.2004

From the Editorial Board: The article corresponds completely to submitted manuscript.