Data Science Journal, Volume 12, 5 March 2013

MATHEMATICAL TOOLS FOR GEOMAGNETIC DATA MONITORING AND THE INTERMAGNET RUSSIAN SEGMENT

Anatoly Soloviev[1]*, Shamil Bogoutdinov[1], Alexei Gvishiani[1], Ruslan Kulchinskiy[1], and Jacques Zlotnicki[2]

1Geophysical Center of Russian Academy of Sciences, 3, Molodezhnaya St., 119296 Moscow, Russia (GC RAS)
**[1]Email:* a.soloviev@gcras.ru
[2]Observatoire de Physique du Globe de Clermont-Ferrand, Campus des Cuzeaux, 24, av. des Landais, 63177 Aubiere Cedex, France
Email: jacques.zlotnicki@wanadoo.fr

ABSTRACT

In this paper, a new approach to the detection of anomalies in geophysical records is connected with a fuzzy mathematics application. The theory of discrete mathematical analysis and collection of algorithms for time series processing constructed on its basis represents the results of this research direction. These algorithms are the consequence of fuzzy modeling of the logic of an interpreter who visually recognizes anomalies in records. They allow analyzing large data sets that are not subjected to manual processing. The efficiency of these algorithms is demonstrated in several important geophysical applications. Plans for an extension of the Russian INTERMAGNET segment are presented.

Keywords: Magnetic field, Fuzzy sets, Time series, Magnetic observatory

1 INTRODUCTION

Detection of anomalies in geomagnetic records is a fundamental task of data analysis. The significance of the process represented by such records is often concentrated in these anomalies. We present a new approach for solving this task that is based on fuzzy logic and fuzzy mathematics application (Zadeh, 1965). The mathematical theory of Discrete Mathematical Analysis (DMA) (Gvishiani, Agayan, Bogoutdinov, & Soloviev, 2010) and collection of algorithms for time series processing (e.g., Gvishiani, Agayan, & Bogoutdinov, 2008; Bogoutdinov, Gvishiani, Agayan, Solovyev, & Kihn, 2010; Soloviev, Chulliat, Agayan, Bogoutdinov, & Gvishiani, 2011) constructed on the basis of DMA represent the results of this research direction. These algorithms are a consequence of fuzzy modelling of the logic of an interpreter who visually recognizes anomalies in records. The goal is its further application to the automated analysis of large data sets that are not subjected to manual processing. Sufficient “flexibility” of the algorithms is provided by a wide range of “rectifications” that arise in the interpreter operation modelling. The efficiency of these algorithms was demonstrated in several important geological, geophysical, and geodynamic applications, including the global real-time monitoring of magnetic storms (Veselovsky, Agayan, Kulchinskiy, Gvishiani, Bogoutdinov, Petrov et al., 2011), the recognition of artificial disturbances in geomagnetic records (Soloviev, Bogoutdinov, Agayan, Gvishiani, & Kihn, 2009), and the monitoring of volcanoes (Zlotnicki, LeMouel, Gvishiani, Agayan, Mikhailov, & Bogoutdinov, 2005). Herein we present an overview of several of these successful applications dealing with magnetic field studies.

The largest global network of on-ground magnetic observations is the International Real-time Magnetic Observatory Network (INTERMAGNET) (Love, 2008). This network consists of more than 110 observatories; however, only 5 of them are located on Russian territory. The Geophysical Center of the Russian Academy of Sciences (GC RAS) has elaborated a plan to extend the Russian INTERMAGNET segment (Soloviev, 2011). In particular, five new INTERMAGNET observatories are being deployed in Russia through the joint efforts of the GC RAS and the institutions of regional RAS branches. A regional geomagnetic data node of the Russian INTERMAGNET segment is being created on the basis of the Russian WDC for Solar-Terrestrial Physics at GC RAS. A particular feature of this node is the introduction of an automated system for recognition of artificial disturbances in incoming preliminary magnetograms.

2 FUZZY MEASURE OF ACTIVITY AND MAGNETIC STORM MONITORING

A geomagnetic field is subjected to the fluctuations of different time scales. In order to describe magnetic activity in the planetary scale, the following geomagnetic indices were established: 24-hour Ci-index, 3-hour Kp-index, 1-hour indices Dst, AE, and others (http://www.ngdc.noaa.gov/IAGA/vdat/). The principal idea of these indices is to give an equal estimation of the relative strength of disturbances at various observatories. However, a more detailed study of the morphology of geomagnetic disturbances and their sources has shown that various indices of geomagnetic activity used nowadays express geomagnetic field activity not on the whole Earth surface but in its separate regions.

For studying the dynamics of geomagnetic disturbances during a storm, it is not enough to use just several standard geomagnetic indices (for example, Cp, AE, Dst, etc.). In the process of solar-terrestrial phenomena studies, the necessity of simultaneous determination of the strength of the geomagnetic disturbances at a maximum number of observatories across the globe has arisen. Currently, the largest global network of geomagnetic field observations is INTERMAGNET (Love, 2008). Such a necessity demands the introduction of new parameters independent of geomagnetic latitudes and longitudes.

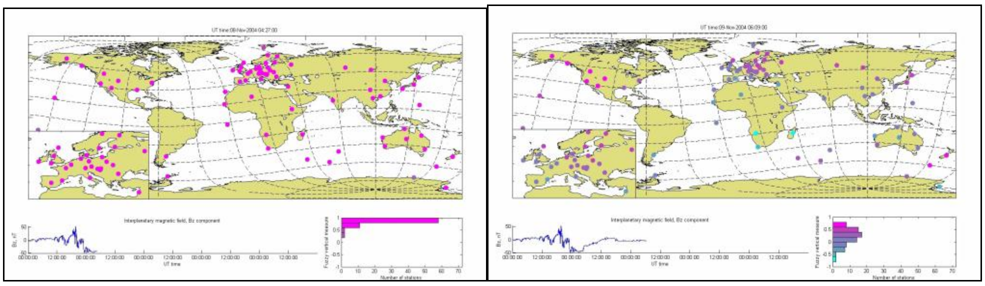
It is very difficult for an expert to manually perform this type of analysis because of the large volume of data involved. To solve this task a new geoinformatics approach based on fuzzy logic methods (Zadeh, 1965) is suggested. In particular we apply the algorithm FCARS (Fuzzy Comparison Algorithm for Recognition of Signals) (Gvishiani et al., 2008) constructed on the basis of DMA (Gvishiani et al., 2010). The application of DMA enables processing and studying multidimensional arrays and time series.

Along with the other DMA algorithms, FCARS represents an attempt to model the logic of a data expert who handles and evaluates data records manually. An expert performs a visual analysis of a record estimating the activity of its small fragments in terms of positive numbers. The elaborated numerical estimates are mentally assigned to the centers of record fragments. In such a manner, the interpreter makes a step to a nonnegative function, which is naturally called straightening the record because more active record points correspond to greater values of this straightening. Next, attempts of the interpreter to search for record anomalies are reduced to the search for elevations in the straightened curve corresponding to the most active intervals of the record. FCARS models this logic at two levels: the local level (realization of the straightening) and the global level (search for elevations on the straightened line). There are several mathematical constructions of the strengthening, and its choice is one of free parameters. The FCARS vertical measure of anomality in point t $\mu^{v}(t)$ answers in the scale [-1,1] the question: "To what degree can a value of a strengthening $\Phi_y(t)$ be considered as large in point t?" This involves so-called fuzzy comparisons of a set of nonnegative numbers (Gvishiani, Agayan, & Bogoutdinov, 2008).

The FCARS vertical measure of anomality allows the introduction of **a measure of geomagnetic activity $\mu(t)$, which estimates the geomagnetic activity at each particular time moment for a particular geomagnetic record. It gives the estimation in a single scale [-1, 1], where value -1 corresponds to a calm state and value 1 corresponds to an anomalous state**. This measure can be applied to the estimation of any magnetic component record with a time step equal to the sampling rate. Testing the measure on magnetograms obtained from several INTERMAGNET observatories showed its high correlation with the regional K-index of geomagnetic activity.

By applying measure $\mu(t)$ to the whole set of magnetograms obtained by all observatories (e.g., the INTERMAGNET network), we can have a snapshot of a storm's distribution over the Earth's surface at each time moment. In the case of INTERMAGNET data, such a picture changes with a time step of 1 minute. To visualize the operation of measure $\mu(t)$, we use GIS technology. Thus, the proposed method for geomagnetic activity monitoring provides a new way of studying the dynamics of spreading geomagnetic disturbances. It allows the performance of geomagnetic activity monitoring in real time mode.

The proposed toolkit was tested on two strong geomagnetic storms observed during the 23-rd solar cycle. In the first case a complicated storm on 8-11 November 2004 consisting of two parts was considered. In the second case an isolated storm on 15 May 2005 was considered. Before applying the method, we studied the selected storms in detail in order to investigate their common and specific features. This study involved INTERMAGNET data, D_{xt}-index values (corrected version of D_{st}) (Mursula, Holappa, & Karinen, 2008), the parameters of solar wind and the interplanetary magnetic field, and data on solar events. Figure 1 illustrates the toolkit application to the first storm monitoring at several time moments.



(a) (b)

$-1 < \mu(t) \leq -0.8$ $-0.1 < \mu(t) \leq 0.1$ $0.8 < \mu(t) \leq 1.0$

Figure 1. Global monitoring of geomagnetic activity during the first storm in continuous time based on INTERMAGNET data (*H* component). The two screenshots correspond to the toolkit operation at two different UT time moments: 8/11/2004 04:27 (a) and 9/11/2004 06:09 (b).

The toolkit visualizes the distribution of the anomality parameter μ(*t*) over the whole set of INTERMAGNET observatories in animation mode on a global map (Figure 1). Below on the right is a dynamic histogram, which reflects the distribution of INTERMANGET observatories according to different values of μ(*t*) from -1 (no anomaly) to 1 (strong anomaly) for each time. Below on the left there is a dynamic plot of the B_z component of the interplanetary magnetic field, which evolves correspondingly. A clear conformity between the histogram of observed anomality and B_z behavior can be seen: each drastic change of B_z behavior leads to the transition of the majority of observations into an anomalous state.
.
The results show that global, regional, and local features of the storms have common and individual particularities depending on the external stimulating conditions in the heliosphere and on the Sun. Based on an analysis of the dynamic distribution of geomagnetic disturbances, it was shown that the ring current is not always the main contributor to the equatorial geomagnetic perturbations during the developmental and main phases of strong geomagnetic storms. This has led to the conclusion that the proposed approach gives a more objective and prompt estimation of geomagnetic activity than do a number of classical indices.

3 DE-SPIKING 1-MINUTE AND 1-SECOND MAGNETIC DATA

Each year data experts at observatories and data centers carry out manual processing and filtering of collected preliminary data sets. The aim of such work is to produce definitive data and make it available to the scientific community worldwide. Despite close cooperation between observatories, approaches to data processing may differ and depend on the subjectivity of this or that expert's evaluation. In this connection a mathematical formalization of the recognition of artificial disturbances could contribute to a significant increase in definitive data quality (Soloviev et al., 2009). In turn, an increase of observed data quality contributes significantly to our knowledge about the Earth's magnetic field.

An important step towards such a mathematical formalization was undertaken by Bogoutdinov et al. (2010). The proposed algorithm "SP" (from SPIKE) was applied to the recognition of artificial spikes (Figure 2) on magnetograms recorded with a 1-minute sampling rate. Since the algorithm operation is adjusted by a set of free parameters, it was first determined based on 1-year preliminary records of the three components and total field intensity obtained in 2008 by seven INTERMAGNET observatories located in different parts of the Northern hemisphere. Then this new algorithm was applied to other 1-year preliminary records of the three components and the total field intensity obtained in 2009 by the same observatories. Further, it was applied to 2-year preliminary records from the same observatories obtained during the increased solar activity period in 2003 and 2005. In all cases the probability of a target miss was between 0% and 1%, whereas the probability of a false alarm varied between 5% and 15%. The probabilities were calculated by comparing the algorithm results with definitive data filtered manually for the same time intervals.

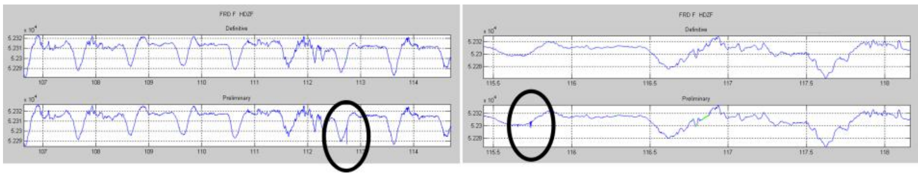


Figure 2. Spikes in 1-minute preliminary data (lower plots) removed manually while producing definitive data (upper plots)

Many observatories now modernize their equipment in order to be able to produce 1-second filtered data. While at many observatories the 1-second data cleaning represents a reasonable amount of work, it becomes a daunting task at some observatories, particularly those installed in remote but important locations where no optimal observatory site could be found. In this case the situation is burdened with the increase of data volume (86400 values per record per day compared to 1440 in 1-minute cases) and the appearance of short-period geomagnetic pulsations similar to artificial spikes and unseen in 1-minute records. Therefore automated de-spiking tools are much in demand in the case of 1-second data acquisition.

For that purpose, Soloviev et al. (2012a; 2012b) developed the “SPs” algorithm (from SPIKEsecond), which is a modified version of the SP algorithm aimed at the recognition of artificial spikes on 1-second magnetograms. As in the case of the SP algorithm, the algorithm SPs was first determined from 20-day (1-20/07/2009) preliminary records and then examined in 10-day (21-31/07/2009) records obtained by the magnetic observatory in Easter Island maintained by the Institut de Physique du Globe de Paris (IPGP), France. The algorithm efficiency was estimated by comparing the results of automated preliminary data processing with definitive data for the same time spans. After that it was applied to other 30-day (1-31/08/2009) records with no definitive data available. The results of the recognition by the algorithm SPs were subsequently evaluated by eye. After a 20-day learning phase in July 2009, the algorithm was able to recognize more than 94% of the spikes on the three components and the intensity recordings in August 2009 while the percentage of false alarms was less than 6%. At all stages, the algorithm showed worse results in processing the vertical component *Z*.

4 THE INTERMAGNET RUSSIAN SEGMENT

Currently, the Russian participation in the INTERMAGNET program is confined to five observatories, which report preliminary data to geomagnetic information nodes (GINs) in Paris (France) and Edinburgh (UK). A weak development of the INTERMAGNET network in Russia and an absence of national GIN induced GC RAS to elaborate a plan to extend the Russian INTERMAGNET segment (Soloviev, 2011). In particular, five sets of geomagnetic equipment compliant with INTERMAGNET standards are ready to be installed in different parts of Russia. Since a network of geomagnetic observations was widely developed in the Soviet Union during the International Geophysical Year in 1957-1958, numerous existing observatories ruled by RAS institutions are considered as possible sites for installing new equipment. These sites include Syktyvkar (Komi Republic), St. Petersburg, Rotkovets (Arkhangelsk region), Novaya Zemlya Islands, Tiksi, and others. Apparently, deploying new observatories in the auroral zone is of the highest priority in the space physics community. However, the major obstacle preventing that placement is a lack of personnel capable of operating observatories in such locations. In this regard, only those places where the operation of observatories is feasible are considered.

Another goal in the framework of the extension of the Russian INTERMAGNET segment is creation of a national geomagnetic data node at GC RAS servicing the Russian INTERMAGNET observatories. A transmission of magnetograms from functioning and future Russian INTERMAGNET observatories to GC RAS will be performed in a real-time mode. A particular feature of this node is an automated system for data quality control, which involves SP and SPs algorithms, applicable to incoming preliminary magnetograms.

5 CONCLUSION

A new way of studying the dynamics of the spread of geomagnetic disturbances is presented in this paper. It is based on fuzzy mathematics and GIS technology and involves data from the whole INTERMAGNET network. The measure of geomagnetic activity $\mu(t)$ is computed according to the FCARS algorithm and allows tracking geomagnetic activity (e.g., magnetic storms) in continuous mode. The FCARS algorithm also serves as the basis for an automated system for processing electrotelluric and electromagnetic observations in the framework of the Russian-French project of monitoring volcanic activity on Reunion Island and in Kamchatka.

Application of the measure μ(*t*) to selected INTERMAGNET records has shown its high correlation with the regional K-index of geomagnetic activity. The proposed method for geomagnetic activity monitoring was tested on two strong geomagnetic storms. Based on the analysis performed, it was shown that the ring current is not always the main contributor to the equatorial geomagnetic perturbations during the development and the main phase of strong geomagnetic storms. This has led to the conclusion that the proposed approach gives a more objective and prompt estimation of geomagnetic activity than do a number of classical indices. Based on this analysis, it can be concluded that geomagnetic proxies could also serve as an important source of indirect information about solar and heliospheric activity in the past when direct observations were not available.

The algorithms SP and SPs based on DMA are specifically aimed at the recognition of singular artificial spikes with a simple morphology on 1-minute and 1-second magnetograms. The algorithms rely on fuzzy mathematics principles. It was shown that after a learning phase these algorithms are able to recognize artificial disturbances efficiently and distinguish them from natural ones, such as short-period geomagnetic pulsations in the 1s-1min period range. This capability is critical and opens the possibility of using the algorithms in an operational environment. The algorithms were tested on real magnetic data. Small probability values for target miss and false alarm were obtained.

The knowledge of experts who carry out geomagnetic data analysis manually is effectively incorporated into all the developed algorithms during their development stages.

Five new INTERMAGNET observatories in Russia are being deployed by the joint efforts of GC RAS and the institutions of regional RAS branches. A regional geomagnetic data node of the Russian INTERMAGNET segment is being created on the basis of the Russian WDC for Solar-Terrestrial Physics at GC RAS. An automated quality control system for on-the-fly processing of incoming magnetograms can significantly facilitate and hasten the transformation of geomagnetic preliminary data into definitive data.

6 ACKNOWLEDGEMENTS

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/).

This work is carried out within the framework of the collaboration program among GC RAS, IPGP, and the Schmidt Institute of Physics of the Earth of Russian Academy of Sciences. It is also supported by the Russian Foundation for Basic Research (grant no. 12-05-00583).

7 REFERENCES

Bogoutdinov, S.R., Gvishiani, A.D., Agayan, S.M., Solovyev, A.A., & Kihn, E. (2010) Recognition of Disturbances with Specified Morphology in Time Series. Part 1: Spikes on Magnetograms of the Worldwide INTERMAGNET Network. *Izvestiya, Physics of the Solid Earth 46* (11), pp 1004–1016.

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

Gvishiani, A.D., Agayan, S.M., Bogoutdinov, S.R., & Soloviev, A.A. (2010) Discrete mathematical analysis and geological and geophysical applications. *Vestnik KRAUNZ. Earth Sciences*, *2*(16), pp 109-125 (in Russian).

Love, J.J. (2008) Magnetic monitoring of Earth and space. *Physics Today 61*, pp 31-37.

Mursula, K., Holappa, L., & Karinen, A. (2008) Correct normalization of the Dst index. *Astrophys. Space Sci. Trans*. *4*, pp 41–45.

Soloviev, A.A., Bogoutdinov, Sh.R., Agayan, S.M., Gvishiani, A.D., & Kihn, E. (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

Soloviev, A. (2011), Artificial intelligence in the Earth's magnetic field study and INTERMAGNET Russian Segment. *ICSU CODATA Newsletter 100*, p 4.

Soloviev, A.A., Agayan, S. M., Gvishiani, A.D., Bogoutdinov, Sh.R., & Chulliat, A. (2012a) Recognition of Disturbances with Specified Morphology in Time Series: Part 2. Spikes on 1-s Magnetograms. *Izvestiya, Physics of the Solid Earth 48* (5) pp 395–409.

Soloviev, A., Chulliat, A., Bogoutdinov, S., Gvishiani, A., Agayan, S., Peltier, A., & Heumez, B. (2012b) Automated recognition of spikes in 1 Hz data recorded at the Easter Island magnetic observatory. *Earth Planets Space 64* (9), pp 743-752, doi:10.5047/eps.2012.03.004.

Veselovsky, I., Agayan, S., Kulchinskiy, R., Gvishiani, A., Bogoutdinov, S., Petrov V., & Yakovchouk, O. (2011) Global, regional and local dynamics of strong geomagnetic storms (#5313). *XXV IUGG General Assembly "Earth on the Edge: Science for a Sustainable Planet" (28 June - 7 July 2011, Melbourne, Australia), Program Book and Abstracts*.

Zadeh, L.A. (1965) Fuzzy sets. *Information and Control 8*, pp 338-353.

Zlotnicki, J., LeMouel, J.-L., Gvishiani, A., Agayan, S., Mikhailov, V., & Bogoutdinov, Sh. (2005) Automatic fuzzy-logic recognition of anomalous activity on long geophysical records. Application to electric signals associated with the volcanic activity of la Fournaise volcano (Réunion Island). *Earth and Planetary Science Letters 234*, pp 261-278.

(Article history: Available online 24 February 2013)