Article

Regression Derivatives and Their Application in the Study of Magnetic Storms

Sergey Agayan[1], Shamil Bogoutdinov[1,2], Roman Sidorov[1], Anatoliy Soloviev[1,2], Dmitriy Kamaev[3], Andron Aleksanyan[3] and Boris Dzeranov[1,*]

[1] Geophysical Center of the Russian Academy of Sciences, Moscow; s.agayan@gcras.ru, shm@gcras.ru, r.sidorov@gcras.ru, a.soloviev@gcras.ru, b.dzeranov@gcras.ru
[2] Schmidt institute of physics of the Earth of the Russian Academy of Sciences
[3] Federal State Budgetary Institution Research and Production Association "Typhoon", Obninsk; post@typhoon.obninsk.ru, a.alexanyan@itresume.ru
* Correspondence: b.dzeranov@gcras.ru

Abstract: Discrete Mathematical Analysis is a data analysis method that uses fuzzy mathematics and fuzzy logic. DMA involves the active participation of the researcher in the study of records, offering technologies and algorithms for analyzing records through the properties of interest to the researcher. In the present work, such properties are related to regression derivatives, and the results obtained are applied to magnetic records. The possibilities of the method in the morphological analysis of geomagnetic storms are demonstrated on the example of three strongest storms that have occurred since the beginning of the current 25th solar cycle.

Keywords: proximity measure; regression derivation; regression smoothing; measures of activity; multi-scale measures of activity

Citation: Agayan, S.; Bogoutdinov, Sh.; Sidorov, R.; Soloviev, A.; Kamaev, D.; Aleksanyan, A.; Dzeranov, B. Regression Derivatives and Their Application in the Study of Magnetic Storms. *Appl. Sci.* **2021**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

1. Introduction

The information-analytical complex MAGNUS was created in the Geophysical Center of the Russian Academy of Sciences (GC RAS) for the collection, storage, processing, and analysis of geomagnetic information [1]. It offers the analysis of geomagnetic records both by standard methods and by new methods aimed at formalizing the experience and knowledge of an expert in working with magnetograms. Such an analysis, in particular, includes the recognition and classification of extreme geomagnetic phenomena, which are closely related to the ability to identify the most significant features in a record, which implies knowledge of its structure (morphology).

This work is devoted to these issues, namely: the study of time series by Fuzzy Logic (FL) methods in the framework of Discreet Mathematical Analysis (DMA) with application to magnetic storms. It is a direct continuation of [2]. In the notation and terminology of the latter, we briefly recall the essence of the matter.

Time series (the record) f, given on a regular grid T, is analyzed by an expert. Concerning f, he is interested in a well-defined local property ξ. The expert's analysis consists of answers to three questions:

- First question: to what extent $\mu\xi_f(t) \in [-1,1]$ the property ξ of interest to the expert is satisfied for the series f at a node $t \in T$?
  The work [2] is almost completely devoted to the answer to this question: it required a fuzzy formalization of the property ξ (formalization of ξ as a fuzzy set on T). The measure $\mu\xi_f$ was its fuzzy membership measure.
  Part of this article is devoted to continuing research on this topic: all the necessary information from [] for this will be given below. And, in general, the authors tried to ensure that this work could be read independently of the previous one, although, of course, the "breath" (influence) [] will be felt everywhere.

- Second question: how does the performance of the property ξ for the series f change over time?
  The answer to this question involves a trend analysis of the dynamics of $\mu\xi_f$ on T, it is the main part of the work and is associated with regression derivatives obtained in DMA relatively recently. Given the importance of regression derivatives, the article provides their detailed description.
- Third question: what are ξ-anomalies for f and how are they constructed?
  In the present work, the first half of the question is not dealt with by the authors: a lot of attention has already been paid to it in the DMA. According to the logic of the interpreter [2], ξ-anomalies on the record f in DMA correspond to stochastic heights on the measure$\mu\xi_f$. The best way to find them so far is to double – use DPS clustering algorithms [3,4].
  The situation with the second half of the question (the structure of ξ-anomalies) is similar to the situation with the first question: it was the subject of the previous work. Here we will show the dynamic progress in it, associated with the regression derivative, which will allow us to better understand the construction of ξ-anomalies, and therefore their coding and classification, incorporated in the MAGNUS system.

The result of the work should be considered as the construction of measures $\mu\xi_f(t,p)$ for the magnetic record f in time t and at different scales p of the manifestation of properties ξ, on it, which have a regression nature.

The measures $\mu\xi_f(t,p)$ are very similar to the wavelet spectrum of the record f (Figure 1), and are a useful and convenient tool for its study.

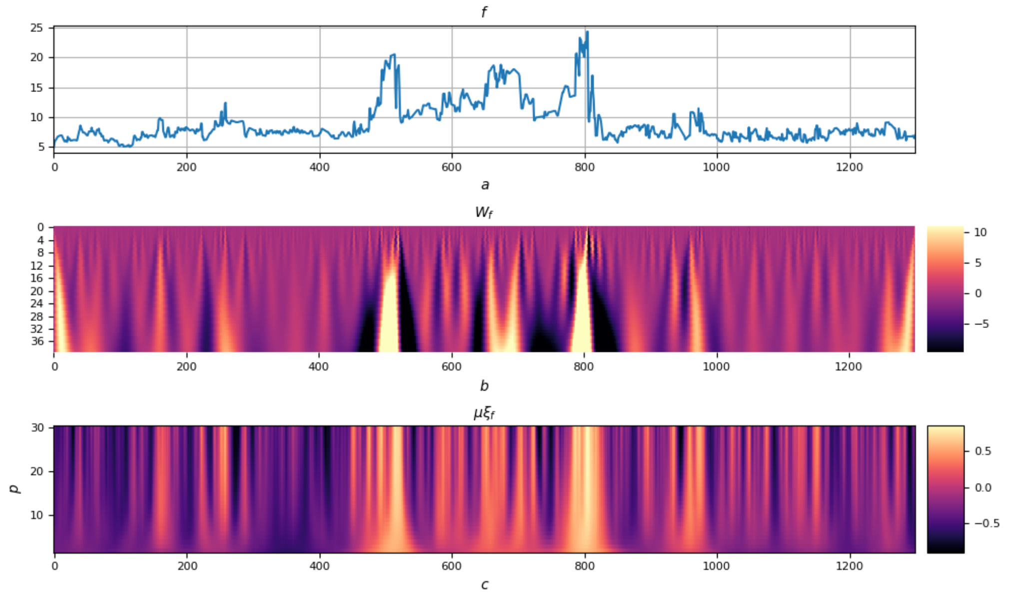

Figure 1. *a* – original record; *b* – activity measure $\mu\xi_f$; *c* – wavelet spectrum W_f

2. Regression derivative and regression smoothing

2.1. General concepts

Let T be a discrete segment $[a,b]$ with nodes t_i, $i = 1,\dots,N$; $t_1 = a$; $t_N = b$. Let's denote by $TS[a,b]$ the space of time series (records) on $[a,b]$. The correspondence $x_i = x(t_i)$ turns $TS[a,b]$ into a N-dimensional space.

Expert analysis of the behavior of a time series involves considering its value not only in a separate node but also simultaneously taking into account the values in some of its neighborhood. This cannot be done without formalizing the concept of localization for the definition domain of the series, i.e. for points of the discrete segment $[a,b]$.

Such localization can be done using a fuzzy structure $\delta_{t_i}(t_j)$ on $[a,b]$, which plays the role of a neighborhood of the node t_i, expresses the property of the proximity of the remaining nodes to the node t_i, is normalized to t_i and decreases with increasing distance from t_i. Thus, $\delta_{t_i} \in \mathrm{Fuzzy}[a,b]$; $\delta_{t_i}(t_j)$ is the degree of proximity t_j to t_i, where $\delta_{t_i}(t_i) = 1$ and

$$(|t_i - t_i| > |t_{\bar{j}} - t_i|) \to \delta_{t_i}(t_{\bar{j}}) < \delta_{t_i}(t_j).$$

A proximity measure δ is a set of fuzzy structures $\delta_{t_i}(t_j)$: $\delta = \{\delta_{t_i}, t_i \in [a,b]\}$. If they are consistent, i.e. $\delta_{t_i}(t_j) = \delta_{t_j}(t_i)$, then δ is called symmetric, which is not always the case.

The measure δ corresponds to the weight matrix $A = A(\delta)$:

$$A = (a_{ij}),\ a_{ij} = \delta_{t_i}(t_j)\left(\sum_{k=1}^{N} \delta_{t_i}(t_k)\right)^{-1};\ i,j = 1,\ldots,N$$

normalized by rows: $\sum_{j=1}^{N} a_{ij} = 1$. It can be represented in the form

$$A = KA_0,\ A_0 = (\delta_{t_i}(t_j)),\ K = \mathrm{diag}\left(\sum_{k=1}^{N} \delta_{t_i}(t_k)\right)^{-1}$$

2.2. Regression characteristics of the fuzzy pattern of the series

If the proximity measure δ is given on $[a,b]$, then for any series $x \in TS[a,b]$ a fuzzy pattern $\mathrm{Im}_{t_i}\, x$ x is defined at any node t_i:

$$\mathrm{Im}_{t_i}\, x = \left\{(x(t_j), \delta_{t_i}(t_j)), t_j \in [a,b]\right\}$$

Linear regression $l_{t_i} \leftrightarrow y(t) = a_i(t - t_i) + b_i$ for $\mathrm{Im}_{t_i}\, x$ will be considered as a tangent for x in t_i: the coefficients a_i and b_i are found from the minimum condition for the functional $J(a_i, b_i)$:

$$J(a_i, b_i) = \sum_{j=1}^{N} a_{ij}\left[x_i - a_i(t_j - t_i) + b_i\right]^2.$$

Definition 1. *1. The value of the variable a_i is called the regression derivative of the series x at the node t_i and is denoted as $Dx(t_i)$.*

2. The value of the variable b_i is called the regression value of the series x at the node t_i and is denoted as $Rx(t_i)$.

The minimum conditions for the functional $J(a_i, b_i)$ lead to the system of equations:

$$\begin{array}{l} \sum_{j=1}^{N} a_{ij}(t_j - t_i)\left[x_i - a_i(t_j - t_i) + b_i\right] = 0 \\ \sum_{j=1}^{N} a_{ij}\left[x_i - a_i(t_j - t_i) + b_i\right] = 0 \end{array}. \tag{1}$$

Let's introduce the operator $M_i(y) = \sum_{j=1}^{N} a_{ij} y_j$ on series $y = (y_j|_1^N)$. It coincides with the mathematical expectation of variable y taking values y_j with probability a_{ij}. Using the operator M_i, the system (1) takes the form

$$\begin{array}{l} M_i(tx) - t_i M_i(x) - a_i M_i(t - t_i)^2 - b_i(M_i(t) - t_i) = 0 \\ M_i(x) - a_i(M_i(t) - t_i) - b_i = 0 \end{array}.$$

Its solution has the form

$$\begin{array}{l} a_i = \frac{M_i(tx) - M_i(t)M_i(x)}{M_i(t^2) - M_i(t)^2} \\ b_i = M_i(x) - \frac{M_i(tx) - M_i(t)M_i(x)}{M_i(t^2) - M_i(t)^2}(M_i(t) - t_i) \end{array}.$$

Note that the denominator $M_i(t^2) - M_i(t)^2$ is the variance $D_i(t)$ of a random variable t taking values t_j with probability a_{ij}. Therefore, it is trivial only in the case of equality $t_1 = \cdots = t_N$, so always $M_i(t^2) - M_i(t)^2 \neq 0$.

The expressions for a_i and b_i can be converted to the form

$$\begin{array}{l} a_i = (D_i(t))^{-1}[M_i(tx) - M_i(t)M_i(x)] \\ b_i = M_i(x) - (D_i(t))^{-1}[M_i(tx) - M_i(t)M_i(x)](M_i(t) - t_i) \end{array}.$$

If we consider in aggregate $(Dx)_i$ and $(Rx)_i$ for all t_i, then we can introduce the operators of regression differentiation and regression smoothing:

$$D : TS[a,b] \to TS[a,b];\ R : TS[a,b] \to TS[a,b].$$

Let us calculate the matrix representations D and R

$$\begin{array}{rcl} a_i & = & (D_i(t))^{-1}[M_i(tx) - M_i(t)M_i(x)] = (D_i(t))^{-1}\sum_{j=1}^N a_{ij}t_jx_j - \\ & - & (D_i(t))^{-1}M_i(t)\sum_{j=1}^N a_{ij}x_j = \sum_{j=1}^N a_{ij}((D_i(t))^{-1}t_j - (D_i(t))^{-1}M_i(t))x_j \\ b_i & = & M_i(x) - (D_i(t))^{-1}[M_i(tx) - M_i(t)M_i(x)](M_i(t) - t_i) = \\ & = & \sum_{j=1}^N a_{ij}x_j - (M_i(t) - t_i)\sum_{j=1}^N a_{ij}((D_i(t))^{-1}t_j - (D_i(t))^{-1}M_i(t))x_j = \\ & = & \sum_{j=1}^N a_{ij}x_j\left[1 - (M_i(t) - t_i)((D_i(t))^{-1}t_j - (D_i(t))^{-1}M_i(t))\right] \end{array}.$$

Let us introduce the notation of diagonal matrices:

$$D = \mathrm{diag}(D_i(t)),\ M = \mathrm{diag}(M_i(t)),\ T = \mathrm{diag}(t_i).$$

From the above formulas follows the matrix form of the representation of the regression value operator

$$\begin{array}{rcl} R(x) & = & [A - D^{-1}(M-T)AT + D^{-1}(M-T)MA]x = \\ & = & [A - D^{-1}(M-T)(AT - MA)]x = \\ & = & [KA_0 - D^{-1}(M-T)(KA_0T - MKA_0)]x \\ R_{ij} & = & K_i\delta_{ij} + (D_i(t))^{-1}K_i\delta_{ij}(t_j - M_i(t))(t_i - M_i(t)) \end{array}$$

and the regression differentiation operator

$$\begin{array}{rcl} D(x) & = & [D^{-1}AT - D^{-1}MA]x = D^{-1}[AT - MA]x = \\ & = & D^{-1}[KA_0T - MKA_0]x \\ D_{ij} & = & (D_i(t))^{-1}K_i\delta_{ij}(t_j - M_i(t)) \end{array}$$

In practical implementation, it is especially convenient to use the matrix forms of writing the formulas discussed above, since this significantly speeds up the calculation process.

2.3. Conclusions and examples

Numerous studies and the examples below give grounds for the following conclusions:

- In contrast to the classical continuous case, discrete tangents l_{t_i} (2.2) at a node t_i are not required to take the value $x(t_i)$, so the operation R can be considered a new smoothing in its universality, not only not inferior to the usual averaging, but also surpassing it in result.

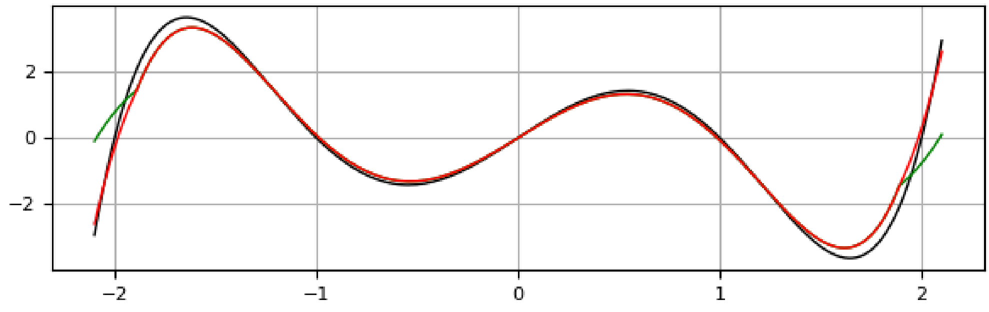

Figure 2. Regular grid. Black color denotes the original function, the green color denotes the moving average, red color is regression smoothing

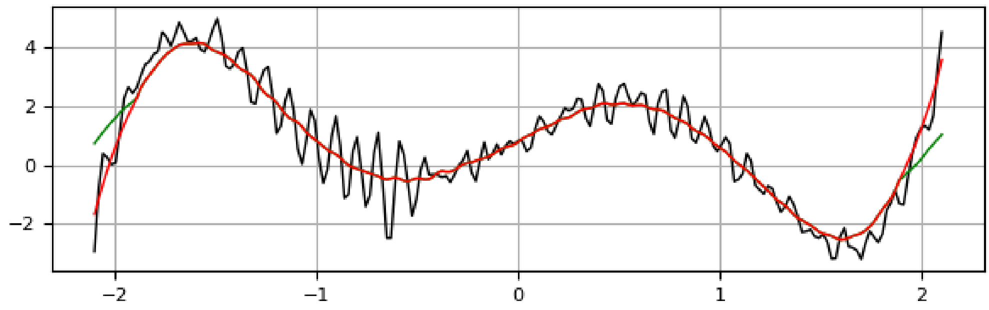

Figure 3. Regular grid. Black color denotes the original function, the green color denotes the moving average, red color is regression smoothing

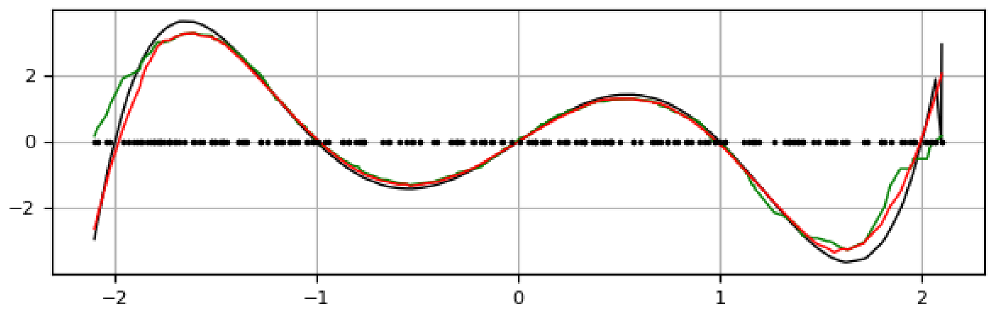

Figure 4. Irregular grid. Black color denotes the original function, the green color denotes the moving average, red color is regression smoothing

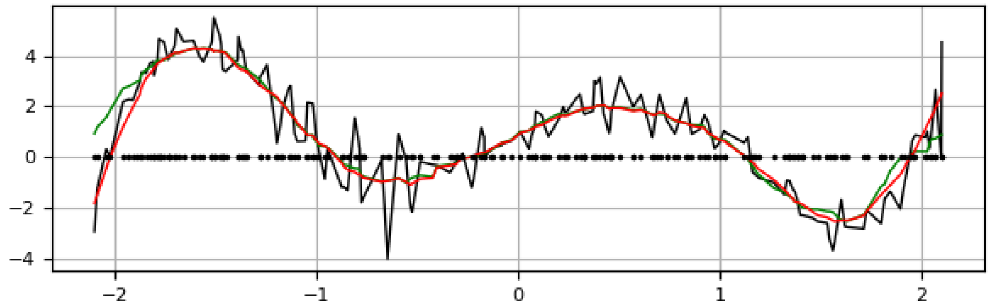

Figure 5. Irregular grid. Black color denotes the original function, the green color denotes the moving average, red color is regression smoothing

- The operator D is closely related to stochastic trends: areas of positive (negative) constancy for Rx correspond to increasing (decreasing) trends for x, and the boundaries between them correspond to its extrema

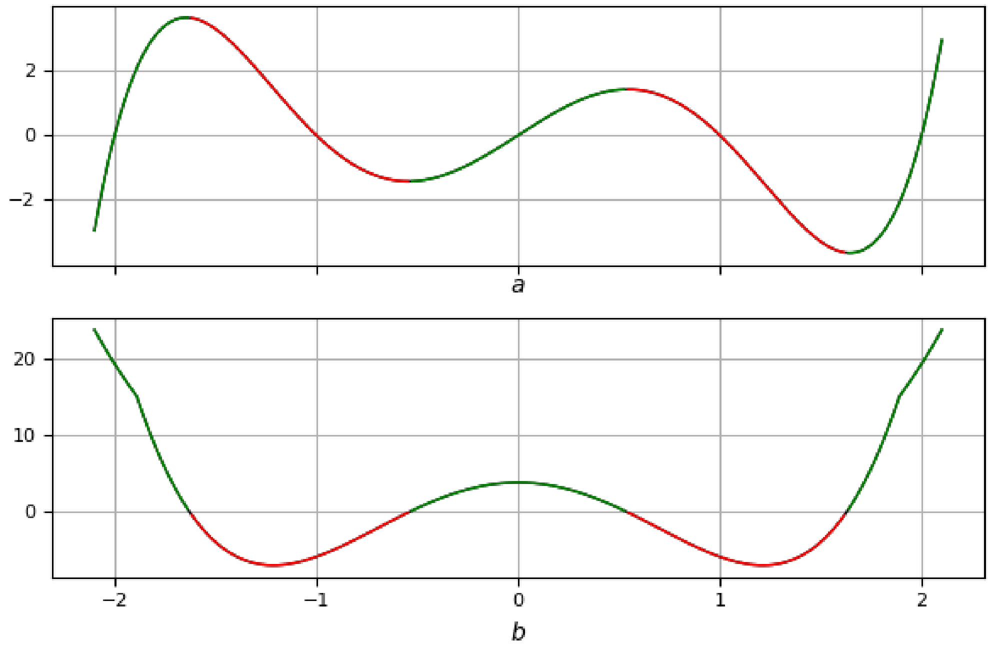

Figure 6. a is the original function, b is its regression derivative. Decreasing zones are marked in red, and increasing zones are marked in green

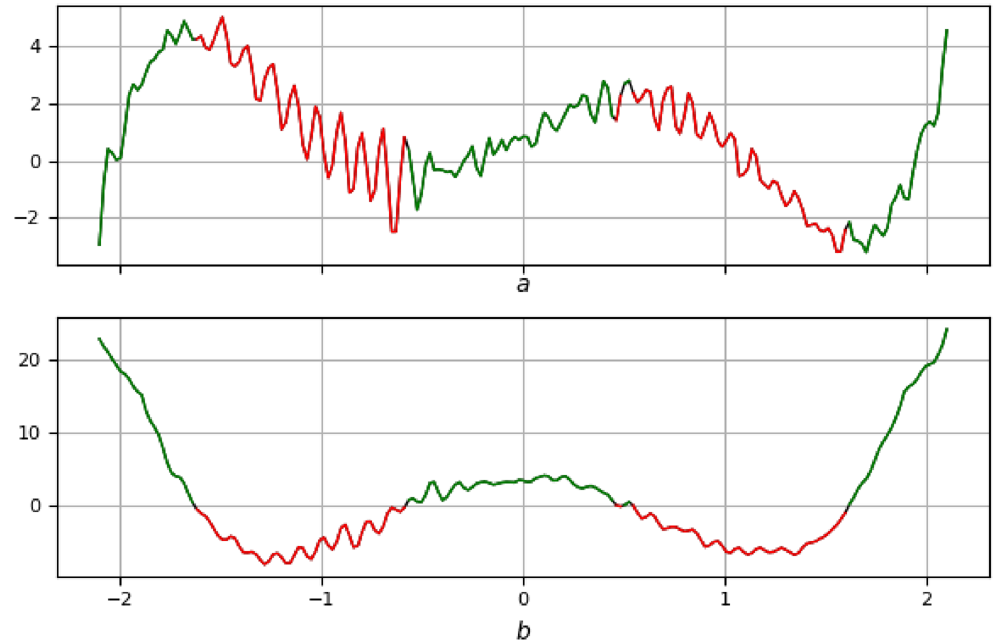

Figure 7. *a* is the original function, *b* is its regression derivative. Decreasing zones are marked in red, and increasing zones are marked in green

- The dependence of the operations D and R on the proximity measure δ on T makes it possible to solve the problems of smoothing and dynamics of time series at different scales. It is this circumstance that underlies their application to the search for anomalies and the study of their morphology.

3. Regression properties and their activity measures

From now on, we will consider a finite regular set of nodes with the discretization parameter h as the observation period of records:

$$T = \{t_1 < \cdots < t_N\},\ t_{i+1} - t_i = h,\ i = 1, \ldots, N-1.$$

We are interested in the local properties of the record associated with its relationship with its regression smoothing, namely: deviation from it and fluctuation around it. The deviation is a regression variant of energy E from [2], and fluctuation is associated with zero-crossing Z, so these properties will be denoted by ER and ZR, respectively.

In addition, regression derivatives will make it possible to obtain new results modulo [2] and, in particular, the answer to the second question of the introduction, namely, to build multi-scale trends both for the record itself and for the anomalous measures associated with it.

3.1. Measures of activity associated with regression smoothing

The construction of activity measures is complete but concise: see [2] for a more detailed explanation

The activity measure$\mu\xi_f$ of property ξ for a record f is obtained in two steps:

- <u>first stage</u> $\leftrightarrow$ quantitative expression ξ on $f \leftrightarrow$ straightening $Q(\xi)f$.
  In our case

$$Q(ER)_f(t_i) = |\Delta_i|,$$
$$Q(ZR)_f(t_{i+1/2}) = \begin{cases} |\Delta_i| + |\Delta_{i+1}|, & \text{if } (\Delta_i)(\Delta_{i+1}) < 0 \\ 0, & \text{if } (\Delta_i)(\Delta_{i+1}) \geq 0 \end{cases}$$

108 and $t_{i+1/2} = \frac{t_i + t_{i+1}}{2}$, $\Delta_i = f(t_i) - Rf(t_i)$.

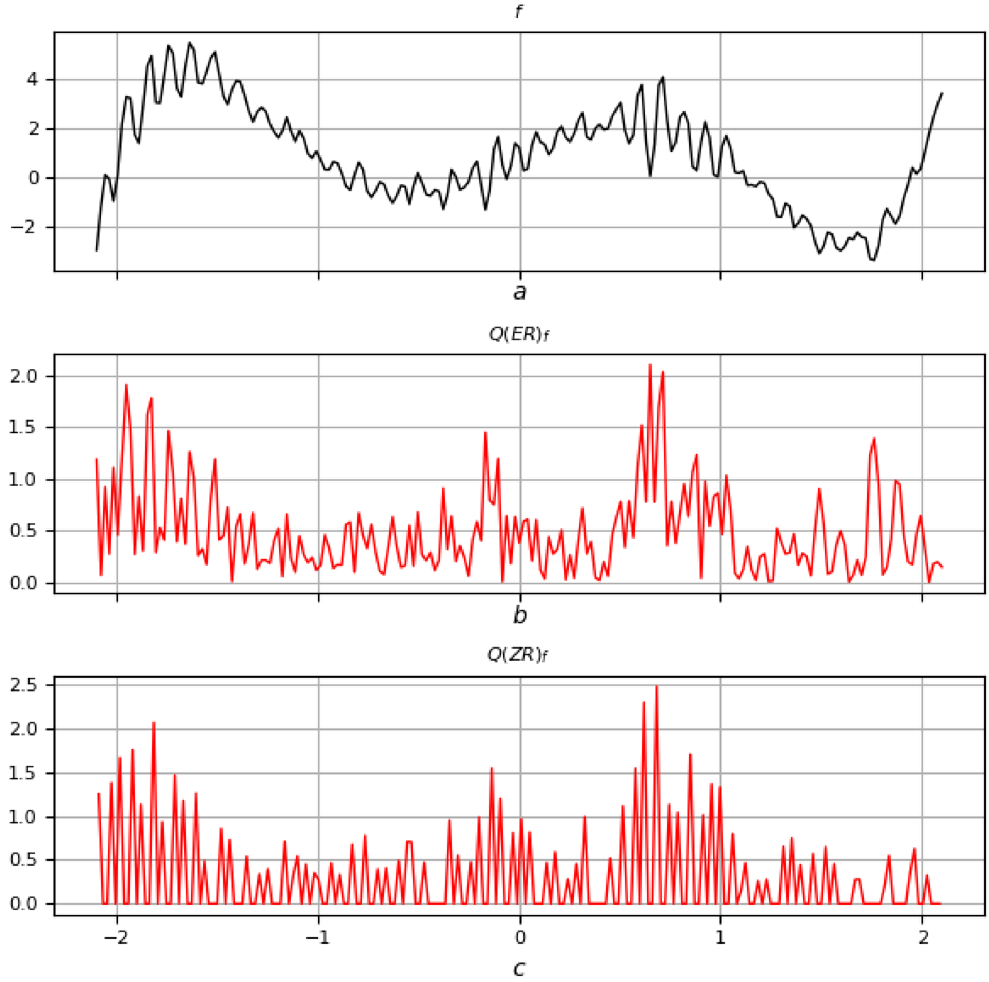

Figure 8. *a* is the original function, *b* is its straightening $Q(ER)_f$, *c* is straightening $Q(ZR)_f$

- <u>second stage</u> $\leftrightarrow$ qualitative expression ξ on f $\leftrightarrow$ activity measure itself $\mu\xi_f$. It is obtained in a node t by comparing the value $Q(\xi)_f(t)$ with the values $Q(\xi)_f(\bar{t})$ at the remaining nodes $\bar{t} \in T \setminus t$ using, so-called measure of maximum mes max:

$$\mu\xi_f(t) = \operatorname{mes\,max}_{Q(\xi)_f(T)} Q(\xi)_f(t).$$

In our case

$$\mu(ER)_f(t) = \frac{Q(ER)_f(t) - M_{Q(ER)_f}(t)}{Q(ER)_f(t) + M_{Q(ER)_f}(t)}$$

where $M_{Q(ER)_f}(t)$ is the Kolmogorov mean of straightening $Q(ER)_f$ with respect to the measure δ_t

$$M_{Q(ER)_f}(t) = \left(\frac{\sum_{\bar{t} \in T} Q(ER)_f(\bar{t})^q \delta_t(\bar{t})}{\sum_{\bar{t} \in T} \delta_t(\bar{t})} \right)^{1/q}.$$

And likewise

$$\mu(ZR)_f(t) = \frac{Q(ZR)_f(t) - M_{Q(ZR)_f}(t)}{Q(ZR)_f(t) + M_{Q(ZR)_f}(t)}$$

where $M_{Q(ZR)_f}(t)$ is the Kolmogorov mean of straightening $Q(ZR)_f$ with respect to the measure δ_t

$$M_{Q(ZR)_f}(t) = \left(\frac{\sum_{\bar{t} \in T} Q(ZR)_f(\bar{t})^q \delta_t(\bar{t})}{\sum_{\bar{t} \in T} \delta_t(\bar{t})} \right)^{1/q}.$$

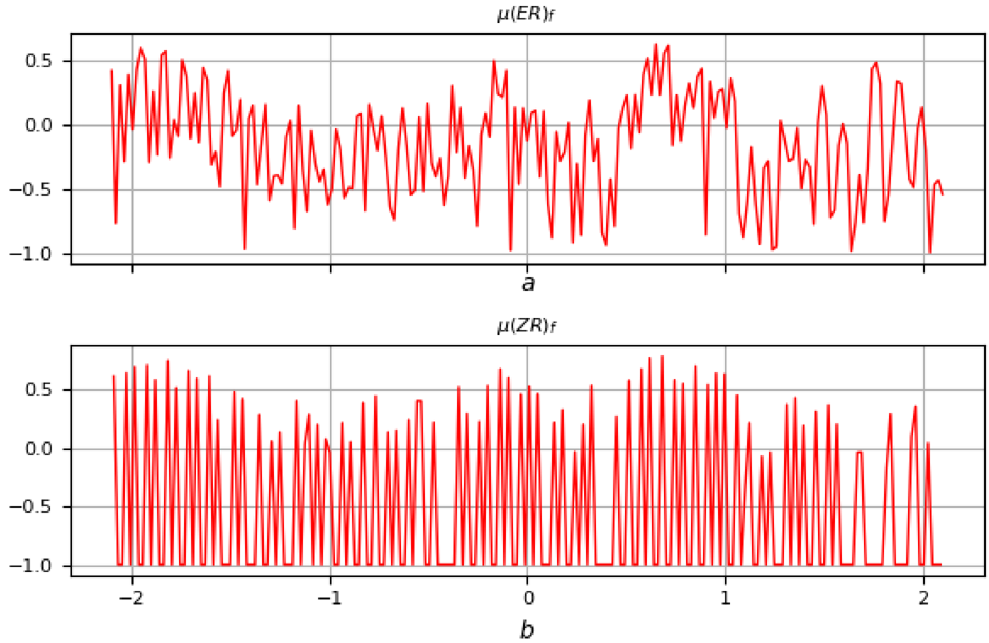

Figure 9. *a* is the measure of anomaly corresponding to the straightening in Figure 8*b*; *b* is the measure of anomaly corresponding to the straightening in Figure 8*c*

109 *3.2. Measures of activity associated with regression differentiation*

In the case of a derivative, the measure of activity takes into account its sign and, therefore, it turns out to be sign-alternating. Let's build it for the original record f. So, $\xi = D$, is the first stage $Q(D)_f = |Df|$, is the second stage

$$\mu(D)_f(t) = \frac{|Df(t)|}{|Df(t)| + M_{|Df|}(t)}$$

where $M_{|Df|}(t)$ is the Kolmogorov mean $|Df|$ regarding measure δ_t

$$M_{|Df|}(t) = \left(\frac{\sum_{\bar{t} \in T} |Df(\bar{t})|^q \delta_t(\bar{t})}{\sum_{\bar{t} \in T} \delta_t(\bar{t})} \right)^{1/q}.$$

110 *3.3. Multi-scale measures of activity*

We need the notion of multi-scale consideration of records: in this paper, it is modeled by a family of localizations $\delta_t(r, p)$, depending on the vision radius r and the localization scale $p \geq 0$ (Figure 10):

$$\delta_t(\bar{t}) = \delta_t(r, p) = \begin{cases} \left(1 - \frac{|\bar{t}-t|}{r}\right)^p, & \text{if } |\bar{t} - t| \leq r \\ 0, & \text{if } |\bar{t} - t| > r \end{cases}. \tag{2}$$

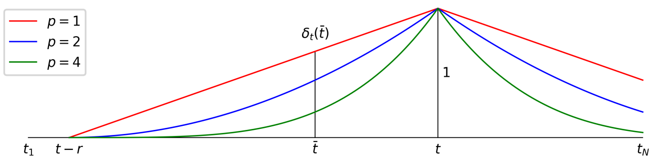

Figure 10. $\delta_t(\bar{t})$ for various values at a fixed radius r

Connecting the activity measure $\mu\xi_f(t)$ with a parametric family of different-scale localizations $\delta_t(r,p)$ (2) we get the activity measure $\mu\xi_f = \mu\xi_f(t,p)$ in the scale interval P.

The function $\mu\xi_f$ is defined on the direct product $T \times P$, and takes values in the interval $[-1,1]$ and represents at each point (t,p) the satisfaction measure of the property ξ for the series f in the node t at the scale of its consideration p.

This is how the measures $\mu ER_f(t,p)$, $\mu ZR_f(t,p)$ and $\mu D_f(t,p)$ are obtained. They will be the main tool in the study of the anomaly on f. Their portraits in the plan give a very clear picture of the structure of the record and help the researcher in its study in a semi-automatic mode (Figure 11).

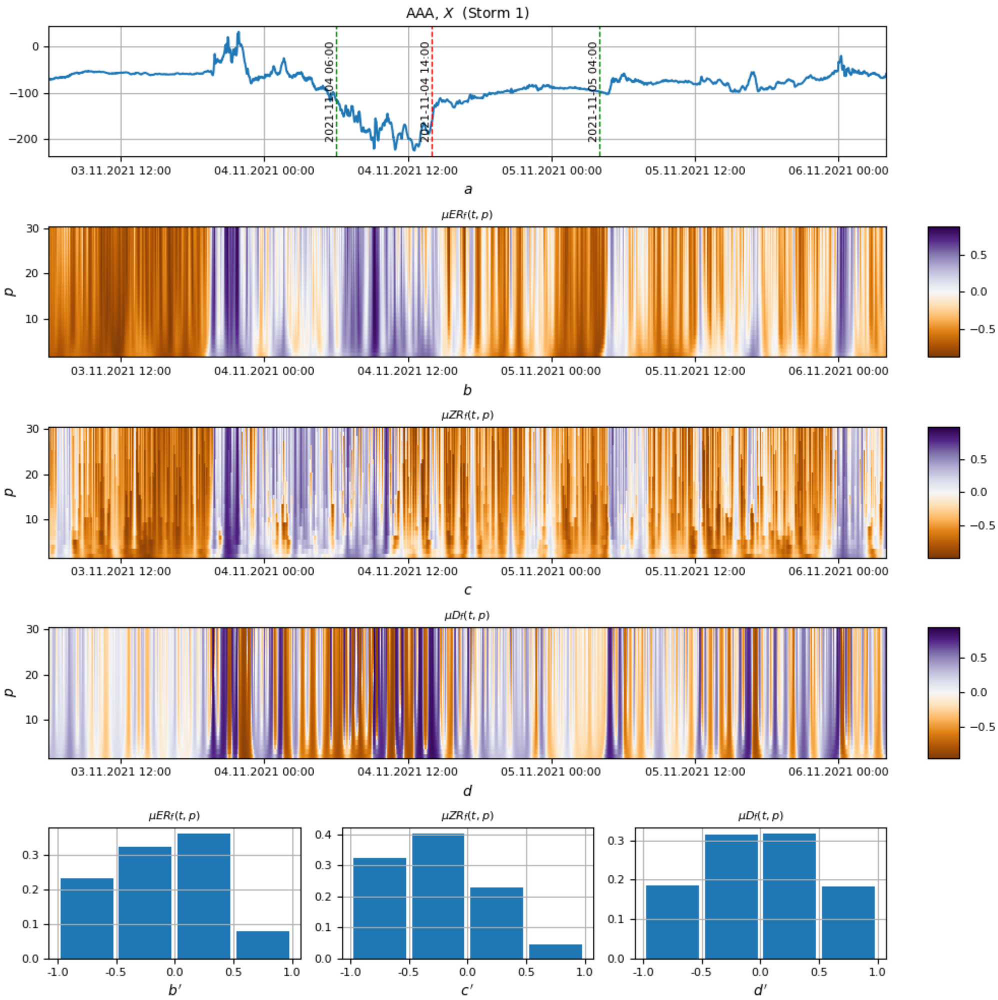

Figure 11. X magnetic component data from AAA geomagnetic observatory for storm 2 (a), the activity measure plots (b-d) and their corresponding histograms (b'-d')

The automatic mode is represented in the work by histograms of activity measures

3.3.1. Coding and analysis of activity measures

The paper proposes the simplest and most natural analysis of $\mu\xi_f$, which is necessary to understand the dynamics of f on T and compare it with the dynamics of other records.

It is related to the division H of the segment $[-1, 1]$ into four segments:

$$H \leftrightarrow [-1, 1] = [-1, -0.5] \vee [-0.5, 0] \vee [0, 0.5] \vee [0.5, 1].$$

Falling of the measure $\mu\xi_f(t, p)$ into these intervals means, respectively, from left to right, a very weak, weak, moderate and strong manifestation of the property ξ on the record f in the node t at the scale p.

The general picture on $T \times P$ of such a quantitative understanding of the situation in (t, p) is given by the histogram $H(\mu\xi_f)$, constructed from partition H for measure $\mu\xi_f$. It continues the work begun by the measure of activity, since it serves as the basis for the four-dimensional coding of anomalies for their further classification.

According to the authors, the measure μER_f gives the most complete idea of the scale of the anomaly, it is specified by the measure μZR_f. The structure inside the anomaly is revealed by the measure μD_f. It is in this order that they are given in the work when analyzing a particular record f.

4. Results

The capabilities of the technique can be displayed by analyzing magnetic observatory data registered during geomagnetic storms. From more than 30 storms that occurred at the beginning of the current 25^{th} geomagnetic activity solar cycle, we selected three storms that took place in November 2021 and during spring of 2022. These three storms are most intense, according to geomagnetic activity indices.

To test the method, data from two geomagnetic observatories were selected: the Borok observatory (IAGA-code BOX, Russia, 58.07° N, 38.23° E), and the Alma-Ata observatory (AAA, Kazakhstan, 43.25° N, 76.92° E). Both observatories belong to the INTERMAGNET network. These two observatories were chosen in order to see how the method works in geomagnetic conditions that differ by latitude. We analyzed the X component as it is most exposed to the external magnetic field during a geomagnetic storm. It is often hard to identify the storm evolution phases onsets and ands using only geomagnetic observatory data, even cleared from possible artificial disturbances and converted into absolute component values, due to intense magnetic activity variations during a magnetic storm evolution. It is even harder to do it if the analyzed data were registered by high- and mid-latitude observatories and less affected by the magnetic field driven by the equatorial Dst current system, whose plot clearly displays the storm onset, minimum and decay. Therefore, for convenience, the moments of storm onsets and ends, as well as the Dst minimum values at the ends of main storm phases, were identified using the Dst index data [ISGI, 2023]. Dst data decrease to the threshold value of -50 nT was considered a storm onset, and its increase during the storm relaxation phase above this value was defined as its end. The information on the storms is placed in Table 1. The time is given in UTC.

Table 1: Information on the storms analyzed using the approach

Start Day & Time	Peak Day & Time	Dst min, nT	End Day& Time	Duration, Hrs	Kp max
04.11.2021 06:00	04.11.2021 14:00	-105	05.11.2021 04:00	23	8-
13.03.2022 19:00	14.03.2022 01:00	-83	14.03.2022 08:00	14	6+
14.04.2022 09:00	14.04.2022 23:00	-86	15.04.2022 03:00	19	6o

After that the corresponding dashed marking lines with captions were superimposed on magnetic data plots for Dst index data (Fig. 12).

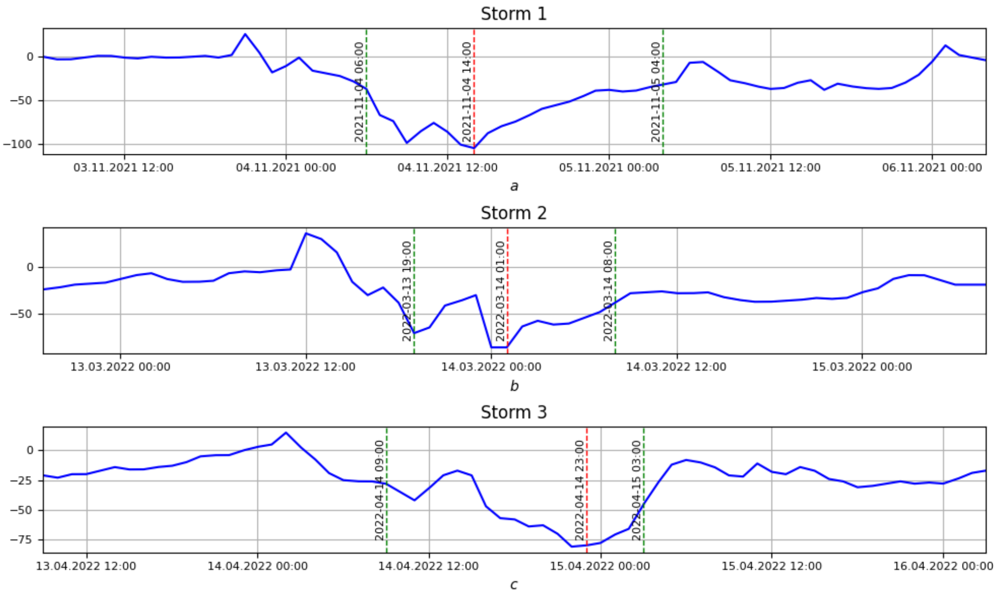

Figure 12. Dst index data for the storm period (see Table 1). Dashed lines mark the storm onsets, peaks and ends

The first of the analyzed magnetic storms is the one that occurred on November 4–5, 2021. This storm, driven by a coronal mass ejection due to a M1.7 solar flare, was the one of the most intense storms since the beginning of the new 25th solar cycle. The studied interplanetary magnetic field (IMF) and solar wind data extracted from NASA/GSFC's OMNI data set through OMNIWeb [SPDF, 2023] generally shows that, despite the fact that the initial interplanetary field state did not look very suitable for a storm generation, the overall energy driven by the coronal mass ejection produced an intense impact of the magnetohydrodynamic shockwave on the Earth's magnetosphere by the end of November 3. As the *Bz* magnetic field component abruptly turned southward, indicating the moment of the shockwave arrival, the particle speed and proton density in the plasma flux also rapidly increased 1.5 times and more than 3 times, respectively. The resulting sudden commencement signal, reaching 40 to 55 nT, is clearly seen on the records of magnetic observatories on November 3 at approximately 19:50 UTC, and also can be identified on the Dst plot (Fig. 12). According to the planetary K-index data, in the evening of November 3 the planetary geomagnetic activity reached 4 points. During November 4 it increased to almost 8 points, which corresponds to strong geomagnetic disturbance. By the end of the main storm phase, the total storm magnitude, as found out by the Dst peak value, was about -105 nT. During November 5, the Kp-index took values from 2 to 4 points (4 points also corresponds to a disturbed geomagnetic situation). The storm recovery phase lasted till November 5, 04:00 UTC.

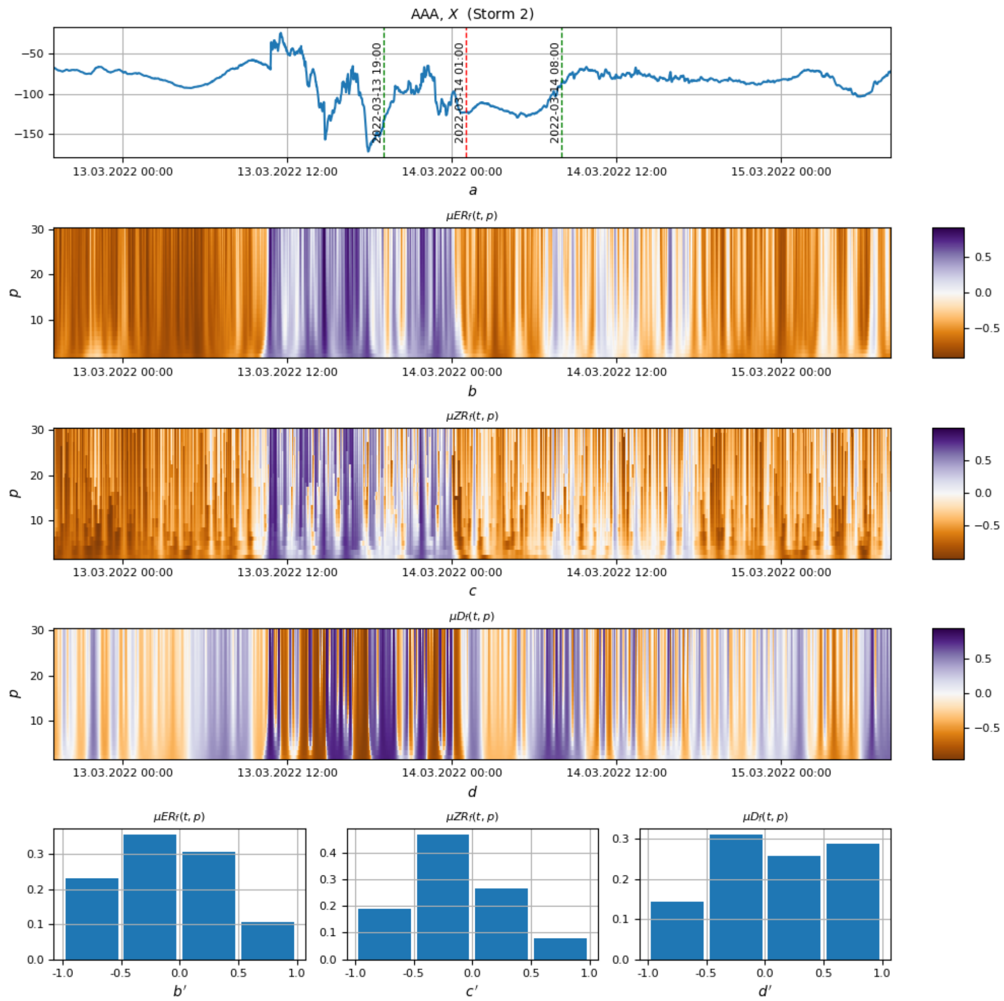

Figure 13. X magnetic component data from AAA geomagnetic observatory for storm 2 (a), the activity measure plots (b-d) and their corresponding histograms (b'-d')

The activity measure plots for this storm for both observatories (Fig. 13, 14) display some similarities, however, the particular periods are reflected by it in a different way. The $\mu ER_f(t,p)$ measure plot clearly displays more high-amplitude elements of the storm signal, such as the sudden commencement beginning and its abrupt decrease, as well as the most intense oscillations on the X component during the main storm phase. The particularity of the behavior of this measure is that most of the anomalous fragments highlighted by it as positive are bounded by abrupt increases and decreases and correspond to changes of physical conditions of the magnetic field of the Earth and its interaction with the solar wind. Therefore, the storm sudden commencement and the main phase are clearly marked, whereas the calm periods of the signal correspond to negative μER_f values. The next plot for the $\mu ZR_f(t,p)$ does not reflect much of the energy of anomalous fragments. Like the μER_f, it also reflects the sudden commencement and most of the main phase period, however, as follows from its definition, it reflects mainly the dynamics typical for oscillations around the regression smoothing. Therefore, it shows the detailed sequence of oscillations of X data for AAA observatory and even more detailed oscillations for BOX observatory during the main phase of the storm (these oscillations are larger for BOX observatory due to higher geomagnetic disturbance levels at the BOX latitude). The μDR_f, related to derivative and therefore with signsl variability, emphasizes smaller oscillations of the initial signal quite similar to theμZR_f

measure; however, relatively large-scale abrupt fragments are reflected in a way close to the μER_f result. Certainly, the μDR_f values are low for fragments with little variability.

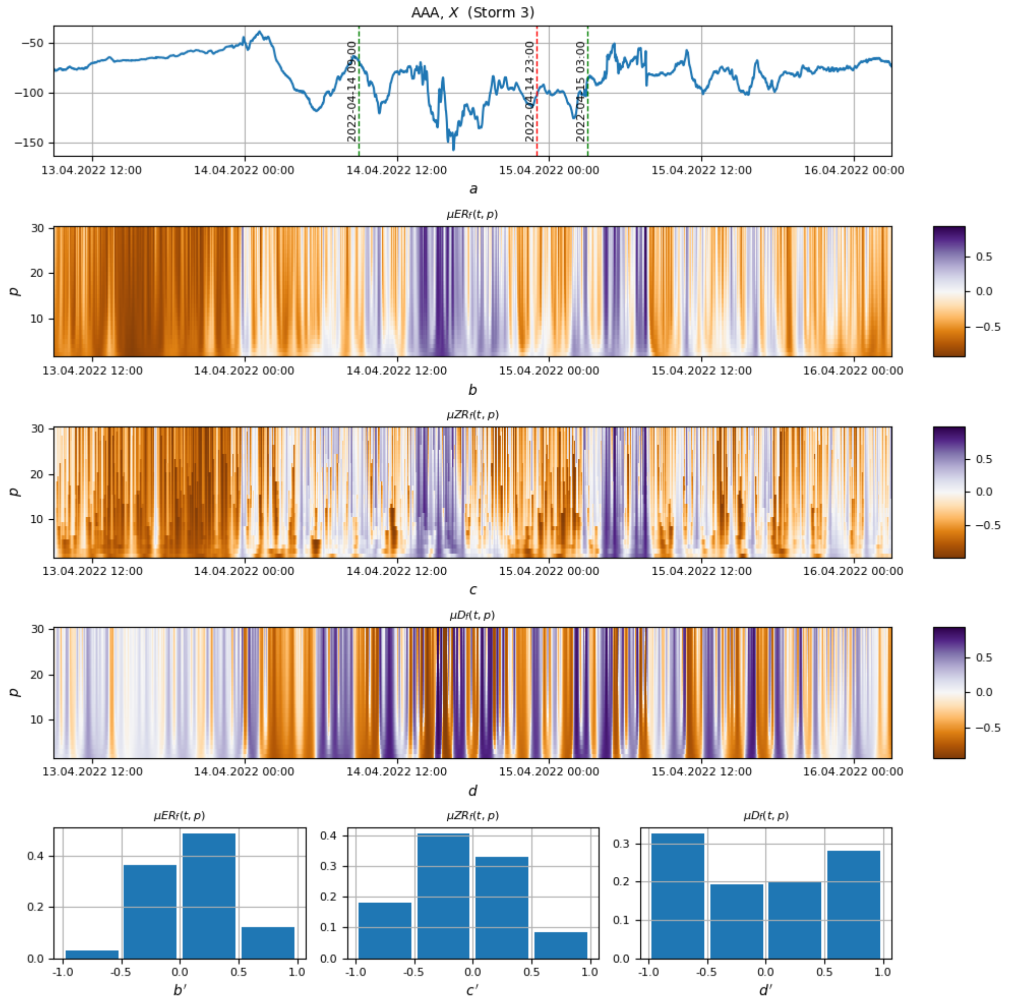

Figure 14. X magnetic component data from AAA geomagnetic observatory for storm 2 (*a*), the activity measure plots (*b*-*d*) and their corresponding histograms (*b*′-*d*′)

The next storm that occurred in March 2022 was quite similar to the previous one in the initial interplanetary conditions of its generation as, according to IMF and plasma data, the coronal mass ejection also caused a large shockwave impact on the magnetosphere, which resulted in a sudden commencement in the middle of March 13; however, due to a series of intense IMF Bz direction alternations, the storm evolution began several hours later at about 19:00 UTC. By the end of the relatively short and intense main phase, the Dst value for this storm reached a minimum of-83 nT (on March 14, 01:00 UTC). The measure plots for both observatories again have similarities, as for both geomagnetic X component signals from them the μER_f shows large positive values related to the sudden commencement moment and the following fragment related to intense alternations of $\underline{Bz}$ and solar wind characteristics. On the μZR_f and μDR_f plots this fragment is not displayed in such a generalized way, as they reflect more small-scale details. However, the μZR_f plot does not reflect any particularities for the final storm phase, whereas theμDR_f and μER_f show the oscillations seen on the X data and possibly related to auroral disturbances.

Before the next storm that occurred in April, the IMF Bz component turned southward on April 13, however, initially the solar wind energy was lower than that of two previous storms, and its impact on the Earth's magnetosphere was too low to produce an abrupt sudden commencement signal. Nevertheless, during the storm evolution, the energy driven by the solar wind plasma resulted in a total storm magnitude of -80 nT, according to Dst index data. As seen, the behavior of each measure is similar to its behavior for previous storms, however, in this case, the μZR_f plot is closer to theμER_f plot in the part related to the main phase of the storm; nevertheless, the μER_f again indicates both large- and small-scale morphological features of the storm in a more optimal way than the other two indicators.

The histograms for μER_f, built as an additional quantitative geomagnetic activity assessment, display the distribution of energy levels within the interval [-1, 1]. As seen, maximal μER_f occurrence is related to the [-0.5, 0.0] interval; therefore, the most fragments that respond to the energy indicator are related to slightly negative values. Strong positive correlation is seen between the histograms for different observatories during the same storm. This confirms that the chosen method allows assessing geomagnetic activity regardless of geomagnetic latitude and even more or less predicting the expected levels of disturbances. Moreover, this strong correlation is seen even for histograms related to different storms. This possibly confirms a reliable connection of the indicator with the physical processes of interactions between the magnetosphere and the solar wind during a geomagnetic storm.

Conclusion

The proposed technique is a new result of implementation of a DMA-based approach to geomagnetic data analysis. This group of methods includes our earlier developments, in particular: the geomagnetic activity measure [5,6], applications to the Earth's main magnetic field variations studies [7,8] and a recent approach to analysis of a geomagnetic storm morphology [2].

Research is expected to continue in two directions: the search for anomaly zones in the spectrum $\mu\xi_f(t, p)$ and their encoding.

The search for anomalous zones is supposed to be carried out by DMA-clustering algorithms [3,4] in three directions: horizontal, vertical and general, two-dimensional.

Horizontal search should give ξ-anomalies of record f in time at the fixed scale of its consideration.

A vertical search should give ξ-anomalies of the record f in frequency at a fixed node of time.

A general search should give connected two-dimensional anomalous regions in the spectrum and, thereby, describe the dynamics of the ξ-anomaly in time and frequency (appearance, development, disappearance).

The encoding of magnetic anomalies is supposed to be carried out primarily based on the measure μER_f. In support of what has been said, we present the results of a comparison of storms based on the ER encoding: it was built for three storms at two stations (AAA, BOX).

For each of the remaining five, the best one was chosen based on the ER encoding. The results are shown in the Figures 16-20.

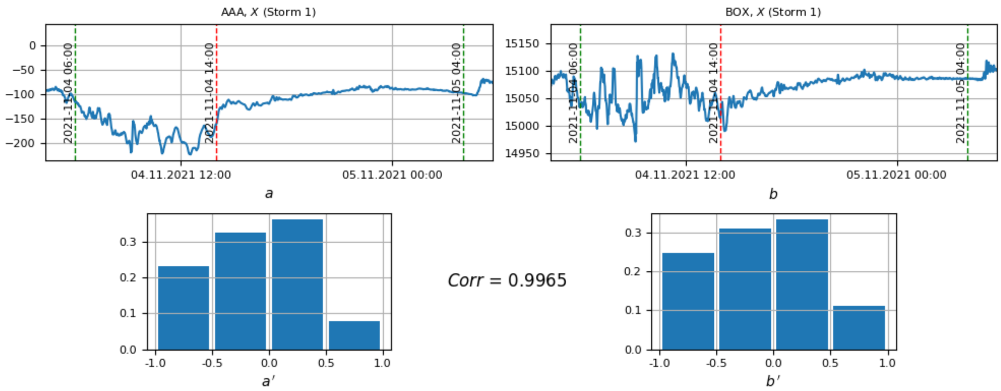

Figure 15. BOX, X (Storm 1) clothest to AAA, X (Storm 1) based on ER-encoding

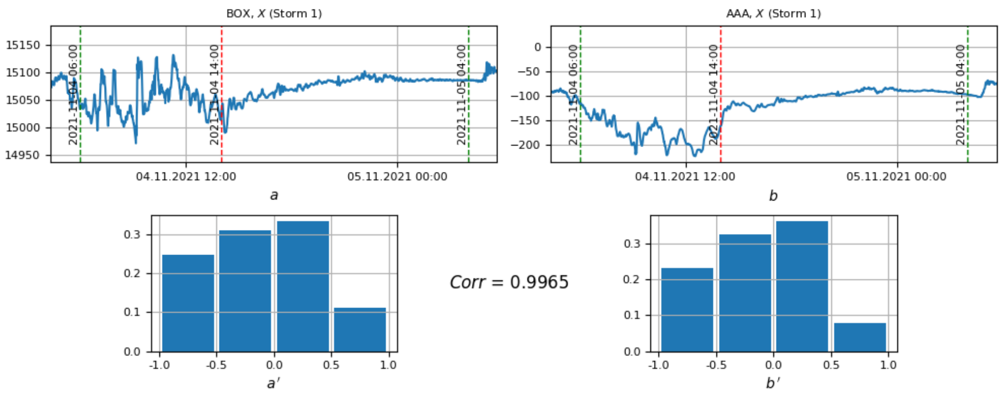

Figure 16. AAA, X (Storm 1) closest to BOX, X (Storm 1) based on ER-encoding

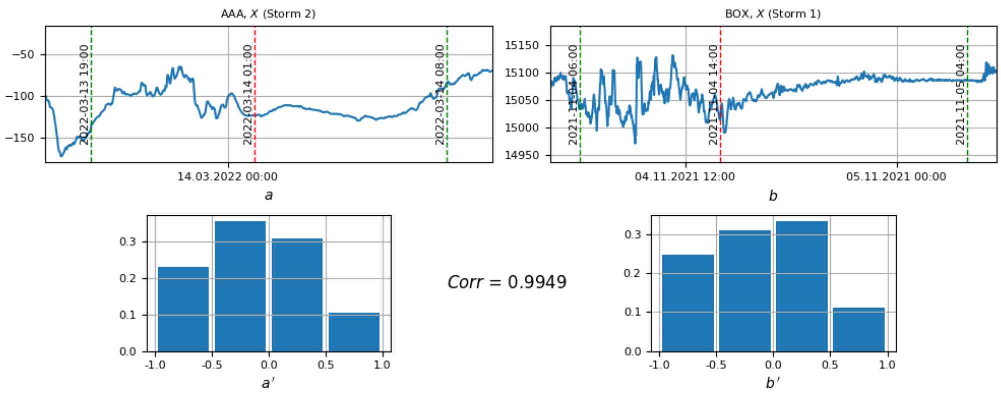

Figure 17. BOX, X (Storm 1) closest to AAA, X (Storm 2) based on ER-encoding

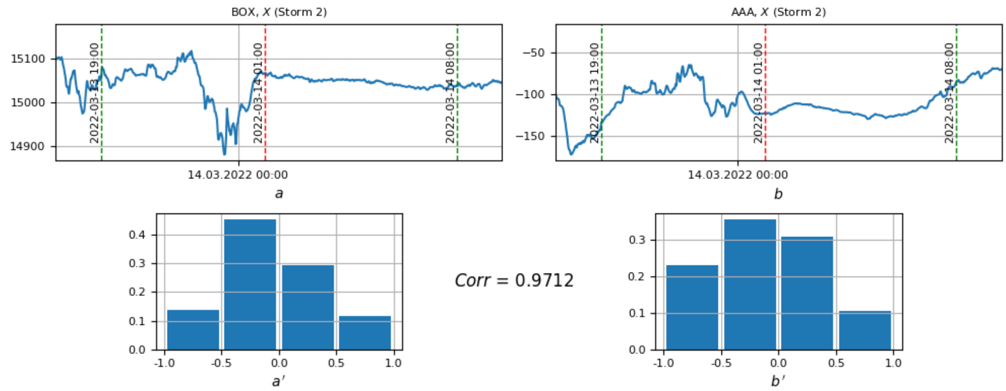

Figure 18. AAA, X (Storm 2) closest to BOX, X (Storm 2) based on ER-encoding

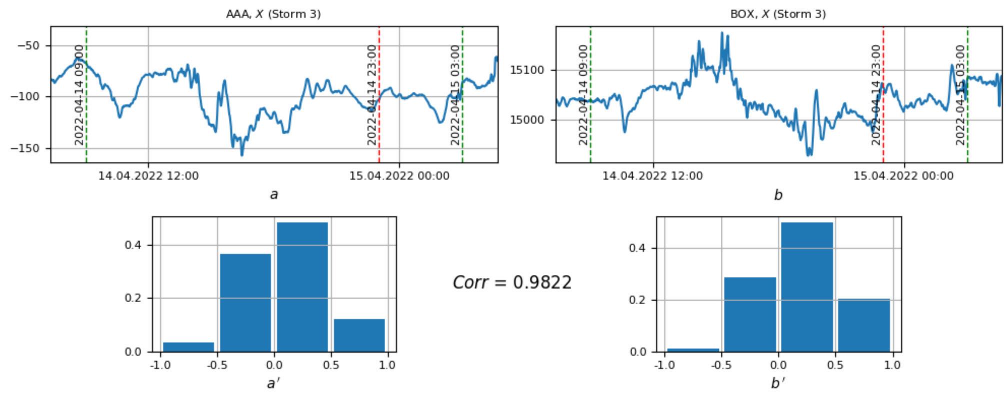

Figure 19. BOX, X (Storm 3) closest to AAA, X (Storm 3) based on ER-encoding

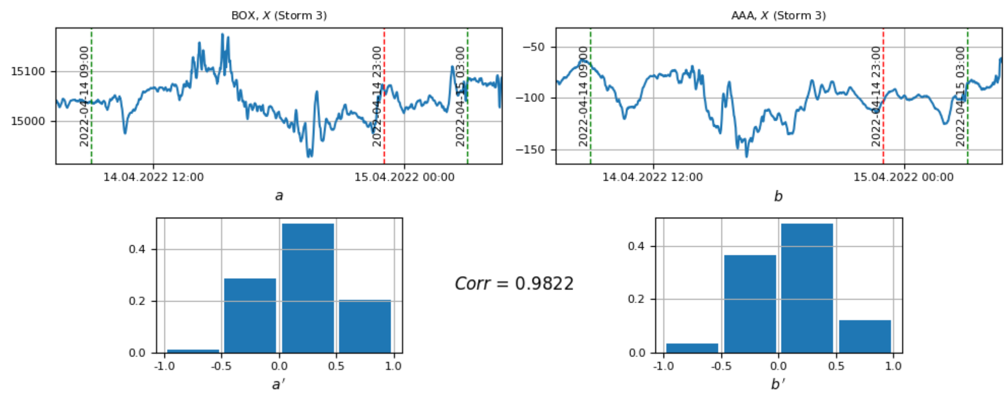

Figure 20. AAA, X (Storm 3) closest to BOX, X (Storm 3) based on ER-encoding

And the last remark: the proposed spectra $\mu\xi_f$ are of universal importance in data analysis, while the mathematical fundamental nature of the property ξ comes to the fore, and in the regression case everything is in order due to the fundamental nature of the regression itself. The geological nature of ξ is subsidiary (conditionality of ξ).

The proposed has a universal meaning in data analysis: the mathematical fundamental nature of the property ξ comes to the fore (and this is all right in the regression case due to the fundamental nature of the regression itself), while the geological nature of ξ fades into the background.

5. Patents

This section is not mandatory, but may be added if there are patents resulting from the work reported in this manuscript.

Author Contributions: All authors contributed to the study conception and design. Conceptualization, original draft preparation: A.S.M., B.Sh.R., K.D.A.; conceptualization, methodology, review and editing and validation: S.R.V., S.A.A.; material preparation, formal analysis, data curation, algorithm development: B.Sh.R., , K.D.A., A.A.O. All authors read and approved the final manuscript.

Funding: This work was conducted in the framework of budgetary funding of the Geophysical Center of RAS, adopted by the Ministry of Science and Higher Education of the Russian Federation.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FL	Fuzzy Logic
DMA	Discrete Mathematical Analysis
DPS	Discrete Perfect Sets

References

1. Gvishiani, A.; Soloviev, A. *Observations, Modeling and Systems Analysis in Geomagnetic Data Interpretation*; Springer International Publishing, 2020. doi:10.1007/978-3-030-58969-1.
2. Agayan, S.; Bogoutdinov, S.; Krasnoperov, R.; Sidorov, R. A Multiscale Approach to Geomagnetic Storm Morphology Analysis Based on DMA Activity Measures. *Applied Sciences* **2021**, *11*, 12120. doi:10.3390/app112412120.
3. Sergey Agayan.; Shamil Bogoutdinov.; Dmitriy Kamaev.; Vladimir Kaftan.; Maxim Osipov.; Victor Tatarinov. Theoretical Framework for Determination of Linear Structures in Multidimensional Geodynamic Data Arrays. *Applied Sciences* **2021**, *11*, 11606. doi:10.3390/app112411606.
4. Agayan, S.M.; Bogoutdinov, S.R.; Dzeboev, B.A.; Dzeranov, B.V.; Kamaev, D.A.; Osipov, M.O. DPS Clustering: New Results. *Applied Sciences* **2022**, *12*, 9335. doi:10.3390/app12189335.
5. Soloviev, A.; Smirnov, A.; Gvishiani, A.; Karapetyan, J.; Simonyan, A. Quantification of Sq parameters in 2008 based on geomagnetic observatory data. *Advances in Space Research* **2019**, *64*, 2305–2320. doi:10.1016/j.asr.2019.08.038.
6. Oshchenko, A.A..; Sidorov, R.V..; Soloviev, A.A..; Soloviev, E.N. Overview of anomality measure application for estimating geomagnetic activity. *Geophysical research* **2020**, *21*, 51–69. doi:10.21455/gr2020.4-4.
7. Soloviev, A.A.; Smirnov, A.G. Accuracy Estimation of the Modern Core Magnetic Field Models Using DMA-Methods for Recognition of the Decreased Geomagnetic Activity in Magnetic Observatory Data. *Izvestiya, Physics of the Solid Earth* **2018**, *54*, 872–885. doi:10.1134/s1069351318060101.
8. Agayan, S.M.; Soloviev, A.A.; Bogoutdinov, S.R.; Nikolova, Y.I. Regression Derivatives and Their Application to the Study of Geomagnetic Jerks. *Geomagnetism and Aeronomy* **2019**, *59*, 359–367. doi:10.1134/s0016793219030022.