

REGRESSION DERIVATIVES AND THEIR APPLICATION IN THE STUDY OF
MAGNETIC STORMSS. Agayan¹ , Sh. Bogoutdinov^{1,2} , R. Sidorov^{*,1} , A. Soloviev^{1,2} ,
D. Kamaev³, A. Aleksanyan⁴, and B. Dzeranov¹ ¹ Geophysical Center of the Russian Academy of Sciences, Moscow, Russia² Schmidt Institute of Physics of the Earth of the Russian Academy of Sciences, Moscow, Russia³ Research and Production Association "Typhoon", Obninsk, Russia⁴ Obninsk Institute for Nuclear Power Engineering, Obninsk, Russia* **Correspondence to:** Roman Sidorov, r.sidorov@gcras.ru.

Abstract: Discrete Mathematical Analysis (DMA) 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 activities.

Citation: Agayan, S., Sh. Bogoutdinov, R. Sidorov, A. Soloviev, D. Kamaev, A. Aleksanyan, and B. Dzeranov (2023), Regression Derivatives and Their Application in the Study of Magnetic Storms, *Russian Journal of Earth Sciences*, 23, ES6001, EDN: PHVZCV, <https://doi.org/10.2205/2023es000889>

1. Introduction

The systematic study of geomagnetic storms as extreme phenomena of geomagnetic activity with a certain morphology began approximately in the middle of the 20th century, although the perturbations of the Earth's external magnetic field were studied even earlier. Today, this is a vast area of research. There are numerous studies related to the geomagnetic storms. Issues under considerations include the onset and evolution of a storm [Akasofu and Chapman, 1963]. Another group of studies is related to the dependence of magnetic storm evolution and its phases from different types of solar wind [Yermolaev et al., 2014; Zhang, 1992]. Some researchers conducted studies of characteristic morphological features of geomagnetic storms and their occurrence over a solar cycle [Pandey and Dubey, 2009; Yokoyama and Kamide, 1997], as well as different aspects of magnetosphere dynamics [Borojev and Vasiliev, 2017; Lazutin, 2012] and the disturbances which take place during the storm phases [Gromova et al., 2016; Mishin et al., 2007; Yermolaev et al., 2012].

Until recent years, a researcher when working with magnetograms could rely only on statistical methods of spectral-temporal analysis. Recently new methods of records analysis have appeared. They are connected with development of artificial intelligence, fuzzy mathematics and allow a researcher to have more active position, to express his experience and his knowledge. This is particularly related to analysis of magnetograms. Research of data and methods of their analysis using fuzzy mathematics has now taken shape as an independent direction, which includes methods of fuzzy regression and analysis of fuzzy time series. We can highlight the main stages of development of this direction.

Research of data and methods of their analysis using fuzzy mathematics has now taken shape as an independent direction, which includes methods of fuzzy regression and analysis of fuzzy time series [Batyrshin et al., 2007; Kacprzyk et al., 2007; Kovalev, 2007;

RESEARCH ARTICLE

Received: 25 October 2023

Accepted: 28 December 2023

Published: 30 December 2023



Copyright: © 2023. The Authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Pedrycz and Smith, 1999; Tanaka et al., 1982; Yarushkina, 2004; Yarushkina et al., 2007]. We can highlight the main stages of development of this direction.

At the initial stage, studies of the fuzzy regression model were carried out. The second stage is the development of soft computing methods, within which a huge number of studies have carried out studies of the effectiveness of soft computing for time series analysis. The third stage was the transition from the analysis of time series using fuzzy mathematics methods to the analysis of fuzzy time series. The development of fuzzy database methods has made it possible to move to the stage of extracting rules from fuzzy (granular) time series. The proposed work should be attributed to the direction of using fuzzy mathematics methods for the analysis of discrete time series, that is, to the second stage.

This work should be viewed precisely from these positions: it proposes a methodology for formalizing the logic of a researcher studying a record, and its implementation within the framework of Discrete Mathematical Analysis (DMA) – a new approach to data analysis, focused on the researcher and occupying an intermediate position between hard mathematical methods and soft intellectual [*Agayan et al., 2016, 2021a, 2018*].

The DMA solution scenario consists of two parts. The first is informal: it explains the researcher's logic, introduces the necessary concepts, and explains the schemes and principles of the solution. The second is of a formal nature: with the help of the DMA apparatus, all concepts receive strict definitions within the framework of Fuzzy Mathematics (FM) and Fuzzy Logic (FL) (since the researcher does not think in numbers, but in fuzzy concepts), and schemes and principles become algorithms.

The result of this work in theoretical terms should be considered the construction for the original record $f = f(t)$ and the property P specified relative to it, of interest to the researcher, a two-dimensional function $\mu P_f = \mu P_f(t, s)$, expressing on the scale of the segment $[-1, 1]$ in at a given node t , a measure of the manifestation of property P on a record f at a given scale of its consideration $s \geq 0$ [*Agayan et al., 2016, 2021b, 2020; Oshchenko et al., 2020; Soloviev et al., 2017*].

Similar to wavelet analysis, the μP_f measure visualizes what is happening at f and allows us to better understand the nature of the process behind it. Formal analysis of μP_f , including the extraction of useful information about f , can be performed using Image Processing, as well as morphological and cluster DMA methods [*Agayan et al., 2021a, 2020, 2022*].

This is the motivation for the research presented in this paper. It continues the DMA study of records that form one of the foundations of the MAGNUS functionality, designed for collecting, processing, storing and analyzing geomagnetic information [*Gvishiani et al., 2016a*]. Before that, the authors were engaged in approaches to recognizing disturbances with known morphology on geomagnetic data, aimed at search for artificial disturbances. With the creation of MAGNUS, the DMA-based approaches were integrated into it. For example, using the local indicators, the analysis of St. Patrick's Day Storm was performed [*Gvishiani et al., 2016b*]. The developed anomalousness measure [*Soloviev et al., 2017*] was applied to geomagnetic activity studies [*Agayan et al., 2016*], in Sq variation analysis [*Soloviev et al., 2019*], and in the global geomagnetic storm analysis at different latitudes during different phases [*Oshchenko et al., 2020*]. The preceding research, also related to the DMA apparatus, was intended to demonstrate the possibilities of activity measures in the analysis of the structure of a magnetic storm [*Agayan et al., 2021b*]. This study continues this approach.

From a practical point of view, the main result of this work should be considered a multiscale morphological geomagnetic analysis of magnetic storms associated with their properties based on regression discrete derivatives.

2. Notations and Conventions

We will consider the observation period T of a record (time series) f to be a finite regular set of nodes on the real line \mathbb{R} with sampling parameter h :

$$T = \{t\} = \{t_1 < \dots < t_N\}; t_{i+1} - t_i = h, i = 1, \dots, N - 1,$$

and the record of f itself as a function on T : $f : t \rightarrow \mathbb{R}$. Let $\mathcal{F}(T)$ denote the N -dimensional space of records on T .

Analysis by a researcher of a record involves consideration of its values not only in a separate node, but also simultaneous consideration of its values in some of its neighborhood. This is why the segment T needs to be localized at each of its nodes t . This can be done using the fuzzy structure δ_t on T , acting as a neighborhood of node t and expressing the proximity of nodes t normalized in \bar{t} : $\delta_t(\bar{t}) \in [0, 1]$ – measure of proximity of \bar{t} to t :

$$\delta_t \in \text{Fuzzy } T : (\delta_t(\bar{t})) \wedge (|\bar{t} - t| < |\bar{t} - t|) \rightarrow \delta_t(\bar{t}) \leq \delta_t(\bar{t}). \tag{1}$$

We will consider the proximity measure δ on T to be a set of fuzzy structures δ_t : $\delta = \{\delta_t, t \in T\}$.

Example 1. $\delta = \delta(p, r)$; s – scale parameter, r – viewing radius (Figure 1)

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

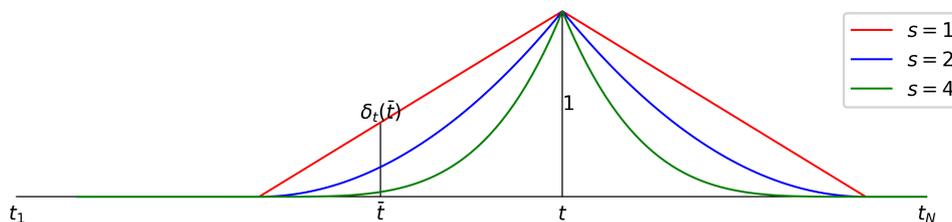


Figure 1. Proximity of $\delta(p, r)$ to node t for different s .

The parameters s and r are chosen by the researcher: the parameter r (view radius) is responsible for the boundaries of the review, and the parameter s (the scale of the review) is for the accuracy of consideration within the boundaries of the review.

3. Records Exploration Using Fuzzy Mathematics

The study of a record f by a researcher presupposes some property of P that interests him: it is the dynamics of the execution of P for f on T that the researcher needs, in particular, the zones in T where P is most pronounced on f and which he considers anomalous (P -anomalous) for f .

The DMA has a recording research program that allows us to formalize and algorithmize the above. It is implemented in the language of FM and FL, contains several stages and includes, in particular, the construction on T of a fuzzy structure (measure) $\mu P_f(t)$, which expresses at node t the degree of manifestation of property P on record f . The measure μP_f should be understood as a fuzzy formalization of the property P on the record f . Its construction involves two stages.

The first stage is called straightening of property P , consists of constructing a quantitative expression of P on record f , which is called straightening of P by f and is denoted by P_f : $P_f : T \rightarrow \mathbb{R}^+, P_f(t) \leftrightarrow$ “quantity of property P on record f at node t ”.

Straightening P_f serves as the basis for the qualitative expression of property P on record f in the form of a fuzzy structure μP_f on T , a measure of property P on f : $\mu P_f \leftrightarrow$ "quality (degree of manifestation) of property P on record f at node t ".

The measure μP_f is a membership function on T to the fuzzy concept "manifestation of P on f ", its construction constitutes the second stage of the DMA research program for studying records. In DMA, μP_f is also called a measure of the activity (anomaly) of property P on a record f .

A fundamental scheme for DMA emerges (Figure 2), which forms the basis of his approach to records.

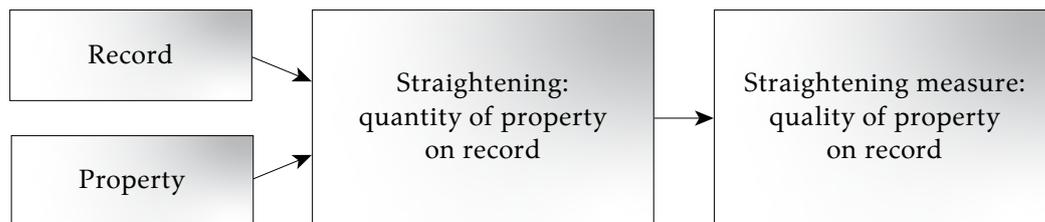


Figure 2. Scheme of DMA emerges.

4. Straightening (Quantity of Properties)

DMA leaves to the researcher the choice of both the property P itself and the construction of its quantitative expression P_f on the record f . Nevertheless, reality has shown the stability of this choice: a circle of basic straightenings has been determined that most researchers would like to deal with. Behind each of them is a fundamental mathematical concept. Let us present the most important of them: energy E (dispersion, continuity), scatter O (Cauchy fundamentality, variation), ruggedness L (frequency, length). Within the framework of this work, two more constructions related to regression derivatives will be added to them.

All listed properties P and their straightenings P_f are local: the quantity of the property $P_f(t)$ is obtained only after localizing $T(t)$ of the space T at node t , which is done using the measure δ_t by fuzzification of T : $T \rightarrow T(t) = T(t|\delta_t) = \{(\bar{t}, \delta_t(\bar{t})), \bar{t} \in T\}$.

Definition 1. Let $\mathbb{F}(T)$ be the space of functions on T .

1. Construction of straightening of the property $P \leftrightarrow$ non-negative functional on T , parameterized by T :

$$P : \mathbb{F}(T) \times T \rightarrow \mathbb{R}^+.$$

2. Straightening the record f based on the construction $P \leftrightarrow$ is a non-negative function:

$$P_f : t \rightarrow P(f, t).$$

The value $P(f, t) = P_f(t)$ is understood as a quantitative assessment of the behavior of record f at node t when P looks at its dynamics. This view is local in all interesting cases, so that everywhere below the straightening construction P is connected to some fixed localization δ_t (1), therefore $P(f, t) = P_f(t|\delta_t)$.

Example 2.

1. Energy (dispersion)

$$E_f(t|\delta_t) = \frac{\sum_{\bar{t} \in T} |f(\bar{t}) - M_f(t|\delta_t)| \delta_t(\bar{t})}{\sum_{\bar{t} \in T} \delta_t(\bar{t})},$$

where

$$M_f(t|\delta_t) = \frac{\sum_{\bar{t} \in T} f(\bar{t}) \delta_t(\bar{t})}{\sum_{\bar{t} \in T} \delta_t(\bar{t})}.$$

2. Length (ruggedness)

$$L_f(t|\delta_t) = \frac{\sum_{\bar{t}, \bar{\bar{t}} \in T: |\bar{t} - \bar{\bar{t}}| = h} |f(\bar{t}) - f(\bar{\bar{t}})| \delta_t(\bar{t}) \delta_t(\bar{\bar{t}})}{\sum_{\bar{t}, \bar{\bar{t}} \in T: |\bar{t} - \bar{\bar{t}}| = h} \delta_t(\bar{t}) \delta_t(\bar{\bar{t}})}$$

5. Measure of Straightening (Quality of Property)

As mentioned above, the measure μP_f of straightening P_f is a function of membership on T to the fuzzy concept “manifestation of property P on record f ”. In DMA there are several designs of transitions from P_f to μP_f . Let us present one of them using fuzzy comparison.

Definition 2. Fuzzy comparison $n(b, a)$ of non-negative numbers a and b measures the degree of superiority of a over b :

$$n(b, a) = \text{mes}(b < a) \in [-1, 1].$$

Comparison n should be understood as a fuzzy binary relation on the half-axis \mathbb{R}^+ , consistent with its natural order. There are many fuzzy comparisons; let’s choose one of them, further considering that $n(b, a) = (a - b)(a + b)^{-1}$. It can be extended (ambiguously) to compare a with the finite collection $B = \{b\} \subset \mathbb{R}^+$. Any such extension $n(B, a)$ will be a membership function on \mathbb{R}^+ for the fuzzy concept “to be large modulo B ”, therefore $n(B, a)$ is also treated as a measure of the maximum of a modulo B and is denoted by $\text{mes max}_B a$:

$$n(B, a) = \text{mes}(B < a) = \text{mes max}_B a \in [-1, 1].$$

Example 3. Binary extension

$$n(B, a) = \frac{\sum_{b \in B} n(b, a)}{|B|}$$

Example 4. Let ψ be a non-negative function on T , then the measure $\text{mes max}_{\text{Im} \psi} \psi(t)$ shows to what extent ψ is large at node t . In this case, it is denoted by $\text{mes max} \psi(t)$, and the image of $\text{Im} \psi$ is omitted in the index.

Figures 3a and 4a show two reliefs, and Figures 3b and 4b show their measures of maximum. The latter make it possible to divide nodes into anomalous and moderate: node t is anomalous (moderate) if $|\text{mes max} \psi(t)| \geq 0.5$ (< 0.5). Anomalous nodes (large – red, small – magenta), in contrast to moderate ones (large – green, small – blue), do not always exist: there are no abnormally large nodes, for example, on the second relief (Figure 4a), which, of course, is right. Note that the traditional probabilistic-stochastic approach to anomaly for ψ based on the distribution function of its image $\text{Im} \psi$ is softer and always effective.

We return to the property P for writing f : it is natural to consider the quality $\mu P_f(t)$ of its manifestation at node t to be the degree of maximum of its quantity $P_f(t)$:

$$\mu P_f(t) = \text{mes max} P_f(t).$$

The quality of manifestation $\mu P_f(t)$ can also be interpreted as the degree of interest of the researcher in the record f at node t in connection with the property P .

Definition 3. The fuzzy quantity $\mu P_f(t)$ on T is called a measure of property P for f .

The measure of a property is dimensionless and does not depend on the nature of the property, because it does not express the property itself, but the degree (quality) of its manifestation. Thus, the transition $P_f(t) \rightarrow \mu P_f(t)$ translates the analysis of the record f into the language of FL and FM: measures of the property μP_f for different straightenings P_f take values on a single scale of the segment $[-1, 1]$ and can be combined in any compositions and

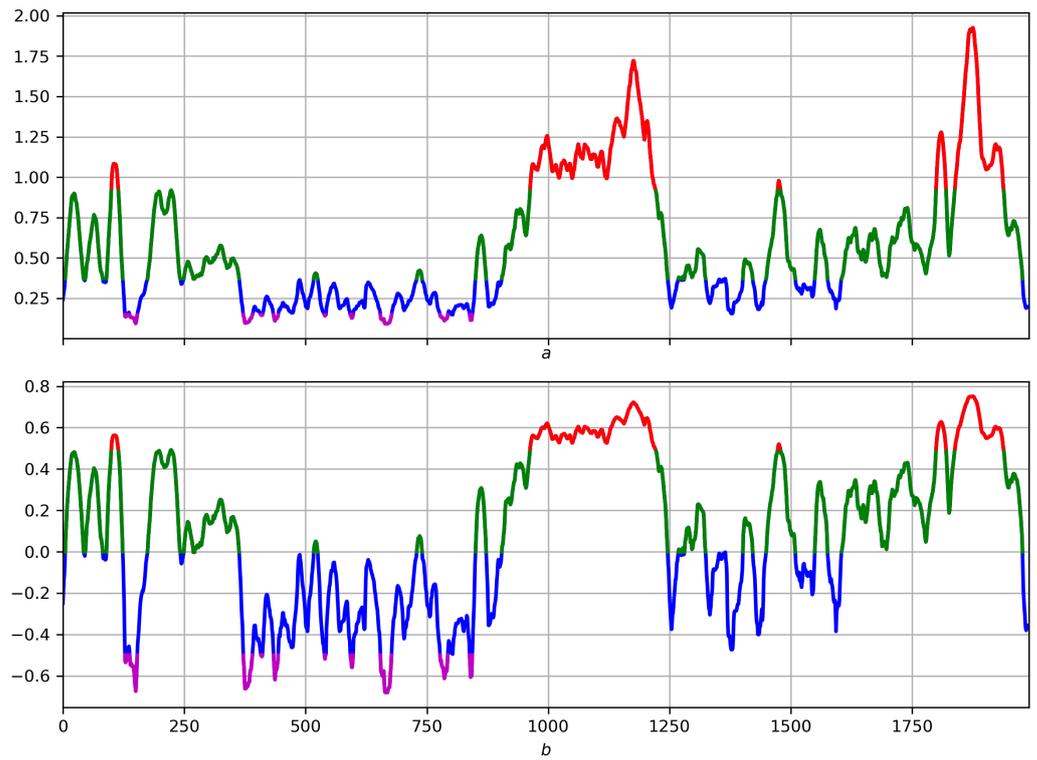


Figure 3. *a* – synthetic function; *b* – maximality measure.

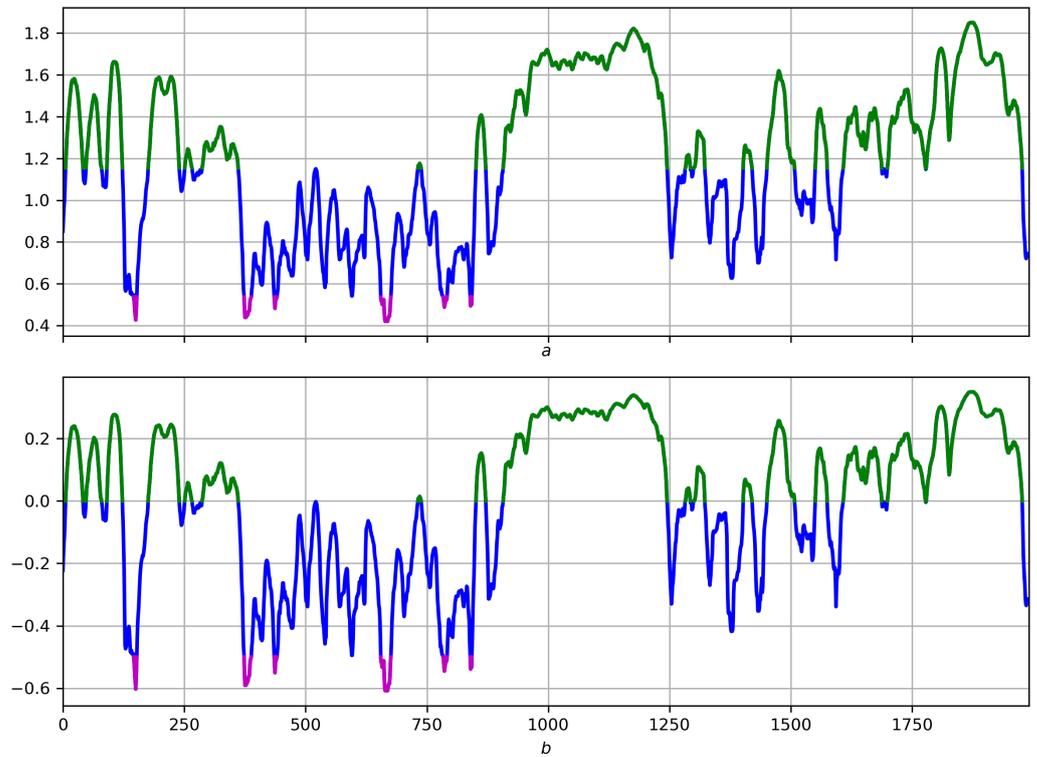


Figure 4. *a* – synthetic function; *b* – maximality measure.

in any quantities using numerous fuzzy logic operations and all kinds of averaging, which are denoted by $*$. These combinations are correct not only from a technical (syntactic) point of view, but also semantically, since all measures of straightening express the same essence – the quality of manifestation of the corresponding property.

It becomes possible to give meaning to a complex approach to record f from a set of straightenings $\mathfrak{P} = \{P\}$

$$\mu(\mathfrak{P})_f(t) = *_{P \in \mathfrak{P}}(\mu P_f(t)).$$

It is precisely this construction that, in the general case, models the researcher's view of the record f (a complex property on f). In many ways, such modeling is an art; it consists of selecting basic straightenings \mathfrak{P} and connecting them correctly $* - *(\mathfrak{P})$. It is definitely not possible without understanding the measures μP_f on the record f for the basic straightening P_f .

This article is devoted to such an understanding in relation to the properties associated with discrete regression differentiation, and in this sense continues the research of the authors in the works [Agayan et al., 2016, 2021b, 2019]. But before we get into that, let's take a quick look at the anomaly-related phase of the DMA research program.

If a fuzzy property is expressed on the scale of the segment $[-1, 1]$, then falling into the segment $[0.5, 1]$ can naturally be considered a strong (anomalous) manifestation of it.

Definition 4. Node t is P -anomalous for record f if $\mu P_f(t) > 0.5$.

The Figure 5 shows the triad: record \rightarrow straightening \rightarrow straightening measure for the ruggedness property L , anomalous nodes are marked in red.

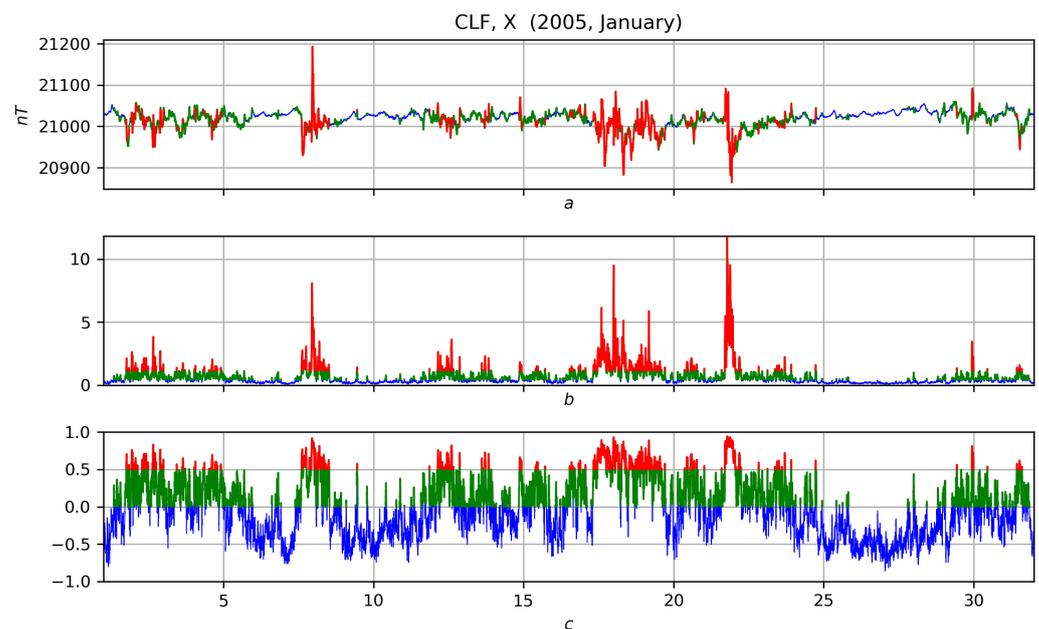


Figure 5. Triad: a – record f ; b – straightening L_f ; c – straightening measure μL_f .

The set of all anomalous nodes in T cannot be considered the final answer to P -anomaly: an anomalous node surrounded by calm nodes loses its anomaly and, conversely, a calm node surrounded by anomalous ones cannot in any way be considered completely calm.

In DMA, a functional clustering algorithm FDPS has been created that is capable of stably identifying the bases of hills on stochastic non-negative reliefs. Figure 6 compares its performance with simple level selection.

The result $FDPS(T, \mu P_f)$ of the FDPS algorithm on the space T with respect to the measure μP_f is the answer about the P -anomaly of the record f .

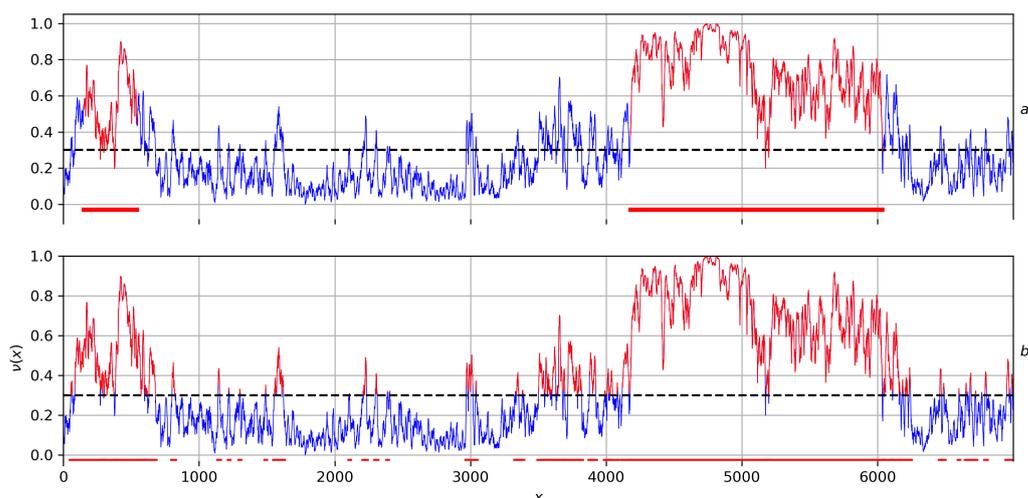


Figure 6. *a* – anomalies identified by the FDPS algorithm with $r = 23.26$; *b* – anomalies identified by level. The dotted line is the specified level $\alpha = 0.3$.

6. Regression Differentiation and Regression Smoothing

Continuous case: let the function f be integrable on an interval I containing zero internally. Then for a sufficiently small $\Delta > 0$ the segment $[-\Delta, \Delta]$ is contained in I . Let us denote by f_Δ the restriction of f to the segment $[-\Delta, \Delta]$: $f_\Delta = f|_{[-\Delta, \Delta]}$ and calculate the projection $\text{pr } f_\Delta$ of the function f_Δ in space $L^2[-\Delta, \Delta]$ into the two-dimensional subspace of linear functions $\text{Lin}[-\Delta, \Delta]$. In [Agayan et al., 2019] proven:

State 1. *If a function f has a tangent at zero, then as $\Delta \rightarrow 0$ the linear projection $\text{pr } f_\Delta$ tends to it.*

The projection $\text{pr } f_\Delta$ is nothing more than a linear regression of f on $[-\Delta, \Delta]$, and therefore the tangent is the limit position of local continuous regressions. This approach to differentiation in the continuous case can be extended to the discrete case, since discrete regressions are as effective and fundamental as continuous ones.

Discrete case. The limit transition $\bar{t} \rightarrow t$ in T is performed by the proximity measure δ_t (1) by fuzzification $T|_{\delta_t} = \{(\bar{t}, \delta_t(\bar{t})), \bar{t} \in T\}$. The above statement gives grounds to consider the tangent $l_\delta f(t) \leftrightarrow l_\delta f(t)(\bar{t}) = a_t \bar{t} + b_t$ for recording f at node f to be a linear regression constructed from the fuzzy image $\text{Im}_\delta f(t) = \{(f(\bar{t}), \delta_t(\bar{t})), \bar{t} \in T\}$. Omitting the standard things associated with linear regressions, we present the formulas for a_t and b_t :

$$a_t = \frac{\begin{vmatrix} \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) f(\bar{t}) & \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) \\ \sum_{\bar{t} \in T} \delta_t(\bar{t}) f(\bar{t}) & \sum_{\bar{t} \in T} \delta_t(\bar{t}) \end{vmatrix}}{\begin{vmatrix} \sum_{\bar{t} \in T} \bar{t}^2 \delta_t(\bar{t}) & \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) \\ \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) & \sum_{\bar{t} \in T} \delta_t(\bar{t}) \end{vmatrix}}, \quad b_t = \frac{\begin{vmatrix} \sum_{\bar{t} \in T} \bar{t}^2 \delta_t(\bar{t}) & \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) f(\bar{t}) \\ \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) & \sum_{\bar{t} \in T} \delta_t(\bar{t}) f(\bar{t}) \end{vmatrix}}{\begin{vmatrix} \sum_{\bar{t} \in T} \bar{t}^2 \delta_t(\bar{t}) & \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) \\ \sum_{\bar{t} \in T} \bar{t} \delta_t(\bar{t}) & \sum_{\bar{t} \in T} \delta_t(\bar{t}) \end{vmatrix}}.$$

Definition 5.

1. The angular coefficient a_t is called the regression derivative of f at t and is denoted by $D_\delta f(t)$.
2. The function $t \rightarrow a_t$ is called the regression derivative f and is denoted by $D_\delta f \in \mathcal{F}(T)$.
3. The functional correspondence $f \rightarrow D_\delta f$ is a linear operator $\mathcal{F}(T)$ and is called regression differentiation D_δ .

Definition 6.

1. The value $l_{\delta}f(t) = a_t t + b_t$ of the regression tangent $l_{\delta}f(t)$ of record f at node t is called the regression value of f at t and is denoted by $R_{\delta}f(t)$.
2. The function $t \rightarrow R_{\delta}f(t)$ is called regression smoothing f and is denoted by $R_{\delta}f \in \mathcal{F}(T)$.
3. The functional correspondence $f \rightarrow R_{\delta}f$ is a linear operator $\mathcal{F}(T)$ and is called regression smoothing R_{δ} .

Everywhere the measure δ is assumed to be from the family $\delta(s, r)$ (2). We denote the corresponding differentiation and smoothing by $D(s, r)$ and $R(s, r)$. Numerous studies provide grounds for the following conclusions:

1. Smoothing $D(s, r)$, not inferior to the usual averaging $M(s, r) = M_{\delta(s, r)}$ in universality, surpasses it in results (Figures 7–8).

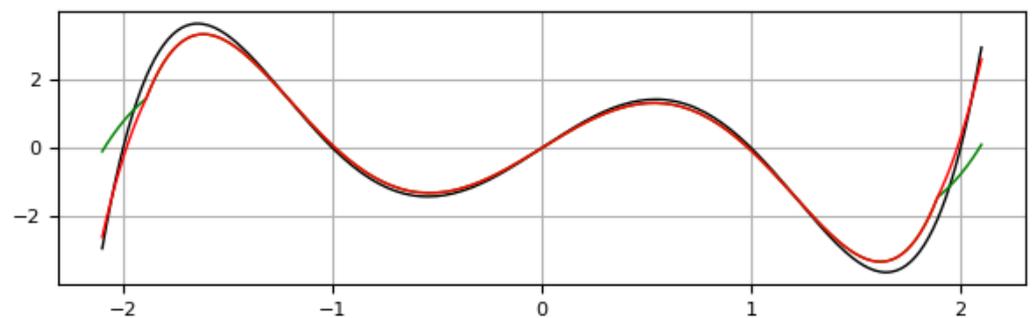


Figure 7. Regular grid. Black color – original function, green color – moving average, red color – regression smoothing.

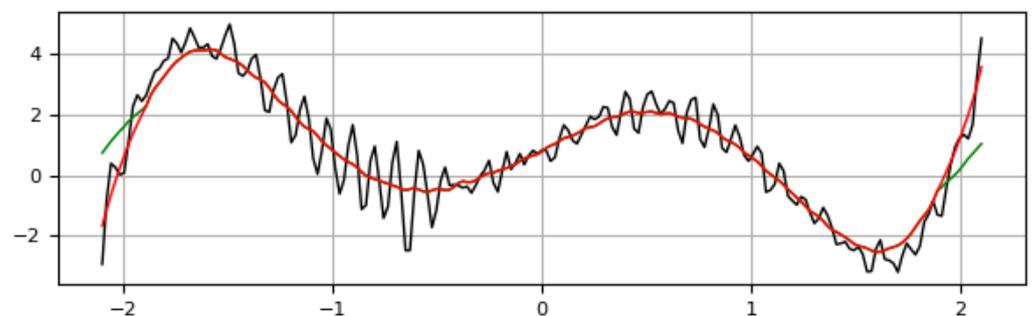


Figure 8. Regular grid. Black color – original function, green color – moving average, red color – regression smoothing.

2. The operator $R(s, r)$ is closely related to stochastic trends: areas of positive (negative) sign for $D(s, r)f$ correspond to increasing (decreasing) trends for f (Figures 9–10).

7. Multi-scale property analysis

Combining the property measure μP_f with a parametric family of different-scale localizations $\delta(s, r)$ (2) gives the spectrum of manifestation of property P on a record f in the interval of scales S :

$$\mu P_f(t, s) = \mu P_f(t | \delta_t(t, s)).$$

The function μP_f is defined on the product $T \times S$, takes values on the scale of the segment $[-1, 1]$ and at each point (t, s) there is a measure $\mu P_f(t, s)$ of the manifestation of property P on the record f at node t at the scale of its consideration s , that is, relative to the localization $\delta_t(t, s)$.

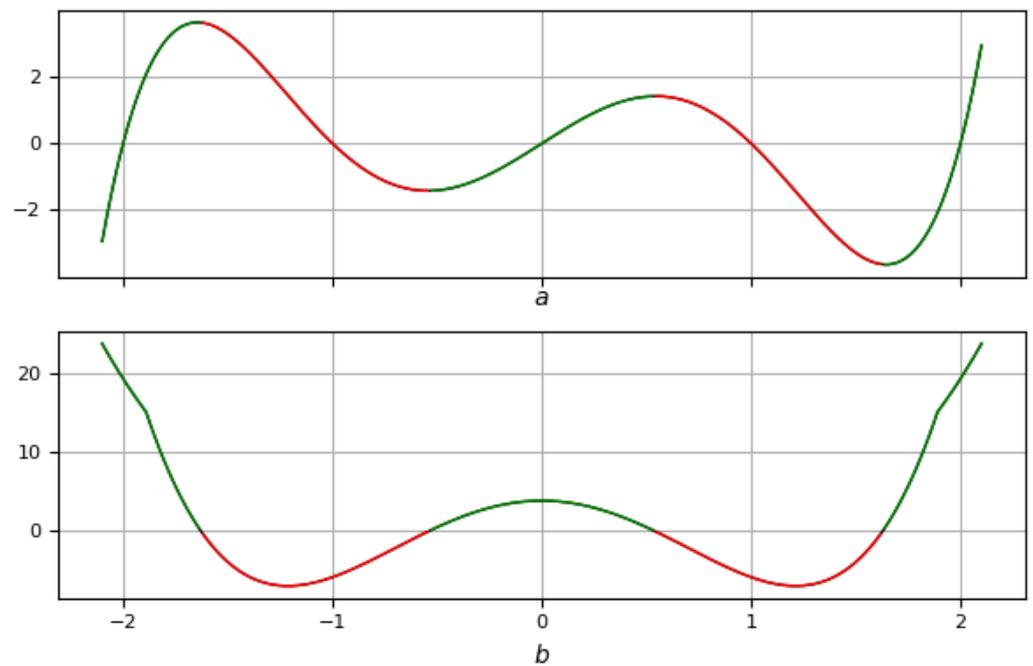


Figure 9. a – original function, b – its regression derivative. Red indicates decreasing zones, green indicates increasing zones.

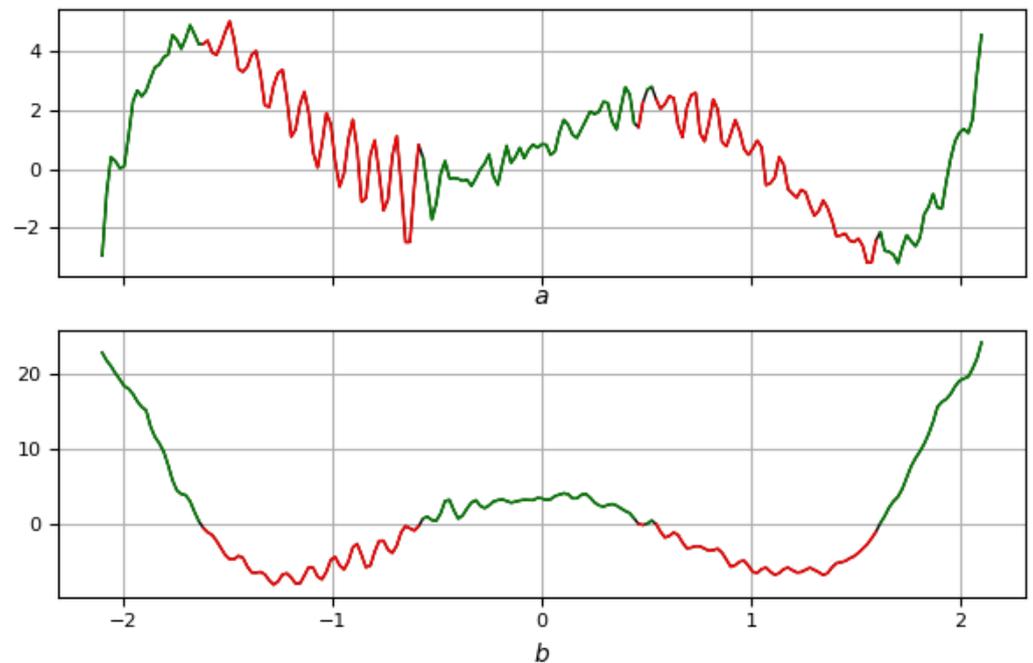


Figure 10. a – original function, b – its regression derivative. Red indicates decreasing zones, green indicates increasing zones.

Thus, the one-dimensional notation of f on T is associated with the two-dimensional function μP_f on $T \times S$. This redundancy makes it possible to better see what is happening at f and understand the nature of the process behind it.

Like the wavelet spectrum, the two-dimensional relief $\mu P_f(t, s)$ provides a visual representation of the dynamics of the emergence, evolution and disappearance of P -anomalies on f at different scales and in time. Formal analysis of the measure μP_f , which consists of extracting useful information from the record f , can be performed using the same methods as for the wavelet spectrum, in particular, Image Processing. Morphological and cluster DMA methods are also suitable. This is the motivation for conducting research at different scales.

8. Work Objectives (Statement and Goals)

One of the most important when analyzing any record is the “discontinuity-discontinuity” connection. In turn, the main design of their quantitative expression (straightening) is associated with the idea of smoothing: one or another smoothing of a record is considered its ideal scenario, the deviation from which quantitatively expresses “discontinuity-continuity” (small deviation \leftrightarrow continuity, large deviation \leftrightarrow discontinuity).

Thus, when constructing the straightening “energy” E_f , the ideal scenario for recording f was considered to be its averaging $M_\delta f$, and the energy itself $E_f(t) = EM_\delta f(t)$ was a local deviation at node t from its “correct” value $M_\delta f(t)$.

Similarly, if the ideal scenario for f is its regression smoothing $R_\delta f$, then the deviation $ER_f = |f - R_\delta f|$ will be a new straightening for f , quantitatively expressing for it the properties of continuity and discontinuity:

$$ER_f(t) = ER_f(t|\delta_t) = \frac{\sum_{\bar{t} \in T} |f(\bar{t}) - R_\delta f(t)\delta_t(\bar{t})|}{\sum_{\bar{t} \in T} \delta_t(\bar{t})}.$$

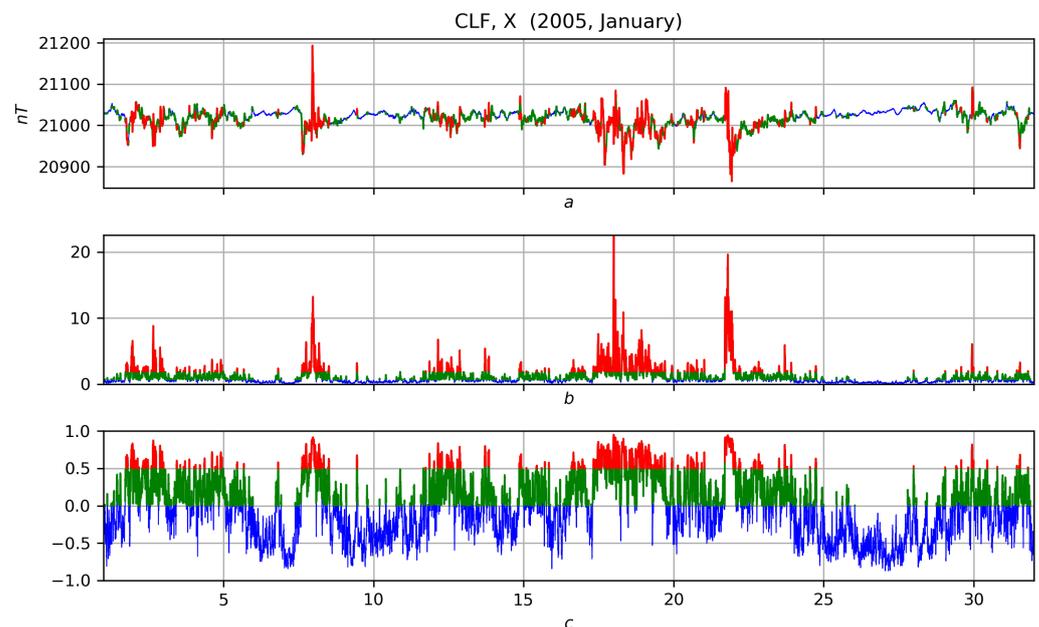


Figure 11. Triad: a – record f ; b – straightening ER_f ; c – maximality measure μER_f .

A multi-scale study of the μER_f measure for magnetic storms f is the first task of this work (Figure 11). The second task is a similar study for magnetic storms f of the measure of their local growth μD_f , based on regression differentiation D_δ (Figure 12):

$$\mu D_f(t) = \mu |D_f|(t) \operatorname{sgn} D_f(t),$$

where $\mu|D_f|(t)$ is a measure of the maximum modulus $|D_\delta f(t)|$ in the positive scale of the interval $[-1, 1]$:

$$\mu|D_f|(t) = \frac{\sum_{\bar{t} \in T} \frac{|D_\delta f(t)|}{|D_\delta f(t)| + |D_\delta f(\bar{t})|}}{|T|}.$$

The study, in addition to the phenomenological geomagnetic analysis of measures, involves their simplest statistical analysis. It is associated with the division H of the segment $[-1, 1]$ into four segments:

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

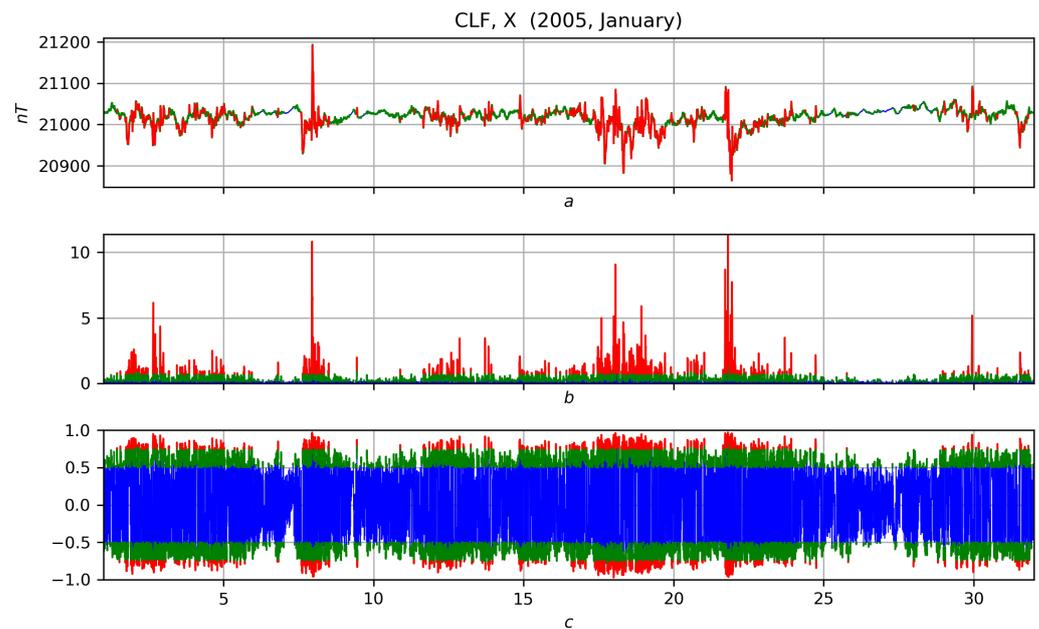


Figure 12. Triad: a – record f ; b – straightening Df ; c – maximality measure μDf .

If the measure $P_f(t, s)$ falls into these intervals, it means, respectively, from left to right, very weak, weak, moderate and strong manifestation of property P on the record f at node t at scale s .

The general picture on $T \times S$ of such a qualitative understanding of the situation in (t, s) is given by the histogram $H(\mu P_f)$, constructed from the partition H for the measure μP_f . It represents a four-dimensional coding f , is carried out in the work and serves as the basis for further studies of records, including their correlation and classification.

9. Application to Geomagnetic Storm Analysis

The capabilities of the technique can be displayed by analyzing magnetic observatory data registered during geomagnetic storms. From more than 50 storms that occurred during the current 25th magnetic activity solar cycle, we selected three storms that took place in November 2021 and during spring of 2022. The information on these storms, including their duration and peak values of geomagnetic activity indices K_p and Dst , is given in Table 1. According to NOAA magnetic storm intensity scale [NOAA, 2023], two of these three storms (the 2nd and 3rd) are moderate in their intensity (G2), and the first one is intense (G3). This also agrees with their intensity classification using the Dst index: the first storm, having a Dst minimum value of -105 nT, refers to intense storms (however it is just 5 nT lower than the intense storm range threshold), and two other storms are moderate. This table actually gives the information about time spans for two storm phases, these are the main phase (MP) and the recovery, or relaxation phase (RP). The main phase onset can

generally be assumed as an exact moment of a whole magnetic storm onset; however, before some storms the sudden commencements take place several hours prior to a storm onset, and in this case, depending on a particular research goals, the sudden commencement time moment can be a point in time that can be considered the beginning of a storm. For convenience, we use the Dst index data [ISGI, 2023] to identify the timestamps of storm onsets and ends, as well as the Dst minimum values at the end of main storm phases. Dst index abrupt decrease to slightly negative values was considered a storm onset, and its increase during the storm recovery phase above -30 nT was defined as its end. The corresponding dashed marking lines with captions were superimposed on magnetic data plots for Dst index data (Figure 13).

Table 1. Information on seismic stations used

MP Start, UTC	Peak (MP End), Peak (UTC)	Peak Dst, nT	RP End, UTC	Duration, Hrs	Kp max
04.11.2021 06:00	04.11.2021 13:00	-105	05.11.2021 04:00	22	8-
13.03.2022 14:00	14.03.2022 00:00	-83	14.03.2022 08:00	18	6+
14.04.2022 15:00	14.04.2022 21:00	-86	15.04.2022 03:00	12	6o

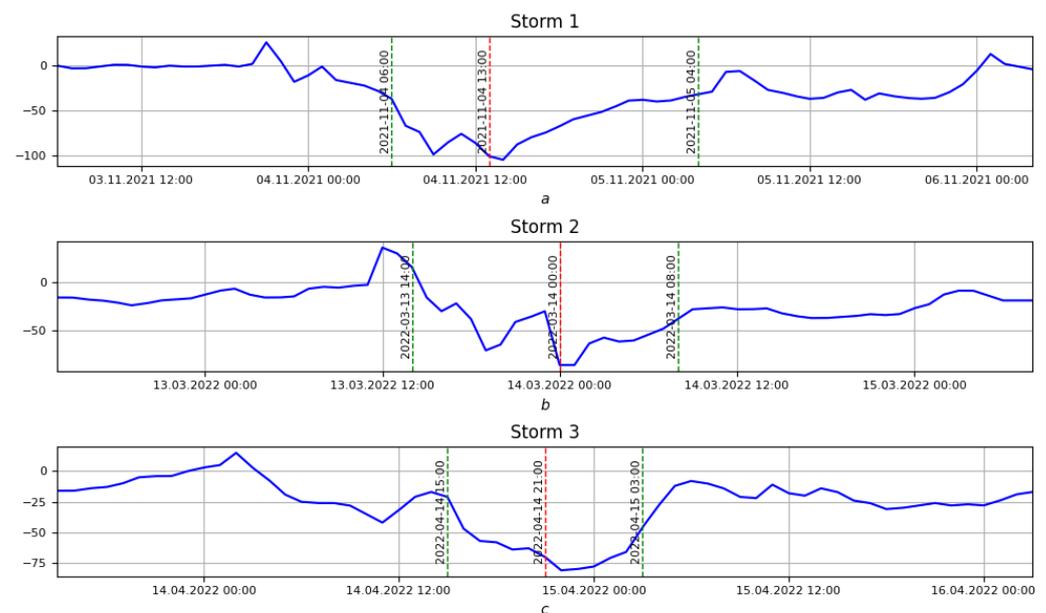


Figure 13. Dst index data for the selected storms (see Table 1). Dashed lines mark the storm onsets, peaks and ends.

To test the general capabilities of the method, data from two magnetic observatories were selected: the Borok observatory (IAGA-code BOX, Russia) and the Alma Ata observatory (AAA, Kazakhstan). The information on these observatories is given in Table 2 and includes their names, geographic coordinates (ϕ , λ) and geomagnetic coordinates (ϕ_M , λ_M). The geomagnetic coordinates were calculated for the corresponding time periods using the web service designed at the World Data Center for Geomagnetism, Kyoto [WDC, 2023]. Both observatories are members of INTERMAGNET network [Intermagnet, 2023], so we use their names as given on the INTERMAGNET website. As these two observatories differ by latitude, they were chosen in order to see how the method works in geomagnetic conditions at different latitudes. 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 time limits of storm evolution phases using only magnetic observatory data, even cleared

from possible artificial disturbances and converted into complete component values, due to intense magnetic activity variations.

Table 2. Information on seismic stations used

IAGA Code	Name	ϕ, λ	ϕ_M, λ_M (2021)	ϕ_M, λ_M (2022)
AAA	Alma Ata	43.25° N, 76.92° E	34.83° N 153.22° E	34.88° N 153.20° E
BOX	Borok	58.07° N, 38.23° E	53.61° N 123.20° E	53.64° N 123.14° E

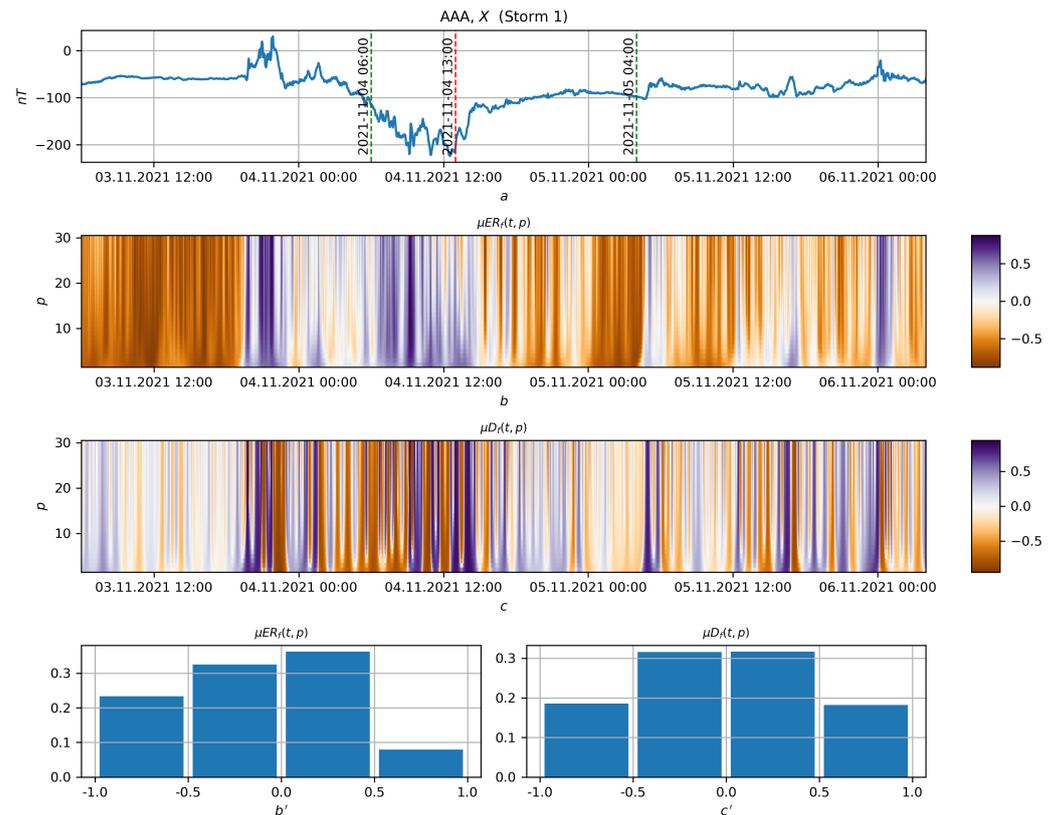


Figure 14. X component recorded at AAA magnetic observatory during Storm 1 (a); the activity measure plots (b–d) and their corresponding histograms (b'–d').

The first storm (November 4–5, 2021, Figures 14, 15) was the one of the first 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 [OMNI, 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 B_z 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 (Figure 13). During November 4, the planetary K-index 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.

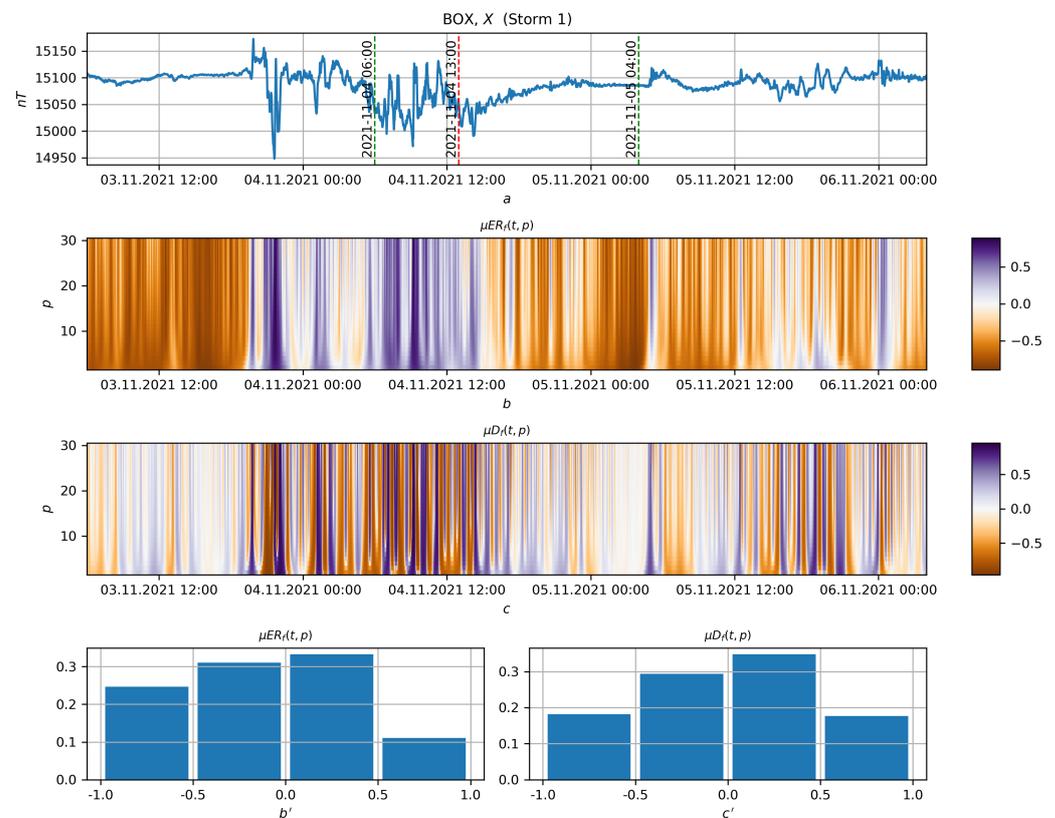


Figure 15. X component recorded at BOX magnetic observatory during Storm 1 (a); the activity measure plots (b–d) and their corresponding histograms (b'–d').

The activity measures for this storm for both BOX and AAA observatories (Figures 16–17) were built not only for the above mentioned storm phases, but also for some time period prior to the MP onset and for some period after the storm recovery. Their plots display some general similarities. However, the particular storm periods are reflected in a different way. The $\mu ER_f(t, p)$ measure plot clearly displays more high-magnitude 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 as fragments within the values are strongly positive. On the contrary, the calm periods of the X component correspond mostly to negative μER_f values (an example is the storm recovery phase fragment). The next measure, μDR_f , related to derivative and therefore to an overall signal variability, emphasizes smaller oscillations of the initial signal; however, relatively large-scale abrupt fragments are reflected in a way close to the μER_f result. Certainly, the μDR_f values appear to be lower for less variable fragments. This can be a tool to highlight the structure of oscillation sequences related to the storm phases. It's important to note that the initial X component magnetic records for AAA and BOX observatories are quite different in their range: the AAA data for the MP ending period has a minimum of more than 200 nT compared to the conditions before the storm (Figure 14a), whereas the corresponding BOX data (Figure 15a) decrease is only about 100 nT. Moreover, the BOX data has disturbances higher in their amplitude

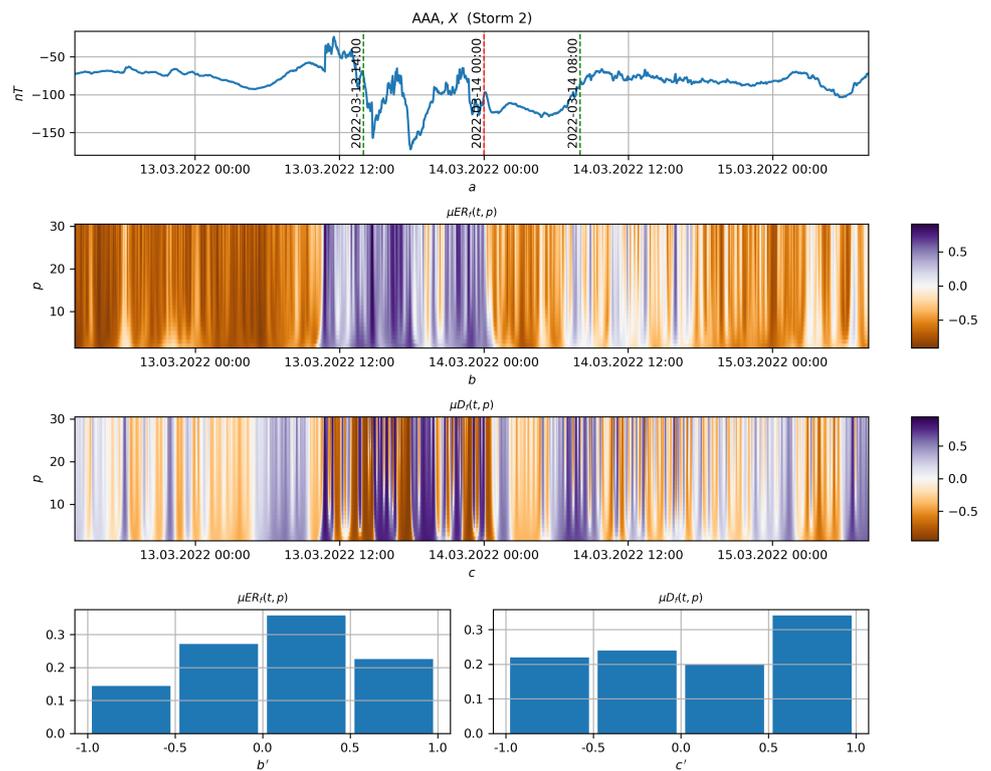


Figure 16. X component recorded at AAA magnetic observatory during Storm 2 (a); the activity measure plots (b–c) and their corresponding histograms (b’–c’).

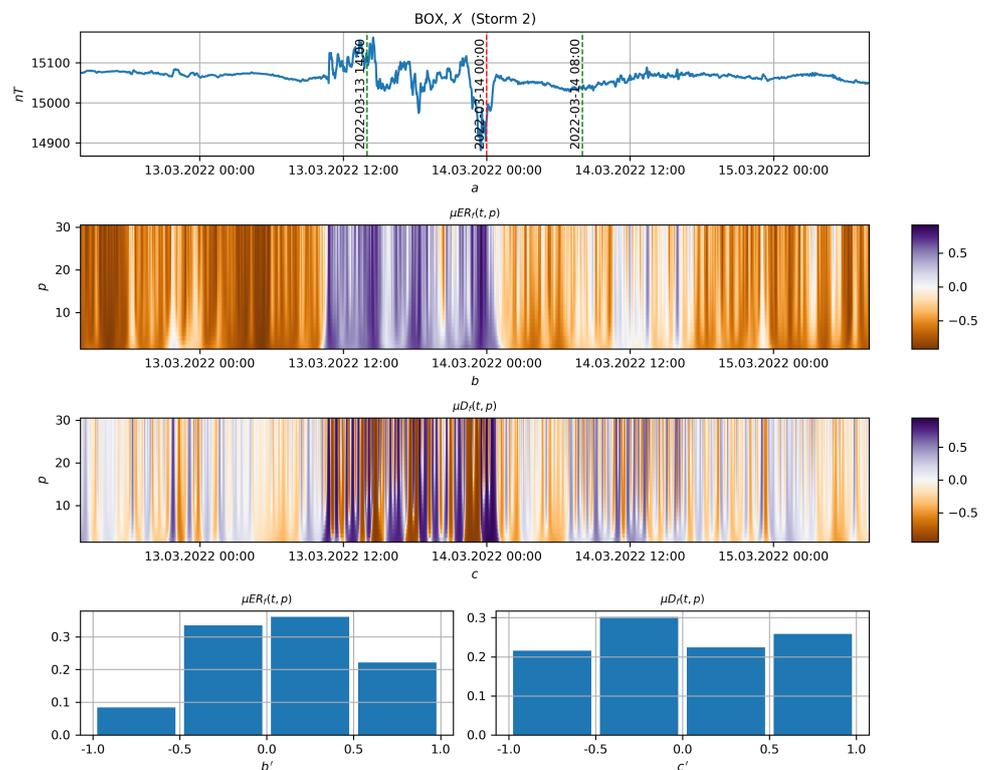


Figure 17. X component recorded at BOX magnetic observatory during Storm 2 (a); the activity measure plots (b–c) and their corresponding histograms (b’–c’).

than the corresponding ones in the AAA data. The reason for both is in the geographical location of these observatories (see Table 2) and, in particular, in the latitudinal distance of these observatories from the equatorial ring current and the auroral zone: the Dst current system contributes more to the overall dynamics of geomagnetic variations at the AAA observatory, whereas the auroral disturbances, including the ones occurring during the recovery phase of the storm, have a stronger impact on the BOX variation data. The next storm that occurred in March 2022 (Figures 16–17) was quite similar to the previous one in the initial interplanetary conditions 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 B_z 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: both μER_f plots show large positive values related to the sudden commencement moment and the following fragment related to intense alternations of B_z and solar wind characteristics during the storm phase that influenced the geomagnetic field registered on the Earth's surface. Both μDR_f and μER_f show the oscillations possibly related to auroral disturbances, however, μDR_f reflects more small-scale details for this phase.

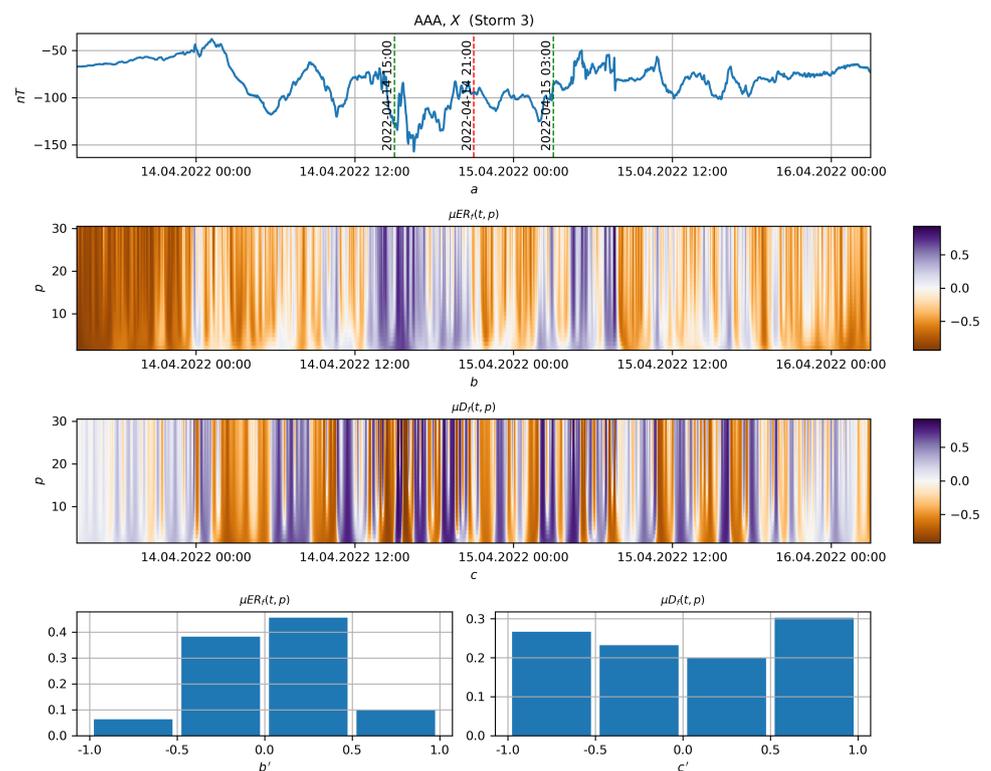


Figure 18. X component recorded at AAA magnetic observatory during Storm 3 (a); the activity measure plots (b–c) and their corresponding histograms (b'–c').

On the eve of the third storm that occurred in April 2022 (Figures 18–19), the IMF B_z 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. Like in the previous cases, the closeness of BOX geomagnetic (as well as geographic) latitude to the auroral oval results in multiple

oscillations caused by related disturbances due to the auroral current system behaviour. Unlike the BOX data, the AAA X data appears to have more long-period disturbances, which are a result of several generations of a storm. As seen from Figures 18–19 (b, c), the behavior of each measure is similar to its behavior for previous storms; nevertheless, the μER_f again indicates both large- and small-scale morphological features of the storm more clearly than the other indicator.

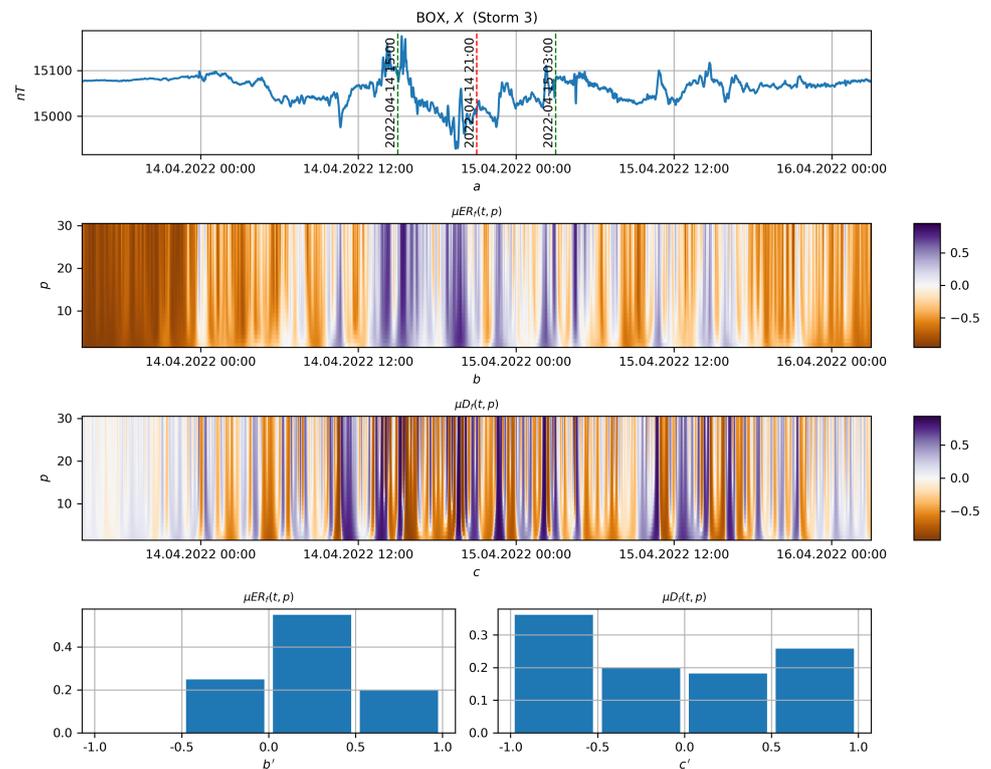


Figure 19. X component recorded at BOX magnetic observatory during Storm 3 (a); the activity measure plots (b–c) and their corresponding histograms (b'–c').

The histograms for μER_f , representing 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 predicting the expected levels of disturbances to a certain extent. Notably, this strong correlation is seen for histograms related to different storms. This suggests a reliable connection of the indicator with the physical processes of interactions between the magnetosphere and the solar wind during a geomagnetic storm.

The indicators also provide an opportunity for spectral decomposition of geomagnetic variations, as the p increase results in adding more high-frequency details to the measure plots.

10. Discussion and Conclusion

In this work, to record $f = f(t)$ and the local property P relating to it, a fuzzy measure (spectrum) $\mu P_f = \mu P_f(t, s)$ of the manifestation of P on f in time t and scale of consideration s is constructed. Thus, the one-dimensional record $f(t)$ is associated with the two-dimensional function $\mu P_f(t, s)$. This redundancy makes it possible to better see what is happening at f and understand the nature of what is behind it. The authors suggest

further improvement of the transition $f \rightarrow \mu P_f$. But the main thing for them today is the similarity of the spectrum of properties and the wavelet spectrum. More precisely, when the property of P is the correlation of the record f with one or another wavelet ψ : $P = \text{cor}_\psi$, then the measure constructed for it is similar to the wavelet spectrum, but more contrasting (Figures 20–21). So in this sense, the property measure generalizes wavelets.

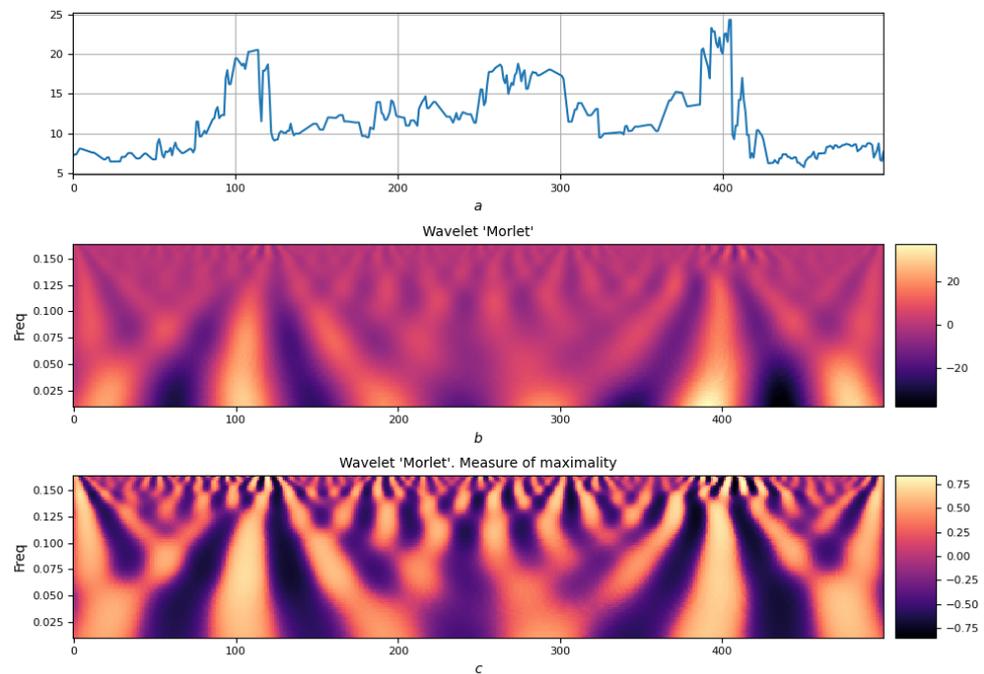


Figure 20. *a* – synthetic record (units in both axes are dimensionless); *b* – wavelet spectrum (Morlet); *c* – maximality measure.

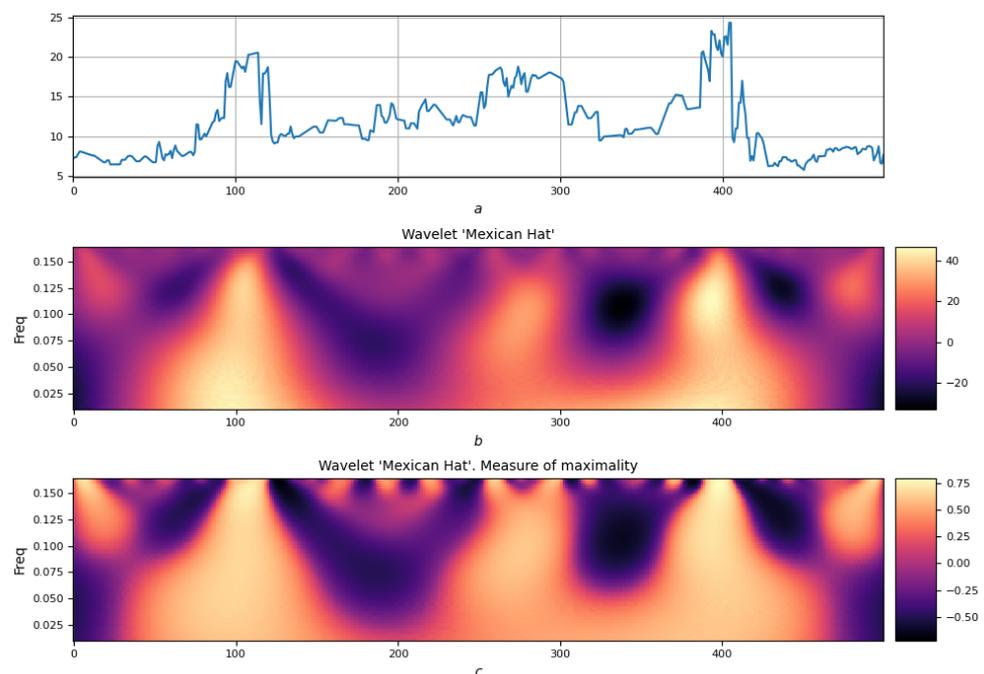


Figure 21. *a* – synthetic record (units in both axes are dimensionless); *b* – wavelet spectrum (Mexican Hat); *c* – maximality measure.

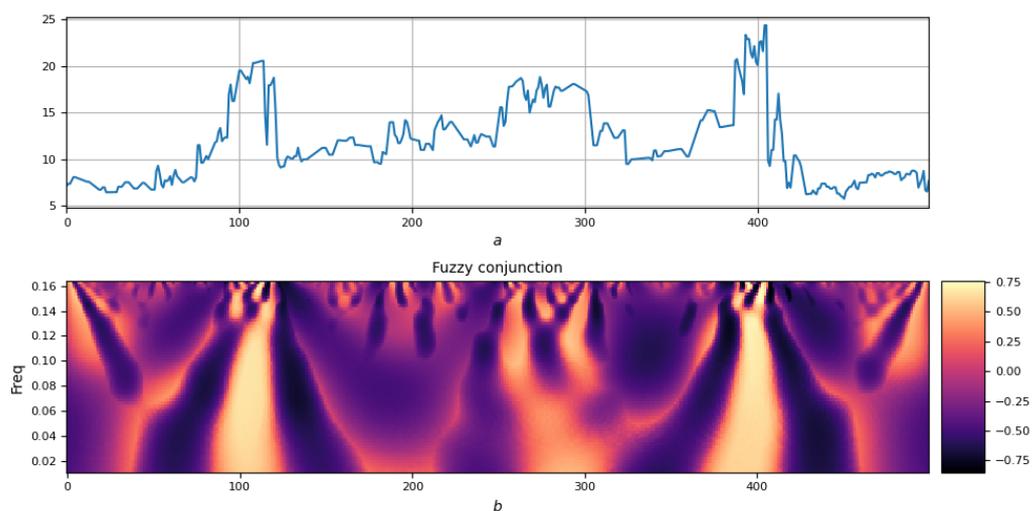


Figure 22. *a* – synthetic record (units in both axes are dimensionless); *b* – connection of measures (Figures 20c and 21c): fuzzy conjunction.

There is one more important circumstance: property measures, being fuzzy structures, can be combined using fuzzy logic operations into new measures (spectra). In projection onto wavelets, this is a completely new thing for them: the wavelet spectrum obtained by such a combination of two traditional wavelet spectra does not have a basis wavelet. Figure 22 shows the conjunction of the wavelet spectra from Figures 20 and 21.

The authors are interested in the possibilities of fuzzy logic in wavelet theory.

Acknowledgments. 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 (grant number 075-00443-24-00). 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 employed facilities and data provided by the Shared Research Facility “Analytical Geomagnetic Data Center” of the Geophysical Center of RAS (<http://ckp.gcras.ru/>).

References

- Agayan, S., S. Bogoutdinov, A. Soloviev, and R. Sidorov (2016), The Study of Time Series Using the DMA Methods and Geophysical Applications, *Data Science Journal*, 15, <https://doi.org/10.5334/dsj-2016-016>.
- Agayan, S., S. Bogoutdinov, D. Kamaev, V. Kaftan, M. Osipov, and V. Tatarinov (2021a), Theoretical Framework for Determination of Linear Structures in Multidimensional Geodynamic Data Arrays, *Applied Sciences*, 11(24), 11,606, <https://doi.org/10.3390/app112411606>.
- Agayan, S., S. Bogoutdinov, R. Krasnoperov, and R. Sidorov (2021b), A Multiscale Approach to Geomagnetic Storm Morphology Analysis Based on DMA Activity Measures, *Applied Sciences*, 11(24), 12,120, <https://doi.org/10.3390/app112412120>.
- Agayan, S. M., S. R. Bogoutdinov, and R. I. Krasnoperov (2018), Short introduction into DMA, *Russian Journal of Earth Sciences*, 18(2), 1–10, <https://doi.org/10.2205/2018es000618>.
- Agayan, S. M., A. A. Soloviev, S. R. Bogoutdinov, and Y. I. Nikolova (2019), Regression Derivatives and Their Application to the Study of Geomagnetic Jerks, *Geomagnetism and Aeronomy*, 59(3), 359–367, <https://doi.org/10.1134/s0016793219030022>.

- Agayan, S. M., V. N. Tatarinov, A. D. Gvishiani, S. R. Bogoutdinov, and I. O. Belov (2020), FDPS algorithm in stability assessment of the Earth's crust structural tectonic blocks, *Russian Journal of Earth Sciences*, 20(6), 1–14, <https://doi.org/10.2205/2020es000752>.
- Agayan, S. M., S. R. Bogoutdinov, B. A. Dzeboev, B. V. Dzeranov, D. A. Kamaev, and M. O. Osipov (2022), DPS clustering: New results, *Applied Sciences*, 12(18), 9335, <https://doi.org/10.3390/app12189335>.
- Akasofu, S.-I., and S. Chapman (1963), The development of the main phase of magnetic storms, *Journal of Geophysical Research*, 68(1), 125–129, <https://doi.org/10.1029/jz068i001p00125>.
- Batyrshin, I. Z., A. O. Nedosekin, and A. A. Stetsko (2007), *Fuzzy hybrid systems. Theory and practice*, 207 pp., Fizmatlit.
- Borojev, R. N., and M. S. Vasiliev (2017), Relationship between Indexes of Geomagnetic Activity on the Main Phase of a Magnetic Storm for Various Types of Solar Wind, *Science and Education*, 1, 67–70, in Russian.
- Gromova, L. I., N. G. Kleimenova, A. E. Levitin, S. V. Gromov, L. A. Dremukhina, and N. R. Zelinskii (2016), Daytime geomagnetic disturbances at high latitudes during a strong magnetic storm of June 21–23, 2015: The storm initial phase, *Geomagnetism and Aeronomy*, 56(3), 281–292, <https://doi.org/10.1134/s0016793216030051>.
- Gvishiani, A., A. Soloviev, R. Krasnoperov, and R. Lukianova (2016a), Automated Hardware and Software System for Monitoring the Earth's Magnetic Environment, *Data Science Journal*, 15, <https://doi.org/10.5334/dsj-2016-018>.
- Gvishiani, A. D., R. V. Sidorov, R. Y. Lukianova, and A. A. Soloviev (2016b), Geomagnetic activity during St. Patrick's Day storm inferred from global and local indicators, *Russian Journal of Earth Sciences*, 16(6), 1–8, <https://doi.org/10.2205/2016es000593>.
- Intermagnet (2023), International Real-time Magnetic Observatory Network, <https://intermagnet.org>, accessed on 1 November 2023.
- ISGI (2023), International service of geomagnetic indices, <https://isgi.unistra.fr>, accessed on 1 November 2023.
- Kacprzyk, J., A. Wilbik, and S. Zadrozny (2007), Linguistic Summarization of Time Series by Using the Choquet Integral, in *Lecture Notes in Computer Science*, pp. 284–294, Springer Berlin Heidelberg, https://doi.org/10.1007/978-3-540-72950-1_29.
- Kovalev, S. M. (2007), Hybrid fuzzy-temporal models of time series in problems of analysis and identification of weakly formalized processes, in *Integrated Models and Soft Computing in Artificial Intelligence. Proceedings of the IVth International Scientific and Practical Conference (Kolomna, May 28-30, 2007)*, Moscow: Fizmatlit., vol. 1, pp. 26–41.
- Lazutin, L. L. (2012), *Global and polar magnetic storms*, 214 pp., Moscow State University, in Russian.
- Mishin, V. M., M. Foerster, T. I. Saifudinova, A. D. Bazarzhapov, Y. A. Karavaev, L. A. Sapronova, and S. I. Solovyev (2007), Spontaneous substorms and ordered type of magnetospheric disturbances during the superstorm of November 20, 2003, *Geomagnetism and Aeronomy*, 47(4), 429–441, <https://doi.org/10.1134/s0016793207040032>.
- NOAA (2023), National Oceanic and Atmospheric Administration, <https://www.noaa.gov>, accessed on 1 November 2023.
- OMNI (2023), OMNIWeb Plus, <https://omniweb.gsfc.nasa.gov>, accessed on 01 November 2023.
- Oshchenko, A. A., R. V. Sidorov, A. A. Soloviev, and E. N. Solovieva (2020), Overview of anomaly measure application for estimating geomagnetic activity, *Geophysical research*, 21(4), 51–69, <https://doi.org/10.21455/gr2020.4-4>.
- Pandey, S. K., and S. C. Dubey (2009), Characteristic features of large geomagnetic storms observed during solar cycle 23, *Indian Journal of Radio & Space Physics*, 38(6), 305–312.
- Pedrycz, W., and M. Smith (1999), Granular correlation analysis in data mining, in *FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems. Conference Proceedings (Cat. No.99CH36315)*, IEEE, <https://doi.org/10.1109/fuzzy.1999.790078>.
- Soloviev, A., S. Agayan, and S. Bogoutdinov (2017), Estimation of geomagnetic activity using measure of anomalousness, *Annals of Geophysics*, 59(6), <https://doi.org/10.4401/ag-7116>.

- Soloviev, A., A. Smirnov, A. Gvishiani, J. Karapetyan, and A. Simonyan (2019), Quantification of Sq parameters in 2008 based on geomagnetic observatory data, *Advances in Space Research*, 64(11), 2305–2320, <https://doi.org/10.1016/j.asr.2019.08.038>.
- Tanaka, H., S. Uejima, and K. Asai (1982), Linear Regression Analysis with Fuzzy Model, *IEEE Transactions on Systems, Man, and Cybernetics*, 12(6), 903–907, <https://doi.org/10.1109/tsmc.1982.4308925>.
- WDC (2023), World Data Center for Geomagnetism, Kyoto, <https://wdc.kugi.kyoto-u.ac.jp>, accessed on 1 November 2023.
- Yarushkina, N. G. (2004), *Fundamentals of the theory of fuzzy and hybrid systems*, 320 pp., Finansy i statistika.
- Yarushkina, N. G., T. R. Yunusov, and T. V. Afanasyeva (2007), Terminal-server traffic modeling based on fuzzy time series trend analysis, *Program products and systems*, (4), 15–19.
- Yermolaev, Y. I., I. G. Lodkina, N. S. Nikolaeva, and M. Y. Yermolaev (2012), Recovery phase of magnetic storms induced by different interplanetary drivers, *Journal of Geophysical Research: Space Physics*, 117(A8), A08,207, <https://doi.org/10.1029/2012ja017716>.
- Yermolaev, Y. I., I. G. Lodkina, N. S. Nikolaeva, and M. Y. Yermolaev (2014), Influence of the interplanetary driver type on the durations of the main and recovery phases of magnetic storms, *Journal of Geophysical Research: Space Physics*, 119(10), 8126–8136, <https://doi.org/10.1002/2014ja019826>.
- Yokoyama, N., and Y. Kamide (1997), Statistical nature of geomagnetic storms, *Journal of Geophysical Research: Space Physics*, 102(A7), 14,215–14,222, <https://doi.org/10.1029/97ja00903>.
- Zhang, G.-L. (1992), Interplanetary disturbance structures and geomagnetic storm, in *AIP Conference Proceedings*, AIP, <https://doi.org/10.1063/1.41724>.