applied sciences

MDPI

*Article*

# A Multiscale Approach to Geomagnetic Storm Morphology Analysis Based on DMA Activity Measures

**Sergey Agayan [1], Shamil Bogoutdinov [1,2], Roman Krasnoperov [1,*] and Roman Sidorov [1]**

[1] Geophysical Center of the Russian Academy of Sciences, 119296 Moscow, Russia; s.agayan@gcras.ru (S.A.); shm@gcras.ru (S.B.); r.sidorov@gcras.ru (R.S.)
[2] Schmidt Institute of Physics of the Earth of the Russian Academy of Sciences, 123242 Moscow, Russia
* Correspondence: r.krasnoperov@gcras.ru

**Abstract:** The article is focused on the approach based on the discrete mathematical analysis conception and continues a series of studies related to the application of the previously developed methodology to geophysical data analysis. The main idea of the study is the modification of earlier conceptions regarding the interpreter's logic that allows introducing a multiscale approach and performing the time series analysis using the activity measure plots, implying the vertical scale. This approach was used to study the morphology of several intense geomagnetic storms at the final stages of the 23rd and 24th solar activity cycles. Geomagnetic observatory data and interplanetary magnetic field parameters as well as the solar wind flux speed and proton density were analyzed for each of the studied storms using the activity measures. The developed methods, applied to geomagnetic storm morphological analysis, displayed good results in revealing the decreases and increases in various durations and intensities during storms, detecting low-amplitude disturbances, and storm sudden commencement recognition. The results provide an opportunity to analyze any physical data using a unified scale and, in particular, to implement this approach to geomagnetic activity studies.

**Keywords:** time series; dynamic indicator; activity measure; geomagnetic storm; geomagnetic variations; interplanetary magnetic field

check for updates

**Citation:** Agayan, S.; Bogoutdinov, S.; Krasnoperov, R.; Sidorov, R. A Multiscale Approach to Geomagnetic Storm Morphology Analysis Based on DMA Activity Measures. *Appl. Sci.* **2021**, *11*, 12120. https://doi.org/10.3390/app112412120

Academic Editor: Jianbo Gao

Received: 15 November 2021
Accepted: 17 December 2021
Published: 20 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. 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/).

## 1. Introduction

One of the directions of the development of discrete methods for data analysis and discrete mathematics in general is associated with modeling the expert's ability to deal with data. An expert is able to highlight the anomalies in physical data of small dimension, switch from the local level of anomalies to the global one for a holistic interpretation, find the signals of the desired morphology on the data records of a small length, and do so much more efficiently than any formal apparatus, but an expert is powerless in cases of large dimensions and volumes.

Our attempts to overcome this contradiction, i.e., to solve the problem of "algorithmization of an Expert", began 20 years ago.

Solution concept: the advantage of an expert over mathematics is that he/she understands the natural manifestations of fundamental properties of proximity, limit, continuity, connectivity, and trend (forming the basis of data analysis) in a more natural and stable way of perception.

Solution technique: the fact that an expert thinks and operates not with numbers, but with fuzzy concepts was immediately taken into account. Therefore, the basis of our approach was formed by fuzzy sets, some of which were models of discrete analogs of the mathematical properties mentioned above as well as fuzzy logic that allows you to combine fuzzy models into data analysis algorithms, in particular, according to the scenarios of classical mathematics.

Solution result: the "discrete mathematical analysis" (DMA)—a new approach to data analysis—is expert-oriented and an intermediate between "hard" mathematical methods and "soft" combinatorial ones [1,2].

The solution to the problem within the DMA framework consists of two parts. The first part is informal. In it, the logic of the expert is analyzed, the necessary concepts are introduced, and the scheme and principles of the solution are explained. The second part has a formal character: with the help of the DMA apparatus, all concepts receive strict definitions in the framework of fuzzy mathematics and fuzzy logic, and the scheme and principles become algorithms.

This is how the DMA-based studies of data records began, the informal basis of which was the logic of an expert looking for anomalies on data records [3]. Let us remind the reader: the expert scans the record, evaluating the activity of small (in his/her opinion) fragments of it, similar to each other, with positive numbers. Thus, from the initial data, the expert proceeds to a non-negative function, which is called the straightening of the record since the more active points of the initial data record correspond to the larger values of this function. Further, the expert's search for an anomaly in the data is reduced to the search for extreme values on its straightening. Thus, the expert operates at two levels: local (straightening of a data record), and global (search for extreme values on straightening).

The formalization of this logic became the main one in the DMA-based data studies at the first stage. The algorithms DRAS, FLARS, and FCARS [4–8] created at that time competed quite adequately with the classical spectral analysis algorithms in searching for anomalies, and a number of effective geophysical applications, implementing the new algorithms, was discovered [1,3,9–11]. At the local stage of constructing straightenings, they acted similarly to each other, and at the global stage of searching for maximums on straightening, they acted in different ways.

The algorithms allowed the experts to freely understand and select a straightening as a quantitative expression of the property of interest. Nevertheless, reality has shown the stability of the following choice: a set of basic straightenings is formed that most experts would like to deal with. There is a fundamental mathematical concept behind each of them, which confirmed the correctness of our guidelines when creating the DMA. Let us name the most important of them today: energy $E$ (dispersion and continuity), spread $O$ (Cauchy fundamentality), jaggedness $L$ (frequency and length), and increase rate $G$ (modulus of the derivative). Their set is open for replenishment. This is determined by practice and gives great flexibility to the analysis of data records using DMA.

The second stage of the DMA-based studies of data records is focused mainly on geomagnetism (in particular, search for the patterns of known waveform in INTERMAGNET observatory data [12]). A meaningful analysis of the geomagnetic data records presupposes taking into account a whole complex of parameters and cannot be reduced to a quantitative approach using a single straightening. It requires a set of straightenings and their integration into a single point of view. This cannot be done mechanically due to the fact that the straightenings of different natures are initially incomparable.

The solution found was to proceed from a record straightening to its activity measure. The activity measure shows the level of the expert's interest in the data record behavior in one or another point, based on the selected straightening [13] within a fuzzy scale ([0, 1] or $[-1,1]$).

This two-step transition "data record $\rightarrow$ straightening $\rightarrow$ activity measure" completely translates the DMA-based studies of data records into the language of fuzzy sets, which immediately opens up new opportunities for these studies. Let us discuss two of them.

The first one: it becomes possible to formalize the expert's complex point of view on the record. The "straightening $\rightarrow$ activity measure" transition translates the data analysis into the fuzzy mathematics language: the measures of activity for different straightenings take values in a single fuzzy scale ([0, 1] or $[-1, 1]$) and can be combined in any compositions and any quantities with the help of numerous operations of fuzzy logic and all kinds of

averaging, and, therefore, are able to model the activity measure for a complicated view of an expert at the record. All of this led to the new geomagnetic activity indicators [13,14].

The second one: measures of activity of various records and, moreover, of a completely different nature, express the same essence, namely, the expert's interest in them. Therefore, if the records cannot be directly compared, then their activity measures can be. This circumstance served as the basis for DMA-based monitoring of geophysical processes, which is understood as a system of geophysical data records [15,16].

Thus, at the second stage, the focus on DMA-based data records studies shifted to the analysis of their activity measures, which, together with parallel achievements in DMA clustering (DPS algorithm), made it possible to largely solve the problem of recognizing anomalies [13].

Summarizing the results of the second stage of the DMA-based studies of records, we should note that DMA was the one of the origins of the functions of the MAGNUS system designed for geomagnetic information acquisition, processing, storage and analysis [17,18].

The third stage: DMA-based studies of data records in the first two stages represent a sequential improvement in the formalization of the expert's logic. The main thing here is the expert with his/her view of the record, including both the property of his/her interest and the scale of its consideration on the record. Therefore, the modeling taking into account this circumstance and its results is necessarily limited.

At the third stage, we remove this restriction by studying the relationship between a property and a record in different scales. The idea of different scales, which came from wavelets and fractals, makes the property, and not the expert, the main thing in the analysis of the recording. This makes it possible to understand the connection between them more fully and objectively.

The first step in this direction is the construction of a multiscale activity measure for a given property of an analyzed record. This is done in the theoretical section of the article. Its practical part (results and discussion) is focused on the geomagnetic storm analysis using such an activity measure for the energy and jaggedness properties. The main opportunities in the application of the method to geomagnetic activity studies are highlighted.

## 2. Materials and Methods

### 2.1. Initial Data and Assumptions

Let us assume that a observation period $T$ of recordings $f$ is a finite regular set of nodes with a sampling parameter $h$:

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

and let us call the record $f$ a function on $T$: $f : T \to \mathbb{R}$ and denote by $\mathcal{F}(T)$ the space of records on $T$.

Let us assume that a localization of space $T$ in a node $t$ is a fuzzy structure $\delta_t$ on $T$, normalized in $t$ and decreasing with increasing distance from $t$:

$$(\delta_t(t) = 1) \wedge (|\bar{\bar{t}} - t| > |\bar{t} - t|) \Rightarrow \delta_t(\bar{\bar{t}}) \leq \delta_t(\bar{t}) \tag{1}$$

Multiple scales in this study are modeled with a group of localizations $\delta_t(r, p)$, depending on the radius $r$ and the scale of localization $p$ (Figure 1)

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

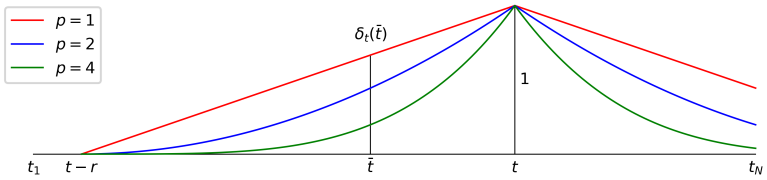



**Figure 1.** $\delta_t(\bar{t})$ for various $p$ values at a fixed radius $r$.

### 2.2. Straightening

**Definition 1.**

1. *A straightening construction $D$ is a non-negative functional on $T$, parameterized by $T$:*

$$D : \mathfrak{F}(T) \times T \to \mathbb{R}^+$$

2. *A straightening of a record $f$ based on the construction $D$ is non-negative function $D_f$: $t \to D(f,t)$.*

The value $D(f,t)$ is denoted as $D_f(t)$ and informally understood as a quantitative estimate of behavior of a function $f$ in a node $t \in T$ within a view $D$ on its dynamics; therefore, the construction of the straightening $D$ necessarily connects with the scale of view on $f$ in $t$, that is, with the closeness measure $\delta_t$, so everywhere further

$$D_f(t) = D(f,t) = D(f,t|\delta_t). \tag{3}$$

In DMA, there is a set of basic straightenings, open for adjunction. Let us give examples for some of them, which are used for magnetic storm analysis.

Energy (dispersion)

$$E_f(t|\delta_t) = \frac{\sum_{\bar{t}\in T} |f(\bar{t}) - M_f(t|\delta_t)|\delta_t(\bar{t})}{\sum_{\bar{t}\in T} \delta_t(\bar{t})} \tag{4}$$

where

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

Length (jaggedness)

$$L_f(t|\delta_t) = \frac{\sum_{\bar{t},\bar{\bar{t}}\in T:|\bar{t}-\bar{\bar{t}}|=h} |f(\bar{t}) - f(\bar{\bar{t}})|\delta_t(\bar{t})\delta_t(\bar{\bar{t}})}{\sum_{\bar{t},\bar{\bar{t}}\in T:|\bar{t}-\bar{\bar{t}}|=h} \delta_t(\bar{t})\delta_t(\bar{\bar{t}})}. \tag{5}$$

Straightenings are associated with very natural and important but simpler points of vision (properties) of a record. A complex property is made up of simple ones, and addition cannot be direct (mechanical). The point is that initially, the straightening $D_f(t)$ cannot be compared with each other at different $D$. Their association requires additional efforts that imply constructing measures for them.

### 2.3. Simple Activity Measure

An activity measure $\mu D_f(t)$ of a straightening $D_f(t)$ (3) is a function of membership in a fuzzy concept "function $f$ activity in a node $t$ according to the indicator $D$" on $T$. It is obtained from the indicator $D_f(t)$ using one or another maximality measure construction mes max. Several such constructions are presented in DMA. In this study, a construction mes max, based on fuzzy comparisons [6], is used.

**Definition 2.** *Fuzzy comparison $n(a,b)$ for non-negative values a and b determines the measure of superiority of b over a with a scale within an interval* $[-1,1]$

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

Further, this is a construction

$$n(a,b) = \frac{b-a}{b+a}. \tag{6}$$

Let $A \subset \mathbb{R}^+$ be the ultimate set of non-negative numbers, $b \in \mathbb{R}^+$. The maximality measure $\mathrm{mes}\max_A(b)$ is a fuzzy structure on $\mathbb{R}^+$, responding to the question: "to what extent is the number $b$ larger than the set $A$ in modulo?"

The maximality measure expands the concept of fuzzy comparison between numbers to a comparison between a set and a number:

$$\mathrm{mes}\max_A(b) = \mathrm{mes}(A < b) = n(A,b) \in [-1,1].$$

In DMA, there are several ways to expand the comparison $n(a,b)$ to the comparison $n(A,b)$. In this study, this is the generalized gravitational approach:

$$n(A,b) = n(M_q(A),b) \tag{7}$$

where

$$M_q(A) = \left(\frac{\sum_{a\in A} n(a,b)}{|A|}\right)^{1/q}$$

is the quasi-arithmetic (Kolmogorov) mean of $A$ with a positive exponent $q$. The usual gravitational continuation is obtained at $q = 1$.

For the measure $\mu D_f(t)$, the role $A$ is played by the image $\mathrm{Im}\, D_f = \{D_f(\bar{t}) : \bar{t} \in T\}$, and the role $b$ is played by the value $D_f(t)$:

$$\mu D_f(t) = \mathrm{mes}\max_{\mathrm{Im}\, D_f} D_f(t) = n(\mathrm{Im}\, D_f, D_f(t)). \tag{8}$$

The record properties that can be expressed using the straightenings are considered simple, and the activity measures were constructed for them.

### *2.4. Complex Activity Measures*

The transition $D_f \to \mu D_f$ (8) translates the analysis of a function $f$ on a language of fuzzy logic and fuzzy mathematics: the activity measures $\mu D_f$ for different straightenings $D$ take on values in a unified scale of the interval $[-1,1]$ and can be combined in any compositions and any quantities using numerous FL operations and every kind of averaging that we denote as $*$.

It becomes possible to give the meaning of complex activity $f$ for the aggregate of straightenings $\mathfrak{D}$ in a node $t \in T$:

$$\mu\mathfrak{D}_f(t) = *_{D\in\mathfrak{D}}(\mu D_f(t)).$$

In general, by this new construction, a complex property $\varepsilon$ of the $f$ dynamics is modeled in node $t$. In many ways, such modeling is art and consists in the selection of basic straightenings $\mathfrak{D}(\varepsilon)$ and their proper connection $*(\varepsilon)$:

$$\mu\varepsilon_f(t) = *(\varepsilon)_{D\in\mathfrak{D}(\varepsilon)}(\mu D_f(t)). \tag{9}$$

Let us consider three important cases of these compounds.

A junction $*$ using fuzzy conjunction $\top$ makes it possible to simultaneously take into account all the basic dynamics from $\mathfrak{D}(\varepsilon)$:

$$\mu\varepsilon_f(t) = \top_{D\in\mathfrak{D}(\varepsilon)}(\mu D_f(t)).$$

A junction $*$ using fuzzy disjunction $\perp$ makes it possible for every basic dynamics from $\mathfrak{D}(\varepsilon)$ to show itself:

$$\mu\varepsilon_f(t) = \perp_{D\in\mathfrak{D}(\varepsilon)}(\mu D_f(t)).$$

A junction $*$ using weighted average makes it possible to take the weights $\omega(D)$ of the basic dynamics $D$ from $\mathfrak{D}(\varepsilon)$ into account:

$$\mu\varepsilon_f(t) = \frac{\sum_{D\in\mathfrak{D}(\varepsilon)}\omega(D)\mu D_f(t)}{\sum_{D\in\mathfrak{D}(\varepsilon)}\omega(D)}.$$

### *2.5. Multiscale Activity Measures and Their Analysis*

By joining the anomalousness measure $\mu\varepsilon_f(t)$ with a parametric group (9) of multiscale localizations $\delta_t(r,p)$ (2), we obtain the activity measure $\mu\varepsilon_f$ in the interval of scales $P$:

$$\mu\varepsilon_f(t,p) = \mu\varepsilon_f(t|\delta_t(r,p)). \tag{10}$$

The function $\mu\varepsilon_f$ is defined on a direct product $T\times P$, and takes on values within an interval $[-1,1]$, representing in each point $(t,p)$ the measure of fulfillment of a property $\varepsilon$ for a time series $f$ at a point $t$ at its vision scale $p$.

Therefore, $\mu\varepsilon_f$ is a multiscale measure of activity of a property $\varepsilon$ on a record $f$. In this study, the simplest and most natural $\mu\varepsilon_f$ analysis is proposed, necessary to understand the $f$ dynamics on $T$ and to compare it with the dynamics of other records.

Let us denote $\mu\varepsilon_{f,p}$ and $\mu\varepsilon_{f,t}$ the coordinate dependencies $t\to\mu\varepsilon_f(t,p)$ and $p\to\mu\varepsilon_f(t,p)$. The first three characteristics represent a general averaging (a value) $\overline{\mu\varepsilon_f}$ of the measure $\mu\varepsilon_f$ on $T\times P$ and the functional dependencies $(\overline{\mu\varepsilon_f})_p$: $p\to\overline{\mu\varepsilon_f(t,p)}$ and $(\overline{\mu\varepsilon_f})_t$: $t\to\overline{\mu\varepsilon_f(t,p)}$, averaging the measure $\mu\varepsilon_f$ by $t$ and $p$ on the layers $(T,p)$ and $(t,P)$, respectively.

The second group of five characteristics is related to the partition $H$ of an interval $[-1,1]$ in four parts:

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

The values of the measure $\mu\varepsilon_f(t,p)$ within these intervals mean, respectively, from left to right, very weak, weak, moderate, and strong occurrence of the property $\varepsilon$ on a time series $f$ in node $t$ at scale $p$.

The overall depiction $T\times P$ of such quantitative understanding of the situation in $(p,t)$ is provided by a histogram $H(\mu\varepsilon_f)$, built along partition $H$ for a measure $\mu\varepsilon_f$. Similarly, the histograms $H(\mu\varepsilon_{f,p})$ and $H(\mu\varepsilon_{f,t})$ do the same on the layers $(T,p)$ and $(t,P)$. The histogram $H((\overline{\mu\varepsilon_f})_p)$ gives a qualitative $H$-understanding of the distribution of horizontal $t$-averagings $\overline{\mu\varepsilon_{f,p}}$ along $p$, and the histogram $H((\overline{\mu\varepsilon_f})_t)$ does the same in a dual way for $t$ for vertical $p$-averagings $\overline{\mu\varepsilon_{f,p}}$.

## 3. Results and Discussion

The practical part is focused on the geomagnetic storm analysis using the activity measure (10) for the "energy" and "length" properties. We recall the equations for them explicitly, as they were mentioned in the theoretical part.

Activity measure "Energy"

$$\mu E_f(t,p) = \frac{E_f(t) - M_q(\operatorname{Im} E_f)}{E_f(t) + M_q(\operatorname{Im} E_f)},$$

where

$$M_q(\mathrm{Im}\, E_f) = \left( \frac{\sum_{\bar{t}\in T} E_f(\bar{t})^q}{|T|} \right)^{1/q}, \quad q > 0;$$
$$E_f(t|\delta_t) = \frac{\sum_{\bar{t}\in T} |f(\bar{t}) - M_q(t|\delta_t)|\delta_t(\bar{t})}{\sum_{\bar{t}\in T} \delta_t(\bar{t})};$$
$$M_q(t|\delta_t) = \frac{\sum_{\bar{t}\in T} f(\bar{t})\delta_t(\bar{t})}{\sum_{\bar{t}\in T} \delta_t(\bar{t})};$$
$$\delta_t(\bar{t}) = \delta_t(r,p)(\bar{t}) = \begin{cases} \left(1 - \frac{|\bar{t}-t|}{r}\right)^p, & \text{if } |\bar{t}-t| \le r \\ 0, & \text{if } |\bar{t}-t| > r \end{cases}.$$

Activity measure "Length"

$$\mu L_f(t,p) = \frac{L_f(t) - M_q(\mathrm{Im}\, L_f)}{L_f(t) + M_q(\mathrm{Im}\, L_f)},$$

where

$$M_q(\mathrm{Im}\, L_f) = \left( \frac{\sum_{\bar{t}\in T} L_f(\bar{t})^q}{|T|} \right)^{1/q}, \quad q > 0;$$
$$L_f(t|\delta_t) = \frac{\sum_{\bar{t},\bar{\bar{t}}\in T:|\bar{t}-\bar{\bar{t}}|=h} |f(\bar{t}) - f(\bar{\bar{t}})|\delta_t(\bar{t})\delta_t(\bar{\bar{t}})}{\sum_{\bar{t},\bar{\bar{t}}\in T:|\bar{t}-\bar{\bar{t}}|=h} \delta_t(\bar{t})\delta_t(\bar{\bar{t}})};$$
$$\delta_t(\bar{t}) = \delta_t(r,p)(\bar{t}) = \begin{cases} \left(1 - \frac{|\bar{t}-t|}{r}\right)^p, & \text{if } |\bar{t}-t| \le r \\ 0, & \text{if } |\bar{t}-t| > r \end{cases}.$$

### *3.1. Initial Geomagnetic and Interplanetary Data*

To display the possibilities of the method in the analysis of geomagnetic storm features, we applied it to several strong storms that took place at the ends of the 23rd and 24th solar cycles. It is known that during the final period of an 11-year solar cycle, the solar flare activity often increases. For instance, in the 23rd cycle [19], the most flare-active periods (leading to geoeffective coronal mass ejections) took place in 2003–2005, and in the 24th cycle in 2015–2020 (after the solar maximum in 2014).

We selected the Honolulu (HON) geomagnetic observatory data for the tests. Due to its latitudinal position being close to equatorial, during geomagnetic storms, the equatorial ring current contributes fully to the geomagnetic data recorded by this observatory. Therefore, in these data, the global geomagnetic activity is fully represented, and the impact of local geomagnetic disturbances occurring in higher latitudes and driven by ionospheric currents is almost absent. The data from the HON observatory were used for *Dst* geomagnetic index calculation [20]. Horizontal components of the magnetic field of the Earth are usually under the most powerful influence of a geomagnetic storm, and their records clearly display the storm evolution and its phases. Thus, to display the storm dynamics using activity measures, we selected the northward ($X$) component of the total geomagnetic field from the definitive geomagnetic data [12] of the HON observatory.

Initially about 10 storms were studied. However, in order not to overload the article with excessive material and charts, here we focus on 4 strong storms, which we consider the most representative and suitable to highlight the possibilities of the method in the analysis of the features of storms during their different phases. The analyzed storms are listed in Table 1. These storms are not the strongest in the mentioned cycles, as we did not analyze some of the most extreme and well-studied events, such as the storm on 20–24 November 2003, or the St. Patrick's Day storm on 17–18 March 2015 [21] that appeared to be the most intense geomagnetic storm in the 24th cycle (*Dst* index $\le -230$ nT).

Prior to three of these eight storms, storm sudden commencements (SSC) occurred, which can be clearly seen on the magnetic variation data. Storm sudden commencements occur due to the magnetohydrodynamic shockwave impact on the Earth's magnetosphere prior to geomagnetic storm onset. The SSC signal can often be clearly identified on $X$ or $H$ magnetic component data. However, sometimes it is not so possible to recognize it. Therefore, as an additional reference, we used the SSC database from the International Service on Rapid Magnetic Variation [22].

**Table 1.** Main features of the geomagnetic storms selected for analysis.

| Storm Number | Solar Cycle | Start Date and Time (UTC) | End Date and Time (UTC) | min Dst (nT) | SSC Time (UTC) |
|---|---|---|---|---|---|
| 1 | 23 | 2005-05-15 06:00 | 2005-05-19 06:00 | −263 | 02:38 |
| 2 | 23 | 2006-12-14 23:00 | 2006-12-16 17:00 | −146 | 14:14 |
| 3 | 24 | 2015-06-22 19:00 | 2015-06-29 01:00 | −204 | 18:33 |
| 4 | 24 | 2018-08-25 19:00 | 2018-08-28 21:00 | −174 | No SSC |

The method was applied first to the data describing the space weather conditions during storms—the interplanetary magnetic field (IMF) $B_z$ and $B_y$ components, the solar wind speed and density, extracted from the NASA/GSFC OMNI data set through OMNIWeb [23] and 1 min averaged. Changes in the direction of $B_z$ IMF component are a key factor in the geomagnetic storm onset and generation [24]. The solar wind speed data were registered by the ACE spacecraft. Solar wind speed and proton density are the characteristics displaying the energy of the coronal mass ejection, as the high-density and high-speed solar wind plasma also affects the storm behavior and possible enhancements [25]. Next, we applied the method to the observatory geomagnetic data to see the geomagnetic storm response on the Earth's surface and juxtaposition between the ground-based and interplanetary measurements.

### 3.2. General Information on the Analyzed Geomagnetic Storms

The first storm on 15–19 May 2005 was the one of the largest of the 23rd solar cycle, with a Dst index reaching −263 nT right on the first day of the storm. The IMF plots (Figure 2b,c) show the major $B_z$ and $B_y$ disturbances, and the SSC signal is seen on the X component data on 02:40 UT, 15 May 2005 (Figure 2a). This is the time moment when the $B_z$ component turned southward first for a relatively small period. Next it returned to low positive values for a while, turned north, reaching even up to 35 nT, and then its rapid decrease began, which nearly matched the main phase of the geomagnetic storm (seen on the $X$ component plot). The disturbances corresponding to the storm generation are also seen on the solar wind flux speed and proton density plots (Figure 2b,c, respectively).

The next one of the geomagnetic storms we selected for the research was the storm on 14–16 December 2006. At the beginning of 14 December, the $B_z$ (Figure 3c) southward direction did not last for long periods, and the $B_y$ component (Figure 3b) was also close to zero. However, the solar wind speed was relatively high right at the beginning of the day (more than 600 km/s). Around 14:14 UT, there was an abrupt $B_z$ decrease to −10 nT and $B_y$ increase to 6 nT. The storm onset is also seen on the proton density (Figure 3e) and flux speed (Figure 3d) plots as an increase from 2 to 6 n/cc and from 600 to nearly 1000 km/s, respectively. The SSC signal is seen on the $X$ observatory data component (Figure 3a) clearer than on the IMF plots. The proton density reached its maximum of about 25n/cc by the end of 14 December and returned to low (undisturbed) values the next day as well as the flux speed that decreased slowly over the next two days. In addition, there was a sudden disturbance during the main phase of the storm which is seen as abrupt $B_z$ and $B_y$ increases and relatively small increases in solar wind plots.

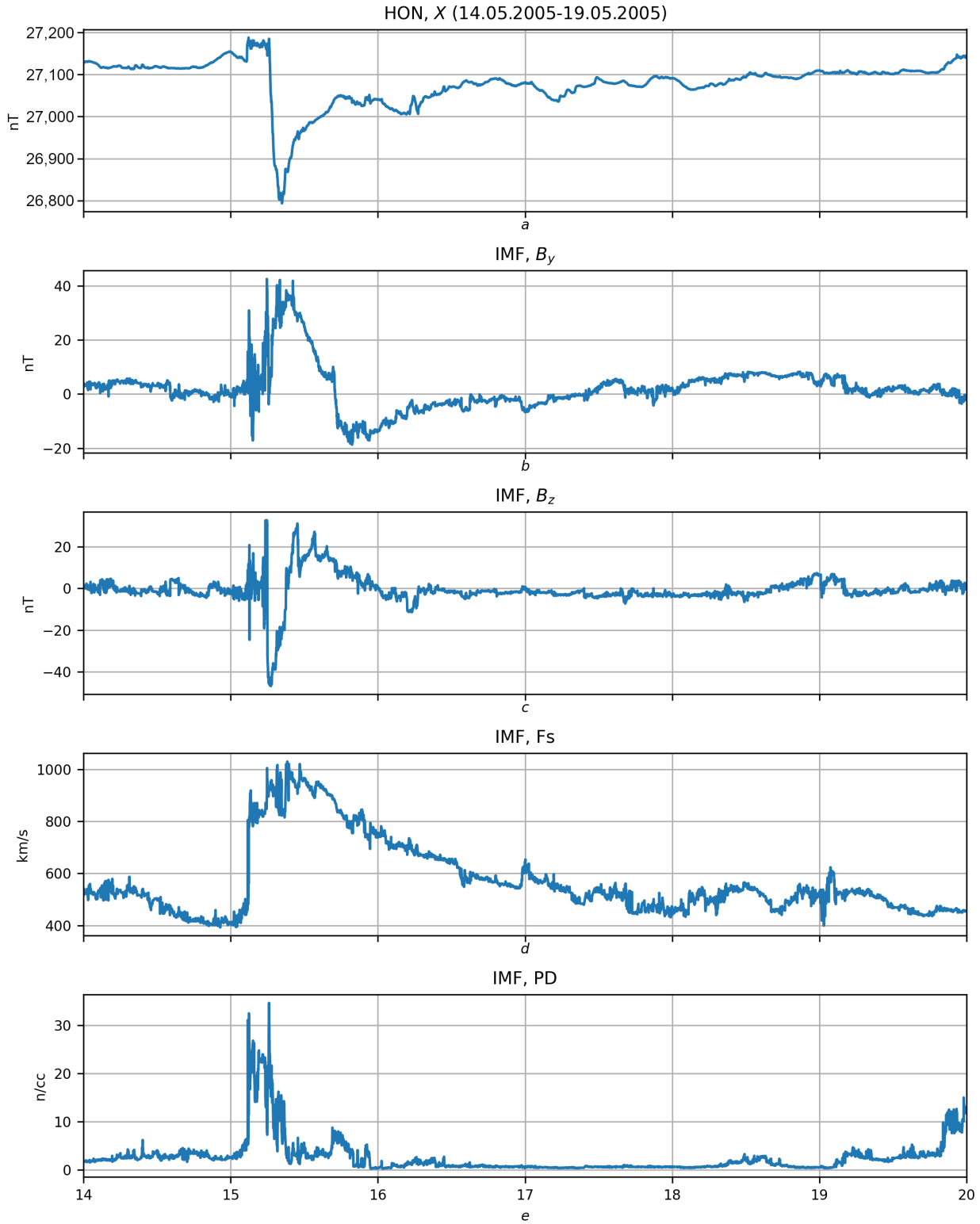



**Figure 2.** The geomagnetic data and solar wind parameters during a geomagnetic storm of May 2005: HON observatory $X$ component (**a**), IMF $B_y$ (**b**) and $B_z$ (**c**) components, solar wind flux speed (**d**) and proton density (**e**).

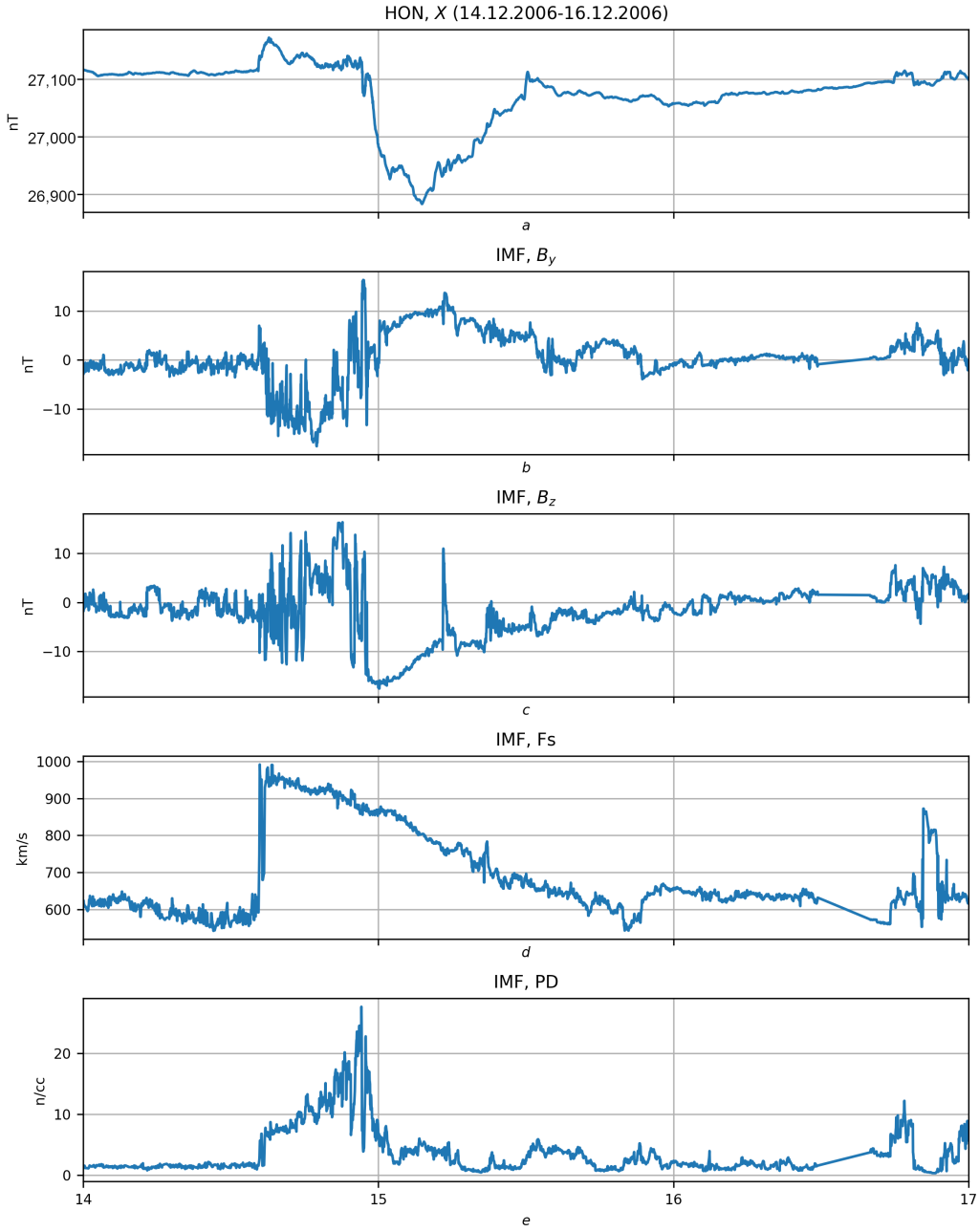



**Figure 3.** The geomagnetic data and solar wind parameters during a geomagnetic storm of December 2006: HON observatory $X$ component (**a**), IMF $B_y$ (**b**) and $B_z$ (**c**) components, solar wind flux speed (**d**) and proton density (**e**).

The magnetic storm on 22–29 June 2015 (Figure 4) was one of the strongest in the 24th solar cycle (minimum Dst index reached −204 nT by the end of the main storm phase). The interplanetary conditions were initially calm, but on 22 June at 18:33 UT, the $B_z$ component of the IMF (Figure 4c) turned to the south, reaching the local peak value of −40 nT. At the same time, the $B_y$ component (Figure 4b) changed first to −25 nT and then to 32 nT. This IMF disturbance was caused by the CME arrival to the magnetosphere frontal part, characterized by the plasma flux speed increase to more than 700 km/s (as seen in Figure 4d) and the proton density increase from 11 to 60–70 n/cc (Figure 4e). The simultaneous abrupt increases in the mentioned IMF and plasma parameters indicate a SSC at 18:33 UT, which is also clearly seen on the $X$ component observatory data plot (Figure 4a). Note that the SSC signal peak is followed by an intense decrease that recovered in a short

time. The storm main phase began on 23 June, when the $B_z$ became negative again reaching −27 nT. The overall storm duration was more than 6 days; some local and low-amplitude disturbances occurred, according to IMF and observatory data on 24–26 June during the storm recovery phase.

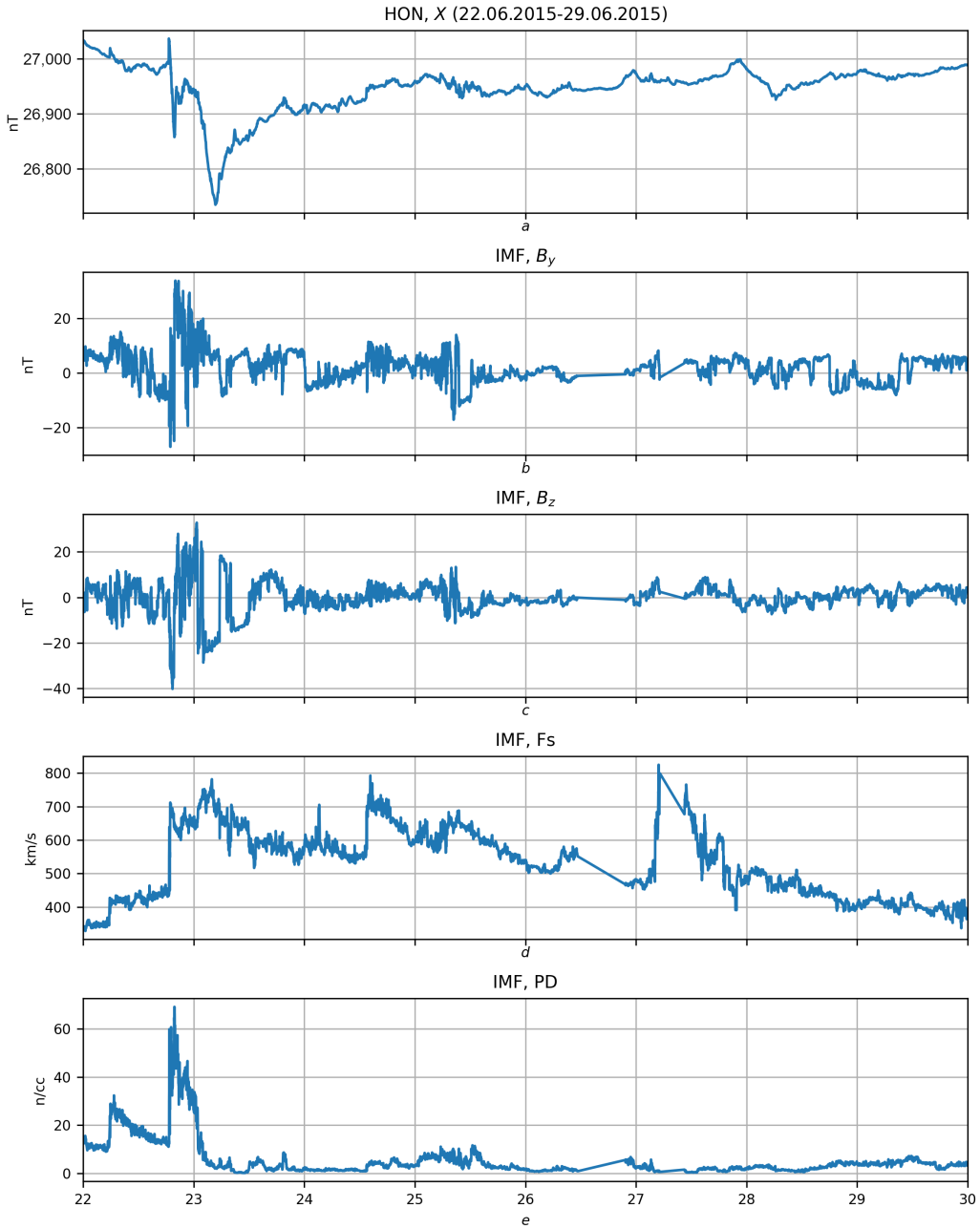



**Figure 4.** The geomagnetic data and solar wind parameters during a geomagnetic storm of June 2015: HON observatory $X$ component (**a**), IMF $B_y$ (**b**) and $B_z$ (**c**) components, solar wind flux speed (**d**) and proton density (**e**).

During the storm that occurred on 25–28 August 2018 (Figure 5), the $B_z$ began pointing southward in the afternoon on 25 August (Figure 5c), when the flux speed rapidly increased from 400 to 450 km/s (Figure 5d). At the same time, a proton density peak occurred (nearly 25 n/cc, as seen in Figure 5e). The amount of energy driven to the Earth's magnetosphere by the solar wind was not enough for a clear SSC signal occurrence. The storm onset at about 19:00 is clearly identified on the $X$ component data plot in Figure 5a. Later, the proton

density increased even more (to 35 n/cc); however, it was during the recovery phase of the storm on 26 August after 06:00 UTC, and the solar wind speed increased more slowly.

For further discussion of the results of the application of activity measures to storm data, we denote the storms as "Storm 1", . . . , "Storm 4", according to their chronological order in Table 1.

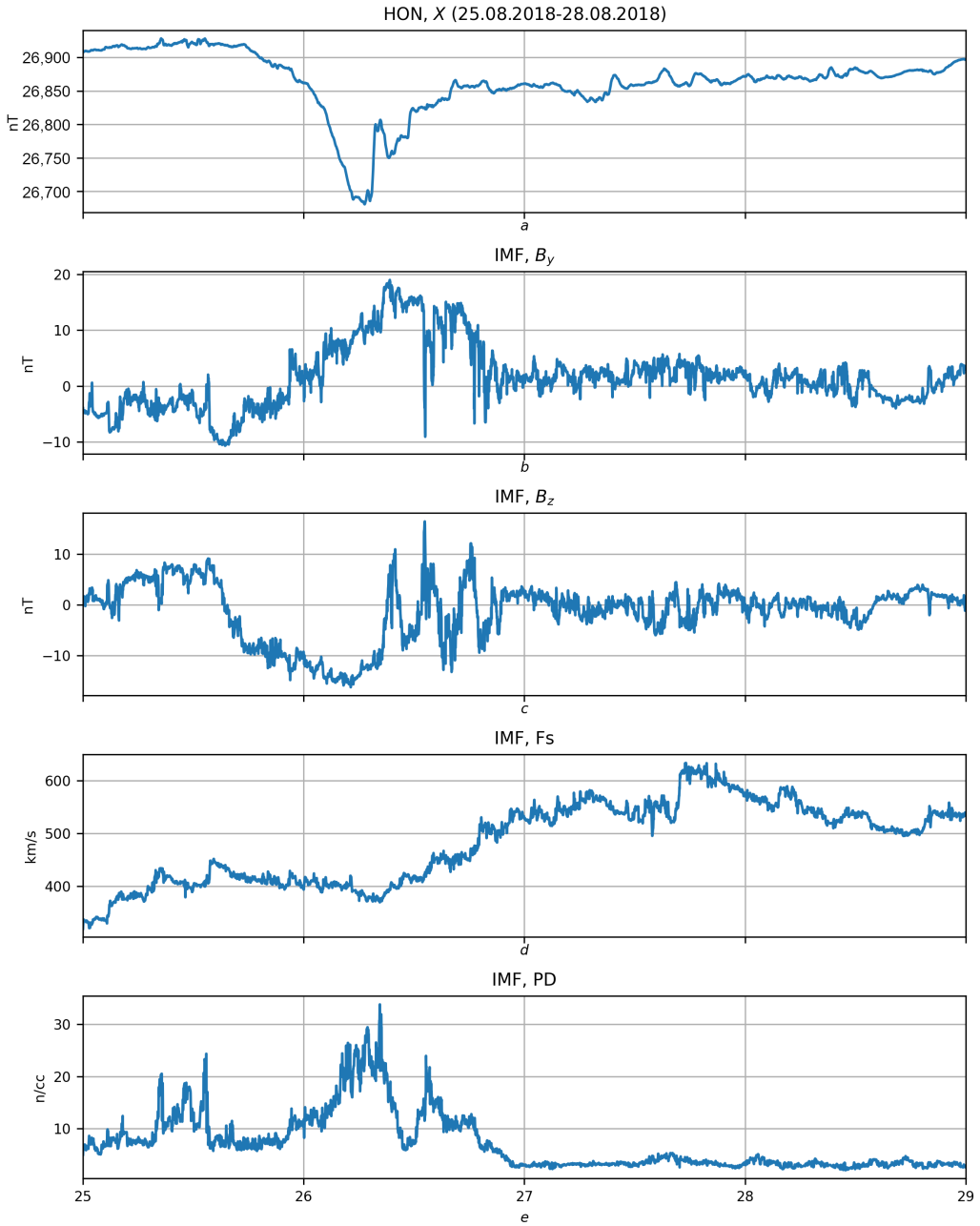



**Figure 5.** The geomagnetic data and solar wind parameters during a geomagnetic storm of August 2018: HON observatory $X$ component (**a**), IMF $B_y$ (**b**) and $B_z$ (**c**) components, solar wind flux speed (**d**) and proton density (**e**).

### *3.3. Discussion*

The 32.71 nT $B_z$ peak at 05:48 UT, 15 May 2005, during the Storm 1 (Figure 6a) is characterized by a positive zone (Figure 6b) on the energy activity measure plot (starting from $p = 10$) and on the length activity measure plot; its consequent abrupt decrease is displayed as a wider zone along all the $p$ values. The corresponding time moment is marked with an arrow. At the same time, the $B_y$ was also subject to disturbances related to

the shockwave arrival to the magnetosphere frontal part that were quite simultaneous with the $B_z$ characteristic disturbances. However, the decreases on it—the first one from 10:08 and the second on, more intensely, from 14:28 to 19:47 UT—generally do not correlate with the $B_z$ behavior (as seen in Figure 2), as the $B_z$ remained positive and slightly decreasing during 15 May 2005 and then was close to zero during the whole recovery phase, except for an abrupt decrease on 16 May 2005 before noon (also marked with an arrow in Figure 6a). Two mentioned $B_y$ decreases correspond to zones of positive sign on the activity measure plots. The first decrease is revealed by a relatively large zone at low $p$ values (from 0 to 10) continuing as a thin zone at higher $p$ values up to $p = 45$. The second decrease is revealed as a wider zone on all $p$ values from 0 to 50, with its widest part at the lowest $p$ values. The SSC signal (seen on the $X$ component in Figure 6c) corresponds to a positive zone on the activity measure plot (Figure 6d) at the moment of the shockwave impact on the magnetosphere front.

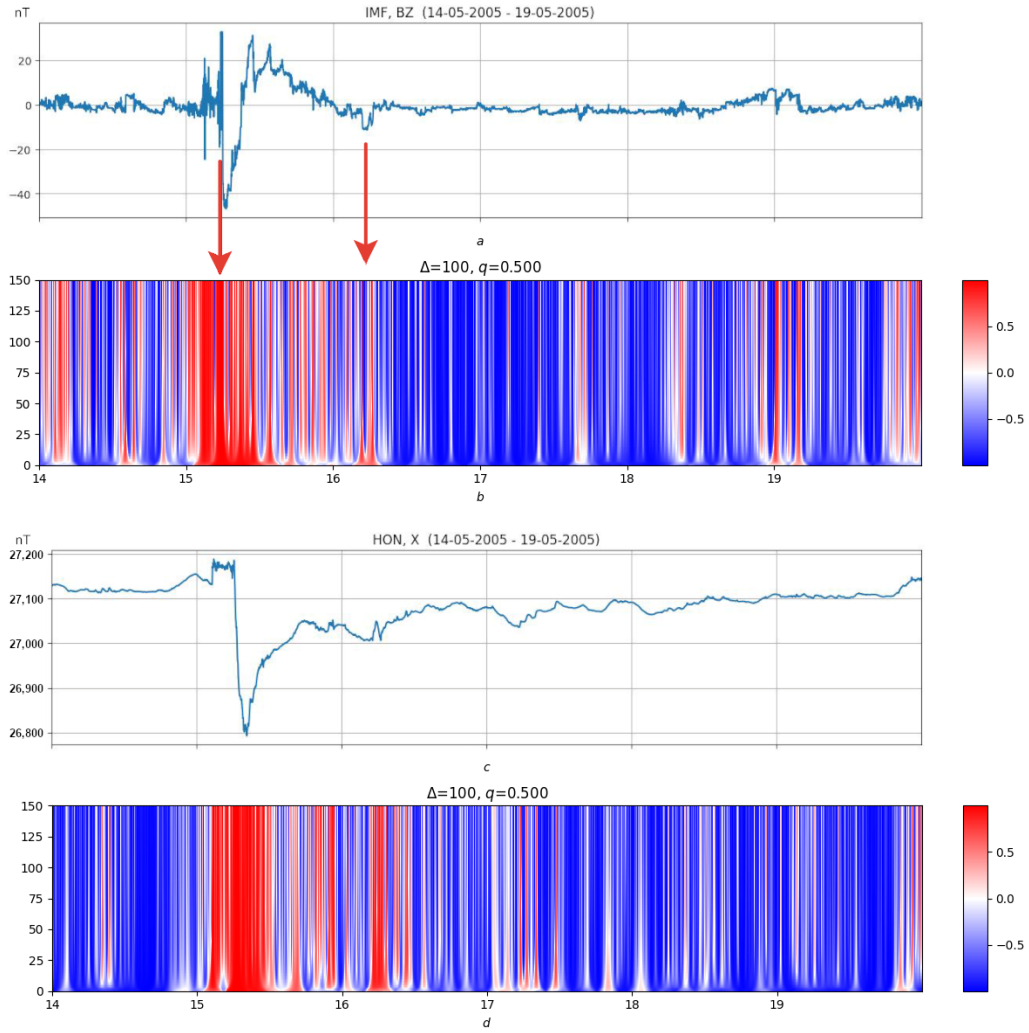



**Figure 6.** The $B_z$ IMF component (**a**) and the $X$ magnetic field component from the HON observatory (**c**) during a geomagnetic storm of May 2005. Below each component is the plot of its activity measure (**b**,**d**) based on the "Energy" indicator.

Similar features of the plasma speed and density data are characterized by similar positive zones on the activity measure plots. The proton density data show the density increase (from 6 to 33 n/cc) at the SSC moment, also matching the flux speed increase from 400 to 800 km/s (and then it gradually increased even to 1029 km/s). The increases in the flux speed are also shown as local positive zones. Slight increases and decreases in the proton density during the main storm phase are also shown as positive and negative

short-time zones, respectively. These plots are not shown in this figure, but we will provide the corresponding plots for other storms to display the activity measure behavior for the plasma speed and proton density data).

The energy activity measure plots for the time series corresponding to Storm 2 (Figure 7a,c) show an agreement for the fragments corresponding to the SSC, the main storm phase and the beginning of the recovery phase. Similar sequences of segments indicating extreme dynamics are seen for $B_z$ (Figure 7b), $B_y$ and proton density plots for these periods. The solar wind speed time series (Figure 7c) displays some differences after the SSC; however, generally the activity measure plot for this time series (Figure 7d) repeats the basic features corresponding to the main phase of the storm and several local disturbances on the recovery phase (15–16 December 2006), marked with arrows, as in the previous case.

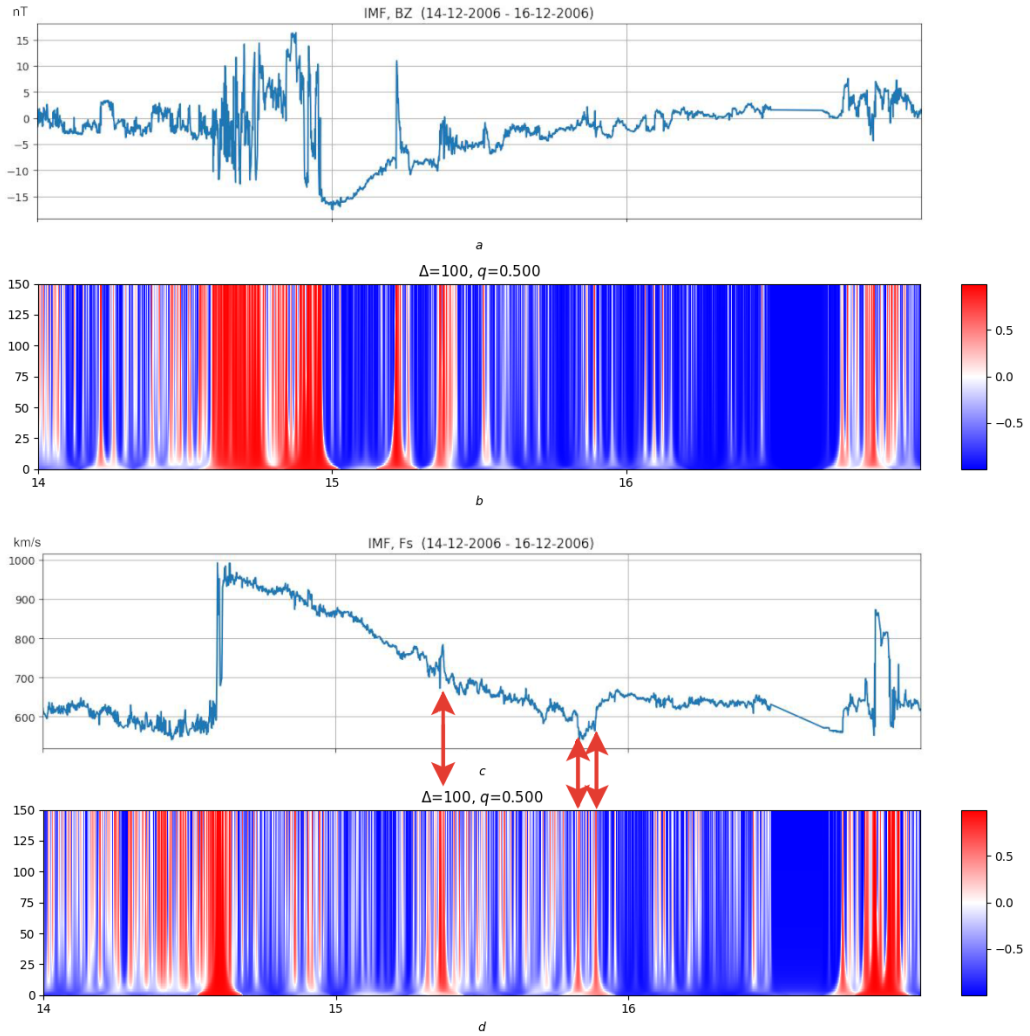



**Figure 7.** The $B_z$ IMF component (**a**) and the $X$ magnetic field component from the HON observatory (**c**) during a geomagnetic storm of December 2006. Below each component is the plot of its activity measure (**b**,**d**) based on the "Energy" indicator.

During Storm 3 (22–29 June 2015), the IMF components alternated within the range of tens of nano-Teslas, but the crucial moments of the storm evolution (the beginning and the end of the main storm phase) are again marked for both $B_z$ and $B_y$ activity measure plots and agree well with the corresponding $X$ data plotted in Figure 4a. The solar wind data time series (especially the flux speed) contain a great amount of small-scale variations that are displayed as separate short-period zones (Figure 4d,e). For this storm, we display a comparison between the activity measure based on the "Energy" indicator (Figure 8b) and the one based on the "Length" indicator (Figure 8c) for the proton density data

to demonstrate the principal difference between these indicators. At the same $\Delta$ and $q$ parameters, the measure peaks look more "combined" on the "Length" plot as if they were parts of one long-period disturbance in the time series, whereas on the "Energy" plot, the separate peaks are emphasized and the difference in their energy is more visible.

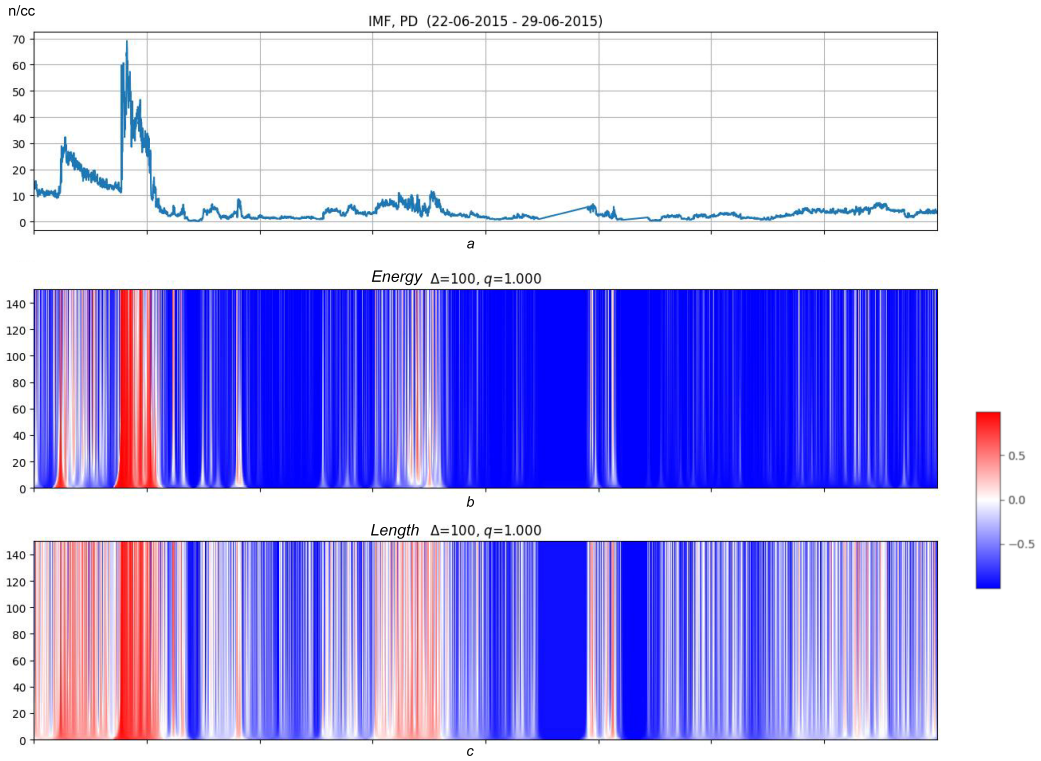



**Figure 8.** The flux speed time series during a storm on 22–29 June 2015 (**a**) and the activity measure plots for this time series based on indicators: "Energy" (**b**), and "Length" (**c**).

For Storm 4 from our selection, the one without a SSC prior to the storm onset, as seen in Figure 9 for the $X$ geomagnetic component analysis, generally, the same fragments of extreme dynamics (the main storm phase and local disturbances in the recovery phase) were detected by the "Energy" and "Length" activity measures as for the previous storms. Abrupt increases and decreases, like in previous cases, are clearly seen as vertical positive (red) or negative zones. Therefore, the activity measure algorithm detects similar physical processes in the magnetosphere and confirms its stable work.

Finally, as a quantitative assessment of the activity for geomagnetic data, a histogram was built, depicting the amount of the measure levels within the interval $[-1, 1]$. The histograms were built for the "Energy" and "Length" indicators applied to the $X$ geomagnetic component at various $\Delta$ and $q$ levels for the geomagnetic storm on 22–29 June 2015. The results show that, both for the "Energy" and "Length" indicators, most of the measure values are concentrated within the interval $[-1, 0]$. This agrees with the common storm-related dynamics seen on the horizontal geomagnetic field components—abrupt decreases related to the storm onset and main phase, and slow increases on the storm recovery phase. Figure 10 displays an example for such histograms for "Energy" and "Length" indicators at $\Delta = 1000$ and $q = 1$. Note that the "Energy" maximal occurrence (Figure 10a) corresponds to $[-1, -0.5]$, whereas the "Length" maximal occurrence (Figure 10b) corresponds to $[-0.5, 0.0]$, which, in our opinion, means that there are generally fewer extreme and dynamically "jagged" segments than the extreme and dynamically "energetic" ones.

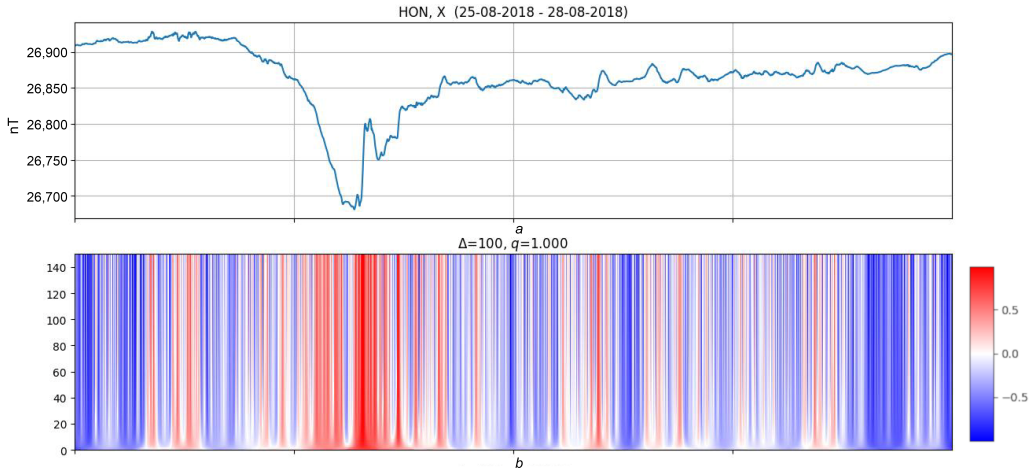



**Figure 9.** The *X* geomagnetic field component data registered during a storm on 26–28 August 2018 (**a**) and the activity measure plots for these data, based on the "Length" indicator (**b**).

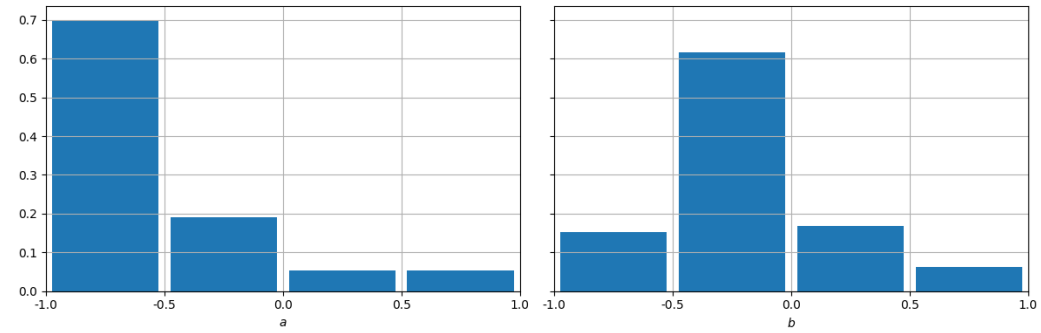



**Figure 10.** Histograms of the activity measure levels for "Energy" (**a**) and "Length" (**b**) indicators built for the *X* component of geomagnetic observatory data during the storm on 22–29 June 2015.

## 4. Conclusions

The present study displays the new applications of discrete mathematics methods, implying the expert's logic modeling, to physical data analysis. Brief information about the basic concept and development of the methods was provided in the introduction. In the "Materials and Methods" section, the methodology was formulated for studying a data record using the analysis of fulfillment of a local property for it at various scales. Two properties of a record were selected as the basic ones: energy (4) (local dispersion ≡ local deviation of a record from its local mean value) and length (5) (describing local jaggedness of a record). The data related to geomagnetic storms (ground-based observatory geomagnetic measurements, interplanetary magnetic field and solar wind data) were selected as records and analyzed using this approach ("Results and Discussion" section). In theoretical terms, the authors see a further continuation of the research in addition to the existing analysis of the regression derivatives recently obtained in the framework of DMA [26,27].

The main conclusion of the study regards the features of the described methods in the magnetic storm morphological analysis using geomagnetic data. The methods based on straightening reveal the decreases and increases in various durations and intensities during the storm. The general features of the storm, such as its onset and the main phase period, are clearly seen on the activity measure plots. Note that the algorithm based on the "Energy" straightening clearly detects even low-amplitude disturbances, such as relatively weak SSC signals or geomagnetic variations during the recovery phase of the storm, while the "Length" straightening is more effective in the detection of sequences of disturbed fragments of a time series. With the increase in the vision radius Δ, the long-term signal fragments can be seen better (such as trends corresponding to the geomagnetic

storm phases), while at small $\Delta$ the short sporadic disturbances are better revealed. This provides an opportunity to successfully implement this group of methods into geomagnetic storm studies.

In all cases, the SSC signals on geomagnetic data plots are denoted by vertically oriented red zones on the activity measure plot, corresponding to all $p$ domains and indicating an abrupt increase in the geomagnetic field intensity. The beginnings of the main phase of the storm on horizontal geomagnetic component data are obviously represented by vertical blue zones, marking the horizontal magnetic field suppressed state according to the ground-based magnetic observatory data.

Modern geoscience and environmental studies imply the application of advanced mathematical techniques focused on complex dataset analysis [28]. This is mainly due to such data features as nonstationarity, sporadic disturbances, large data sets, etc. Our experience in geophysical data studies makes it possible to conclude that the DMA-based approaches to data analysis have also proven their efficiency and robustness.

Another feature of the method is the possibility to perform the combined space weather analysis based on it. For example, the developed techniques can be used in weak SSC signal recognition using a combination of space weather data time series ($B_z$ and $B_y$ IMF components and plasma characteristics, such as flux speed and proton density) and ground-based magnetic observatory data. The use of various mutually complementary space weather data will provide a comprehensive study of extreme phenomena in near-Earth space. This also provides an opportunity to analyze any physical data from various sources in a unified way, using a unified scale.

**Author Contributions:** All authors contributed to the study conception and design. Conceptualization, original draft preparation, S.A. and S.B.; conceptualization, methodology, review and editing and validation, R.K. and R.S.; material preparation, formal analysis, data curation, and algorithm development, S.B. and R.S. All authors have read and agreed to the published version of the 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.

**Institutional Review Board Statement:** Not applicable.

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

**Data Availability Statement:** All the datasets on which the results are based are available free. Ground-based geomagnetic observatory data are available at [12], and interplanetary magnetic field and solar wind speed and density data can be found at [23].

**Acknowledgments:** 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 [12]. We also thank the International Service on Rapid Magnetic Variations, Observatori de l'Ebre [22]. We acknowledge use of NASA/GSFC's Space Physics Data Facility's OMNIWeb service [23].

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

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| DMA | Discrete Mathematical Analysis |
| DPS | Discrete Perfect Sets |
| SSC | Storm Sudden Commencement |
| IMF | Interplanetary Magnetic Field |

## References


1. Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R.; Solovief, A.A. Discrete mathematical analysis and geological and geophysical applications. *Bull. Earth Sci.* **2010**, *16*, 109–125. (In Russian)
2. Agayan, S.M.; Bogoutdinov, S.R.; Krasnoperov, R.I. Short introduction into DMA. *Russ. J. Earth Sci.* **2018**, *18*, ES2001. [CrossRef]
3. Naimark, B.M. Algorithm for detecting a seismic signal against the background of microseisms. *Comput. Seismol.* **1966**, *1*, 5–9. (In Russian)
4. Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R.; Ledenev, A.V.; Zlotnicki, J.; Bonnin, J. Mathematical methods of geoinformatics. II. Fuzzy logic algorithms in problems of anomaly detection on time series. *Cybern. Syst. Anal.* **2003**, *39*, 103–111. (In Russian) [CrossRef]
5. Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R.; Tikhotsky, S.A.; Hinderer, J.; Bonnin, J.; Diament, M. Algorithm FLARS and recognition of time series anomalies. *Syst. Res. Inf. Technol.* **2004**, *3*, 7–16.
6. Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R.; Zlotnicki, J.; Bonnin, J. Mathematical methods of geoinformatics. III. Fuzzy comparisons and recognition of anomalies in time series. *Cybern. Syst. Anal.* **2008**, *44*, 309–323. (In Russian) [CrossRef]
7. Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R. Fuzzy recognition of anomalies in time series. *Dokl. Earth Sci.* **2008**, *421*, 838–842. [CrossRef]
8. Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R.; Graeva, E.M.; Zlotnicki, J.; Bonnin, J. Recognition of anomalies from time series by fuzzy logic methods. *Russ. J. Earth Sci.* **2008**, *10*, ES1001. [CrossRef]
9. Zlotnicki, J.; Le Mouel, J.-L.; Gvishiani, A.; Agayan, S.; Mikhailov, V.; Bogoutdinov, S.; Kanwar, R.; Yvetot, P. Automatic fuzzy-logic recognition of anomalous activity on long geophysical records: Application to electric signals associated with the volcanic activity of La Fournaise volcano (Reunion Island). *Earth Planet. Sci. Lett.* **2005**, *234*, 261–278. [CrossRef]
10. Bogoutdinov, S.R.; Agayan, S.M.; Gvishiani, A.D.; Graeva, E.M.; Rodkin, M.V.; Zlotnicki, J.; Le Mouel, J.L. Fuzzy logic algorithms in the analysis of electrotelluric data with reference to monitoring of volcanic activity. *Izv. Phys. Solid Earth* **2007**, *43*, 597–609. [CrossRef]
11. Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R. Discrete mathematical analysis and monitoring of volcanoes. *Eng. Ecol.* **2008**, *5*, 26–31. (In Russian)
12. INTERMAGNET. Available online: https://intermagnet.org (accessed on 3 November 2021).
13. Agayan, S.; Bogoutdinov, S.; Soloviev, A.; Sidorov, R. The Study of Time Series Using the DMA Methods and Geophysical Applications. *Data Sci. J.* **2016**, *15*, 16. [CrossRef]
14. Soloviev, A.; Agayan, S.; Bogoutdinov, S. Estimation of geomagnetic activity using measure of anomalousness. *Ann. Geophys.* **2016**, *59*, 1–17. [CrossRef]
15. Kulchinskiy, R.G.; Kharin, E.P.; Shestopalov, I.P.; Gvishiani, A.D.; Agayan, S.M.; Bogoutdinov, S.R. Detection and analysis of geomagnetic anomalies using fuzzy logic methods. *Russ. J. Earth Sci.* **2010**, *11*, RE4003. (In Russian) [CrossRef]
16. Soloviev, A.; Bogoutdinov, S.; Gvishiani, A.; Kulchinskiy, R.; Zlotnicki, J. Mathematical Tools for Geomagnetic Data Monitoring and the INTERMAGNET Russian Segment. *Data Sci. J.* **2013**, *12*, WDS114–WDS119. [CrossRef]
17. Gvishiani, A.D.; Soloviev, A.A. *Observations, Modeling and Systems Analysis in Geomagnetic Data Interpretation*; Springer Nature Switzerland AG: Cham, Switzerland, 2020. [CrossRef]
18. Gvishiani, A.; Soloviev, A.; Krasnoperov, R. Automated Hardware and Software System for Monitoring the Earth's Magnetic Environment Automated Hardware and Software System for Monitoringthe Earth's Magnetic Environment. *Data Sci. J.* **2016**, 15, 1–24. [CrossRef]
19. Ishkov, V.N. Properties of the Current 23rd Solar-Activity Cycle. *Sol. Syst. Res.* **2005**, *39*, 453–461. [CrossRef]
20. Sugiura, M. Hourly Values of Equatorial Dst for the IGY. *Ann. IGY* **1964**, *35*, 9–45.
21. Gvishiani, A.D.; Sidorov, R.V.; Lukianova, R.Y. Geomagnetic activity during St. Patrick's Day storm inferred from global and local indicators. *Russ. J. Earth Sci.* **2016**, *16*, ES6007. [CrossRef]
22. International Service on Rapid Magnetic Variations. Available online: http://www.obsebre.es/en/rapid (accessed on 1 November 2021).
23. OMNIWeb Plus. Available online: https://omniweb.gsfc.nasa.gov (accessed on 1 November 2021).
24. Dungey, J.W. Interplanetary magnetic field and the auroral zones. *Phys. Rev. Lett.* **1961**, *6*, 47–48. [CrossRef]
25. Adachi, H.; Sakurai, T.; Marubashi, K. Geomagnetic effects of high-density plasma with southward magnetic field in the interplanetary coronal mass ejection observed on May 2–3. *Earth Planets Space* **2006**, *58*, 315–322. [CrossRef]
26. Agayan, S.M.; Soloviev, A.A.; Bogoutdinov, S.R.; Nikolova, Y.I. Regression derivatives and their application to the study of geomagnetic jerks. *Geomagn. Aeron.* **2019**, *59*, 359–367. [CrossRef]
27. Agayan, S.M.; Bogoutdinov, S.R.; Kamaev, D.A.; Dobrovolsky, M.N. Stochastic trends based on fuzzy mathematics. *Chebyshevskii Sb.* **2019**, *20*, 92–106. (In Russian) [CrossRef]
28. Chakraborty, S.; Chattopadhyay, S. Exploring the Indian summer monsoon rainfall through multifractal detrended fluctuation analysis and the principle of entropy maximization. *Earth Sci. Inform.* **2021**, *14*, 1571–1577. [CrossRef]