Estimation of geomagnetic activity using measure of anomalousness

Anatoly Soloviev[1,2,*], Sergey Agayan[1], Shamil Bogoutdinov[1,2]

[1] *Geophysical Center, Russian Academy of Sciences, Moscow, Russian Federation*

[2] *Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Moscow, Russian Federation*

ABSTRACT

Herein, we present a newly developed indicator for estimating geomagnetic activity. It is based on the magnitude of measure of anomalousness (MA) of magnetometer recordings at a given time or interval. It is intended for automated estimation of geomagnetic activity level in the area of a specific magnetic station or in a given region using data of a set of stations. It reflects geomagnetic activity level at different observatories in a single scale [−1, 1], regardless of their latitudinal location and consequently typical disturbance amplitudes. To a certain extent MA indicator is an analog of traditional K index. However, a well-known shortcoming of the latter is its long, 3-hour update rate. Moreover, K index calculation requires subtraction of Sq variation that also causes delays. At the same time there is a demand for operational geomagnetic indices that have maximal time resolution and are available in near real-time. The proposed MA indicator aims to address the shortcomings of the traditional K index. The MA calculation may be implemented automatically with the same time resolution as the initial data are recorded.

1. Introduction

Over the last few decades in geophysics and related sciences a rapid growth of volumes of incoming data on the processes in the Earth's interior, at the surface and in the near Earth space has been taking place. Ground geophysical observation networks constantly expand and perform transition to higher data sampling rate standards to meet the needs of a wider range of scientific interests. For instance, SDO (solar dynamics observatory) satellite transmits daily about 1.5 terabytes of data on the solar atmosphere at small spatial and temporal scales, and in many wavelengths simultaneously (http://sdo.gsfc.nasa.gov/). Currently the magnetic observatory network of the highest standard (INTERMAGNET) [Love and Chulliat 2013] contains more than 120 observatories all over the world. That is a huge constantly growing amount of data, barely subjected to manual continuous analysis. Besides, the collected archives of historical environmental data are being digitized in the data centers all over the world for several decades. Consequently, the role of adequate methods for big digital data analysis has arisen. Since digital data is discrete by definition, it makes development of discrete mathematics, which is aimed at studying discrete mathematical objects as opposed to continuous objects, much more relevant. Taking into account the present progress in data acquisition, it becomes possible and seems more efficient to use complete set of data sources and apply multi-value estimation of observed process that reflects its both temporal and spatial variability.

Planetary and regional indices are widely used for geomagnetic activity estimation based on ground observations [Mayaud 1980, Rangarajan 1989, Menvielle 1990, Siebert 1996]. Most indices reflect specific physical processes taking place in near-Earth space, originating disturbances observed on the ground. The first C index was introduced in 1906 [Lincoln 1967], and since then the principles of their implementation and use have not changed much. Unfortunately, even now many indices allow the activity to be estimated only retrospectively, with several hours delay. Traditionally indices are calculated using a strictly fixed number of observatories (<20) and designed to quantify the total strength of magnetic storms, substorms, their triggering sources or general activity level. The regional choice of observatories is made with an assumption of predefined and localized geographical distribution of the effects under consideration. Such approach seems to be a bit outdated and not efficient enough, given the rapid growth of data volumes. At the same time, such indices as Kp and Dst have to be kept unchanged to allow consistency over several decades. Moreover, it seems that nowadays application of such measures,

that integrally quantify regional or global processes by a single number, turns to be outdated too. Another disadvantage of the majority of traditional indices is their low time resolution of 1-3 hours that sometimes conceals more rapid but still significant disturbances. No wonder that the traditional approach to geomagnetic activity evaluation is being constantly improved, thus, addressing the demands for:

- higher time resolution of regional [e.g., Della-Rose et al. 1999] and planetary [e.g., Iyemori et al. 1999, Chambodut et al. 2015] indices,

- real-time estimation of the geomagnetic activity [e.g., Stankov et al. 2011],

- increase of the informative value of extracted signals [e.g., Mandrikova et al. 2013, 2014] and others.

Herein, we attempt to overcome these limitations with respect to the family of K indices. We've constructed a new geomagnetic activity indicator MA using discrete mathematical analysis (DMA) approach [e.g., Gvishiani et al. 2008a, 2008b, 2010; Bogoutdinov et al. 2010; Soloviev et al. 2012a, 2012b] which complements the existing indices through a number of improvements. It depends on specific magnetic record and time and provides transformation of initial recordings to continuous interval [−1, 1]. That normalized scale reflects a measure of activity at a given time and associates values close to −1 with complete absence of activity and values close to 1 with extreme activity. Normalized scale of abnormality estimation provides independence of latitudinal and longitudinal location of observatory, whose data are being studied. The MA time resolution equals to sampling rate of initial magnetic record (1-minute for INTERMAGNET data), that potentially gives a possibility for real-time estimation of geomagnetic activity. MA requires little in the way of subjective judgment by the user, except to specify a duration time to calculate the "straightening", and setting thresholds that indicate if an event of interest has occurred. These features are the principal advantages of the method comparing to traditional K and Kp index estimation techniques for dealing with modern-day real time data streams.

In this paper, we give the description of the method, report on the results of MA application to real data and demonstrate the method's advantages. The calculation of the MA is briefly presented in Section 2 and its detailed mathematical description is given in Annex 1. An example of magnetic activity recognition using the MA in February 2014 is also given. We then compare the MA with K index (Section 3) and show that it does not contradict to traditional estimation techniques. Moreover, it surpasses K index because of the higher time resolution, normalized estimation scale and no need of Sq calcula-

tion. Then we determine the optimal graded scale of the MA for evaluation of various manifestations of geomagnetic activity (Section 4). That scale allows to distinguish between background and anomalous segments of initial record and specifies several degrees of abnormality. In Section 5, we demonstrate application of the MA to geomagnetic activity studies using data from observatories located in different parts of Russia. We take data sets recorded during January 7, 2015, and March 17, 2015, magnetic storms as examples. In particular, these applications demonstrate potential MA applicability for recognition of magnetic storm precursors, such as sudden commencements. Then we discuss the MA functioning in continuous mode for efficient operational detection of magnetic activity (Section 6). In Section 7, we apply the MA to global magnetic activity studies and show that sometimes traditional indicators do not adequately evaluate global perturbations. We demonstrate it by the example of substorms that occurred in the period of November 8-11, 2004. They caused geomagnetic disturbances all over the globe that were successfully recognized using the MA. This fact confirms the universality of the proposed method.

2. Methodology

The developed MA, that we also denote as $\mu(t)$, is intended for estimation of geomagnetic activity using ground observation recordings. It represents modification of FCARS (fuzzy comparison algorithm for recognition of signals) algorithm and it was partly described in [Gvishiani et al. 2008 , Soloviev et al. 2013]. In this paper we intend to give a complete review of the MA for the first time, including its detailed definition and various applications to geomagnetic data.

The detailed mathematical description of $\mu(t)$ calculation is given in Annex 1. In this section, we briefly outline the procedure and give some examples. Firstly, the non-negative functional, called "straightening", is computed in a sliding time-duration window Δ using initial measurements, that are defined on an interval (registration period) T. "Straightening" can be considered to be successfully specified, if the anomalies in a raw record coincide with uplifts on the straightened record (Figure 1). Different constructions of "straightening" functional are given, for example, in [Gvishiani et al. 2008b, 2014]. Hereafter, we use functional "Length", which integrates the absolute values of first differences over time window Δ = 21 samples (INTERMAGNET data are sampled every minute). The selected functional reflects "jaggedness" of a fragment of the original time series. Due to this property "Length" allows to detect fragments, that contain both high-frequency (e.g., geomagnetic pulsations accompanying magnetic
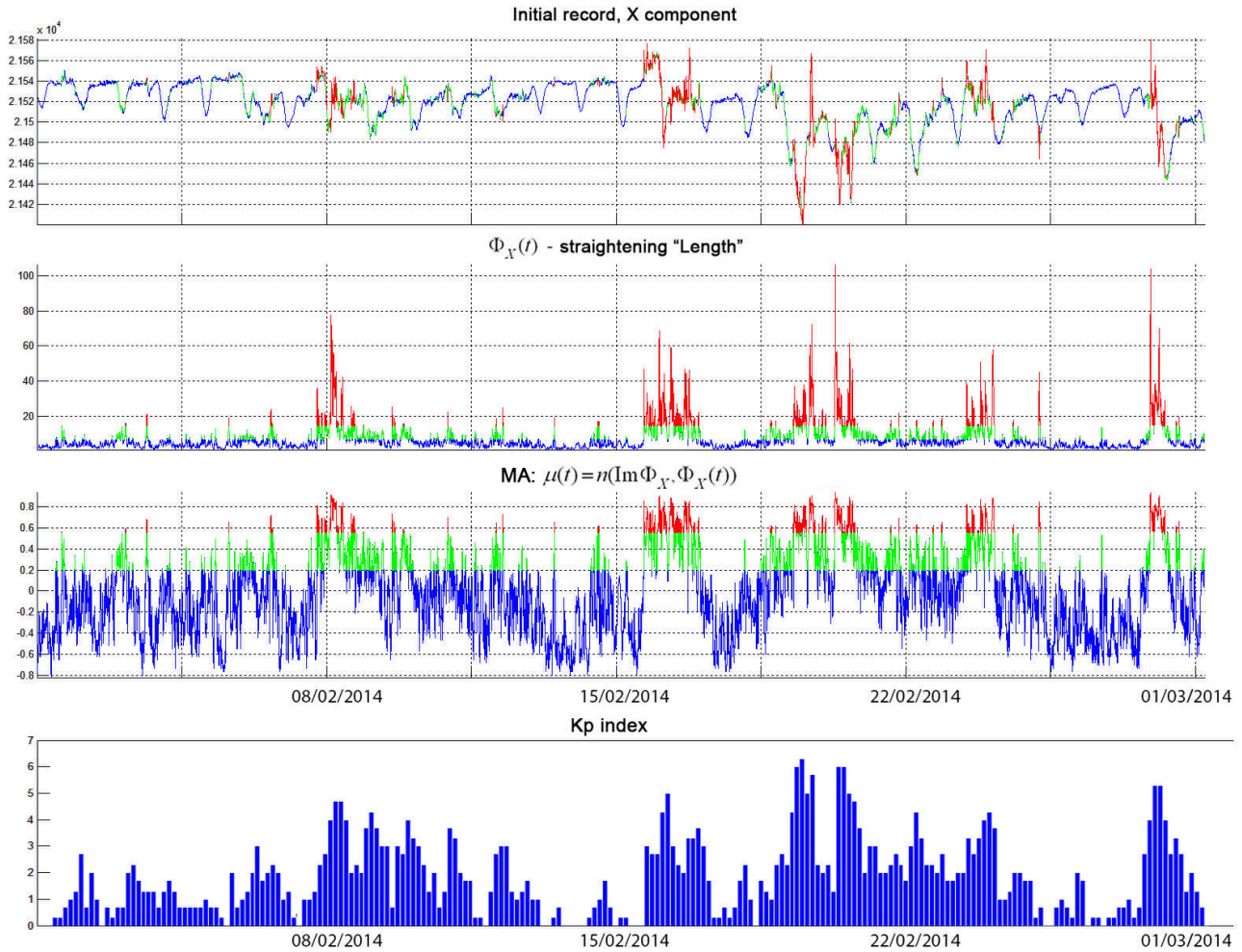
Figure 1. X component recorded at PET observatory in February 2014 (1-st plot), "straightening" Φ_X of the record (2-nd plot), MA $\mu(t) \in [-1, 1]$ (3-rd plot) and Kp index (4-th plot). Background values of νB set ($\mu(t) < 0.2$) are marked with blue, potentially anomalous values of νP set ($\mu(t) \in [0.2, 0.55)$) are marked with green and anomalous values of νA set ($\mu(t) \geq 0.55$) are marked with red in plots 1-3. Date format: dd/mm/yyyy.

storms and substorms) and long-term high-amplitude disturbances, that are essential components of increased geomagnetic activity. This universal feature of the functional "Length" along with several tests indicates its more efficient applicability in recognition of extreme natural geomagnetic events comparing to other types of "straightening". Sums of absolute first differences are used in other indices such as inter-hourly variability (IHV) index, that is derived from hourly means of H component over a 7-hour interval centered on local midnight [Svalgaard and Cliver 2007]. However, it is intended as a long-term index for determination of solar wind speed variability over long time intervals (e.g., last ~130 years), whereas MA applies to minute (or second) values and estimates more rapid variations of magnetic field.

The next step is identification of both absolutely significant and localized but small uplifts of the "straightening". We use "fuzzy comparisons" [Gvishiani et al. 2008b], that is another fundamental element of the DMA approach, to estimate significance of a given "straightening" data point relative to the image of the "straightening". The latter is essentially the set of all val-

ues of the "straightening" functional, obtained from initial record over registration period T. It results in a new time-series of mapped values that fall within $[-1, +1]$, which we call MA. This mapping assigns higher values (closer to $+1$) to local extremes of the "straightening" and consequently to anomalous measurements in initial record, and lower values (closer to -1) to background measurements. The resulting MA time-series has the same domain discreteness as initial record.

In the given definition of the MA (see Annex 1) the initial values are classified into three categories - background, potentially anomalous and anomalous. Such fragmentation of initial record automatically results from fragmentation of the corresponding MA record using β_w and β_s values: β_w separates background and potentially anomalous segments and β_s separates potentially anomalous and anomalous segments in the range of MA. In general, any number of abnormality levels (e.g., absence of disturbance, weak disturbance, moderate disturbance, strong disturbance, anomaly, strong anomaly) or continuous $[-1, 1]$ scale can be applied. Choice of a proper threshold for β_w, that sepa-

rates background and active signals, is essentially a choice of the method's "sensitivity". One can obtain, for example, the best match between μ 3-hour means and local K index [Bartels et al. 1939] by varying this parameter for given observatory. Another way is to define a single β_w threshold for all the observatories.

We give below an example of MA application to magnetic activity estimation based on a single observatory magnetogram. We denote by νB the set of background points, by νA the set of anomalous points and by νP the set of potentially anomalous points, defined according to β_w and β_s values. They provide fragmentation of the registration period $T = \nu B + \nu A + \nu P$. Figure 1 shows the MA calculated for north component (X) recorded at Paratunka observatory (IAGA code PET) in February 2014. The figure contains the initial record plot, its "straightening" and MA values with 3-level color gradation of abnormality. In this example threshold values β_w and β_s were taken equal to 0.2 and 0.55 respectively (later we define this gradation more precisely).

Geomagnetic storm took place on February 19-20, 2014, according to the diagram of the planetary Kp

index that is aimed to globally estimate magnetic activity of different natural origins (4-th plot in Figure 1). Besides, an increased magnetic activity was observed on February 15-16, 2014. The mentioned events were classified as anomalous and marked with red in Figure 1. We can also see less significant activity, that was reasonably marked with green (potentially anomalous events) in the upper plot.

3. Comparison with K index

The MA was compared with geomagnetic activity indices, in particular, with 3-hour K index [Bartels et al. 1939, Menvielle and Berthelier 1991], to see how it correlates with traditional approaches to geomagnetic activity estimation. The MA values were averaged beforehand over 3-hour intervals and normalized to scale [0, 1] for a proper comparison. The first comparison was performed using geomagnetic data recorded in the period of increased magnetic activity at Chambon-la-Forêt mid-latitude observatory (CLF) in January 2005 (Figure 2a). The second one was based on 1-month record obtained at Sodankyla high-latitude observatory
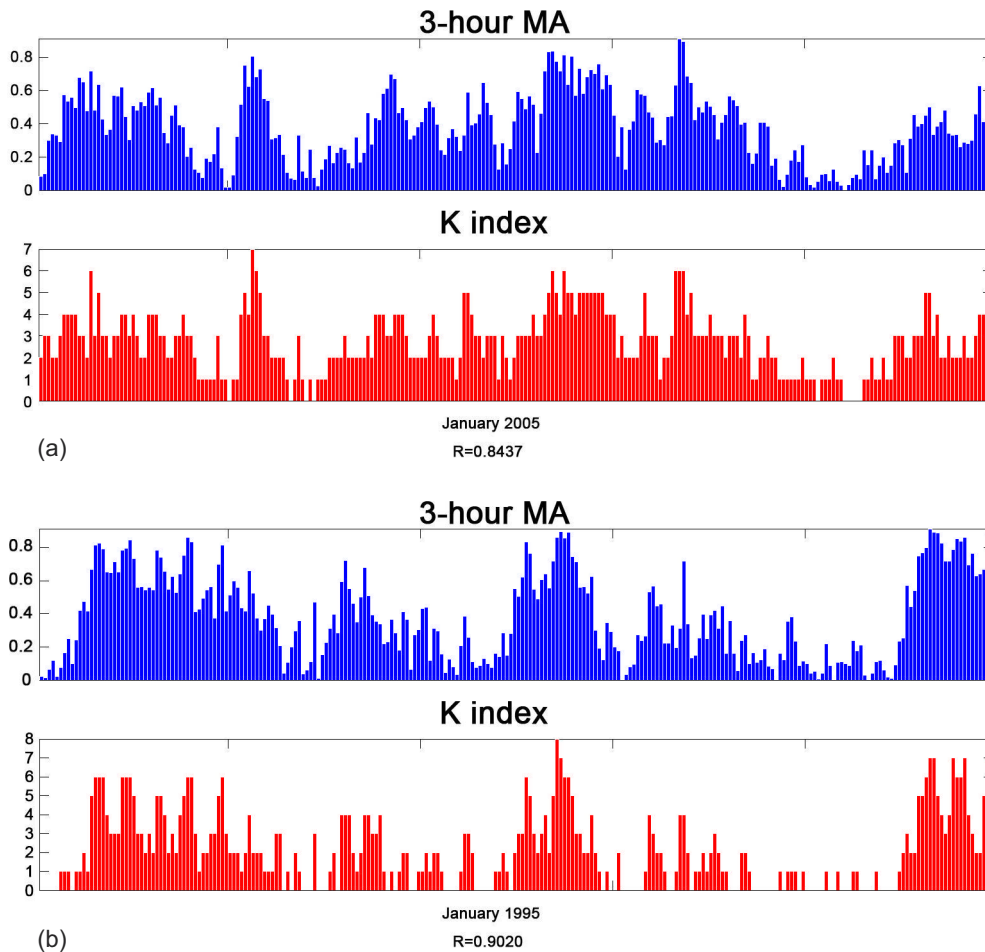

Figure 2. Comparison between the 3-hourly averaged minute values of MA and the regional 3-hour *K* index values *K*(*t*) for the disturbed magnetic records of the north component: Chambon-la-Forêt observatory (CLF), January 2005 (correlation coefficient *R* is 0.84) (a); Sodankyla observatory (SOD), January 1995 (correlation coefficient *R* is 0.90) (b).

(SOD) during January 1995. The record contained many highly active events due to auroral location, although the period under consideration was close to the minimum of solar activity. The comparison has demonstrated an exceptional similarity between the two methods with the Pearson correlation coefficient (R) equal to 0.90 (Figure 2b). It is seen from Figure 2 that the MA plot is more sensitive to magnetic disturbances while K plot is rougher, despite the artificial 3-hour averaging of MA. This is due to the use of a discrete set of resulting values in the K index definition. Generally, the MA can be calculated with a minimal time resolution equal to the sampling rate of initial recordings (1 minute in the case of INTERMAGNET observatories).

Large calculation statistics provides us with the information on how close MA 3-hour means follow K index depending on latitude and season. To evaluate latitudinal variation of correlation coefficient R we considered data from 13 observatories located at latitudes from 71°N to 65°S. All of them except one are situated in the western hemisphere. For each observatory, K and MA were compared for the period of April 1-30, 2006, as it contains phases of both decreased and increased magnetic activity. The results are given in Figure 3. It shows that R remains quasi constant with latitude change and varies between 0.80 and 0.90 with the average of 0.86.

Figure 4 illustrates seasonal variability of correlation coefficient R that was calculated using 2003 data from high-latitude SOD observatory and mid-latitude Memambetsu (MMB) observatory. One can see that R does not fall below 0.77 in summer and reaches 0.92 in winter with the average of 0.86 at SOD and varies between 0.70 and 0.90 with the average of 0.79 at MMB. Cases of low correlation were examined separately. It was found that lower correlation values conform to periods of more homogeneous field behavior, e.g. globally undisturbed periods. The reason for that is clear - the MA is more sensitive to low-amplitude rapid disturbances, such as continuous pulsations, rather than K index. That is due to the "Length" functional construction, used in the MA definition. It results in lower linear correlation between MA and K index during magnetically quiet periods that contain minor field perturbations. Several examples of such periods can be clearly seen in Figure 2 (those where K index is close to zero). The results show that in principle the MA doesn't contradict to traditional way of geomagnetic activity estimation and supplements the family of K indices.

4. Graded scale of activity

Now we have to define a proper threshold β_w within $[-1, 1]$ for the MA values that would help us to distinguish between background ($\mu(t) < \beta_w$) and anom-
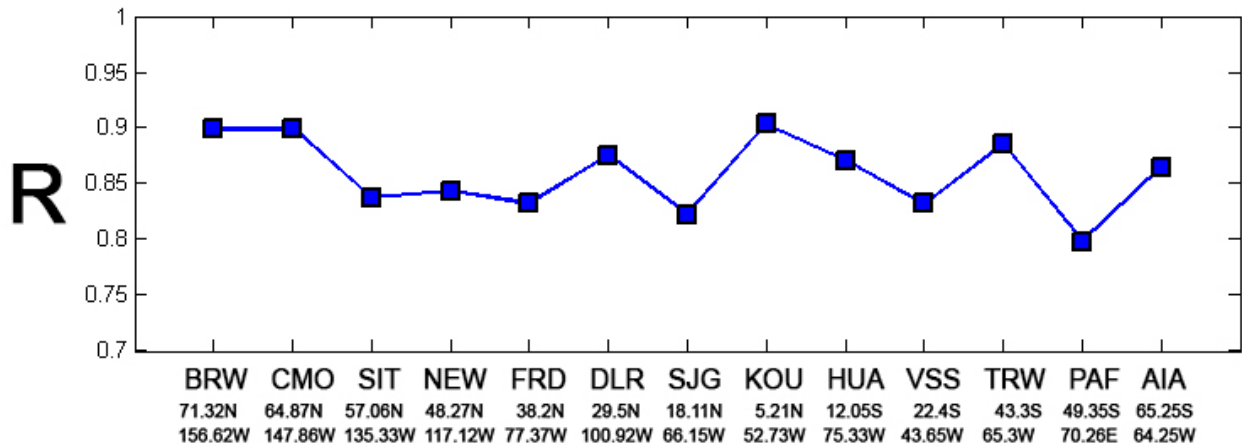

Figure 3. Latitudinal variation of correlation coefficient R between K index and MA for April 2006.
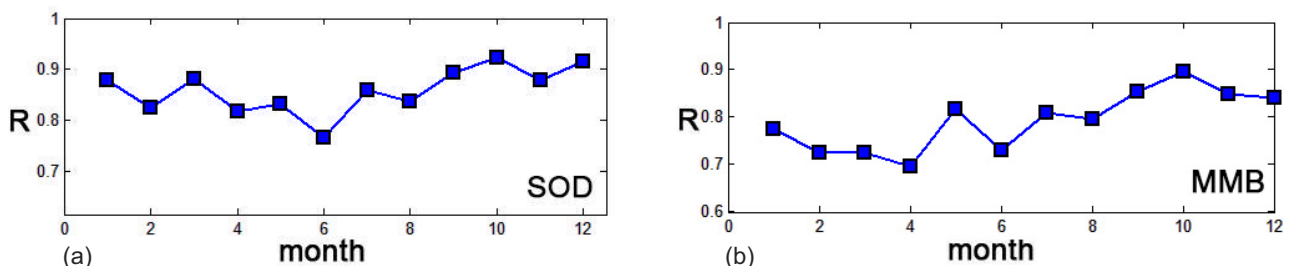

(a)

(b)

Figure 4. Seasonal change of the correlation coefficient R between K index and MA in 2003 at SOD observatory (a) and MMB observatory (b).
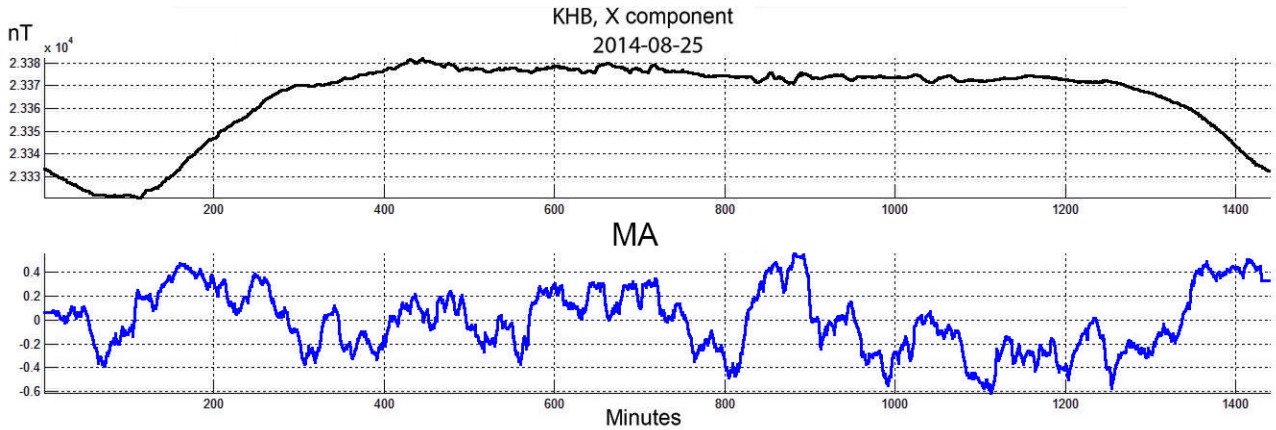
Figure 5. Activity estimation using MA in *X* component, recorded at Khabarovsk observatory (upper plot) on August 25, 2014 (lower plot).

alous ($\mu(t) \geq \beta_w$) events. Let us analyze MA application to magnetic measurements recorded during decreased geomagnetic activity. Herein we study the magnetogram recorded at Khabarovsk observatory (KHB) on August 25, 2014. This one is among the 5 quietest days of August 2014 according to official IAGA classification (International Q-Days and D-Days: http://www.gfz-potsdam.de/en/section/earths-magnetic-field/data-products-services/kp-index/qd-days/) and it is estimated with 0 and 1 values of planetary Kp index. MA application to magnetic activity estimation based on the north component of the magnetic field is given in Figure 5.

This example shows that during quiet period MA values practically do not exceed 0.4 except for three fragments of short duration. Many similar tests were made using data from other observatories located at different latitudes, which we do not present in this paper; herein we limit ourselves to only one example. It followed from all the examples that MA threshold should be taken from the interval [0.4, 0.5] for local estimation of geomagnetic activity. For recognition of very small variations the threshold value should be reduced. Nevertheless, in the present study such disturbances are not considered as anomalous events.

We can now introduce several grades of abnormality for those geomagnetic recordings $y(t)$ that are associated with $\mu(t) \geq \beta_w$. The following MA graded scale was empirically elaborated based on numerous tests mentioned above:

1) $-1 \leq \mu(t) < 0.4 \Rightarrow t$ corresponds to background event;
2) $0.4 \leq \mu(t) < 0.55 \Rightarrow t$ corresponds to weakly anomalous event;
3) $0.55 \leq \mu(t) \leq 0.7 \Rightarrow t$ corresponds to anomalous event;
4) $0.7 \leq \mu(t) \leq 1 \Rightarrow t$ corresponds to strongly anomalous event.

Let us demonstrate the MA application with the chosen scale by example of the record that contains both quiet and disturbed fragments. September 11-15, 2014, is among such periods (Indices of Global Geomagnetic Activity: http://www.gfz-potsdam.de/en/kp-index/). Figure 6 demonstrates the results of geomagnetic activity recognition using Irkutsk observatory (IAGA code IRT) data for the given time interval. Recordings, that were obtained after September 13, are classified as background (blue color). During that period decreased magnetic activity was observed. On the contrary, strong disturbances were recognized on September 12 and 13 (purple and red colors). High values of planetary Kp index (Indices of Global Geomagnetic Activity) are associated with that days as well. This example again shows that the selected threshold value $\mu(t) = 0.4$ is adequate to distinguish between background and anomalous geomagnetic events in magnetograms. All other examples given in the paper (Figures 7-10) also argue for the proper selection of the MA graded scale.

MA allows to detect active fragments in initial record and estimate their relative intensity. Quite often there is a need to estimate absolute intensity of disturbed fragments to be able, for example, to compare different anomalous events for different time periods by their intensity. In that case integral characteristics of "straightening" values could serve as a full intensity indicator of recognized anomaly, since its uplifts are supposed to be anomalies in initial record. Such integral characteristics could be expressed using, for example, mean and maximum values of the "straightening" within each anomalous time interval.

5. Multi-observatory geomagnetic activity estimation

Below we give examples of recognition and estimation of geomagnetic activity caused by moderate magnetic storm on January 7, 2015, and the strongest March 17, 2015, storm so far observed during the current solar cycle. Herein we use spatially distributed data sets obtained from several Russian observatories.
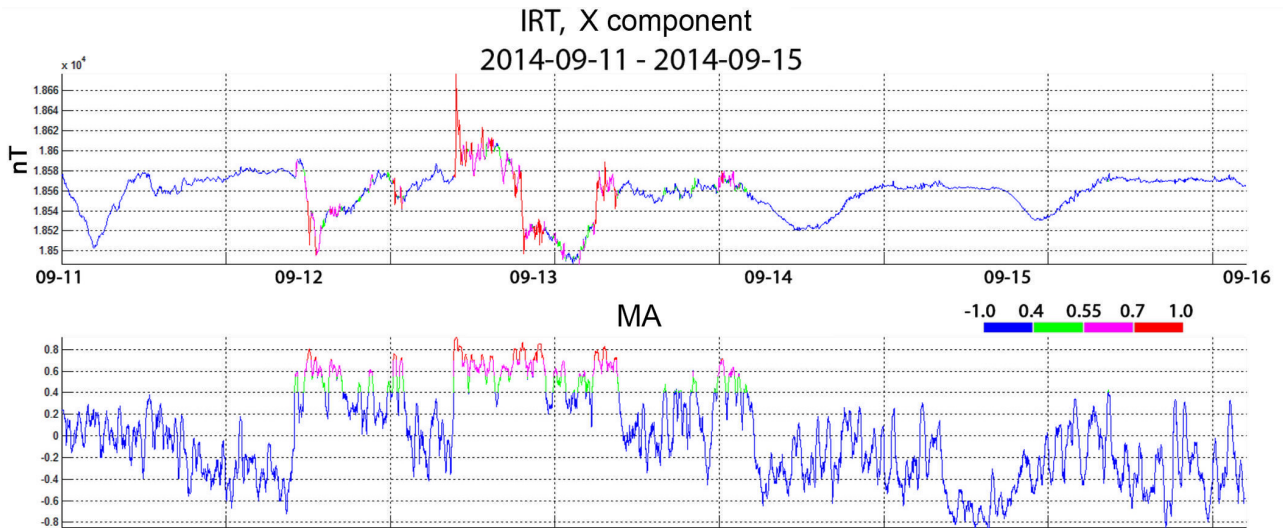
Figure 6. Estimation of geomagnetic activity using *X* component record, obtained at Irkutsk observatory on September 11-15, 2014 (upper plot) using MA (lower plot). In the both plots background values ($\mu(t) < 0.4$) are marked with blue, weakly anomalous values ($\mu(t) \in [0.4, 0.55)$) are marked with green, anomalous values ($\mu(t) \in [0.55, 0.7)$) are marked with purple and strongly anomalous values ($\mu(t) \geq 0.7$) are marked with red.

We considered a 3-day period from January 6, 2015, 00:00 UT, to January 8, 2015, 23:59 UT, during which the moderate storm took place. The results of geomagnetic activity recognition for the given period are presented in Figure 7. We used data recorded at Russian geomagnetic observatories that are located at different geographic latitudes: Cape Schmidt (IAGA code CPS, 68.9N), Klimovskaya (IAGA code KLI, 60.9N) [Geomagnetic data 2015, Soloviev et al. 2015], Saint Petersburg (IAGA code SPG, 60.5N) [Geomagnetic data 2016], Moskva (IAGA code MOS, 55.5N) and Novosibirsk (IAGA code NVS, 54.9N). Such selection allows to observe the diversity of magnetic activity signatures in different parts of the country.

The signal intensity, that corresponds to the storm, expectedly decreases with decrease of observatory latitude (Figure 7). Also, one can see the change of the signal morphology that is typical for different latitudinal belts. For example, signals at KLI and SPG ($\sim$61°N) observatories have similar spike-like shape, whereas signals at MOS and NVS ($\sim$55°N) observatories have more complicated morphology.

Magnetic storm onset and main phase were recognized as anomalous (purple) and strongly anomalous (red) events in the whole set of studied magnetograms. Besides, the algorithm recognized a disturbance, that synchronously, but with varying intensity, affected the records in the afternoon on January 6, 2015. The disturbance was classified as weakly anomalous (green) event in all magnetograms under consideration. The magnetograms also contain small geomagnetic pulsations of Pc5 type, that are represented by regular oscillations during the recovery phase of the storm and later on. With the chosen MA parameters such events are not classi-fied as anomalous and thus they are not recognized.

During the storm of March 17, 2015, the deviations of more than 2000 nT were observed in magnetograms, recorded at a number of Russian observatories. The storm was caused by the coronal mass ejection from the Sun, that reached the magnetosphere on March 17 at 04:30 UT [Kataoka et al. 2015]. After that it took several hours for the storm to develop with the planetary index value Kp = 8 (Planetary K-index: http://www.swpc.noaa.gov/products/planetary-k-index). The storm was so strong that aurora could be seen even in Moscow. So far it is the strongest storm in the current solar cycle.

The results of recognition of extreme geomagnetic events, that accompanied the March 17 storm, are given in Figure 8. The algorithm was applied to the data from the Russian observatories located at different geographical latitudes: Klimovskaya (IAGA code KLI, 60.9N), Magadan (MGD, 60.1N), Moskva (MOS, 55.5N), Novosibirsk (NVS, 54.9N), Paratunka (PET, 53.0N), Irkutsk (IRT, 52.3N) and Khabarovsk (KHB, 47.6N). As in the previous examples, the observatories in the figure are given in the descending order according to their latitudes.

Notably, this strongest storm was accompanied with the clear sudden commencement (SC) signature originated from the sharp compression of the magnetosphere during the collision with interplanetary cloud. SC is an important precursor of geomagnetic storms in the analysis of ground magnetograms, which makes it possible to predict strong magnetic disturbances several hours in advance. However, the ultimate factor, that defines the storm development, is the direction of the interplanetary magnetic field (IMF) vertical component. In
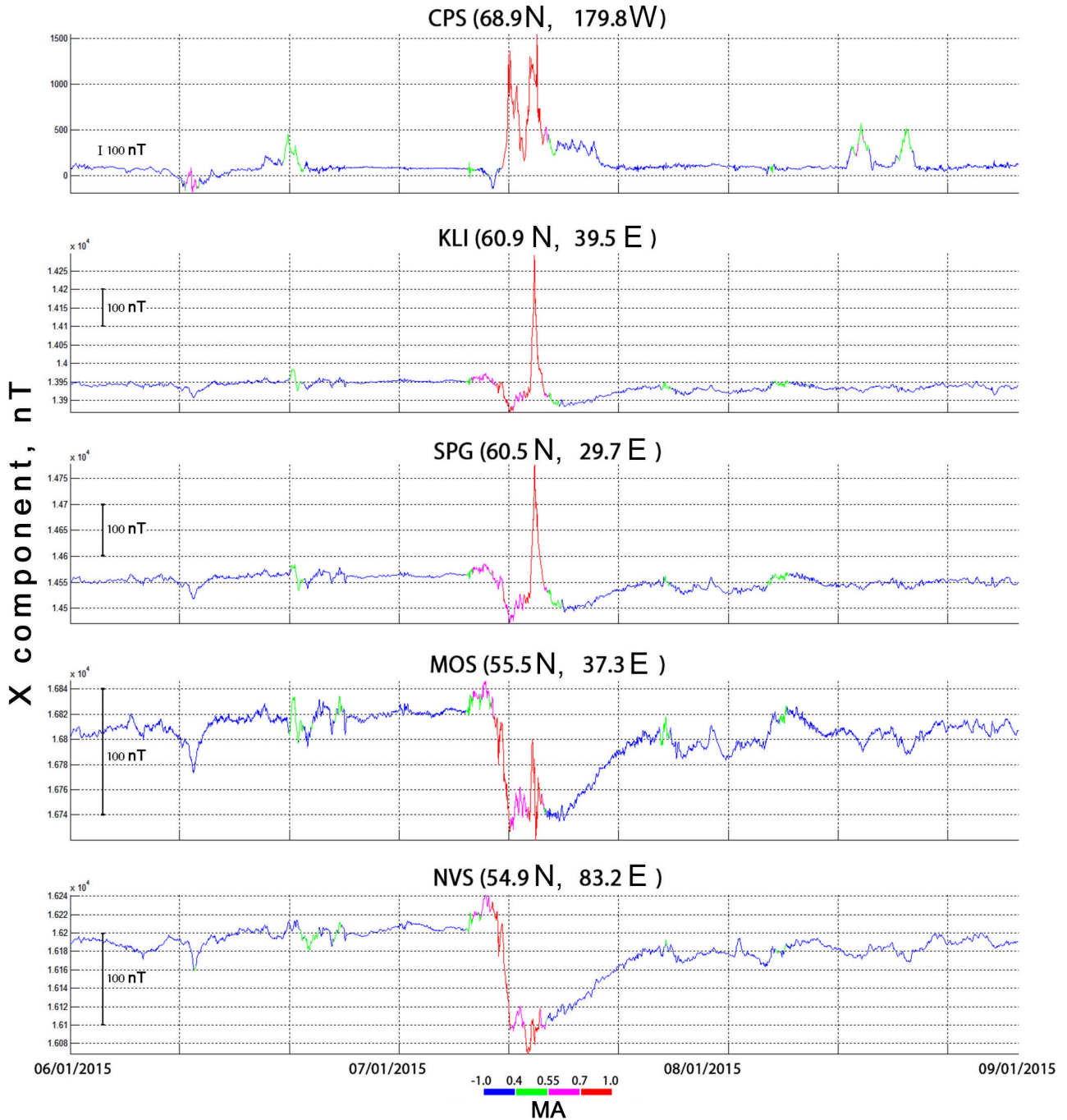
Figure 7. Recognition of geomagnetic activity before, during and after magnetic storm of January 7, 2015, using data of selected Russian geomagnetic observatories. Date format: dd/mm/yyyy.

Figure 8 SC time is indicated with grey strip. The SC was successfully recognized using MA as weakly anomalous and anomalous events in all magnetograms. Its typical simultaneous occurrence in all magnetic records over the world helps to separate it from other local signals estimated by the same magnitude of MA.

6. Continuous estimation of magnetic activity

MA allows to estimate geomagnetic activity simultaneously on a set of magnetograms, as it was shown above. One of its principal advantages is the ability to operate in real-time, as its time resolution depends on the initial data sampling rate only. In the case of INTERMAGNET observatories it equals to 1 minute. Frequently, observatory data transmission is delayed (30 minutes, 24 hours, etc.), which is usually dictated by observatory technical capacities. In this case, the recognition is performed with the same time delay. In the previous section we analyzed the MA application to different data segments of 1 day, 5 days and 1 month. The given examples prove the robustness of the MA being applied to measurements for different registration periods. A question of the adequate registration period selection (from the current moment to the past) for the
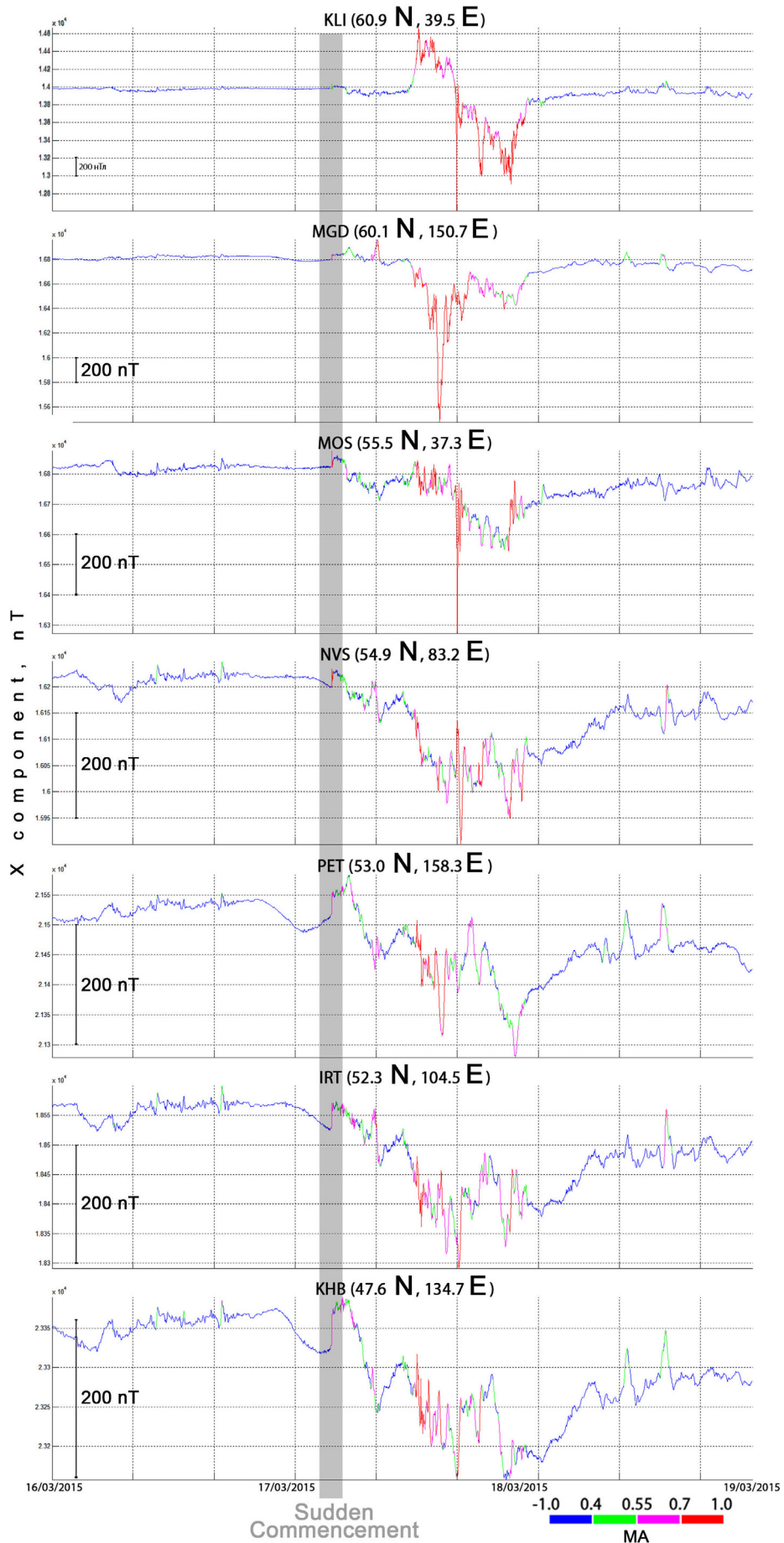
Figure 8. Simultaneous recognition of geomagnetic activity before, during and after geomagnetic storm on March 17, 2015, using data from Russian magnetic observatories. Time of the storm's sudden commencement is indicated with grey strip. Date format: dd/mm/yyyy.
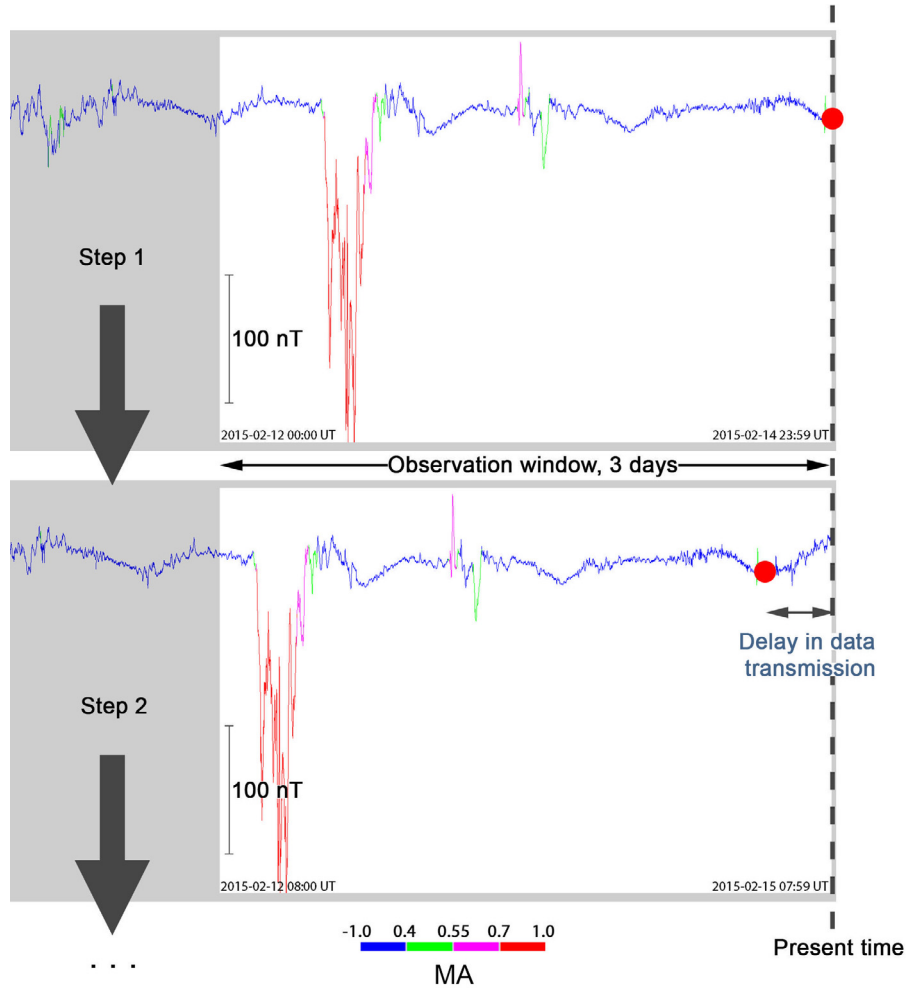
Figure 9. Sequential and continuous recognition of geomagnetic activity within 3-day registration period using MA by example of the data from "Cape Schmidt" (IAGA code CPS) observatory. Observation window shift between previous ('Step 1') and subsequent ('Step 2') steps of the algorithm operation corresponds to 8-hour delay of data transmission. Red point represents the timestamp of the most recently transmitted data at step 1 (right border of observation window). Its shift at step 2 is naturally caused by new 8-hour data input.

automated routine operation of the algorithm arises.

Let us remind that the straightening is calculated in a sliding time window Δ within the registration period T that is under consideration. Δ always includes 21 samples and its shift step is equal to 1 sample. The chosen size of time window Δ has turned to be optimal, as smaller size leads to higher MA values associated with insignificant fluctuations in original record (e.g., isolated short-period pulsations), whereas bigger window size results in low MA values associated with relatively significant, and thus important, but isolated fluctuations in original record (e.g., storm sudden commencement). The result of straightening is referred to the center of the window Δ. In real-time applications we apply $T = [t_c - 3 \text{ days}, t_c]$, where t_c is the latest moment of data transmission. The chosen observation window of 3 days corresponds to a typical time interval that contains onset, development, maximum intensity and decrease of the majority of geomagnetic disturbances. Figure 9 illustrates automated on-the-fly detection of increased geomagnetic activity in the north

component (X) record transmitted from the Cape Schmidt (CPS) observatory to the Russian-Ukrainian Geomagnetic Data Center (http://geomag.gcras.ru/). The figure shows, how two sequential 3-day periods from February 12, 2015, 00:00 UT, to February 14, 2015, 23:59 UT (upper panel, 'Step 1'), and from February 12, 2015, 08:00 UT, to February 15, 2015, 07:59 UT (lower panel, 'Step 2'), are analyzed by the algorithm. The time shift of 8 hours, illustrated with red point in upper and lower panels, is due to 8-hour time delay in data transmission from CPS observatory. As in the previous examples, here the initial record is split into four types of fragments as the result of the MA application: background (blue, $-1 \leq \mu(t) < 0.4$), weakly anomalous (green, $0.4 \leq \mu(t) < 0.55$), anomalous (purple, $0.55 \leq \mu(t) \leq 0.7$) and strongly anomalous (red, $0.7 \leq \mu(t) \leq 1$).

7. Recognition of geomagnetic activity caused by substorms

Parameters of the solar wind and IMF are likely the main sources of information on the forthcoming

geomagnetic disturbances on Earth, that are caused by the solar activity. These data are registered by satellite located in Lagrange point 1 in the Sun-Earth system. However, ground observations of the Earth's magnetic field also provide valuable information for recognition and analysis of magnetic activity, including substorms. Synchronous data transmission from several observatories enables monitoring of spatial distribution of geomagnetic activity and estimation of its intensity in different regions of the Earth. Let us demonstrate it by the example of the magnetic storm, that took place on November 8-11, 2004, using the developed GIS-based software [Kulchinsky et al. 2010, Soloviev et al. 2013, Gvishiani et al. 2014]. This software enables application of time-dependent MA to magnetograms of all observatories of the INTERMAGNET network for specified period in animated mode. As a result, one gets visualization of the MA value distribution, indicating storm-related disturbances, over the Earth's surface at each moment of time. In case of INTERMAGNET data, the time resolution of the software operation is 1 minute. The software also provides dynamic histogram of MA values and IMF vertical (Bz) component plot.

We selected four time stamps, when according to the MA an increased geomagnetic activity was observed at almost all observatories of the global INTERMAGNET network (Figure 10), while solar wind and IMF were undisturbed:

1) 18:54 UT, November 8, 2004;
2) 21:09 UT, November 8, 2004;
3) 00:24 UT, November 9, 2004;
4) 11:15 UT, November 9, 2004.

The solar wind parameters under consideration included: speed (V), density (n) and temperature (T). Besides, the IMF components (Bx, By, Bz) and geomagnetic planetary indices Ap and Dst were analyzed. Ap index is interchangeable with planetary Kp index: it reflects observed deviation magnitude in nT, that corresponds to dimensionless intensity defined by Kp index. Dst index reflects intensity of ring current encircling the Earth in the near-equatorial plane. Time variations of the mentioned parameters are given in Figure 11. The four selected time stamps are marked with vertical blue lines. All of them correspond to nearly constant values of speed, density and temperature of the solar wind. Variations of IMF components are close to zero and Ap and Dst values are low, that reflects decreased geomagnetic activity. However, at the given times MA values, calculated for H component, are close to 1 for majority of observatories (Figure 10).

We additionally studied time variations of AE geomagnetic index from November 8 to November 10, 2004, to explain the observed effect. This index is a quantitative measure of magnetic activity in auroral zone of the northern hemisphere (~60°-80°N). It reflects intensification of the ionospheric currents streaming along the boundary of the auroral oval (electrojets). Electrojets intensify during substorms and normally affect geomagnetic recordings obtained at the high-latitude stations and observatories. AE index plots are given in Figure 12.

The considered four time stamps coincided with the recovery phase of the storm, accompanied with large AE index values. It can be concluded that auroral electrojets were so intense that their magnetic variations were observed also at lower latitudes. Geomagnetic storms are always followed by substorms, that in turn often do not require any triggers in solar wind, since they originate in the magnetotail, where large amounts of energy are accumulated and released. Thus, MA demonstrates universality, being capable to detect geomagnetic anomalous events of different origin with the use of ground observations.

In the same manner we studied time stamps, which corresponded to undisturbed solar wind, IMF and small values of planetary indices and, at the same time, to globally increased geomagnetic activity during the magnetic storm of May 15, 2005. Again, it was concluded that the global magnetic disturbance, observed on the Earth, resulted from highly intensive auroral electrojets.

8. Conclusions

Space weather is defined by the solar activity change, affecting the processes in the Earth's magnetosphere, ionosphere and atmosphere, as well as on the Earth's surface. Performance and reliability of the various space and ground-based technological systems, human health and safety depend on the state of the space weather; its extreme events may cause malfunctions and accidents. Ground and space geomagnetic observations are the most accessible and, at the same time, informative means to monitor space weather conditions.

We developed a new indicator "measure of anomalousness" for evaluation of activity of geomagnetic record at a given time or interval. The MA is fully based on DMA approach, that has found numerous successful applications in volcanic activity monitoring [Zlotnicki et al. 2005], defining locations and characteristics of anomalous bodies in the crust using data of aero magnetic [Mikhailov et al. 2003] and gravitational surveys [Widiwijayanti et al. 2003], recognition of earthquake-prone areas [Gvishiani et al. 2013], monitoring of geomagnetic activity [Soloviev et al. 2013] and dynamic evaluation of global disturbance in real time [Kulchinsky et al. 2010], recognition of anthropogenic anomalies in 1-minute and 1-second magne-
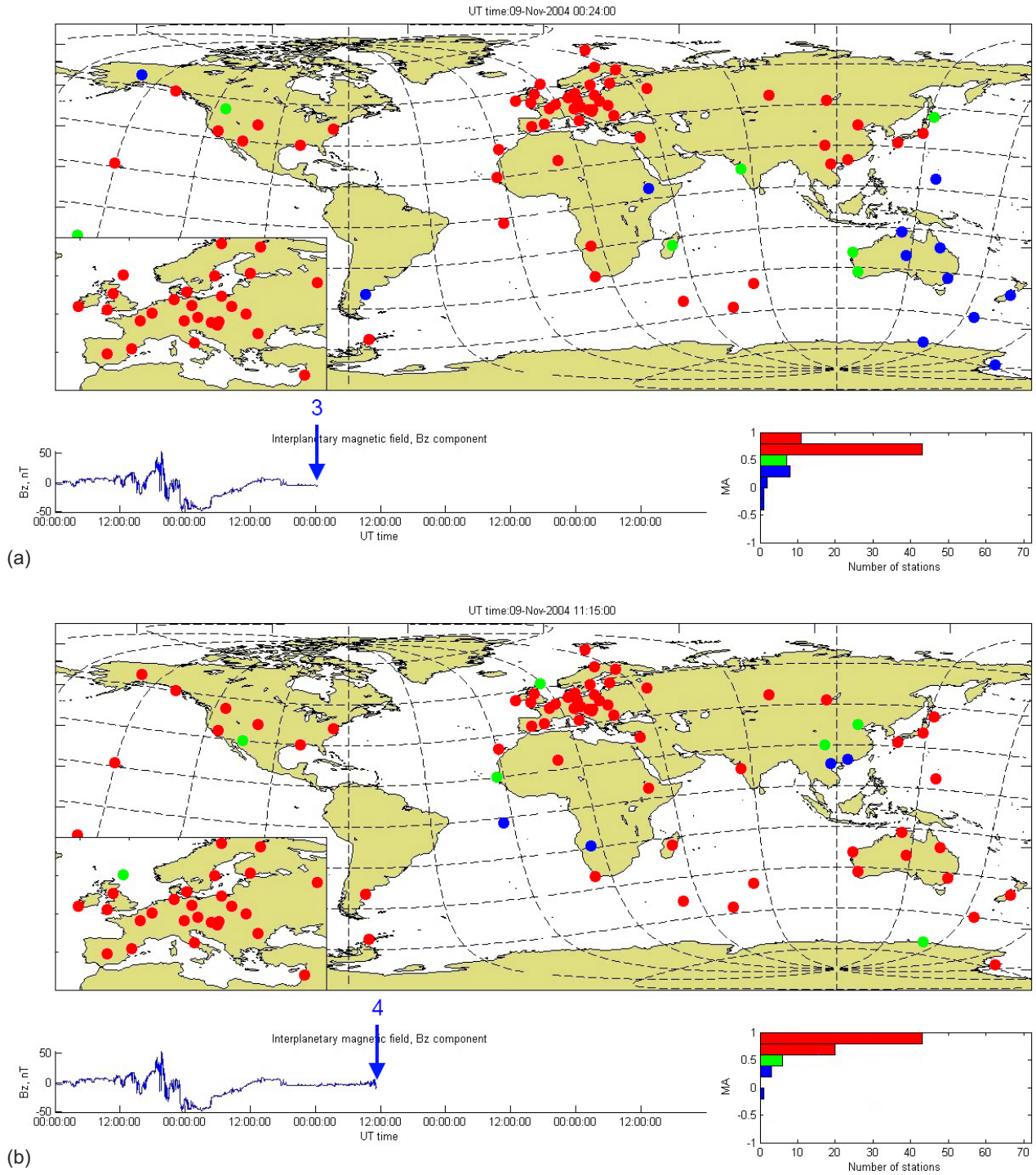
Figure 10. An example of the GIS-based software operation for global monitoring of MA, calculated using H component, over the globe. Dynamic histogram with MA value distribution depending on the number of observatories is given below on the right and dynamic plot with IMF Bz component is given below on the left. These two snapshots were taken at times 3 (a) and 4 (b) marked with blue arrow (similar snapshots are available at times 1 and 2).

tograms [Bogoutdinov et al. 2010; Soloviev et al. 2012a, 2012b] and others. Formal description of the MA involves the notion of fuzzy comparison. Adequate modelling of local activity is based on a set of "straightenings", that are different constructions of non-negative functionals of initial record. "Straightening" uplifts correspond to anomalous fragments of initial record.

As a result, the MA enables automated estimation of abnormality level either in the neighborhood of a given observatory or in a given region, using a set of observatories, as well as estimation of global distribution of magnetic disturbance. Being applied to recordings obtained at different latitudes, the MA provides estimation in a single scale $[-1, 1]$ and takes into account typical level of disturbance in a region and, thus, local features of the magnetic field behavior. Its time resolution is equal to sampling rate of initial record, comparing to most of the traditional indices that are
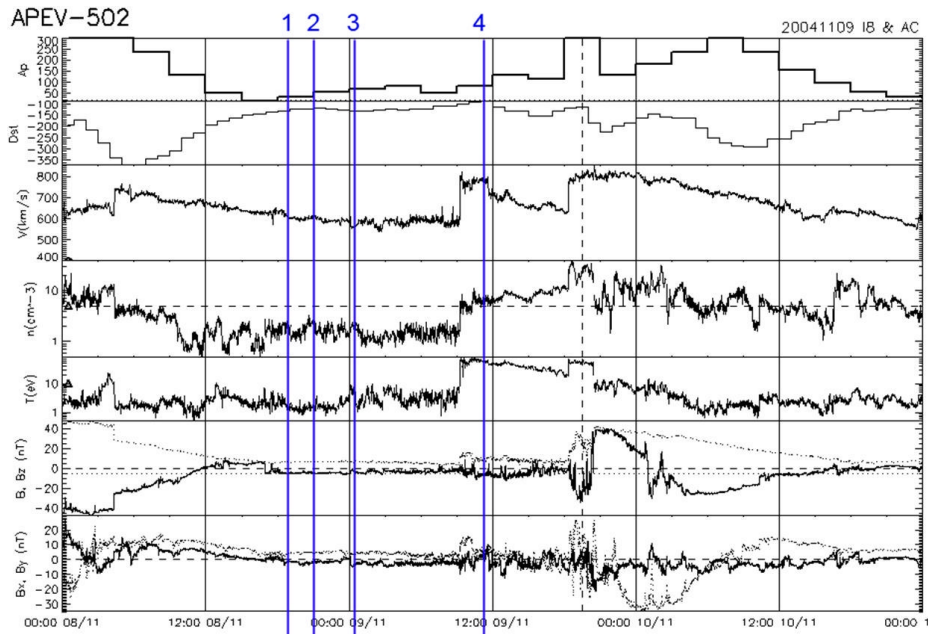
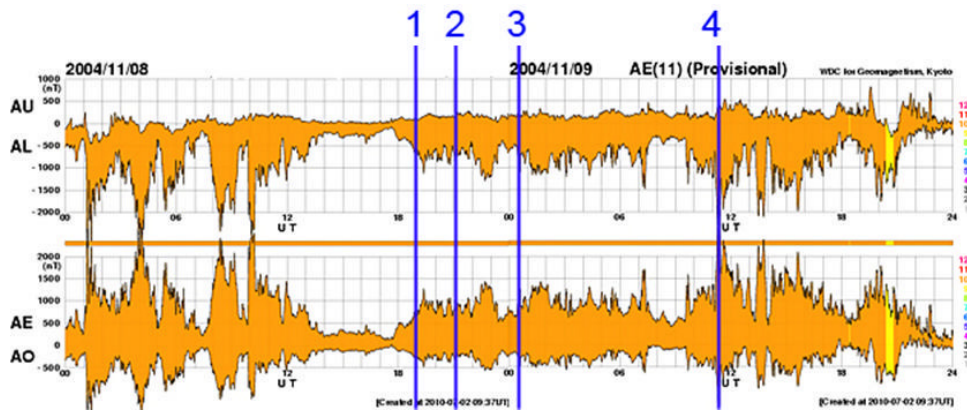**Figure 11.** Time variations of solar wind, IMF parameters and geomagnetic indices Ap and Dst from November 8 to November 11, 2004 (http://dbserv.sinp.msu.ru/apev/).

Figure 12. Time variations of AE index from November 8 to November 10, 2004 (http://wdc.kugi.kyoto-u.ac.jp/). The four selected time stamps are marked with vertical blue lines.

averaged over longer periods. This feature makes the MA valid in real-time applications, where time delay minimization is crucial.

The MA can be treated as a representative of a new generation in the family of K indices, that are traditionally calculated at observatories. Such indicators are aimed at estimation of disturbance level individually at a specific station or observatory, which seems to be more objective way for studying geomagnetic activity structure and particularly magnetic storms [e.g., Yakovchouk et al. 2012], comparing to the usage of spatially averaged characteristics (e.g., planetary indices such as Dst and Kp). Various tests have shown that coefficient of correlation between traditional K index and the MA, being artificially averaged over 3-hour intervals, is persistently in average 0.8-0.9. At the same time, the MA is much more sensitive to geomagnetic disturbances due to its higher time resolu-

tion and more sophisticated mathematical construction. Besides, the MA does not require time-consuming recognition and subtraction of Sq variation.

Our studies demonstrate diverse applications of the MA to geomagnetic records from numerous spatially distributed observatories during the periods of decreased and increased geomagnetic activity. A choice of optimal MA threshold value β_w to distinguish between background and anomalous events in magnetograms was justified. Defining a single threshold for all the observatories is enough for efficient detection of magnetic activity worldwide. The reason is that the MA represents relative measure of activity and always evaluates typical variations during quiet periods by lower values. This made it possible to introduce general graded scale of abnormality according to MA. The stability of the activity estimation using the selected graded scale was confirmed by all examples given in

the paper (Figures 5-9).

An essential feature of the developed MA is its ability to deal continuously with real time data streams. A chosen time window size of 3 days (from the current moment to the past) appears to be optimal for routine recognition and evaluation of geomagnetic activity. Examples of synchronous estimation of magnetic activity using multiple magnetograms were presented. They demonstrate, that MA provides an opportunity for continuous monitoring and tracking the distribution of disturbances over the globe, as well as studying the spatiotemporal structure of individual geomagnetic storms. That becomes possible by continuous processing of the whole set of available observations (e.g., around 140 observatories of INTERMAGNET network worldwide). It complies with the increasing necessity of simultaneous determination of the intensity of geomagnetic disturbances over the globe.

Using the MA we performed recognition and estimation of geomagnetic activity caused by magnetic storm on November 8-11, 2004, moderate magnetic storm of January 7, 2015, and the strongest magnetic storm of March 17, 2015, so far occurred during the current solar cycle. The result analysis showed potential possibility of the MA to be used for detecting precursors of strong magnetic disturbances using ground observations. It also argues for MA universality in monitoring of extreme geomagnetic events of different nature, including storms and substorms.

In addition to recognition of increased magnetic activity, the proposed MA can be dually used for defining quiet periods. In particular, this feature makes MA applicable for selecting quiet days and consequently determination of Sq variations for a specified observatory with a small time delay. Our studies, that are beyond the scope of this paper, have already shown high correlation rate between quiet days selected according to classical IAGA approach and MA. From this perspective, MA can be used for selection of data, obtained during expansive geomagnetic survey over periods of varying geomagnetic conditions.

Future plans include definition of full intensity indicators for recognized disturbed fragments, study of systematic latitudinal/longitudinal differences in the MA behavior and comparison of storms and substorms of different intensities according to MA using the global observatory network data.

Data and sharing resources

The results presented in this paper rely on data collected at magnetic observatories. We thank the national institutes that support them, INTERMAGNET for promoting high standards of magnetic observatory practice (http://www.intermagnet.org) and Russian-Ukrainian Geomagnetic Data Center (http://geomag.gcras.ru) for making the data freely available online.

Catalogue of Space Storms: http://dbserv.sinp.msu.ru/apev/.

Indices of Global Geomagnetic Activity: http://www.gfz-potsdam.de/en/kp-index/.

International Q-Days and D-Days: http://www.gfz-potsdam.de/en/section/earths-magnetic-field/data-products-services/kp-index/qd-days/.

International Real-time Magnetic Observatory Network: http://www.intermagnet.org.

Planetary K-index: http://www.swpc.noaa.gov/products/planetary-k-index.

Russian-Ukrainian Geomagnetic Data Center: http://geomag.gcras.ru.

World Data Center for Geomagnetism, Kyoto (Japan): http://wdc.kugi.kyoto-u.ac.jp/.

References

Bartels, J., N.H. Heck and H.F. Johnson (1939). The three-hour-range index measuring geomagnetic activity, Terrestrial Magnetism and Atmospheric Electricity, 44, 411-454.

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

Chambodut, A., A. Marchaudon, C. Lathuillère, M. Menvielle and E. Foucault (2015). New hemispheric geomagnetic indices   with 15 min time resolution, Journal of Geophysical Research - Space Physics, 120, 9943-9958; doi:10.1002/2015JA021479.

Della-Rose, D.J., J.J. Sojka and L. Zhu (1999). Resolving geomagnetic disturbances using "K-like" geomagnetic indices with variable time intervals, Journal of Atmospheric and Terrestrial Physics, 61, 1179-1194.

Geomagnetic data (2015). Geomagnetic data recorded at Geomagnetic Observatory Klimovskaya (IAGA code: KLI), Geophysical Center of the Russian Academy of Sciences, 2015; http://doi.org/10.2205/kli2011.

Geomagnetic data (2016). Geomagnetic data recorded at Geomagnetic Observatory Saint Petersburg (IAGA code: SPG), ESDB repository, Geophysical Center of the Russian Academy of Sciences, 2016;

doi:10.2205/SPG2012.

Gvishiani, A.D., S.M. Agayan and Sh.R. Bogoutdinov (2008a). Fuzzy recognition of anomalies in time series, Doklady Earth Sciences, 421 (1), 838-842.

Gvishiani, A.D., S.M. Agayan, Sh.R. Bogoutdinov, J. Zlotnicki and J. Bonnin (2008b). Mathematical Methods of Geoinformatics: III. Fuzzy Comparisons and Recognition of Anomalies in Time Series, Cybernetics and Systems Analysis, 44 (3), 309-323.

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

Gvishiani, A.D., M. Dobrovolsky, S. Agayan and B. Dzeboev (2013). Fuzzy-based clustering of epicenters and strong earthquake-prone areas, Environmental Engineering and Management Journal, 12 (1), 1-10.

Gvishiani, A.D., R. Lukianova, A. Soloviev and A. Khokhlov (2014). Survey of Geomagnetic Observations Made in the Northern Sector of Russia and New Methods for Analysing Them, Surveys in Geophysics, 35 (5), 1123-1154; doi:10.1007/s10712-014-9297-8.

Iyemori, T., T. Araki, T. Kamei and M. Takeda (1999). Mid-latitude Geomagnetic Indices "ASY" and "SYM" for 1999 (Provisional); http://swdcwww.kugi.kyoto-u.ac.jp/aeasy/asy.pdf.

Kataoka, R., D. Shiota, E. Kilpua and K. Keika (2015). Pileup accident hypothesis of magnetic storm on 17 March 2015, Geophysical Research Letters, 42, 5155-5161; doi:10.1002/2015GL064816.

Kulchinsky, R.G., E.P. Kharin, I.P. Shestopalov, A.D. Gvishiani, S.M. Agayan and Sh.R. Bogoutdinov (2010). Fuzzy logic methods for geomagnetic events detections and analysis, Russian Journal of Earth Sciences, 11, RE4003; doi:10.2205/2009ES000371.

Lincoln, J.V. (1967). Geomagnetic indices, In: S. Matsushita and W. H. Campbell (eds.), Physics of Geomagnetic Phenomena, vol. 1, Academic, San Diego, Calif., p. 67.

Love, J.J., and A. Chulliat (2013). An international network of magnetic observatories, Eos, Transactions AGU, 94 (42), 373-384; doi:10.1002/2013EO420001.

Mandrikova, O.V., I.S. Solovyev, V.V. Geppener, D.M. Klionskiy and R.T. Al-Kasasbeh (2013). Analysis of the Earth's magnetic field variations on the basis of a wavelet-based approach, Digital Signal Processing, 23 (1), 329-339; doi:10.1016/j.dsp.2012.08.007.

Mandrikova, O.V., I.S. Solovev and T.L. Zalyaev (2014). Methods of analysis of geomagnetic field variations and cosmic ray data, Earth, Planets and Space, 66, 148; doi:10.1186/s40623-014-0148-0.

Mayaud, P.N. (1980). Derivation, meaning and use of geomagnetic indices, Geophysical Monograph Series 22, AGU, Washington, D.C., 154 p.

Menvielle, M. (1990). About the derivation of geomagnetic indices from digital data, In: K. Kauristie, C. Sucksdorff and H. Nevanlinna (eds.), Proc. International Workshop on Observatory Data Acquisition and Processing, Geophysical Publications No. 15, Finnish Meteorological Institute, Helsinki, 117-126.

Menvielle, M., and A. Berthelier (1991). The K-derived planetary indices: description and availability, Reviews of Geophysics, 29 (3), 415-432.

Mikhailov, V., A. Galdeano, M. Diament, A. Gvishiani, S. Agayan, Sh. Bogoutdinov, E. Graeva and P. Sailhac (2003). Application of artificial intelligence for Euler solutions clustering, Geophysics, 68 (1), 168-180.

Rangarajan, G.K. (1989). Indices of geomagnetic activity, In: J.A. Jacobs (ed.), Geomagnetism, vol. 3, Academic Press, London, 323-384.

Siebert, M. (1996). Geomagnetic activity indices, In: W. Dieminger, G.K. Hartmann and R. Leitinger (eds.), The Upper Atmosphere - Data Analysis and Interpretation, Springer, Berlin, 887-911.

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

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

Soloviev, A., S. Bogoutdinov, A. Gvishiani, R. Kulchinskiy and J. Zlotnicki (2013). Mathematical Tools for Geomagnetic Data Monitoring and the INTERMAGNET Russian Segment, Data Science Journal, 12, WDS114-WDS119; doi:10.2481/dsj.WDS-019.

Soloviev, A., M. Dobrovolsky, D. Kudin, R. Sidorov (2015). Minute values of X, Y, Z components and total intensity F of the Earth's magnetic field from Geomagnetic Observatory Klimovskaya (IAGA code: KLI), Geophysical Center of the Russian Academy of Sciences; http://doi.org/10.2205/kli2011min.

Stankov, S., K. Stegen and R. Warnant (2011). K-type geomagnetic index nowcast with data quality control, Annals of Geophysics, 54 (3), 285-295; doi:10.4401/ag-4655.

Svalgaard, L., and E.W. Cliver (2007). Interhourly variability index of geomagnetic activity and its use in deriving the long-term variation of solar wind speed, Journal of Geophysical Research, 112, A10111; doi:10.1029/2007JA012437.

Widiwijayanti, C., V. Mikhailov, M. Diament, C. Deplus, R. Louat, S. Tikhotsky and A. Gvishiani (2003). Structure and evolution of the Molucca Sea area: constraints based on interpretation of a combined sea-surface and satellite gravity dataset, Earth and Planetary Science Letters, 215, 135-150.

Yakovchouk, O.S., K. Mursula, L. Holappa, I.S. Veselovsky and A. Karinen (2012). Average properties of geomagnetic storms in 1932-2009, Journal of Geophysical Research, 117, A03201; doi:10.1029/2011 JA017093.

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

*Corresponding author: Anatoly Soloviev,
Geophysical Center, Russian Academy of Sciences,
Moscow, Russian Federation; email: a.soloviev@gcras.ru.

Annex 1. Calculation of MA

Consider evenly sampled data points with sampling interval h and a finite time series $y = \{y_t = y(th)\}$, where $t \in \mathbb{N}$ is measurement (natural) number. y is defined on an interval (registration period) $T \subset \mathbb{R}_h^+$, where $\mathbb{R}_h^+$ is a set of positive real numbers with discretization step h. We introduce a local scan parameter $\Delta > 0$, that is a multiple of h: $\Delta = kh$, $k \in \mathbb{N}$. We define a local scan fragment of a record y with its center at $th \in T$ as a segment of the record:

$$\Delta^t y = \{y_{t-k}, ..., y_t, ..., y_{t+k}\}.$$

Let us define the superposition $t \to \Delta^t y \to \Phi(\Delta^t y)$ that we denote $\Phi_y(t)$ and call it the "straightening" of y on the basis of Φ; the mapping Φ would be the "straightening" functional. In the present work as we use functional "Length", which is defined as follows:

$$L(\Delta^t y) = \sum_{j=t-k}^{t+k-1} |y_{j+1} - y_j|. \qquad (1)$$

As a result,
$$\Phi_y(t) = L(\Delta^t y) \geq 0. \qquad (2)$$

Fuzzy comparison $n(a, b)$ of real numbers a and b measures the degree of superiority of b over a in the sign-alternating scale of the segment $[-1, 1]$: $n(a,b) \in [-1, 1]$, where $n(a,b)$ close to 1 means that b is significantly larger than a, $n(a,b) = 0$ means that b is equal to a and $n(a,b)$ close to -1 means that b is significantly smaller than a [Gvishiani et al. 2008b]. Most of the DMA algorithms including the present one use the construction

$$n(a,b) = \frac{b-a}{\max(a,b)} \qquad (3)$$

that is defined only for non-negative a, b. It seems to be sufficient, because a record is processed by means of the analysis of its "straightening", which is always non-negative (2).

For a fuzzy comparison $n(A, b)$ of an arbitrary real number $b \in \mathbb{R}$ with an arbitrary finite subset $A \subset \mathbb{R}$, $A = \{a_i\}$, $i = 1, \dots N$, we use a binary construction:

$$n(A,b) = \frac{\sum_{i=1}^{N} n(a,b)}{N} \in [-1,1]. \qquad (4)$$

Note that $n(A, b)$ is not linear, but monotonically increasing function with the fixed A and increasing b.

The measure of anomalousness (MA) $\mu(t) \in [-1,1]$ at point $t \in T$ is defined as a fuzzy comparison (4) of the image $\operatorname{Im}\Phi_y = \Phi_y(T)$ of "straightening" Φ_y (see (1) and (2)), considered over the whole registration period T,

with its value $\Phi_y(t)$ at point t:

$$\mu(t) = n(\operatorname{Im}\Phi_y, \Phi_y(t)). \qquad (5)$$

Using MA (5) we can now classify initial recordings depending on their activity rate. An example of such classification is given below.

Let β_s, $\beta_w \in [-1,1]$ be the specified levels of strong and weak abnormality in the range of MA, respectively. Then α_s and α_w, which are solutions of $n(\operatorname{Im}\Phi_y, \alpha_s) = \beta_s$ and $n(\operatorname{Im}\Phi_y, \alpha_w) = \beta_w$, respectively (4), will be strong and weak abnormality levels in the range of $\operatorname{Im}\Phi_y$. Thus, we assume that

1. t is background point, if $\mu(t) < \beta_w$, which is equivalent to $\Phi_y(t) < \alpha_w$;

2. t is anomalous point, if $\mu(t) \geq \beta_s$, which is equivalent to $\Phi_y(t) \geq \alpha_s$;

3. t is potentially anomalous point, if $\mu(t) \in [\beta_w, \beta_s)$, which is equivalent to $\Phi_y(t) \in [\alpha_w, \alpha_s)$.