

RESEARCH PAPER

Automated Hardware and Software System for Monitoring the Earth's Magnetic Environment

Alexei Gvishiani, Anatoly Soloviev, Roman Krasnoperov and Renata Lukianova

Geophysical Center of the Russian Academy of Sciences, Molodezhnaya St., 3, 119296, Moscow, Russia

Corresponding author: Roman Krasnoperov (r.krasnoperov@gcras.ru)

The continuous growth of geophysical observations requires adequate methods for their processing and analysis. This becomes one of the most important and widely discussed issues in the data science community. The system analysis methods and data mining techniques are able to sustain the solution of this problem. This paper presents an innovative holistic hardware/software system (HSS) developed for efficient management and intellectual analysis of geomagnetic data, registered by Russian geomagnetic observatories and international satellites. Geomagnetic observatories that comprise the International Real-time Magnetic Observatory Network (INTERMAGNET) produce preliminary (raw) and definitive (corrected) geomagnetic data of the highest quality. The designed system automates and accelerates routine production of definitive data from the preliminary magnetograms, obtained by Russian observatories, due to implemented algorithms that involve artificial intelligence elements. The HSS is the first system that provides sophisticated automatic detection and multi-criteria classification of extreme geomagnetic conditions, which may be hazardous for technological infrastructure and economic activity in Russia. It enables the online access to digital geomagnetic data, its processing results and modelling calculations along with their visualization on conventional and spherical screens. The concept of the presented system agrees with the accepted 'four Vs' paradigm of Big Data. The HSS can increase significantly the 'velocity' and 'veracity' features of the INTERMAGNET system. It also provides fusion of large sets of ground-based and satellite geomagnetic data, thus facilitating the 'volume' and 'variety' of handled data.

Keywords: Earth's magnetic field; absolute magnetic observations; geomagnetic activity; monitoring system; data mining; system analysis; big data

Introduction

Russian geomagnetic observatories and the low orbiting satellites of the European Space Agency (ESA) Swarm mission carrying the high-precision magnetometers provide constantly growing volumes of high-quality information that forms a basis for complex monitoring of the Earth's electro-magnetic environment. Better understanding of the underlying physical processes is extremely important from both fundamental research and applied field perspectives. For example, the processes caused by solar-terrestrial interaction considerably affect the functioning of modern technological infrastructure. On the other hand, continuous growth of the Earth's magnetic environment observations and big data phenomenon in general require adequate methods for their processing and analysis, otherwise such data can't be justified and are simply useless. Such demand becomes one of the most important and widely discussed issues among data science experts and multi-disciplinary communities. Recent advances in system analysis approach and data mining techniques address this challenge to a certain extent. In particular, the methods are universal enough to be applied to any time-series (the typical representation of the dynamic geophysical parameters) of any volume in the pattern recognition problem. Such methods are aimed at modelling the logic of an interpreter, who studies data 'by eye', and transferring this knowledge in the domain of huge data sets with their subsequent automated analysis and proper decision-making. In the field of geomagnetic studies, they enable formalized

multi-criteria detection of extreme magnetic events of different origin and dually normal state of magnetic environment. Complex and timely analysis of observatory and satellite data allows for modelling the structures of the internal and external parts of the Earth's magnetic field with a short time delay.

Nowadays, all existing geomagnetic data centers are divided into four categories: depositaries (e.g., Data centers of the World Data System (WDS) of the International Council for Science (ICSU)), advanced depositaries (e.g., SuperMag and Space Physics Interactive Data Resource (SPIDR)), centers for real-time monitoring (e.g., NOAA's National Centers for Environmental Information (NCEI)) and geomagnetic information nodes (International Real-Time Magnetic Observatory Network (INTERMAGNET)). The first two types of data centers aim at archiving historical data, usually in the form of flat files. Monitoring data centers aim at real-time acquisition of operative information on various electromagnetic parameters of the Earth's environment for providing interested parties with timely forecasts; the decision-making is usually carried out by human beings. INTERMAGNET Geomagnetic Information Nodes (GINs) collect data from the highest-quality observation network comprising around 140 observatories worldwide. Historically magnetograms are stored as daily/monthly text/binary files and distributed on DVDs and through the web-site. The high quality of the resulting measurements stems from manual and scrupulous data processing, carried out by observatory and GIN operators. None of the mentioned data centers provides coordinated storage of ground and satellite observations and continuous automated intellectual analysis of the data streams, available for the global community.

This paper presents an innovative analytical/engineering solution and corresponding holistic hardware/software system (HSS) developed for efficient retrieval, storage, processing, and intelligent analysis of geomagnetic data with automatization of the majority of the data management processes. It represents a core of the Russian national observatory network and provides coordinated storage of ground observations with magnetic measurements obtained by ESA Swarm satellite mission using advanced data storage system. The raw 'preliminary' geomagnetic data provided by ground observatories contain anthropogenic noise and represent relative variations of the magnetic field. Thus, they require filtering and rectification to obtain the status of 'definitive' magnetograms, which completely describe magnetic field as full vector time variations. The existing practice of definitive geomagnetic data creation shows its highly time- and labor-consuming status. Furthermore, certain operations are still performed manually. On the contrary, the HSS consists of several integrated hardware and software units combined within a single logical and operational structure. The system includes software modules for automated filtering of the observatory data from technological noise and data verification in compliance with the highest INTERMAGNET standards. The latter enables the production of quasi-definitive and subsequent definitive data.

The HSS provides also sophisticated automatic detection and classification of extreme geomagnetic conditions, which may be hazardous for technological infrastructure and economic activity within the regions of Russia. The most innovative system analysis methods and data mining algorithms that involve artificial intelligence elements are integrated in the HSS operation. Besides, the HSS enables the online access to digital geomagnetic data and its processing results along with its visualization on conventional and spherical screens.

Description of the HSS is presented in this paper in the tree-like manner. Initially, general functionality and major components of the HSS are discussed. The following description clarifies the performance features of the HSS subsystems and subcomponents.

Absolute Observations of the Magnetic Field

The sources of the Earth's quasistationary magnetic field are located inside the body of our planet (**Figure 1a**). Internal geomagnetic field changes on a large timescale. The sources of external geomagnetic field are located in the near-Earth space. These sources are complex, variable, three-dimensional systems of electric currents flowing within the magnetosphere and ionosphere of the Earth (**Figure 1b**).

The magnetic effect of magnetosphere-ionospheric currents is being observed as sporadic deviations of geomagnetic elements from the quiet level at the timescale from seconds to tens of hours. The quiet level is formed by diurnal and seasonal variations, which result from variations of the ionization rate and conductivity of ionosphere under solar UV-radiation. Relatively short fluctuations, induced by current systems of magnetic storms, substorms, and other nonstationary phenomena, caused by solar activity, form geomagnetic activity. A high level of geomagnetic activity and rapid geomagnetic variations may be dangerous for modern technological systems. Monitoring of the geomagnetic field and detection of the extreme geomagnetic phenomena is one of the most significant problems of operational control and reliable performance of modern

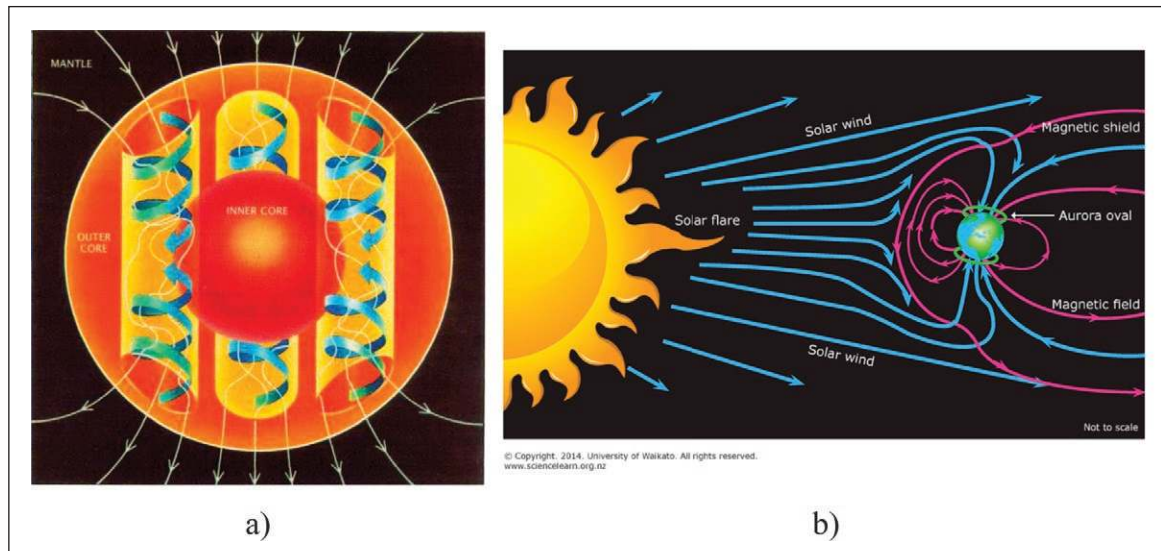


Figure 1: **a)** Internal sources of the Earth's magnetic field (the axis of rotation is vertical and centered in the image) (Bloxham & Gubbins 1989). **b)** Solar wind plasma and the Earth's magnetosphere (not in scale) (University of Waikato 2014).

technological infrastructure. Long-term observations of internal field evolution are of great importance for better understanding the causes of its global changes.

Analysis of the Earth's magnetic field is based on continuous measurements of the components and module of the three-dimensional magnetic field vector (\mathbf{B}), which is proportional to the vector of magnetic field strength (\mathbf{H}): $\mathbf{B} = \mu_0 \mathbf{H}$, where μ_0 – magnetic permeability.

The instantaneous geomagnetic field vector \mathbf{B} is completely described by three independent components. With reference to the preferable coordinate system (Cartesian, cylindrical or spherical), the components are: X – North component, Y – East component, Z – vertical component (positive downwards), H – horizontal component, D – magnetic declination angle (positive east of the North), I – magnetic inclination angle (positive below horizontal), F – geomagnetic field total intensity (module of the geomagnetic field vector, $F = |\mathbf{B}|$). Relations between the components of the geomagnetic field vector are given in **Figure 2**.

Continuous measurements of the geomagnetic field vector are performed by means of magnetometers, installed at observatories and stations. Modern geomagnetic observatories are globally distributed sophisticated facilities that provide accurate monitoring of the geomagnetic field in fixed locations within significant timespans. They provide determination of both secular and short-period variations (Love 2008; Matzka et al. 2010). The most important coordinated international network of geomagnetic observatories, providing geomagnetic data in compliance with the highest standards of quality, is the INTERMAGNET (Mandea & Papitashvili 2009). This network possesses around 140 observatories worldwide (as of September 2016) that transfer geomagnetic data to five GINs (INTERMAGNET 2013b).

The observatory instrumentation is being constantly improved. Now its optimal composition includes: scalar magnetometers (measuring the geomagnetic field total intensity F), fluxgate 3-axial variometers (measuring variations of the geomagnetic vector components), and fluxgate declinometers/inclinometers (measuring absolute values of D and I , required for variational data baseline correction). The final product provided by geomagnetic observatories is the 1-minute (at some observatories 1-second) geomagnetic variations referred to the absolute measurements (Jankowsky & Sucksdorff 1996; Newitt, Barton & Bitterly 1996; Rasson 2007).

In recent years considerable progress has been achieved in the area of ground geomagnetic observations in Russia. One of the key roles in this activity is played by the Geophysical Center of the Russian Academy of Sciences (GC RAS). Several existing geomagnetic observatories have been upgraded to the INTERMAGNET standards – 'Arti' (ARS), 'Saint Petersburg' (SPG) (ESDB 2016), 'Bor/Podkamennaya Tunguska' (POD) (Gvishiani & Lukianova 2015; Kaftan & Krasnoperov 2015; Kaftan, Krasnoperov & Tertyshnikov 2015), and a brand new 'Klimovskaya' (KLI) observatory (ESDB 2015; Krasnoperov, Sidorov & Soloviev 2015; Soloviev et al. 2016). In April 2016 the 'Saint Petersburg' observatory was officially assigned the status of INTERMAGNET observatory (INTERMAGNET 2016). The map of the Russian geomagnetic observatory network is given in **Figure 3**.

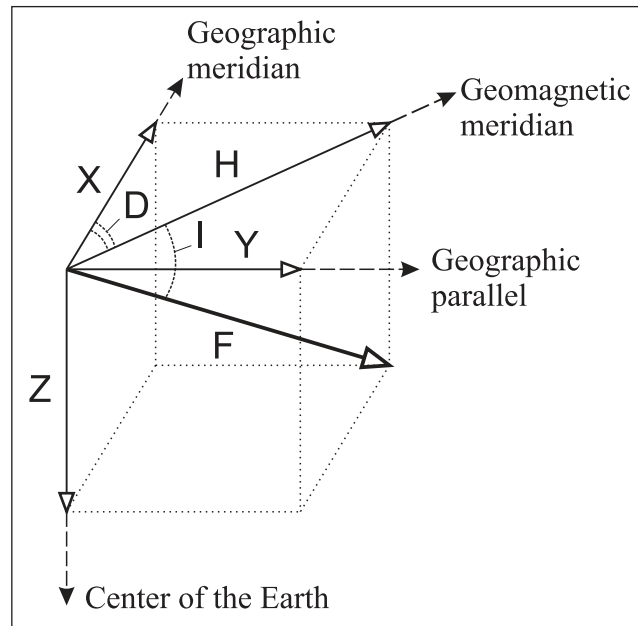


Figure 2: Decomposition of the geomagnetic field vector using the Cartesian (X, Y, Z), cylindrical (D, H, Z) and spherical (D, I, F) coordinate systems (Soloviev et al. 2013b).

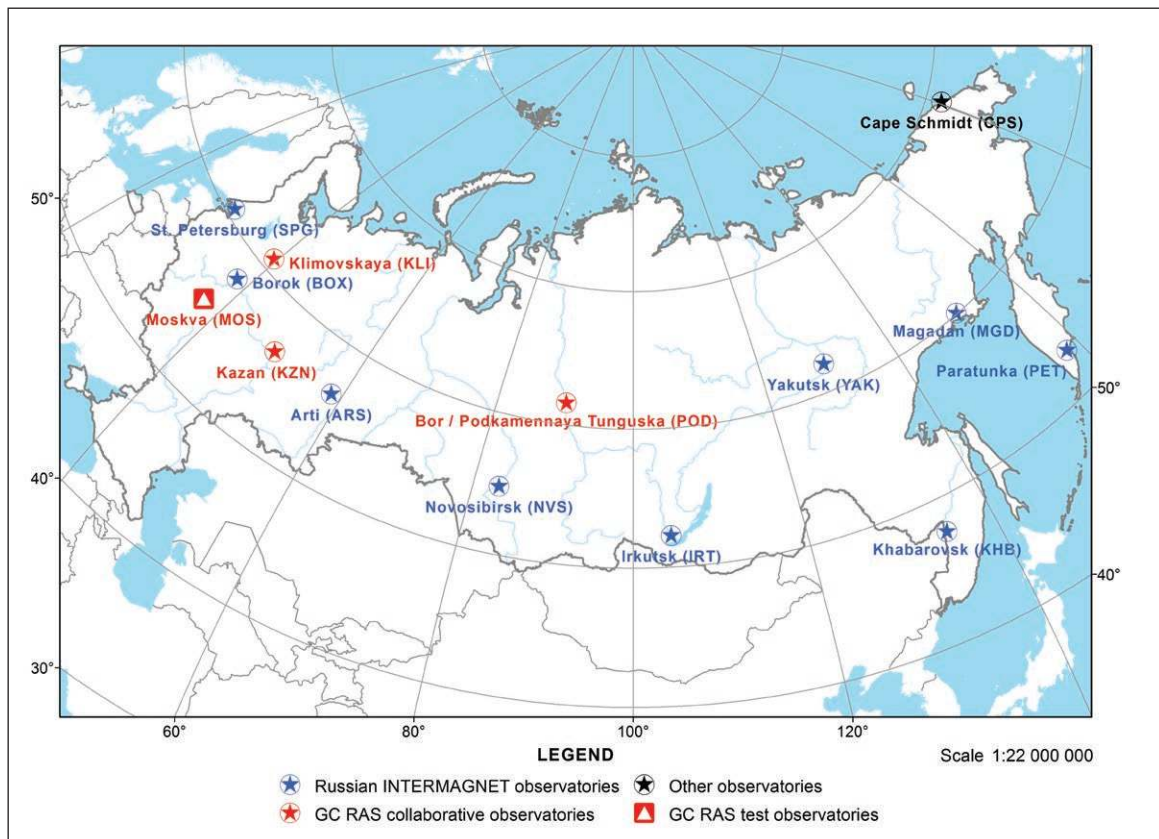


Figure 3: Map of the Russian geomagnetic observatory network. The 14 observatories that transfer geomagnetic data to the HSS: blue stars – INTERMAGNET observatories, red stars – GC RAS collaborative observatories, black stars – other observatories. The ‘Moscow’ observatory is used as a testing ground for educational and experimental purposes.

Operational geomagnetic data, transferred by the INTERMAGNET observatories via telecommunication lines, have the status of *preliminary data*. These raw magnetograms have not been corrected for baseline variations. They also may contain spikes (artificial disturbances) or missing values. The preliminary data are available for users with a short time delay.

The observatory data that have undergone the process of baseline correction, spike removal, and filling in the gaps (where possible), are assigned the status of *definitive data*. Preparation of the definitive data is a complicated time- and labor-consuming procedure that is performed mostly manually by skilled magnetologists at the observatories or GINs. The definitive data for a given INTERMAGNET observatory are prepared annually with a delay up to several years (St-Louis et al. 2012).

To overcome this problem several years ago a new type of observatory geomagnetic data was introduced – the *quasi-definitive data* (Peltier & Chulliat 2010). The quasi-definitive magnetograms are corrected using temporary baselines, and all spikes and gaps are also eliminated. The difference between the definitive and quasi-definitive magnetograms should be within 5 nT (Reay et al. 2011). Preparation of the quasi-definitive data normally is also based on manual data processing and, thus is implemented only at some observatories. The HSS is capable to provide quasi-definitive data in automatic mode with a minimal delay. In this aspect, the HSS is an instrument that can significantly increase the 'velocity' of the INTERMAGNET data provision.

Global and uniform coverage with geomagnetic measurements are also obtained by LEO (low Earth orbit) satellites. The latest satellite mission aimed to perform monitoring of the Earth's magnetic field is the Swarm mission, launched in November 2013 within the ESA Earth Explorer Opportunity Program. The main objective of the mission is to measure the magnetic signals that stem from the Earth's core, mantle, lithosphere, oceans, ionosphere and magnetosphere.

The Swarm constellation comprises of three identical satellites in two different near-polar orbital planes, which provide the global coverage. Two satellites (Swarm A and C) were launched in a similar orbital plane with inclination of 87.4° . This pair of satellites flies at a mean altitude of 450 km. The third satellite (Swarm B) was launched in a circular orbit with inclination of 88° and a mean altitude of 530 km. The orbital plane of the Swarm B satellite has a continuous drift. In four years after the start of the mission it will cause the difference of nine hours in local time with the Swarm A and C satellites (**Figure 4**).

The payload of the three satellites is completely identical. It includes: Vector Field Magnetometer (VFM), Absolute Scalar Magnetometer (ASM), Electric Field Instrument (EFI), and auxiliary equipment (accelerometer, GPS receivers, star trackers, laser retro reflectors). The geomagnetic vector measurements (with 1 Hz and 50 Hz frequency) are complemented by precise navigation, accelerometer, plasma and electric field observations (Haagmans, Bock & Rider 2013; Olsen et al. 2013).

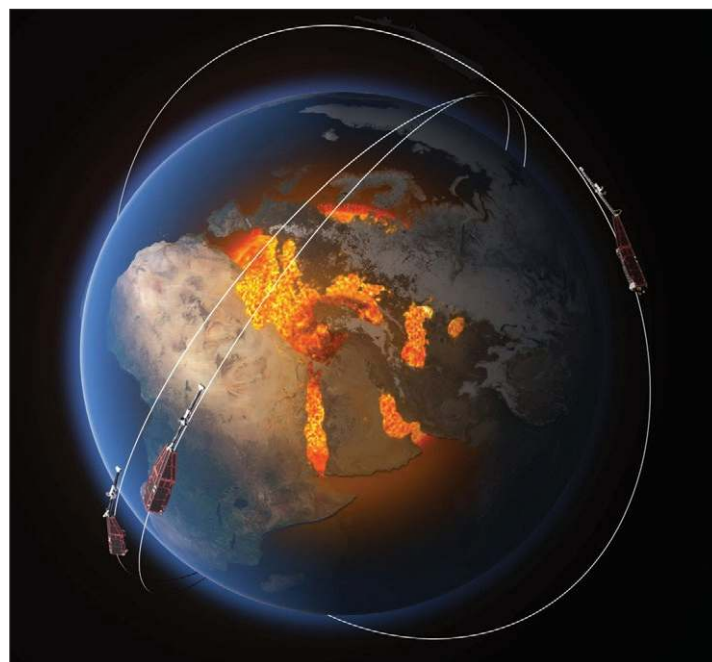


Figure 4: The Swarm constellation (not in scale) (Haagmans, Bock & Rider 2013).

Incorporation of the Swarm data into the HSS makes it a unique cutting-edge instrument for coordinated management and analysis of on-ground and satellite geomagnetic data. Thus, providing fusion of information from fundamentally different sources and broadening the HSS areas of applicability.

Scope of the HSS

The HSS forms a basis for Geomagnetic Analytical Data Center of the Russian INTERMAGNET segment. Advantages of the HSS rest upon effective implementation of amassed experience in geophysical monitoring and intellectual data analysis. The HSS is based on the unified information and telecommunication modular infrastructure that provides data transfer, storage, and multi-criteria processing. The HSS infrastructure remains open to installation of additional software components.

Storage of ground and satellite geomagnetic data (initial and processed) in the HSS is organized within a relational geomagnetic database (DB). It is a significant advantage in comparison to existing INTERMAGNET GINs, which mainly utilize traditional file databases.

The HSS innovative analytical unit is based on formalized integration of recognized knowledge and experience, gained by data experts in the area of recognition and studying of natural extreme events and anthropogenic disturbances on magnetograms. Elements of artificial intelligence are involved in corresponding algorithms that are incorporated as natural parts of the HSS.

The HSS is aimed to do the following tasks: retrieval and systematization of initial ground and satellite observations of the Earth's magnetic field; automated filtering of observatory data from technological noise and data verification in compliance with the INTERMAGNET standards; detecting, classification and coding of the data on extreme geomagnetic events, hazardous for technological infrastructure and economic activity; modelling calculations (hourly interpolation of total intensity deviations over Russia, operative modelling of Sq daily variations for observatories, modelling of geomagnetic effects of large-scale atmospheric dynamics); providing interactive access to initial data, information on extreme events, and modelling results; visualization of the geomagnetic data.

The data flow within the HSS and the system's functions are schematically presented in **Figure 5**.

The input HSS magnetograms are divided into three groups: measurements automatically and continuously recorded and transmitted by the observatories in quasi real-time mode, absolute measurements of geomagnetic field, manually carried out at the observatories, and operational satellite magnetic observations.

Operational data, registered at the Russian geomagnetic observatories, are continuous time series of variations of the three components and total intensity of the geomagnetic field vector with UTC time-synchronization, recorded with 1-minute (optionally, 1-second) sampling rate. Operational data are transmitted in one of the three formats: IMF1.23 (INTERMAGNET 2013a), IAGA-2002 (V-DAT 2011) or Mingeo (MinGeo Ltd. 2016). Depending on the telecommunication capabilities, data are transferred to the HSS from the observatories via e-mail or FTP-protocol with a certain delay (from 10 minutes up to 2 days).

Absolute observations are performed several times per month and provided to the HSS as text files or using the web interface. After the measurement results are uploaded, the variometer baseline values and absolute values of the magnetic field components are automatically calculated and recorded into the HSS DB.

Continuous recordings, carried out by the three satellites of the Swarm constellation (Friis-Christensen, Lühr & Hulot 2006), represent time series of physical parameters observations in the near-Earth environment. They include the measurements of the full geomagnetic field vector components, registered with 1 Hz sampling rate and stored in the binary Common Data Format (CDF). Every time new satellite data are available, geomagnetic field recordings are automatically extracted from the Swarm file storage and stored in the HSS DB. The complete content of the HSS-acquired 1 Hz satellite data is the following (Haagmans, Bock & Rider 2013): values of the three orthogonal components of the geomagnetic field vector in the variometer reference frame; values of the three orthogonal components of the geomagnetic vector in the NEC (North-East-Center) reference frame; values of the geomagnetic vector total intensity; UTC referencing; geographic coordinate referencing of the registered values. The HSS output parameters, provided to the end user, include the raw data (preliminary magnetograms) and results of their processing (quasi-definitive magnetograms) (**Figure 5**).

The results of the automated data processing and analysis include: recognized anthropogenic disturbances; absolute values of geomagnetic field; adopted baseline values; baseline adjusted and verified (quasi-definitive) data; geomagnetic events detected in preliminary data.

The end user is also provided with the results of modelling calculations, performed by the corresponding unit of the HSS. Modelling results include maps of interpolated hourly deviations of the geomagnetic vector

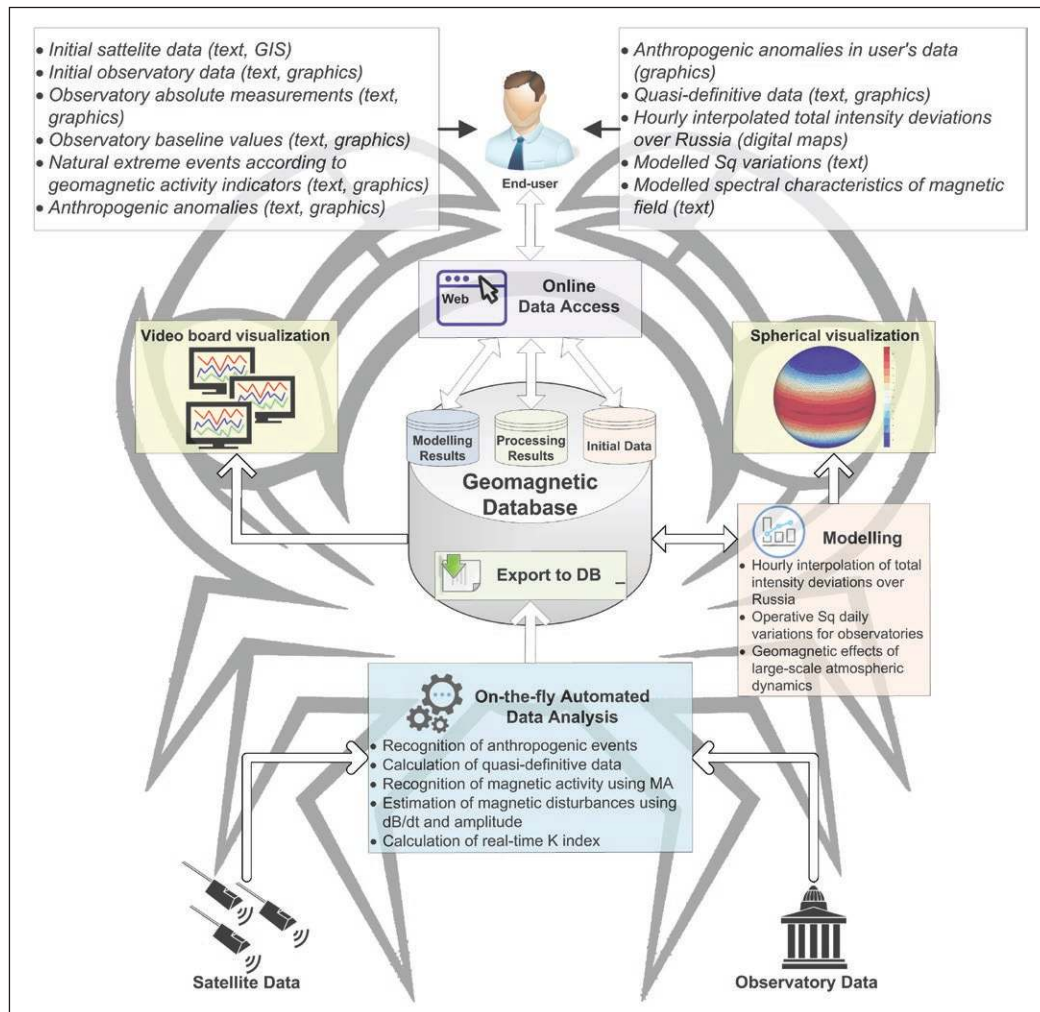


Figure 5: Data flow within the HSS.

total intensity over the territory of Russia, values of Sq variation, calculated in near real time for each observatory, and spectral characteristics of geomagnetic data time series.

All the data (magnetograms and model calculation results), accumulated in the HSS, are available to users through a web interface. The requested data are provided both in digital and graphical formats (**Figure 5**).

HSS Software and Hardware Components

Software components of the HSS are divided into six units: 'Data retrieval', 'Data export into the database', 'Data analysis', 'Model calculations', 'Data access', 'Data visualization'.

The 'Data retrieval' unit contains software modules providing retrieval of operational data and absolute measurements from the observatories, as well as the satellite operational data. Accordingly, the 'Data export into database' unit consists of the software modules for uploading the observatory and satellite data, into the HSS geomagnetic DB.

'Data analysis' software modules are aimed at on-the-fly processing and analysis of the data continuously transmitted to the HSS. These modules perform: automated recognition of anthropogenic disturbances in magnetograms, calculation of the quasi-definitive data, recognition of anomalous geomagnetic events using measure of anomaly (MA), estimation of the rate of change (dBdt) and the amplitude (Amp) of the geomagnetic events and operational calculation of K-index (Kind) of geomagnetic activity.

The 'Model calculations' unit provides: interpolation of hourly deviations of the geomagnetic vector total intensity over the territory of Russia; calculation of Sq variation values; and modelling of geomagnetic data time series' spectral characteristics. This unit performs calculations every time when new data become available in the HSS.

The 'Data access' unit consists of the software modules, which provide online access to geomagnetic data, stored in the HSS DB. The end user is given access to the results of geomagnetic observations: initial

(preliminary magnetograms) and processed (quasi-definitive and definitive magnetograms) observatory data, and Swarm observations. The results of modelling calculations that are stored in the HSS DB are made available with this unit, as well.

Finally, the 'Data visualization' unit is designated for visual representation of HSS operation results. It provides the following capabilities: online data visualization on the Web, quasi real time data visualization on a video board and visualization of spatial characteristics of the Earth's magnetic field using a spherical display.

The hardware components of the HSS include four servers (Mail Server, FTP-Server, DB-Server, and Web-Server), a video board with a controller for geomagnetic data visualization, and a system administrator's PC.

The scheme of the HSS components interaction is shown in **Figure 6**.

The Mail Server is required for receiving e-mail messages from the observatories that contain fragments of continuous recordings with their subsequent transmission to the FTP-Server. The latter is designated for storing geomagnetic data files in their initial form and hosts software for retrieval of observatory and satellite data via FTP protocol. The FTP-Server also hosts a backup copy of the geomagnetic DB, which is regularly synchronized with the DB's primary copy in automatic mode.

The Database Server (DB-Server) is designated for storing the observatory and satellite observations (preliminary magnetograms) along with the data processing results (quasi-definitive magnetograms). It stores the data using relational database management system (DBMS). The DB-Server hosts the software modules of the 'Data analysis' and 'Data export into database' units.

The Web-Server hosts all the web content and the software modules, which provide users with online interactive access to HSS and its functionality. This includes web applications for database queries, software for modelling calculations and online visualization, auxiliary PHP-scripts, which provide interaction between the HSS web interface and server applications.

Initial geomagnetic data, recorded by the observatories and the satellites, are retrieved via the Mail, FTP-, and Web-Severs. The acquisition procedure is controlled by the software modules of 'Data retrieval' unit hosted on the FTP- and Web-Severs. These modules allocate newly transmitted data in their initial format in the File Storage (FS) at the FTP-Server. Data from the FS are automatically converted into the elements of the HSS geomagnetic DB by means of the software modules of 'Data export into database' unit hosted on the Web- and DB-Server. The latter also hosts the DB itself. Geomagnetic data from the FS are processed by the software modules of 'Data analysis' unit.

End users access the HSS using the software modules of the 'Data access' and 'Data visualization' units. The data access modules are hosted on the Web-Server, and data visualization modules are distributed between the Web-Server, video board controller and spherical visualization system. When a user requests modelling results, the corresponding 'Model calculations' module is executed by means of PHP-script on the Web-Server and the results are sent back to the user by means of 'Data access' modules. Geomagnetic data array, stored in the DB, are initial ones for 'Model calculations', 'Data access' and 'Data visualization' software modules.

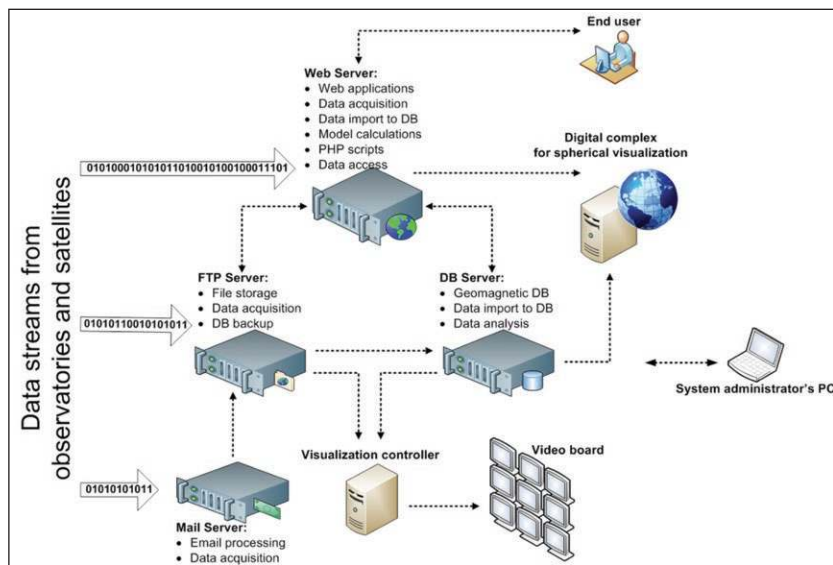


Figure 6: HSS components, their functions, and interaction.

The DB- and FS-stored geomagnetic data are visualized online upon users' requests via the Web-Server, on the video board in continuous, quasi real time mode. The digital projection systems with a spherical screen are used for visualization of the results of model calculations stored as digital maps on the Web-Server.

System administrator's computer, as well as all other computers and servers of the HSS are connected by a local area network. All servers except the Web-Server are protected by a firewall.

Data Retrieval and Storage Structure

The geomagnetic data from an observatory, automatically transferred to the HSS via e-mail or FTP, include: the geomagnetic vector total intensity (F), registered by the observatory scalar magnetometer; variations of the geomagnetic vector orthogonal components ($\Delta X, \Delta Y, \Delta Z$ or $\Delta H, \Delta D, \Delta Z$), registered by the observatory variometer with a 1-second sampling rate; values of the geomagnetic vector orthogonal components (X, Y, Z or H, D, Z) averaged on a 1-minute interval with a Gaussian filter, which is required by the INTERMAGNET standards (St-Louis et al. 2012); temperature variations of the variometer's sensor and electrical unit (t° sensor and t° electronics).

An example of the plots of preliminary magnetograms and auxiliary data, received by the HSS from the 'Klimovskaya' observatory (recently deployed in the Arkhangelsk region) is shown in the **Figure 7**.

All these geomagnetic measurements are synchronized with the UTC time-scale by an observatory data acquisition system using a built-in GPS-receiver.

The results of absolute measurements are registered into paper forms and sent to the HSS as scan-copies or text-files via e-mail or FTP-protocol. These data require further manual processing by the system administrator in order to upload the data into the geomagnetic DB. An option is a special web form, which is the part of the 'Data retrieval' unit. It allows the observatory magnetologists to enter the registered data directly into the HSS DB. Since two major absolute measurement techniques exist (Krasnoperov, Sidorov & Soloviev 2016; Mingeo Ltd. 2006; Rasson 2005), there are two types of web forms and web applications for absolute data uploading and processing. The 'Data retrieval' unit also provides automated acquisition of Swarm satellite data from the ESA data archive. Checking for new data is performed automatically in the beginning of each calendar day. At first, new daily data are temporarily stored in the CDF format and then, after unpacking, they are uploaded into the HSS DB.

The main storage for the incoming and processed geomagnetic data is the HSS relational geomagnetic DB under control of the DBMS, deployed on the DB-Server using MySQL. It stores the following pieces of information on geomagnetic field: reference information on the Russian observatories; preliminary magnetograms continuously transmitted from the observatories; recognized anthropogenic events in preliminary data; results of the absolute measurements regularly measured at the observatories; regular time series of the baseline values calculated for each observatory; quasi-definitive data, automatically derived from preliminary and baseline data; definitive observatory data; information on geomagnetic disturbances and extreme events recognized in observatory data using several indicators; Sq-variations for each observatory; supplementary data from observatories (temperature variations, failure log, etc.); geomagnetic measurements recorded by the Swarm constellation satellites.

The advantages of implementing a DBMS-controlled relational database are unified, structured, and effective storage of data. All the data (initial and processed magnetograms) are stored in one place in a single format. The HSS DB provides multiple variants for accessing the data by means of the structured query language (SQL). The end user may form various custom multi-criteria requests of any desired complexity. The HSS also provides the high level of integrity and safety of the stored data. The whole DB is constantly synchronized with its copy on a different physical server and with the HSS FS.

For proper elaboration of the DB structure several tests were done to choose the optimal way of the data storing. As initial data for the tests, 1-second recordings of the three components and the total intensity of the geomagnetic field vector were used. Additional reason for these tests was introduction of the 1-second registration standard by the INTERMAGNET, which created higher complexity of the query execution as compared to 1-minute data. The main criterion for the DB structure optimization was a query processing speed to a random data set. The tests were made basing on a generated table, which contained 400 million random values. It approximately corresponds to 1-second data volume, obtained from three observatories for the period of four years. Three variants for data storage were considered (**Table 1**): without grouping (1 value per table record); with grouping by days (86,400 values per table record); with grouping by hours (3600 values per table record).

It was obtained that the optimal DB performance is achieved with data grouping by hours. In particular, maximum speed of a query processing for daily data set was reached by using such grouping for both

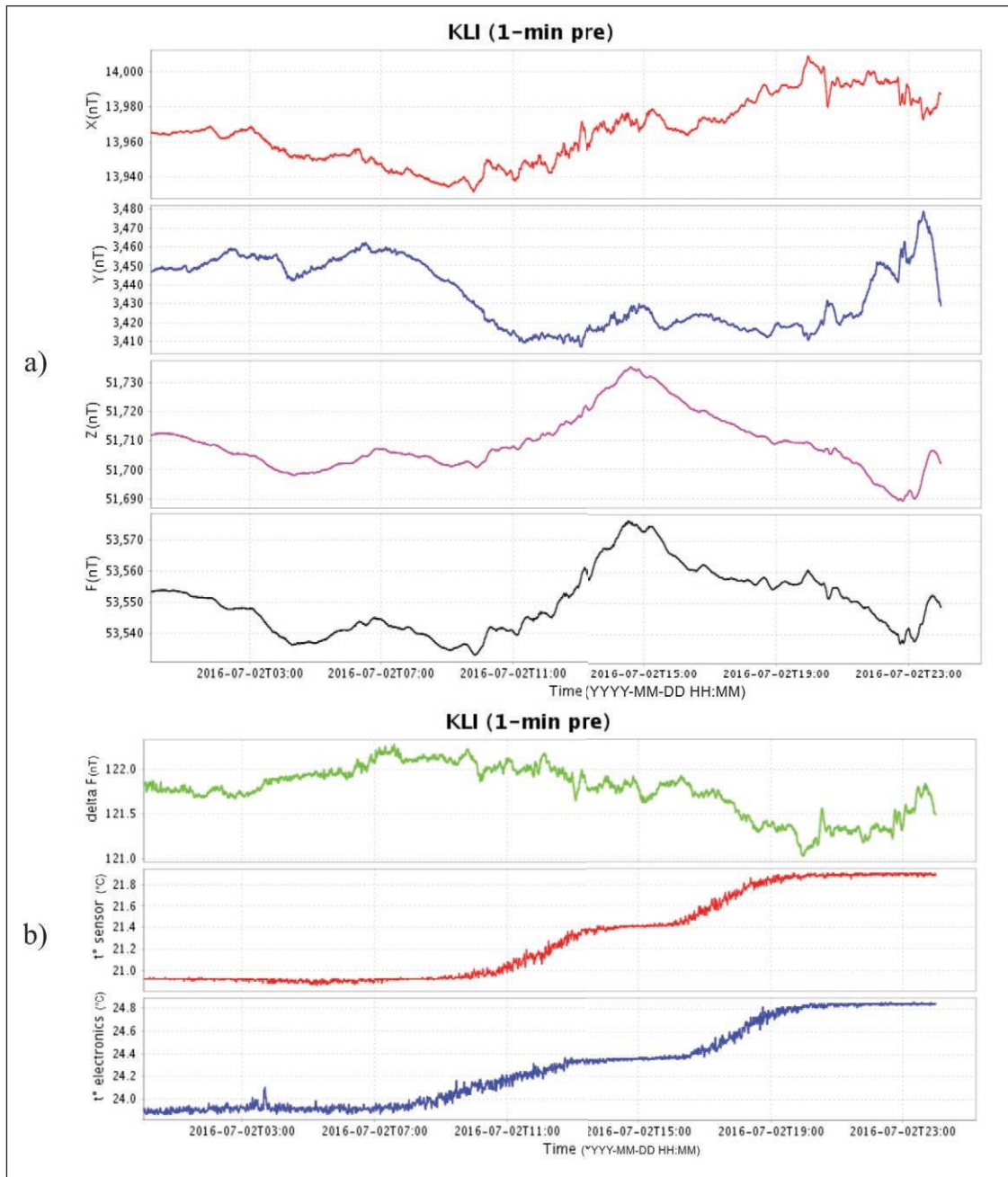


Figure 7: Magnetograms, registered at ‘Klimovskaya’ observatory: **a)** 1-minute magnetograms of the geomagnetic vector orthogonal components (X, Y, Z) and its measured total intensity (F); **b)** difference (delta F) between the measured (Fs) and calculated (Fv) values of the geomagnetic vector total intensity; temperature variations of the observatory variometer sensor (t° sensor) and electronical unit (t° electronics). The plots are generated, using the HSS data plotting tool.

Criterion	No grouping	Daily grouping	Hourly grouping
Access speed (daily data), s	2.0	0.25	0.1
DB volume, GB	16.3	3.72	3.72
Index volume, GB	7.36	0	0.001
Record length, KB	0.03	1097	46

Table 1: Testing DB with table for storing 1-second data, which contains 400 million random values.

1-minute and 1-second observatory data, as well as for recognized anthropogenic events in the DB, stored along with the initial observatory data. Each set of geomagnetic vector components and total intensity for a given time stamp is assigned a 4-digit combination of binary values, where 1 corresponds to anthropogenic disturbance presence and 0 to its absence. The complete scheme of the DB tables is given in **Figure 8**.

Technical solutions, implemented in the HSS, have significant advantages in comparison to the existing INTERMAGNET GINs. For instance, in the Paris GIN all incoming data from more than 15 INTERMAGNET and several other observatories worldwide (BCMT, 2014) are stored in text ASCII format as plain files, that has a number of obvious shortcomings: it is inefficient in terms of the occupied space. Furthermore, its capabilities for performing data search are severely limited due to complexity and inflexibility of request types and slowness of the request's performance. At the Kyoto GIN (WDC for Geomagnetism, Kyoto, 2016) the data are stored in both daily text ASCII format and monthly binary files, which are regularly generated automatically using text files. The presence of binary files reduces the occupied volume; however, it does not provide enough flexibility and simplicity in the data search queries.

The HSS geomagnetic DB facilitates the high level of data interoperability. It stores the initial data, which can be independently assessed by users, with the results of data processing and model calculations. The HSS provides open and versatile access to the whole data array, stored in the system.

Online Processing of Geomagnetic Data

As it was mentioned before, existing INTERMAGNET GINs provide definitive geomagnetic data for the scientific community with a significant time delay. It causes serious constraint for secular variations modelling and creation of operational models of the Earth magnetic field. The latter is required for navigation, directional drilling and exploration geophysics.

Such delays result from the complexity of the data-processing technique that lacks automation and generally relies on the skills of the GINs' specialists. The latter aspect is a serious problem, since the qualification of the specialists may vary, and this leads to the different levels of credibility of the disseminated data. Furthermore, such expertise is rare and takes significant time to acquire, which results the problem of reproducibility of the INTERMAGNET GINs.

A key feature of the HSS is automation of the real-time recognition of anthropogenic anomalies in the incoming magnetograms, as well as baseline correction, and following calculation of the quasi-definitive data. Such automation drastically increases the efficiency of the definitive magnetograms preparation from the preliminary records. In this way, HSS helps in solving the above mentioned problems.

The most indicative anthropogenic anomalies on geomagnetic records have the form of 'spikes' and 'jumps'. Spikes usually occur on initial records of the geomagnetic field variations. Jumps, as a rule, are recognized in baseline values.

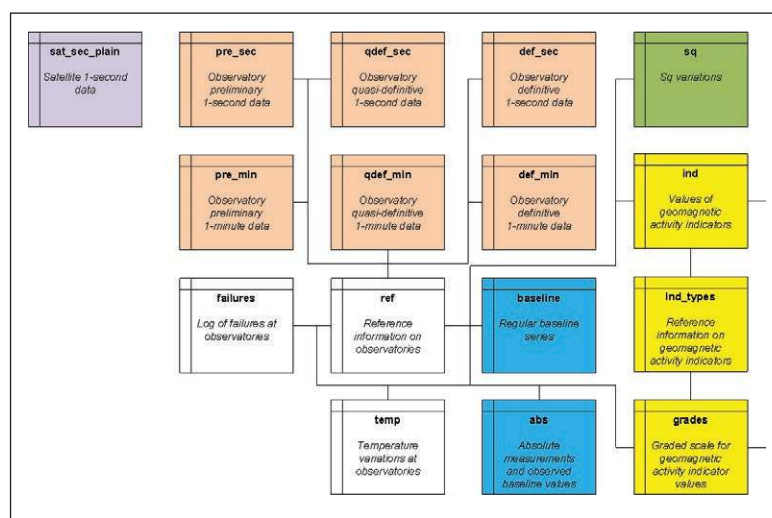


Figure 8: Scheme of the geomagnetic DB tables for storing initial ground and satellite geomagnetic data and their processing results. Tables with orange filling refer to initial observatory data, blue – to absolute observatory data, violet – to initial satellite data, green – to modelling results and yellow – to geomagnetic activity indicators.

Recognition of spikes is performed every time when new geomagnetic data arrive into the HSS. For a given observatory and magnetogram's component a 24-hour data fragment is acquired from the DB. The newly submitted data are the part of this fragment. The software module for spike recognition is then executed. The algorithm defines a spike as a magnetogram fragment having a 'tip', where two opposite sharp 'slopes' meet, surrounded by quiet 'wings'. At first, a possible tip is determined, then the sharpness of its slopes is evaluated. Next, if quiet wings to the left and to the right of the slopes are detected the considered fragment is recognized as a spike. The algorithm has several free parameters (Bogoutdinov 2010; Soloviev et al. 2012a). Such parameters are defined for each observatory individually during the algorithm learning process. Once the algorithm defines the time intervals, which correspond to the anomalous values, the recognition results are stored in the DB (**Figure 9**).

Following the elimination of artificial disturbances, the baseline correction of geomagnetic data is performed. It consists of the two main stages:

1. processing of the geomagnetic vector data stream, recognition and elimination of failures (anthropogenic anomalies) in magnetograms following by necessary quality control. If a backup variometer is available at the observatory, the recognized anthropogenic anomalies are substituted with fragments of correctly registered data. Otherwise – they are deleted.
2. determination of the baseline values using absolute observation results and calculation of the full geomagnetic field vector components.

In this way, the HSS performs automated calculation of the approximating baseline and corresponding quasi-definitive data every time, when new absolute observations are submitted. Consequently, continuous series of the quasi-definitive data are constructed. Thus, the HSS provides baseline corrected definitive data with a minimal time delay. The sequence of the quasi-definitive geomagnetic data automated calculation is presented in **Figure 10**.

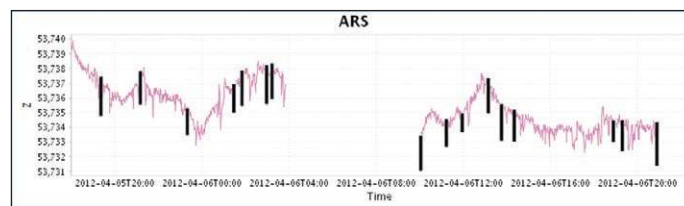


Figure 9: Example of recognition of anthropogenic spikes in the incoming data. Purple line: 1-minute geomagnetic data (Z component). Recognized spikes are marked with black. Observatory 'Arti'.

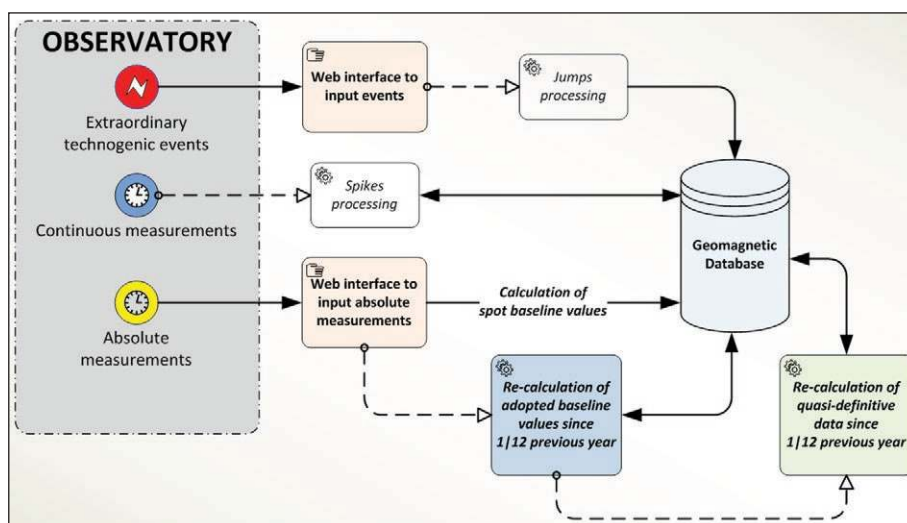


Figure 10: The HSS automated unit for calculation of the quasi-definitive geomagnetic data within the 'Data analysis' module of the HSS.

After adoption of the baseline for a given observatory the previously obtained quasi-definitive data are used for calculation of the definitive magnetograms. The comparison of preliminary and quasi-definitive magnetograms is given in **Figure 11**.

The HSS also provides recognition and assessment of geomagnetic activity of natural origin by means of the original processing algorithms, created by the authors, and implemented into the HSS: MA – recognition of natural anomalies using the measure of anomaly; dBdt – estimation of natural geomagnetic disturbance using the rate of change of the magnetic field; Amp – estimation of the maximal amplitude of geomagnetic disturbances; Kind – K-index operational calculation. We shall call the algorithms MA, dBdt, Amp, and Kind herein as 'indicators'. Each of the indicators is applied to the incoming magnetograms with a specified time resolution (e.g. 1 hour, 3 hours, 1 day). The MA, dBdt, and Amp indicators are used for processing the magnetograms of separate geomagnetic vector components (**Table 2**).

Natural extreme events are defined using original scales of anomaly elaborated for each indicator. Corresponding algorithms were implemented as software modules of the 'Data analysis' unit and deployed on the DB-Server. Automated data processing and analysis based on these algorithms are performed according to a single procedure. After retrieval of new data, a corresponding software module is executed. Then, a certain data interval is extracted from the DB, according to the specified timespan, ending at the current moment. The data fragment is processed by the algorithms. Finally, the recognition results are classified, encoded, and saved in the DB. The HSS DB also holds information on all the indicators in use which allows users to classify geomagnetic disturbances using DB-queries.

'Measure of anomaly' (MA) indicator provides evaluation of geomagnetic activity on multiple magnetograms transmitted to the HSS from the observatories. Significant advantage of the MA is its ability to

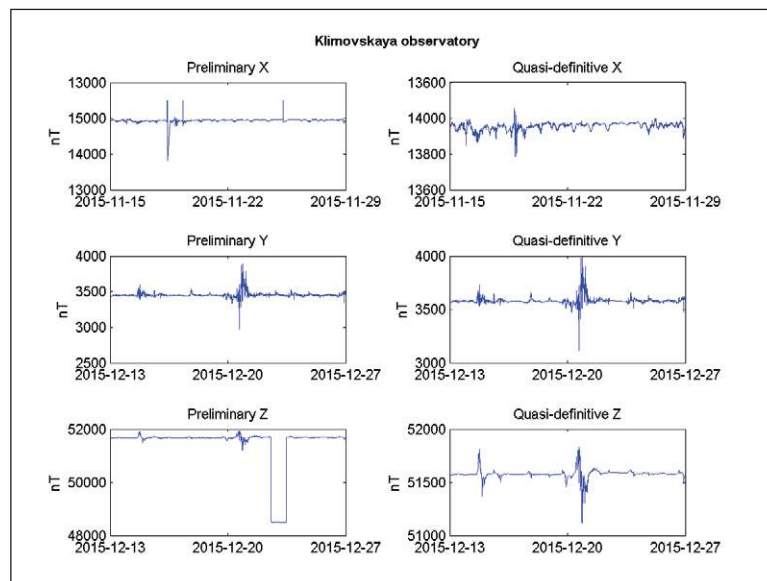


Figure 11: Preliminary and quasi-definitive magnetograms for the 'Klimovskaya' observatory for the period 01.11.2015–31.12.2015. The preliminary magnetograms of the three components demonstrate the following disturbances: X-component – spikes; Y-component – scale factor change; Z-component – jumps.

Indicator	Time interval	Resulting values
MA	1 minute	Value of the measure of anomaly for each of the geomagnetic vector components (X, Y, Z, F or H, D, Z, F) for a 1-minute time interval
dBdt	1 hour; 3 hours	Maximal value of the rate of change for each of the geomagnetic vector components (X, Y, Z, F or H, D, Z, F) within the corresponding time interval
Amp	1 hour; 3 hours; 1 day	Maximal value of disturbance amplitude for each of the geomagnetic vector components (X, Y, Z, F or H, D, Z, F) within the corresponding time interval
Kind	10 minutes; 3 hours	K-index values for the corresponding time intervals

Table 2: The indicators of natural geomagnetic activity: time resolution and resulting values.

operate with a minimal time delay, since the measure's time resolution depends solely on the sampling rate of the initial data. It is 1 sample per minute (optionally 1 sample per second) in case of the INTERMAGNET observatories. However, sometimes data transmission is performed with delays of 30 minutes, 1 day etc., due to technical constraints of a particular geomagnetic observatory or station. It naturally causes the same delays in the MA calculation.

Evaluation of geomagnetic activity by calculations of MA is performed every time when new fragments of 1-minute or 1-second data are received at the HSS. Data are analyzed within a 3-day time interval of the record from current moment backwards. Such interval is considered the most feasible since typical, long-duration magnetic disturbances (initiation, evolution, and maximal intensity increase) keep within it. As a result, every value of initial record is classified according to its measure of anomaly in the scale $[-1, 1]$, where -1 corresponds to background values and 1 corresponds to the most extreme values. The anomaly scale used in the recognition does not depend on the location of a given observatory. The scale includes four grades: $[-1; 0.4]$ is 'background'; $[0.4; 0.55]$ is 'weak anomaly'; $[0.55; 0.7]$ is 'anomaly'; $[0.7; 1]$ is 'strong anomaly' (Soloviev, Agayan & Bogoutdinov 2016). Spike and MA algorithms are based on discrete mathematical analysis (DMA) approach (Gvishiani et al. 2013; Gvishiani et al., 2014; Kulchinsky et al. 2010; Mikhailov et al. 2003; Soloviev et al. 2012a; Soloviev et al. 2012b; Soloviev et al. 2013a; Widiwijayanti et al. 2003; Zlotnicki et al. 2005).

The dBdt indicator of geomagnetic activity estimates the rate of the geomagnetic field change dB/dt and consequently the induced electrical field $E \sim dB/dt$. dBdt is calculated for each component and total intensity. Its maximum values obtained within 1- and 3-hour intervals are classified and stored in the DB. These results are used for calculation of extreme rates of geomagnetically induced currents (GICs). The latter can be used by the network operators for taking sufficient measures for emergency risks mitigation.

The indicator of intensity of geomagnetic disturbances (Amp) is based on the monitoring of amplitude of the field variations. This indicator facilitates quick and clear estimation of a data record anomaly rate from a specified observatory. Such approach requires preliminary definition of sporadic variations amplitude range. According to their values the intensity scale is formed. In the 'Data analysis' HSS unit the incoming data are automatically compared with this scale. This indicator is calculated for all three components and total intensity of the geomagnetic field vector as the maximal amplitude of disturbances within the three time intervals: 1, 3, and 24 hours. The results of the application of the dBdt and Amp algorithms are classified within a 10-grade scale (0–9), where ≤ 3 is 'background'; 4 is 'weak anomaly'; 5–7 is 'anomaly'; 8–9 is 'strong anomaly'.

The 'Data analysis' HSS unit also includes the algorithm (Kind) for real time calculation of the K-index. Classical approach to calculation of the K-index (Bartels 1938; Bartels, Heck & Johnson 1939) occurs to be insufficient for the HSS operational monitoring of geomagnetic environment due to its 3-hour delay. Kind algorithm provides calculation of the K-index with a 10-minute interval, what significantly increases the geomagnetic disturbance response time.

Three component recordings of geomagnetic field for the previous 30 days are used as input data. Five days with minimal diurnal disturbances are chosen, and quiet field is calculated as horizontal component value, averaged for these 5 days. Then the difference between maximal and minimal deviation of the measured horizontal component from the quiet level is calculated for the current 10-minute interval of the current day. The obtained value of the horizontal component is transferred into the operational value of the K-index according to the logarithmic scale and stored in the DB. The results of the Kind algorithm application are also classified within a 10-grade scale (0–9), where ≤ 3 is 'background'; 4 is 'weak anomaly'; 5–7 is 'anomaly'; 8–9 is 'strong anomaly'.

For a better illustration of the magnetic field monitoring using the implemented indicators, we consider periods of increased and decreased magnetic activity that fall within time interval from 19 to 22 December 2015. In this example the magnetogram recorded at 'Magadan' (MGD) observatory and transmitted to HSS was used. The beginning of this period is magnetically quiet, which is expressed in close to 0 values of global Kp- and regional K-indices. The main phase of geomagnetic storm took place on 20–21 December. It is characterized by Kp-index values up to 7 (GFZ 2016) and regional K-index close to 8 (**Figure 12**).

The geomagnetic storm started on 20 December and was accompanied by the storm sudden commencement (SC) that took place several hours in advance late on 19 December. SC is a result of magnetosphere hit by interplanetary shock, delivered towards the Earth by the solar wind. When present, SC manifests as a small-amplitude impulse in magnetogram several hours before geomagnetic storm onset. Often, such geomagnetic signature is clear and global being an important precursor of geomagnetic storm. It is significant, that in the given example SC was recognized as an anomalous event according to both MA (μ_{\min_v1}) and Amp (amp_h_v1), **Figure 12**.

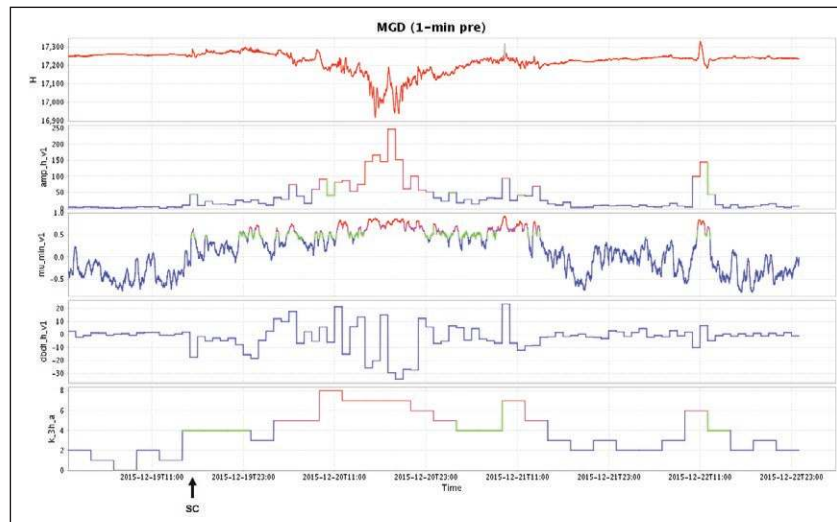


Figure 12: Graphical representation of H component variations (upper plot) from ‘Magadan’ observatory (MGD) and derived geomagnetic activity indicators for the period 19–22 December 2015: amp_h_v1 – Amp hourly maximum amplitudes of H values, mu_min_v1 – MA minute values based on H record, dbdt_h_v1 – dBdt hourly maximum values of the H component rate of change, k_3h_a – 3-hour K index values. Different colors in plots with geomagnetic activity indicators reflect different magnetic field conditions: blue – ‘background’, green – ‘weak anomaly’, purple – ‘anomaly’ and red – ‘strong anomaly’. Time of storm sudden commencement (SC) is marked with a vertical arrow.

The storm-related disturbances in H component were also characterized by anomalous values of MA, amplitude and K index (k_3h_a). On the contrary, due to mid-latitude location of the ‘Magadan’ observatory relatively small values of dBdt (dbdt_h_v1) during the storm were not classified as anomalous. The storm was recorded at the observatories located at higher latitudes, such as ‘Klimovskaya’ (KLI), ‘Saint-Petersburg’ (SPG), and ‘Yakutsk’ (YAK) (see **Figure 3**). Corresponding records contain much stronger changes dBdt indicator reached anomalous values of up to 150 nT/min for these observatories.

Access to Initial Geomagnetic Records and Processing Results and their Visualization

Interactive online access to observatory and satellite geomagnetic data is performed by the ‘Data Access’ software modules. These modules provide: online access to preliminary observatory data and Swarm satellite data; online access to results of anthropogenic disturbance detection; determinations of extreme geomagnetic events with accordance to specified criteria; online service for detection of anthropogenic anomalies in user data; online access to modelling results (hourly deviations of the field total intensity, modelling of Sq daily variations for observatories, modelling of geomagnetic effects of large-scale atmospheric dynamics).

Online access to the observatory data is provided by Java-servlets using HTTP queries. The queries are formed via web interface (**Figure 13a**), with which the user is interacting. Variational and absolute geomagnetic data along with metadata are available for each of the observatories that send data to the HSS.

To form a query, user should specify its parameters. At first, a certain observatory should be selected using an interactive map or the drop-down list. Then the three components and full intensity of the geomagnetic field vector should be specified (X/H, Y/D, Z, F). On the next step the type of requested geomagnetic data should be selected. The available variants are: 1-minute or 1-second preliminary, quasi-definitive or definitive data. A time period for the query should be, of course, defined as well. In this way, it is possible to specify for a query a certain period or choose data available for the previous month, week or day. The output magnetograms can be presented in IAGA-2002 (**Figure 13b**) or comma-separated values (CSV) formats upon a choice of a user. CSV representation of query results contains information on the recognized anthropogenic disturbances.

The HSS also provides access to Swarm data. It is performed using HTTP queries and Java-servlets via special HTML web interface for satellite data queries. Two variants exist for spatial selection of satellite magnetic data. The first one is to select a circular region with a given radius around an observatory. The second

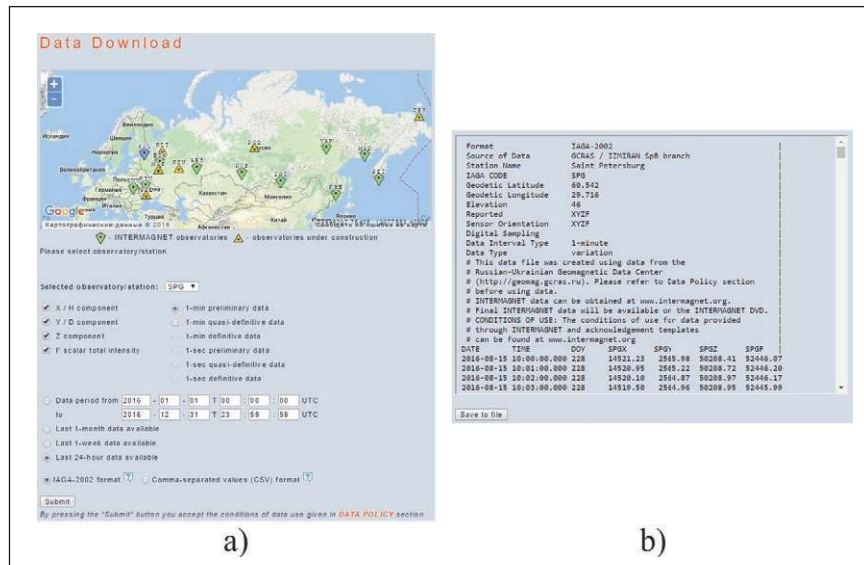


Figure 13: 'Data Access' modules: **a)** web interface for online access to observatory data; **b)** results of a DB query in the IAGA-2002 format. Observatory 'Saint Petersburg'.

possibility is to specify a rectangular geographic region. Since the Swarm constellation includes three satellites (Haagmans, Bock & Rider 2013) one of them (A, B or C) should be selected. Finally, the parameters of time interval are specified. The results are presented in a text format.

The HSS also possesses an interactive tool for recognition of anthropogenic anomalies (spikes) in user-provided magnetograms. The user's data are uploaded into the HSS in the standard IAGA-2002 format. Then the algorithm's free parameters are defined by the user. The query is addressed to the server via a PHP-script and the algorithm is applied (Bogoutdinov 2010; Sidorov, Sploviev & Bogoutdinov 2012; Soloviev et al. 2012a; Soloviev et al. 2012b). Finally, the recognition results are provided to the user as a graphic plot of the initial magnetogram with marked spikes (Figure 14). In this way, it occurs that HSS in its application does not restrict itself to the data provided by the observatories shown in Figure 3. The system can explore data in the same objective and responsible way from any observatory around the world. This opens wide opportunities to use the HSS in global magnetic data studies.

An important part of the HSS functions are innovative visualization modules: online visualization of the geomagnetic data; visualization of data on a video board; visualization of data on a spherical screen.

Functioning of the visualization software module is also based on HTTP queries and Java-servlets. This module possesses an interactive web interface (Figure 15), which allows, for a selected geomagnetic observatory, performing queries with a wide range of parameters.

A certain observatory can be selected using an interactive map or corresponding drop-down list. User can also choose preferable components (X/H, Y/D, Z, F) of the geomagnetic vector. The temperature of the observatory variometer sensor and its electronics unit is also available.

Types of the geomagnetic data and their sample rates, stored in the DB, are also selectable in the HSS. One can choose 1-minute or 1-second preliminary, quasi-definitive or definitive data. User can either specify a custom time period or choose one of the three requests: available data for the last month, week or day. Furthermore, it is possible to specify the size of the output graphic image in pixels. The respond to the user's query is returned as a downloadable PNG-image (Figure 9). Anthropogenic disturbances, recognized by the system, are marked in the output plots.

One of the important HSS hardware components is video board. It includes an array of LCD displays connected to a PC with the GeForce GTX Ti graphics accelerator compatible with four displays. The video board displays the data as magnetogram plots in the real-time mode as soon as new information becomes available. They are synchronized in time between all the observatories. This module deals with the preliminary magnetograms stored in the FS on the FTP-Server, being independent of the HSS DB. Therefore, the video board (Figure 16) can be used as an independent tool for the HSS functioning control.

An advanced approach to geoscience data visualization is implementation of digital spherical screen projection systems. Such projection devices are based on innovative technology, which allows visualization of raster images, animation or video, converting them for projection on a spherical display in a real-time mode

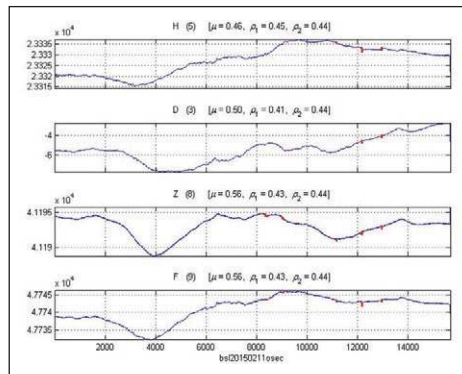


Figure 14: Results of recognition of artificial disturbances (marked with red) on user's magnetograms (H, D, Z, F components) from observatory 'Belsk' (BEL), Poland.

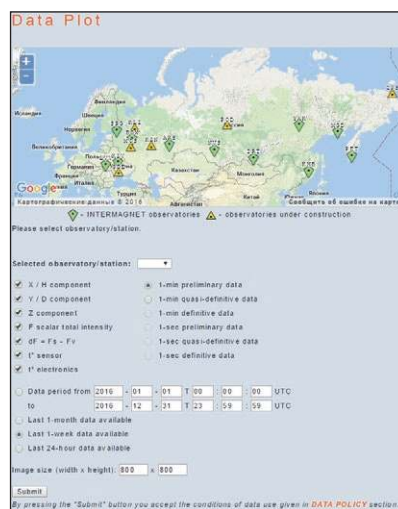


Figure 15: Web interface of the HSS online visualization of geomagnetic data.



Figure 16: HSS visualization of incoming geomagnetic data from the observatories on the video board.

(Berezko et al. 2011; Rybkina et al. 2012; Rybkina et al. 2013; Rybkina et al. 2015). Visualization of data and knowledge on the spherical screen drastically improves clarity of perception and understanding of fundamental processes occurring on, inside and around the Earth.

Geophysical Center of the Russian Academy of Sciences achieved significant results in development of such projection systems and its applications (Rybkina et al. 2015). The projection system consists of a spherical screen of the diameter 61 or 78 cm, digital projector in metal chassis, catadioptric system, and a PC workstation with an original software package 'Orbus' (Rybkina et al. 2014). It provides interactive visualization of any kind of georeferenced graphic data. In the HSS the spherical screen is used for displaying the modelling results that represent digital maps of spatial characteristics of the geomagnetic field, as well as the tracks of geophysical satellites that provide geomagnetic data to the HSS (**Figure 17**).

Access to Modelling Results

The functionality of the HSS includes software 'Model calculations' modules. These modules include interpolation of the hourly deviations of the geomagnetic vector total intensity over the territory of Russia; calculation of the Sq variation (diurnal variation of the external geomagnetic field due to the solar zenith angle change) for each observatory; and calculation of the values of geomagnetic data time series' spectral characteristics.

The HSS includes a software module for calculation of the hourly mean deviations of the observed total intensity of magnetic field from modelled values (Soloviev et al. 2013b). The calculated values are interpolated on a regular geographical grid using the Delaunay triangulation (Rybkina et al. 2013). As a result, for each hour the HSS generates a digital map of the spatial distribution of the calculated deviations with 4096×2048-pixel resolution for the territory, provided with observatory measurements. These maps are stored in the FS on the HSS FTP-Server. The sequences of hourly maps are used for visualization and analysis of the geomagnetic field dynamics and are returned to a user upon requests using an interactive web-form. The visualization of disturbances occurred in the course of the main and recovery phases of the magnetic storm of 17–18 March 2015 is shown in **Figure 18**.

This particular storm was the strongest one within the current Solar cycle and was generated by a series of solar flares occurred on 11–15 March 2015. The significant decrease of the horizontal geomagnetic component was registered first by the most eastern observatories (2015-03-17 13:00–14:00 UT) and then the disturbances propagate westward and then finally, at 2015-03-17 17:00–18:00 UT, they were detected by the most western Russian observatories. The maximal total reduction of the horizontal geomagnetic component over the Russian longitudinal sector due to this storm was registered at 2015-03-17 21:00–00:00 UT. After that the recovery phase of the storm started. The recovery of the field was at first registered by the eastern observatories (2015-03-18 02:00–03:00) and then propagated westward. The total recovery of the geomagnetic field was registered by 2015-03-18 10:00–11:00.

Another software module of the 'Model calculations' unit is implemented for operational modelling of the Sq variation of the geomagnetic field. This module is executed at the beginning of every new day. On the first step the module calculates the filtering coefficients. It also selects the K-index values for the previous 30 days from the HSS DB. Analyzing the K-index values, the module determines quiet days and then marks out the corresponding time intervals on the magnetograms. The hourly averaged values of the

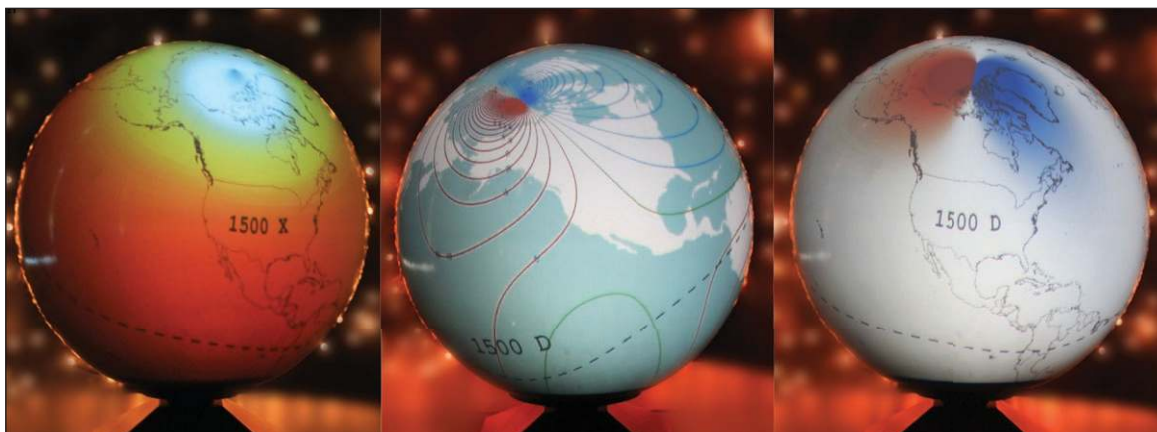


Figure 17: Visualization of geomagnetic data on a spherical screen (Berezko et al. 2011).

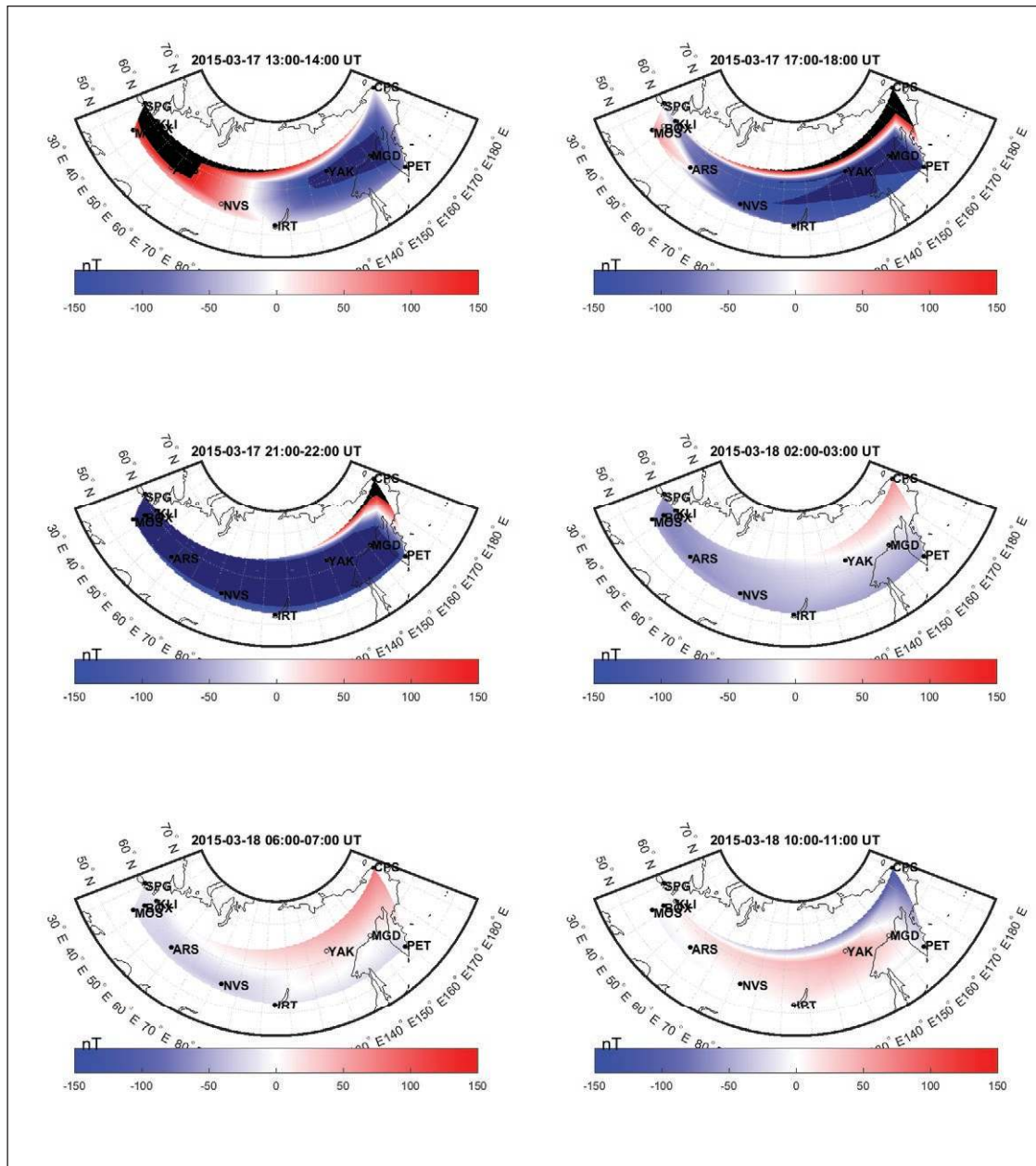


Figure 18: Visualization of interpolated hourly deviations of the total intensity of geomagnetic field for the territory of Russia during the main and recovery phases of the magnetic storm of 17–18 March 2015 (2015-03-17 13:00–2015-03-18 08:00 UT). Interpolation is performed for the region, provided with observatory measurements.

components for these intervals are calculated, combined and smoothed. The results are stored in the FS at the FTP-Server.

Also, the 'Model calculations' unit provides access to the results of estimation of the geomagnetic effects of large-scale atmospheric dynamics as represented by the planetary wave modes, which are very pronounced in mid- latitudes. The fact that atmospheric oscillations have an influence on geomagnetic variations due to a resulting dynamo action in the E-layer of the ionosphere provides a physical basis of the analysis. Extracting planetary waves from geomagnetic time series is performed using empirical mode decomposition method by which the geomagnetic time series of daily mean values of horizontal geomagnetic component are transformed into spectral components. The decomposition procedure results in the amplitude (in nT) of oscillating modes with estimated mean periods. For each observatory that sends data to HSS the corresponding analysis is performed systematically with a 6-month cadence as a sufficiently long data series are collected. The results are stored in DB and may be returned to a user upon request via a common interface as the ASCII time series. **Figure 19** displays an example of the specific mode of the H-component decomposition that is related to a mean period of 16 days.

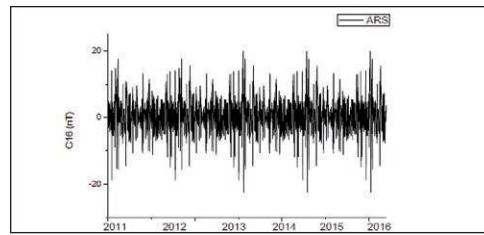


Figure 19: Time series of the amplitude of oscillating 16-day modes in the H geomagnetic component at observatory 'Arti' in January 2011–March 2016.

Discussion on Applications and Conclusion

Geomagnetic observation systems on the Earth and in the near-Earth space are the basis for fundamental research in geomagnetism. The HSS is an efficient tool that helps to explore and promote several long-term problems of magnetic data acquisition, storage, management, processing and analysis. Among them is acceleration of creation and routine production of quasi-definitive and definitive observatory data from the preliminary ones. A full-scale geomagnetic observatory provides automatic registration of the geomagnetic field vector total intensity and its components' variations, as well as regular absolute geomagnetic observations. Currently most of INTERMAGNET observatories produce definitive data, representing time series of complete values of magnetic vector components, in situ. The delay in definitive data production is normally 1–2 years. Before the creation of the HSS, Russian geomagnetic observatories represented scattered observation sites that transmitted preliminary 'raw' data to foreign INTERMAGNET GINs for their further processing. The created HSS for the first time provides a centralized accumulation of Russian observations and their bringing to the highest INTERMAGNET data standards using automatized system analysis methods. Routine operations with magnetic data streams, accepted worldwide, were converted from manual to automated analytical mode with the assistance of mathematical elements of artificial intelligence. It results in improving the accuracy of the definitive data and reduction of delays in their preparation in comparison with foreign data centers. This makes the HSS an indispensable instrument that connects the Russian observatories in a single network and unifies the data analysis procedures. The latter aspect is highly important, since the HSS facilitates the production of high-quality quasi-definitive and definitive data for a vast territory. In addition to fundamental research, seamless and reliable provision of quasi-definitive data is essential in the industry sector as well. It is highly demanded for the support of the high-technology processes of directional drilling for oil and gas exploration and extraction (Gvishiani & Lukianova 2014; Gvishiani et al. 2015).

In this way, the HSS creates a new paradigm in the system of magnetic information acquisition, handling and purification. This paradigm is linked to the big data concept (Roberts 2016; Science International 2015), which is interpreted as four Vs: Volume, Variety, Velocity and Veracity. Velocity (the third V) of the veracious (the fourth V) data availability to research community is a key feature of the magnetometric science. From this point of view, the HSS occurs to be an efficient systems analysis tool that enables to drastically increase the velocity and veracity features of the whole INTERMAGNET system, as well as of its national and regional subnetworks. At the same time, the HSS provides retrieval, management, processing and analysis of the satellite geomagnetic data recorded by the ESA Swarm spacecraft. In this capacity, the HSS is the first holistic systems analysis tool, both in Russia and worldwide, that handles varied (the second V) observatory and satellite magnetic data together at high level of the coordination. Finally, the volume feature (the first V) is gained due to huge data sets, handled by the system. Indeed, real time data are streamed continuously from constantly growing national geomagnetic network, comprising 14 observatories now, with increasing sampling rate, varying from 1/60 Hz to 10 Hz, as well as from the three identical Swarm satellites that provide recordings sampled every 1 second and 1/50 seconds.

The HSS has significant advantages in comparison to the existing GINs data storage approaches. The HSS provides continuous retrieval and storage of geomagnetic data in advanced relational database. This approach implements efficient coordination of ground-based and satellite geomagnetic measurements within a single data management system. In particular, coordination is associated with unification of heterogeneous geomagnetic data that originally are spatially distributed, registered in various formats and with different sampling rates.

The HSS implements the system approach to scientific analysis of geomagnetic data. Its data processing modules employ traditional and original mathematical methods, which provide multi-criteria and complex data analysis for automated recognition of natural anomalies in observatory data. All the automated

operations are performed in near real-time mode, which makes it possible to estimate magnetic activity and make subsequent forecasts with minimum time delay. It should be stressed that all results of continuous data processing are stored together with initial recordings in the unified database.

The HSS possesses the integrated online tools for processing and analysis of the geomagnetic data, including user-uploaded. This makes the HSS an efficient and flexible instrument for the geomagnetism researchers, facilitating operative modeling of the Earth's magnetic field. The functionality of the HSS may be utilized by the experts and decision-makers who need this information for the assessment and mitigation of the risks of extreme geomagnetic conditions. The approaches that form the basis of the HSS provide its easy replicability, which makes it a standardized solution.

List of Acronyms

ASM – Absolute Scalar Magnetometer (part of the Swarm satellite payload)
CDF – common data format
CSV – comma-separated values
DB – geomagnetic database
DBMS – database management system
EFI – Electric Field Instrument (part of the Swarm satellite payload)
ESA – European Space Agency
FS – file storage
GIC – geomagnetically induced current
GIN – Geomagnetic Information Node
HSS – hardware/software system
IAGA – International Association of Geomagnetism and Aeronomy of the International Union of Geodesy and Geophysics
ICSU – International Council for Science
INTERMAGNET – International Real-time Magnetic Observatory Network
LEO – low Earth orbit
MA – measure of anomaly
NCEI – National Centers for Environmental Information
NOAA – United States Department of Commerce National Oceanic and Atmospheric Administration
SC – sudden commencement
SPIDR – Space Physics Interactive Data Resource
SQL – Structured Query Language
UTC – Coordinated Universal Time
VFM – Vector Field Magnetometer (part of the Swarm satellite payload)
WDS – World Data System of the International Council for Science

Acknowledgements

This project was conducted in the framework of the Federal Target Program of the Ministry of Education and Science of the Russian Federation, agreement No. 14.607.21.0058, ID No. RFMEFI60714X0058. This paper contributes the final stage of creation of the experimental prototype of the hardware and software complex for monitoring and detection of extreme geomagnetic events using ground and satellite data.

The results presented in this paper rely on data collected at magnetic observatories. The authors would like to express their gratitude to the national institutions that support them. The authors are also grateful to 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. The authors are also much obliged to the ESA Swarm mission team.

Competing Interests

The authors have no competing interests to declare.

References

- Bartels, J** 1938 Potsdamer erdmagnetische Kennziffern. *Zeitschrift für Geophysik*, 14: 68–78.
Bartels, J, Heck, N H and Johnson, H F 1939 The three-hour-range index measuring geomagnetic activity. *Terrestrial Magnetism and Atmospheric Electricity*, 44: 411–454. DOI: <https://doi.org/10.1029/TE044i004p00411>

- BCMT** 2014 Bureau central de Magnetisme Terrestre. Available at: <http://www.bcmt.fr/> (Last accessed 20 June 2016).
- Berezko, A, Krasnoperov, R, Kedrov, E, Pyatygina, O and Shibaeva, A** 2011 Visualization of earth main magnetic field maps on a spherical display. *Russian Journal of Earth Sciences*, 12(2): 1–3. DOI: <https://doi.org/10.2205/2011ES000508>
- Bloxham, J and Gubbins, D** 1989 The evolution of the Earth's magnetic field. *Scientific American*, 261: 30–37. DOI: <https://doi.org/10.1038/scientificamerican1289-68>
- Bogoutdinov, Sh R., Gvishiani, A D, Agayan, S M, Solovyev, A A and Kihn, E** 2010 Recognition of disturbances with specified morphology in time series. Part 1: Spikes on magnetograms of the worldwide INTERMAGNET network. *Izvestiya, Physics of the Solid Earth*, 46(11): 1004–1016. DOI: <https://doi.org/10.1134/S1069351310110091>
- ESDB** 2015 Geomagnetic data recorded at Geomagnetic Observatory Klimovskaya (IAGA code: KLI). *ESDB repository. Geophysical Center of the Russian Academy of Sciences*. DOI: <https://doi.org/10.2205/KLI2011>
- ESDB** 2016 Geomagnetic data recorded at Geomagnetic Observatory Saint Petersburg (IAGA code: SPG). *ESDB repository. Geophysical Center of the Russian Academy of Sciences (GC RAS)*. DOI: <https://doi.org/10.2205/SPG2012>
- Friis-Christensen, E, Lühr, H and Hulot, G** 2006 Swarm: A constellation to study the Earth's magnetic field. *Earth, Planets and Space*, 58: 351–358. DOI: <https://doi.org/10.1186/BF03351933>
- GFZ, Helmholtz Centre Potsdam** 2016 *Indices of Global Geomagnetic Activity*. Available at: <http://www.gfz-potsdam.de/en/kp-index/> (Last accessed 25 September 2016).
- Gvishiani, A, Dobrovolsky, M, Agayan, S, and Dzeboev, B** 2013 Fuzzy-based clustering of epicenters and strong earthquake-prone areas. *Environmental Engineering and Management Journal*, 12(1): 1–10.
- Gvishiani, A, Lukianova, R, Soloviev, A and Khokhlov, A** 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: <https://doi.org/10.1007/s10712-014-9297-8>
- Gvishiani, A D and Lukianova, R Yu** 2014 Geomagnetic support of the down the hole controlled drilling in the Arctic region. *The Arctic Herald*, (4): 114–123.
- Gvishiani, A D and Lukianova, R Yu** 2015 Geoinformatics and observations of the Earth's magnetic field: The Russian segment. *Izvestia-Physics of the Solid Earth*, 51(2): 157–175. DOI: <https://doi.org/10.1134/S1069351315020044>
- Gvishiani, A D, Lukianova, R Y and Soloviev, A A** 2015 Geomagnetic field analysis and directional drilling problem in the Arctic region. *Gornyi Zhurnal (Mining Journal)*, (10): 94–99. DOI: <https://doi.org/10.17580/gzh.2015.10.17>
- Haagmans, R, Bock, R and Rider, H** 2013 Swarm: ESA's magnetic field mission. In: Fletcher, K (prod. ed.) *ESA communications, BR-302, September 2013, 2nd edn*. ISBN 978-92-9221-067-0.
- INTERMAGNET** 2013a *IMFV1.23 INTERMAGNET GIN Dissemination Format for Minute Values*. Available at: <http://intermagnet.org/data-donnee/formats/imfv123-eng.php> (Last accessed 20 June 2016).
- INTERMAGNET** 2013b *List of IMOs and responsible GINs*. Available at: <http://intermagnet.org/imos/imotblobs-eng.php> (Last accessed 13 September 2016).
- INTERMAGNET** 2016 *Saint Petersburg (SPG)*. Available at: http://intermagnet.org/imos/imos-list/imos-details-eng.php?iaga_code=SPG (Last accessed 27 May 2016).
- Jankowsky, J and Sucksdorff, C** 1996 *Guide for magnetic measurements and observatory practice*. Warsaw: IAGA, 238 p.
- Kaftan, V I and Krasnoperov, R I** 2015 Geodetic observations at geomagnetic observatories. *Geomagnetism and Aeronomy*, 55(1): 118–123. DOI: <https://doi.org/10.1134/S0016793215010065>
- Kaftan, V I, Krasnoperov, R I and Tertyshnikov A V** 2015 Observation with the use of global navigation satellite systems at geomagnetic stations and observatories: applied and fundamental aspects. *Heliogeophysical research*, 12: 1–10. Available at: <http://elibrary.ru/item.asp?id=23802395>.
- Krasnoperov, R I, Sidorov, R V, and Soloviev, A A** 2015 Modern Geodetic Methods for High-Accuracy Survey Coordination on the Example of Magnetic Exploration. *Geomagnetism and Aeronomy*, 55(4): 547–554. DOI: <https://doi.org/10.1134/S0016793215040076>
- Krasnoperov, R I, Sidorov, R V and Soloviev, A A** (ed.) 2016 Operation manual for absolute measurements using fluxgate declinometer/inclinometer. *Geoinformatics Research Papers*, 4(1). DOI: <https://doi.org/10.2205/2016BS038>

- Kulchinsky, R G, Kharin, E P, Shestopalov, I P, Gvishiani, A D, Agayan, S M and Bogoutdinov, Sh R** 2010 Fuzzy logic methods for geomagnetic events detections and analysis. *Russian Journal of Earth Sciences*, 11(4): RE4003. DOI: <https://doi.org/10.2205/2009ES000371>
- Love, J J** 2008 Magnetic monitoring of Earth and space. *Physics Today*, 61(2): 31–37. DOI: <https://doi.org/10.1063/1.2883907>
- Mandea, M and Papitashvili, V** 2009 Worldwide geomagnetic data collection and management. *Eos, Transactions American Geophysical Union*, 90(45):409–424. DOI: <https://doi.org/10.1029/2009EO450001>
- Matzka, J, Chulliat, A, Mandea, M, Finlay, C C, and Qamili, E** 2010 Geomagnetic Observations for Main Field Studies: From Ground to Space. *Space Science Reviews*, 155(1–4): 29–64. DOI: <https://doi.org/10.1007/s11214-010-9693-4>
- Mikhailov, V, Galdeano, A, Diament, M, Gvishiani, A, Agayan, S, Bogoutdinov, Sh, Graeva, E and Sailhac, P** 2003 Application of artificial intelligence for Euler solutions clustering. *Geophysics*, 68(1): 168–180. DOI: <https://doi.org/10.1190/1.1543204>
- MinGeo Ltd.** 2006 *MINGEO DIM Declination/Inclination Theodolite Operation Manual. Version 1.1f*. Budapest: Mingeo Ltd., 17 p.
- MinGeo Ltd.** 2016 *The magnetic data collection platform with Maglin, the acquisition software*. Available at: <http://www.mingeo.com/prod-magrec4b.html> (Last accessed 21 June 2016)
- Newitt, L R, Barton, C E and Bitterly, J** 1996 *Guide for Magnetic Repeat Station Surveys*. Boulder, CO: International Association of Geomagnetism and Aeronomy. 129 p.
- Olsen, N, Friis-Christensen, E, Floberghagen, R, Alken, P, Beggan, C D, Chulliat, A, Doornbos, E, da Encarnação, J T, Hamilton, B, Hulot, G, van den IJssel, J, Kuvshinov, A, Lesur, V, Lühr, H, Macmillan, S, Maus, S, Noja, M, Olsen, P E H, Park, J, Plank, G, Püthe, C, Rauberg, J, Ritter, P, Rother, M, Sabaka, T J, Schachtschneider, R, Sirol, O, Stolle, C, Thébaud, E, Thomson, A W P, Tøffner-Clausen, L, Velínský, J, Vigneron, P and Visser, P N** 2013 The Swarm satellite constellation application and research facility (SCARF) and swarm data products. *Earth, Planets and Space*, 65(11): 1189–1200. DOI: <https://doi.org/10.5047/eps.2013.07.001>
- Peltier, A, and Chulliat, A** 2010 On the feasibility of promptly producing quasi-definitive magnetic observatory data. *Earth, Planets Space*, 62(2): e5–e8. DOI: <https://doi.org/10.5047/eps.2010.02.002>
- Rasson, J** 2007 Observatories, instrumentation In: Gubbins, D and Herrero-Bervera, E (Eds.) *Encyclopedia of Geomagnetism and Paleomagnetism*. New York: Springer, pp. 711–713. DOI: https://doi.org/10.1007/978-1-4020-4423-6_225
- Rasson, J L** 2005 About Absolute Geomagnetic Measurements in the Observatory and in the Field. In: Malcorps, D H (Ed.) *Publication Scientifique et Technique No. 040*. Brussels: L'Institut royal meteorologique de Belgique. 43 p.
- Reay, J, Herzog, D, Sobhana, A, Kharin, E, McLean, S, Nose, M and Sergeyeva, N** 2011 Magnetic observatory data and metadata: types and availability. In: Mandea, M and Korte, M (eds.) *Geomagnetic Observations and Models*. The Netherlands: Springer Science+Business Media B.V, pp. 149–181. DOI: <https://doi.org/10.1007/978-90-481-9858-0>
- Roberts, F S** 2016 What is Big Data and how has it changed? In: Kedrov, E (Ed.) Book of Abstracts of the International Conference 'Data Intensive System Analysis for Geohazard Studies', Sochi region, Mountain cluster, Russia, 18–21 July 2016, BS4002. DOI: <https://doi.org/10.2205/2016BS01Sochi>
- Rybkina, A I, Bobkov, A E, Nikiforov, O V and Pyatygina, O O** 2014 Program for visualization of data for spherical displays 'ORBUS'. *Certificate of registration for a program No. RU 2014618293*. Moscow: Rospatent.
- Rybkina, A I, Bobkov, A E, Nikiforov, O V and Pyatygina, O O** 2015 Hardware and software system for visualization of geophysical data on a spherical screen. *Scientific Visualization*, 7(2): 38–49. Available at: <http://sv-journal.org/2015-2/04.php?lang=en>.
- Rybkina, A I, Pyatygina, O O, Shibaeva, A A, Nikiforov, O V and Krasnoperov, R I** 2012 Digital Projection Systems with a Sphere Shaped Screen as a New Tool for Geoinformatics In: Proceedings of the 1st EAGE International Conference on IT for Geosciences, Dubna, Russia, 7–9 December 2012. DOI: <https://doi.org/10.3997/2214-4609.20142817>
- Rybkina, A I, Soloviev, A A, Kagan A I, Shibaeva, A A, Pyatygina, O O and Nikiforov, O V** 2013 Interpolation of observatory measurements data and visualization of total intensity of the Earth magnetic field. *Vestnik Otdelenia nauk o Zemle RAN*, 5(3): NZ3002. DOI: <https://doi.org/10.2205/2013NZ000116>
- Science International** 2015 *Open Data in a Big Data World*. Paris: International Council for Science (ICSU), International Social Science Council (ISSC), The World Academy of Sciences (TWAS), InterAcademy Partnership (IAP), 17 p.

- Sidorov, R V, Soloviev, A A and Bogoutdinov, Sh R** 2012 Application of the SP algorithm to the INTERMAGNET magnetograms of the disturbed geomagnetic field. *Izvestiya, Physics of the Solid Earth*, 48(5): 410–414. DOI: <https://doi.org/10.1134/S1069351312040088>
- Soloviev, A, Agayan, S and Bogoutdinov, S** 2016 Estimation of geomagnetic activity using measure of anomalousness. *Annals of Geophysics*, 59(6).
- Soloviev, A, Bogoutdinov, Sh, Gvishiani, A, Kulchinskiy, R and Zlotnicki, J** 2013a Mathematical Tools for Geomagnetic Data Monitoring and the INTERMAGNET Russian Segment. *Data Science Journal*, 12: WDS114–WDS119. DOI: <https://doi.org/10.2481/dsj.WDS-019>
- Soloviev, A, Chulliat, A, Bogoutdinov, S, Gvishiani, A, Agayan, S, Peltier, A and Heumez, B** 2012a Automated recognition of spikes in 1 Hz data recorded at the Easter Island magnetic observatory. *Earth, Planets and Space*, 64(9): 743–752. DOI: <https://doi.org/10.5047/eps.2012.03.004>
- Soloviev, A, Khokhlov, A, Jalkovsky, E, Berezko, A, Lebedev, A, Kharin, E, Shestopalov, I, Manda, M, Kuznetsov, V, Bondar, T, Mabie, J, Nisilevich, M, Nechitailenko, V, Rybkina, A, Pyatygina, O and Shibaeva, A** 2013b *The Atlas of the earth's magnetic field. Edited by Gvishiani, A, Frolov, A and Lapshin, V.* Moscow: Geophysical Center of the Russian Academy of Sciences. 361 p. DOI: https://doi.org/10.2205/2013BS011_Atlas_MPZ
- Soloviev, A A, Agayan S M, Gvishiani, A D, Bogoutdinov, Sh R and Chulliat, A** 2012b 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. DOI: <https://doi.org/10.1134/S106935131204009X>
- Soloviev, A A, Sidorov, R V, Krasnoperov, R I, Grudnev, A A and Khokhlov, A V** 2016 Klimovskaya: A New Geomagnetic Observatory. *Geomagnetism and Aeronomy*, 56(3): 342–354. DOI: <https://doi.org/10.1134/S0016793216030154>
- St-Louis, B J, Sauter, E A, Trigg, D F, Coles, R L and Regimbald, D** (eds.) 2012 *INTERMAGNET Technical Reference Manual, Version 4.6, 2012.* Edinburgh: INTERMAGNET. 100 p. Available at: http://www.intermagnet.org/publications/intermag_4-6.pdf.
- University of Waikato** 2014 *Solar wind plasma and the magnetosphere.* Available at: <http://sciencelearn.org.nz/Contexts/Gases-and-Plasmas/Sci-Media/Images/Solar-wind-plasma-and-the-magnetosphere> (Last accessed 24 September 2016).
- V-DAT, IAGA Working Group** 2011 *IAGA-2002 exchange format.* Available at: <http://www.ngdc.noaa.gov/IAGA/vdat/IAGA2002/iaga2002format.html> (Last accessed 20 June 2016).
- WDC for Geomagnetism, Kyoto** 2016 *World Data Center for Geomagnetism, Kyoto.* Available at: <http://wdc.kugi.kyoto-u.ac.jp/> (Last accessed 20 June 2016).
- Widiwijayanti, C, Mikhailov, V, Diamant, M, Deplus, C, Louat, R, Tikhotsky, S and Gvishiani, A** 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. DOI: [https://doi.org/10.1016/S0012-821X\(03\)00416-3](https://doi.org/10.1016/S0012-821X(03)00416-3)
- Zlotnicki, J, LeMouel, J.-L, Gvishiani, A, Agayan, S, Mikhailov, V and Bogoutdinov, Sh** 2005 Automatic fuzzy-logic recognition of anomalous activity on long geophysical records. Application to electric signals associated with the volcanic activity of la Fournaise volcano (Réunion Island). *Earth and Planetary Science Letters*, 234: 261–278. DOI: <https://doi.org/10.1016/j.epsl.2005.01.040>

How to cite this article: Gvishiani, A, Soloviev, A, Krasnoperov, R and Lukianova, R 2016 Automated Hardware and Software System for Monitoring the Earth's Magnetic Environment. *Data Science Journal*, 15: 18, pp. 1–24, DOI: <http://dx.doi.org/10.5334/dsj-2016-018>

Submitted: 24 October 2016 **Accepted:** 17 November 2016 **Published:** 16 December 2016

Copyright: © 2016 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.