Assessment Information

Repository: WDC for Solar–Terrestrial Physics, Moscow
Website: http://www.wdcb.ru/stp/index.en.html
Certification Date: 06 November 2020

This repository is owned by: **Geophysical Center of the Russian Academy of Sciences**

Core Trustworthy Data Repository Requirements

BACKGROUND INFORMATION

Context

R0. Please provide context for your repository.

Repository Type.

Domain or subject-based repository

National repository system; including governmental

Publication repository

Library/Museum/Archives

Research project repository

Other (please describe)

Global, Multidisciplinary

Reviewer 1:

Accept

Reviewer 2:

Accept

Comments

Domain or subject-based repository - Data collections are created in accordance with types of observations for each solar-terrestrial physics discipline. For example, in section "Geomagnetic Variations" there are separate collections: magnetograms, indices, hour and minute values of geomagnetic field elements, etc. Therefore, the repository can be considered as subject-based repository.

National repository - In the last years the WDC for STP mainly concentrates on the collection and dissemination of the national and CIS countries data. One of the goals of the repository is the creation of the most comprehensive collection of national data.

Publication repository – WDC's archive contains a number of publications. Previously (until the 1990s), WDC for STP in Boulder (USA) and WDC for STP in Moscow (USSR) issued publications with various data sets from their archives for distribution and storage, for example http://www.wdcb.ru/stp/solar/solar_publ.html/.

Electronic journal "Russian Journal of Earth Sciences" http://rjes.wdcb.ru/ is published in the Geophysical Center of the Russian Academy of Sciences (host organization) and the database of publications is also stored and available for use.

Archive - One of the functions of the WDC for STP is the guaranteed long-term storage of geophysical data for access and use by a designated community.

Research project repository - During the participation in the IPY 2007-2008 project the website with data relevant to the polar regions of the Earth was created in the Center http://www.wdcb.ru/WDCB/IPY/IPY.html

Global data repository - The World Data Center for Solar-Terrestrial Physics, Moscow (WDC for STP) collected Earth data from stations, institutes, centers and agencies around the world since 1957 until the termination of the WDC System. Geophysical data from over 50 countries are stored in the Center that provides scientists with access to this data.

Multidisciplinary repository - WDC for STP deals with data on the solar-terrestrial physics' disciplines: geomagnetic variations, ionosphere phenomena, cosmic rays, solar activity and interplanetary medium.

Reviewer 1:

Accept

Reviewer 2:

Accept

Brief Description of the Repository's Designated Community.

WDC for STP has a mission to serve the international science community as custodian of global geophysical information. The Center serves both national and international users. The Designated Community includes scientists engaged in basic and applied scientific research in the field of geophysics, geology and related fields of science, students, graduate students and teachers of higher educational institutions. WDC for STP provides assistance to educational programs. WDC for STP is targeted on the scientific organizations, separate researchers, universities and students in the different fields of sciences both in Russia and abroad.

Reviewer 1:

Accept

Reviewer 2:

Accept

Level of Curation Performed.

A. Content distributed as deposited

B. Basic curation – e.g. brief checking; addition of basic metadata or documentation

C. Enhanced curation – e.g. conversion to new formats; enhancement of documentation

D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

Reviewer 1:

Accept

Reviewer 2:

Accept

Comments

A.   WDC for STP is responsible for the collection, guaranteed long-term storage and dissemination of geophysical data beginning from the International Geophysical Year (1957-1958). The Center manages a large amount of geophysical data from foreign and domestic scientific institutions, stations and observatories accumulated over the period since 1957 as a result of the exchange between World Data Centers. A significant part of the data collected in the Center is distributed as deposited. This mainly refers to historical data.

B.   Basic curation – Each dataset entering the Center is checked and analyzed. Metadata and necessary accompanying documents (for example, format description) are prepared in the Center in their absence.

C.   Enhanced curation – In some cases, the dataset is converted to another format (for example, into WDC format, or IAGA2002 format), in others a paper-printed form is converted into electronic form (PDF file or ASCII text file). Sometimes the supporting documents are improved.

D.   Data-level curation – In some cases, the data requires editing. This is done in consultation with the author/contributor of data. In fact, the Center helps the author in preparing the dataset for long-term storage in the Center and usage.

Reviewer 1:

Accept

Outsource Partners. If applicable, please list them.

WDC for STP cooperates with Research Institute of the Hydrometeorological Information - World Data Center (RIHMI-WDC) http://meteo.ru/mcd/index_e.html on the basis of the "Agreement on partnership, cooperation and scientific exchange between GC RAS and RIHMI-WDC". In the frame of these relationships the RIHMI-WDC provides long-term storage of backup copies of the WDC for STP's archive datasets in its robotic library (robotic IBM system Storage 3500 Tape Library).

Our partners include the Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation RAS (IZMIRAN) https://www.izmiran.ru/?LANG=en, WDC for Geomagnetism, Kyoto http://wdc.kugi.kyoto-u.ac.jp/index.html and WDC for Geomagnetism, Edinburgh http://www.wdc.bgs.ac.uk/, WDC for Geoinformatics and Sustainable Development, Ukraine http://wdc.org.ua/ as well as observatories transmitting data to the Center.

Reviewer 1:

Accept

Reviewer 2:

Accept

Other Relevant Information.

The World Data Center for Solar-Terrestrial Physics, Moscow is operated by the Geophysical Center of the Russian Academy of Sciences.

The WDC for STP was founded in 1957 as a part of the WDC-B in the process of establishing the World Data Center System of the International Council of Scientific Unions (ICSU) to preserve and provide access to available data obtained during the International Geophysical Year (1957-1958). The WDC-A (USA) and WDC-B (USSR) exchanged all data (in accordance with the "Guide to the World Data Center System" developed by the Panel on WDCs, ICSU) so that their archives were equivalent, to ensure the safety of data in the event of any unforeseen disasters. Thus a large amount of foreign and domestic data has accumulated in the WDC for STP.

The Center manages, maintains and provides services for the archive of historical and modern results of global observations related to the solar-terrestrial physics' disciplines, obtained during the International

Geophysical Year (1957-1958) and subsequent international projects, such as the "International Year of Quiet Sun,", "International Year of the Active Sun", "International Polar Year 2007-2008" et al., results of land and sea expeditions, launches of satellites, special experiments, results of geophysical observations on global networks of observatories (geomagnetic, ionospheric, solar and cosmic ray stations).

The Center maintains extensive archive of the data on different media including paper, film and electronic media and provides all conditions for their long-time preservation.

Nowadays, the project "Preservation of old data" is performed in the WDC for STP with the aim to transfer data from paper medium to electronic form, to increase data sets in electronic form, to prevent the loss of valuable historical data, and, lastly, to provide free on-line access to them for more efficient use. For example, values of geomagnetic K and C indices according to the data of 40 observatories of Russia and the CIS countries and stations "North Pole" are converted from paper to electronic form and published on the website http://www.wdcb.ru/stp/data/K_indices/; hourly mean values of the Earth's magnetic field elements recorded by the geomagnetic observatories of Russia and the CIS countries are converted to electronic form from paper and published on the website http://www.wdcb.ru/stp/data/magn_hour.val/ (work in progress).

The Center implements data publishing procedures using digital object identifier DOI ("Earth Science DataBase" project) http://esdb.wdcb.ru/. Our datasets are used and cited by the international community, as evidenced by citations in peer-reviewed articles and other media.

WDC for STP is listed as a trusted research data repository by re3data.org http://doi.org/10.17616/R3DC7Z.

In 2016-2017, we completely redesigned our website http://www.wdcb.ru/stp/index.en.html. Separate pages for each type of data with detailed information about data, format descriptions links to organizations producing data have been made.

Currently, the creation of a database "Geomagnetism" is carried out in the Center http://stp.wdcb.ru/. Hourly and minute values of geomagnetic field elements and values of K index according to the data of observatories in Russia and the CIS countries are loaded into the database. The data is thoroughly checked.

Reviewer 1:

Accept

Reviewer 2:

Accept

ORGANIZATIONAL INFRASTRUCTURE

I. Mission/Scope

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R1. The repository has an explicit mission to provide access to and preserve data in its domain.

The mission of the WDC for STP is the collection, guaranteed long-term storage and dissemination of geophysical data on the following disciplines of solar-terrestrial physics: geomagnetic variations, ionospheric phenomena, solar activity and interplanetary medium, and cosmic rays. The Center is responsible for data management. At present, more attention is being paid to the collection of national data.

Nowadays, the WDC for STP is guided by the principles of the ICSU-WDS Constitution and completely supports and agrees with the WDS data policy defined by the WDS Data Sharing Principles http://www.icsu-wds.org/services/data-sharing-principles. The Center provides full, open, timely, nondiscriminatory, and unrestricted access to data, metadata, products, and information, at no cost or at the cost of dissemination. Data, metadata, products and information in the WDC for STP repository are open and are available for scientific and educational use without restriction and free of charge. The Center aims to provide users with access to all data, metadata and products with minimum time delay. These statements can be found on the Center website http://www.wdcb.ru/stp/index.en.html.

The primary role of the Center is to maintain and support the availability and reuse of data and related information. The Center accepts research data and relevant documents from researchers and institutions in Russia and abroad and ensures that data is securely stored so that it can be used in the future.

The statement of responsibility for fulfilling the WDC functions and data management is written down in the Charter of the Geophysical Center of the Russian Academy of Sciences (host organization) – GC RAS http://www.gcras.ru/doc/Ustav_2018.pdf (in Russian).

WDC for STP prepares annual report on activities and every three years final report and presents them at the meetings of the GC RAS Scientific Council http://www.gcras.ru/eng/acouncil.php. The Scientific Council evaluates the work, progress and achievements of the WDC for STP. Always the WDC for STP receives a positive evaluation. The Report of Geophysical Center of RAS for 2017, results of the State Task (in Russian)

can be viewed on the website http://www.gcras.ru/ (the link at the top right) or http://www.gcras.ru/doc/yearreports/2017.pdf. Report on the implementation of research work within the framework of the State Task in 2018 http://www.wdcb.ru/stp/documents/Report_2018.pdf.

Reviewer 1:

Accept

Reviewer 2:

Accept

II. Licenses

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

Data collected in the WDC for STP are completely open. We provide data in accordance with the fundamental principle of the full and open exchange of data and information for scientific and educational purposes.

The Center is responsible for the safety of data transferred to the center for guaranteed unlimited storage and for international exchange. This responsibility covers data availability, integrity, persistence, secure storage. To encourage the use and reuse of data, WDC for STP provides free access to data through the website www.wdcb.ru/stp/index.en.html. The conditions and regulations of data distribution and license for data usage are determined by the data contributor in the deposit agreement. In case of unavailability of the contributor we agree with the relevant institution.

We place the corresponding license on the webpage with data and the user must follow the rules of this license. When we distribute data of the partners, we use a license defined by the partner. We distribute data under the international open access data license Creative Commons CC BY or CC BY-NC, corresponding to the definition of an open license given in article 1286.1 of the Civil Code of the Russian Federation http://www.gk-rf.ru/statia1286_1. There is only a restriction on the use of data for commercial purposes.

Until now, we did not get notice of any violations of these licenses by users. We currently have no special means of identifying non-compliance unless it is brought to our attention by others. We can only check the publications, remind of the citation. In case of violation of the license terms by the user, we will try to resolve this situation together. We plan to implement control by email address before the user can get the data to have better control over correct citations. We are working on this.

Reviewer 1:

It is recommended that the deposit agreement template is made publicly available.

Reviewer 2:

Accept

III. Continuity of access

Compliance Level: 3

Reviewer 1:

3

Reviewer 2:

3

R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

WDC for STP is responsible for the guaranteed long-term storage of geophysical data and ensuring access to data holdings for more than 60 years. Many historical data are available to users only in the WDC for STP repository. That is, in this case, the Center is the only custodian of such data. The Center is responsible for their safety and does everything to preserve and provide access to this data.

Since its creation in 1957, by the decision of the government of the former USSR, the WDC for STP is operated by the Geophysical Center of the Russian Academy of Sciences (GC RAS) which is funded by the Ministry of Science and Higher Education of the Russian Federation. WDC for STP is implemented as an integral part of the GC RAS infrastructure – it is a part of the Laboratory of geophysical data (See the Structure of the GC RAS http://www.gcras.ru/eng/structure.php). GC RAS has existed as an independent institution since 1954, and it is very unlikely that it will be closed.

Unless serious unforeseen conditions arise, we will continue our mission to ensure constant access to our data repositories and ensure their long-term storage. These plans are embedded in GC RAS Development Program for 5 years and in the long-term strategy for 10 years. The GC RAS management attempts to find a mechanism to better protection of its activities by searching for alternative sources of funding including support for the WDC for STP.

In case of unforeseen circumstances, the Russian Academy of Sciences will make a decision about where WDC for STP should be transferred to continue its activity and transfer its archives and functions to another body of the RAS, or the data will be transferred to the RAS Archive. In confirmation the Center has a letter from the Russian Academy of Sciences signed by Acting Chair of Earth Science Department of the RAS.

The WDC for STP has existed since 1957 and it is very unlikely that the Center will be closed.

Reviewer 1:

A continuity plan should be made publicly available in the future.

IV. Confidentiality/Ethics

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.

WDC for STP is guided by the provisions of Federal laws of the Russian Federation Federal law "On science and the state scientific and technical policy" http://docs.cntd.ru/document/9028333 (in Russian) and "On information, information technologies and information protection" https://fzrf.su/zakon/ob-informacii-149-fz/ (in Russian).

The WDC for STP repository contains only open data that can be freely distributed and used for research and educational purposes. WDC for STP provides open access to all data. There are no confidential information, data and personal data in the Center.

The data is accepted by the WDC only if they have received the status of the data intended for the international exchange/open access by the decision of expert commission of the organization that produces data.

A deposit agreement with a contributor reflects the conditions of data storage, distribution and the type of license relating to the data set. A license is provided for users, it defines the conditions for using the data. Information about the author/producer of the data or contributor is indicated on the data webpages. If the data of partners producing the data is published on the website, then the name of the producing organization and the partner's license is indicated.

The Center asks users to site data in their publications. On the webpages of the site there are samples for citation. For data with an assigned DOI, each landing page also contains instructions for data citation with its unique DOI.

Information about author/producer or contributor, sample for citation, DOI is reflected in the metadata.

Reviewer 1:

Accept

Reviewer 2:

Accept

V. Organizational infrastructure

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.

WDC for STP is hosted by the Geophysical Center of the Russian Academy of Sciences (GC RAS) and is incorporated into Laboratory for Geophysical Data, which is a structural section of the GC RAS. GC RAS is the Federal State Budgetary Institution of Science that receives funding from the state through the Ministry of Science and Higher Education of the Russian Federation. GC RAS has the state license on their activities.

The statement of responsibility of fulfilling the functions of the WDC is written down in the Charter of the GC RAS http://www.gcras.ru/doc/Ustav_2018.pdf (in Russian). The WDC for STP activity is based on the three-year research programs and annual business plan approved by the Scientific Council of the GC RAS and the Ministry of Science and Higher Education of the Russian Federation.

The staff of the WDC for STP consists of 3 research workers and two technical. The research workers are specialists of high qualification in geophysics. Each of them curates any discipline of solar-terrestrial physics and is responsible for all operations with data from collection to long-term storage and publication in the network. Their functions include communication with data authors, contributors and data users. The technical workers deals with the maintenance of the archive in good condition, including backup, and data registration. They are involved in the implementation of the project "Preservation of old data".

Every five years, scientific and technical personnel are certified by a commission created by the Director of the GC RAS, which includes representatives of the institution's administration, leading employees from each division. The activity of the GC RAS is reviewed by the Commission of the RAS.

The scientific seminar for the staff of the GC RAS on the topics of modern scientific problems and achievements is held every two weeks. This is useful for professional development of the staff. There is no dedicated permanent funding for staff training.

Sometimes the GC RAS provides an opportunity for the WDC's staff to participate in conferences and meetings.

The decrease in funding of WDC for STP is unlikely. The GC RAS has permanent funding and alternative sources of funding to ensure sustainable work of WDC for STP.

Reviewer 1:

Accept

Reviewer 2:

Accept

VI. Expert guidance

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either in-house, or external, including scientific guidance, if relevant).

Geophysical Center of the Russian Academy of Science (GC RAS) implements fundamental and applied research in the field of geophysics and geoinformatics. The main directions of its activity are study of geophysical fields and processes in various spheres of the Earth; development of new approaches to processing large volumes of complex geophysical databased on geographic information systems (GIS) technologies and methods of fuzzy logic and fuzzy mathematics; development of geological and geophysical applications of artificial intelligence methods and geoinformation systems; information support of scientific research.

The GC RAS functions as a basic institute for the Russian National Geophysical Committee http://ngc.gcras.ru/ and the Russian Committee of CODATA https://codata.org/russian-federation/. The GC RAS takes an active part in many international projects and national programs. Scientists and specialists of the GC RAS are members of international scientific organizations: CODATA (https://codata.org/about-codata/executive-committee/ Alena Rybkina Vice President 2018-2022), IUGG, IAGA (http://www.iaga-aiga.org/index.php?id=commission-histroy Anatoly Soloviev, Chair of the IAGA Interdivisional Commission on History). Employees of the GC RAS attend meetings within these groups whenever possible

Based on the above the WDC for STP has good expert and advisory support within the GC RAS. In addition, the Centre has external experts - some members of the National Geophysical Committee http://ngc.gcras.ru/sostav_eng.html and scientists of Pushkov Institute of Terrestrial Magnetism, the Ionosphere and Radio Wave Propagation of the RAS (IZMIRAN) https://www.izmiran.ru/?LANG=en. All experts are from Russia.

The repository communicates with the community through a message, an action or an event. ... In the presence of feedback, communication is transformed into a two-way process, in which it is possible to adjust the goals and behavior of communication participants.

Reviewer 1:

Accept

Reviewer 2:

Accept

DIGITAL OBJECT MANAGEMENT

VII. Data integrity and authenticity

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R7. The repository guarantees the integrity and authenticity of the data.

WDC for STP receives data from the geomagnetic observatories, ionospheric stations and world-famous research institutes which are partners of WDC for STP for many years. These organizations as contributors guarantee the integrity and authenticity of their data.

Data transmitted to the Center must meet certain requirements in order to be accepted for storage in repository and subsequently placement on the Center's website.

Each dataset and related documents submitted to the Center are analyzed to obtain the complete information for dataset registration and to produce the most complete metadata. Data quality and its conformity with the established standards are checked. These procedures are described in the internal instructions.

Complete information about the dataset is entered into the registration database http://www.wdcb.ru/stp/documents/Registration_Template.pdf .

The original data, metadata, and submitted documents are stored in their initial formats. All subsequent changes made to the dataset for preservation are documented in the log. All changes in data and metadata are documented in the WDC for STP, which guarantees the integrity of the data received for storage.

Complete information about each dataset is stored in the registration database on local computer. Metadata files and log are stored on the computer in the internal system of the Center. All versions of data archived are saved.

The Center ensures the safety of the original data and does not violate their integrity and authenticity in the event of any transformations (for example, when changing the format of data representation).

The dataset accepted for storage is published on the website. Information about dataset is published with the metadata, placed in metadata catalogue on the server http://www.wdcb.ru/metadata/stp/.

Regularly original data and its versions are checked to verify that they have not been altered or corrupted. The result of the check is written to the log.

When forming the database "Geomagnetism" the data is checked before loading, and the result of loading into the database is checked by comparing the downloaded and exported files with the same data. This verifies the data authenticity in database.

Verification of data for loading to the database is performed using special software.

We are in the process of assigning digital object identifiers (DOI) to data uploaded to the Center's website using Crossref agency. Metadata is used for this process. At the first stage, the DOI is assigned not to all datasets but only if there is a request for DOI assignment or we find it useful. Priority is given to geomagnetic and solar activity data. Information about assigning DOI to data sets and usage is available on the webpage with a description of the data http://www.wdcb.ru/stp/data.html.

Reviewer 1:

Accept

Reviewer 2:

Accept

VIII. Appraisal

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.

WDC for STP following Data Policy accepts all data relevant to the Center disciplines and corresponding to legal, qualitative and technical criteria. WDC for STP receives data from the Russian and the CIS countries observatories and research institutes usually on a permanent basis in the accepted international formats or formats used in the WDC. These data before transferring it to the Center are reviewed by experts to decide that data can be used freely for scientific and educational purposes.

Incoming data is evaluated by specialist-curators who know the methods and workflows used to create the data and decide whether the data set is suitable for storage. They perform quality control of data and checks the contents of the accompanying descriptive documents and accompanying metadata. If necessary, the contributor is asked to provide further information, for example about the tools and methods used to create the dataset. The presence of all necessary information for the completeness of metadata is checked. Data is checked for compliance with accepted formats and file structures (granularity). A list of preferred and accepted formats for each type of data is available on the website. If the repository has already related data, the new data is either added to existing data or replaced it, as noted in the metadata.

Some datasets are unique and prepared in special formats. The decision to accept such data is made on the basis of recommendations of experts. Such data is published on the website together with the unique format description and accompanying texts.

The metadata received along with the data from the contributor is supplemented after data analysis. The metadata of the ISO 19115 or DIF standard for each data set is compiled in the Center.

The Center receives data in scanned or paper form. They also undergo analysis and registration procedures. The catalogs of the availability of such data for different types of observations are supported. They are located on the relevant pages of the site. Data can be ordered by e-mail.

Individual data, after converting to electronic form and including into the general repository, are published freely on website in accordance with the license specified by data provider.

Formats for geomagnetic indices:

http://www.wdcb.ru/stp/geomag/format_K_index.html

http://www.wdcb.ru/stp/geomag/format_Kp_ap_Ap_ind.html

http://www.wdcb.ru/stp/geomag/format_AE_AU_AL_A%D0%9E_ind.html

http://www.wdcb.ru/stp/geomag/format_Dst_ind.html

Formats for 1-minute, hour mean and annual values of the geomagnetic field elements:

http://www.wdcb.ru/stp/geomag/format_minute.html

http://www.wdcb.ru/stp/geomag/format_hourly.html

http://www.wdcb.ru/stp/geomag/geomagnetic_annual.html

Format for ionospheric data http://www.wdcb.ru/stp/ionosphere/format_ion.ru.pdf

Formats for cosmic rays data http://www.wdcb.ru/stp/data/cosmic_rays/A_readme/.

Geomagnetic data can be selected from the database "Geomagnetism" (http://stp.wdcb.ru/index.php) in both 'wdc' format and in the alternative generally accepted format 'IAGA 2002'.

Reviewer 1:

Accept

Reviewer 2:

Accept

IX. Documented storage procedures

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R9. The repository applies documented processes and procedures in managing archival storage of the data.

WDC for STP stores data in accordance with documented processes and procedures. All procedures with data are managed in accordance with the instructions (data management plan) (R12) and are documented in the special log in the internal system of the Center. The instructions are developed based on long-term experience of data management in the WDC for STP. The instructions are reviewed regularly and can be changed if necessary. Each new dataset received by WDC for STP is considered as the original.

The original dataset and its copy are placed in appropriate sections of the repository for the long time preservation without any changes. Other copy of original dataset and all subsequent versions of the dataset and their copies are stored in another part of the repository. Regularly original dataset and all its copies are checked to verify that they have not been altered or corrupted and are moved on new storage media. The result of the evaluation is written to the log.

All these procedures allow you to restore any version of data at any time. In order to restore data on the server from short-term disasters their backup is used. Only curators have access to the internal archive storage. Secure password-protected access to the server is available only to the administrator and curators.

Storage management is carried out by the IT service of GC RAS, which provides regular backup, recovery (restoration), security. This provides our ongoing services to users.

Storage media and servers are located in fireproof, secure rooms.

The security policy and procedure for responding to information security incidents belongs to the GC RAS.

The IT service performs all the functions necessary to ensure the activities of the WDC for STP on the basis of the order of the GC RAS administration – an internal document defining the scope of tasks of this division (IT service), and official instructions for the IT service employees.

Reviewer 1:

Accept

Reviewer 2:

Accept

X. Preservation plan

Compliance Level: 3

Reviewer 1:

3

Reviewer 2:

3

R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

WDC for STP in Moscow since 1957 provides guaranteed long-term storage of accumulated archives and since the 1990s digital arrays of data on the disciplines of solar-terrestrial physics in accordance with its mission and Data Policy http://www.wdcb.ru/stp/policy.pdf.

The agreements between the data contributor and WDC for STP and licenses define the authority of the Center to manage data and perform the necessary actions with the data: storage, copying, transformation, and providing free access.

In each case of data transfer, the contributor and the repository agree on the format of the transferred data, metadata, and in what format and with what license the data will be stored in the repository and published on the website.

The Center takes all measures for reliable data storage. Data storage procedures are described in the data management instructions for personnel (R12).

Simultaneously with data storage, the Center provides free access to data via FTP server, through webpages, and from database "Geomagnetism" with a user-friendly interface to ensure ease of use http://stp.wdcb.ru/.

Data descriptions are provided on web pages. References to the author / contributor of the data and citation samples are also given there.

All changes made to the dataset are reflected in the log. All versions of the dataset are saved. Monitoring is regularly carried out against unplanned changes in data archives that are not stored in the database system.

With regard to non-digital data (paper documents), there is a plan for converting them into electronic form and make them available on the network.

From the very beginning of the WDC System, it was assumed that the WDCs would collect data and ensure their guaranteed long-term storage (for an indefinite period of time). The WDC for STP performs its functions in the belief that it should take care of the data in the long term perspective.

We don't have a documented preservation plan, but we will develop one.

Geophysical data does not become obsolete, both modern and old (historical) results of observations of processes on Earth and around are valuable. Storage and processing technologies and the forms of data presentation adopted for use may change (which has already happened more than once). We will take this into account in our conservation plan, which will include a description of all the steps from data collection to dissemination performed at the Center.

Reviewer 1:

A preservation plan should be made publicly available in the future.

Reviewer 2:

Accept

XI. Data quality

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.

All data is subject to technical and content verification by specialist-curators or external expert review before distribution. We cannot check the scientific quality of unique data sets, for example, author's catalogs of solar events, etc. In this case, we rely on external experts.

A significant part of the coming to the Center data is the result of long-term observations at stationary stations and observatories. These data go through a quality control process before transmission to the Center for long-term storage.

In any case, the technical condition and quality of the data is checked. Specially created software is used. First of all, the dataset is checked for format compliance. Then the data is checked for inadmissible characters, compliance with the chronological order of the records, if it should be so. Different parameters are checked for compliance with the established limits, whether the missing values are recorded accordingly, average sums of rows are checked, missing or duplicated lines are recognized. Analysis of dataset includes determination of the number of records in dataset, definition of a record's length etc. Finally the protocol of verification and analysis is formed.

In the case of detection of errors in the dataset, the Center in conjunction with the contributor makes a decision about the necessary changes in dataset and data correction is being implemented together with him. If it cannot be done, a notice about the error is recorded in metadata and in log.

If the format of the dataset does not match the accepted international format or used in the WDC, the dataset is converted into necessary format that is recorded in the log.

Metadata containing sufficient information for a complete description of the data and citing of the resource is compiled on the basis of the accompanying documents and analysis of the data entered into the repository. All datasets are accompanied by format descriptions. In addition to metadata, some data is

accompanied by descriptions to help interpret and use the data without referring to the data authors. We recommend contributors to submit documents describing their datasets.

The landing page for each DOI-assigned dataset represents metadata in accordance with international requirements. Mandatory metadata helps researchers evaluate data quality.

Data users have the ability to comment and evaluate data and metadata via email.

Reviewer 1:

Accept

Reviewer 2:

Accept

XII. Workflows

Compliance Level: 3

Reviewer 1:

3

Reviewer 2:

3

R12. Archiving takes place according to defined workflows from ingest to dissemination.

WDC for STP manages data stored in the Center for a long time, based on the established Data Policy http://www.wdcb.ru/stp/policy.pdf and a set of procedures recorded in the instructions for curators and determined the workflow from receiving to disseminating data – data management plan. All processes with data are documented in the journal and corresponding marks are made in the registration database.

At the time of data adoption for storage, possible changes and the conditions of data distribution, are reconciled with the contributor.

The workflow consists of the following procedures that are performed on a dataset:

(There is a separate instruction for each type of data. the basic procedures included in each instruction are listed here).

1.   The compliance of the dataset with the discipline and type of observation in the repository structure is determined.

2.   The dataset is analyzed (region/country, name/observatory, author/contributor, number of records, file size) and the storage location is determined.

3.   The information about dataset is entered into the registration database http://www.wdcb.ru/stp/documents/Registration_Template.pdf

4.   It is our responsibility to back up of all existing data. Two copies of a dataset are created.

5.   The original dataset and its first copy are placed in the repository for permanent storage. An entry is made in the registration database about their storage location.

        Further work is carried out with the second copy of dataset.

6.   Quality control of the dataset is carried out. The software package is specially worked out to perform this procedure. If errors are found in datasets, their correction is solved in coordination with the author/contributor or in cooperation with relevant experts.

The changes are incorporated into copy of the dataset. The relevant record in the log and a notation in the registration database are made.

7.   If necessary, the data is reformatted into the commonly accepted format. A notation in the registration database and record in the log are made.

8.   Corrected and reformatted copy of original dataset is duplicated and both are placed in the repository. Information on their storage location is entered into the registration database.

9.   The dataset is named according to the internal structure of the archive. An entry is made in the log and in the registration database. If necessary, the data file is divided into parts and names are assigned to each part.

10.   Metadata as XML file is prepared for this dataset. At this step, the information received from the contributor is important. It ensures that accurate and useful metadata is created.

11.   A hypertext document is prepared/modified for the dataset publication on the website.

12.   The publication of the dataset on the website. The date is entered in the registration database.

13.   The Procedures for assigning a DOI to a dataset are performed.

Instruction for metadata preparing for assigning DOI (in Russian)

http://www.wdcb.ru/stp/documents/DOI_manual.ru.pdf

Particular attention is paid to the rules of log filling, where all actions performed with each dataset and its copies are recorded. The log is stored in the internal system of the Center.

WDC for STP manages, maintains and provides services for the extensive archive of historical data on such media as paper, film and microfiches and provides all conditions for their long-time preservation. User can obtain information about these data from the inventory catalogues on the website.

Reviewer 1:

Accept

Reviewer 2:

Accept

XIII. Data discovery and identification

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.

The WDC for STP provides access to the datasets through webpages http://www.wdcb.ru/stp/data.html, where for each discipline there is a list of all types of observations and all data sets:

Geomagnetic Variations http://www.wdcb.ru/stp/geomag/geomagnetic_variations.html;

Ionospheric Phenomena http://www.wdcb.ru/stp/ionosphere/ionospheric_phenomena.html;

Solar Activity and Interplanetary Medium http://www.wdcb.ru/stp/solar/solar_activity.html;

Cosmic Rays http://www.wdcb.ru/stp/cosmic_rays/cosmic_rays.html.

The data archive is structured. As a rule, data sets are formed by observatories, by time (year or month) (for example http://www.wdcb.ru/stp/geomag/magnetogr_list.en.html). The webpage for dataset contain time period, sample of data citation, reference on author/ contributor, additional information, link to format description (for example http://www.wdcb.ru/stp/geomag/geomagn_Kp_ap_Ap_ind.html), DOI, directing the user to a landing page containing metadata about the data resource (for example http://www.wdcb.ru/stp/solar/solar_proton_events.html).

For historical data stored in the archive on paper or film, catalogs are compiled and placed on the webpages. (For example: Catalog of copies (microfilms and microfiches) of analog ionograms of the World Network of Ionospheric Stations http://www.wdcb.ru/stp/ionosphere/ionograms_catalogue.pdf, link on the page http://www.wdcb.ru/stp/ionosphere/ionospheric_phenomena.html; Catalogue of Copies (microfilms and microfiches) of analog magnetograms of the World Network of Geomagnetic Observatories http://www.wdcb.ru/stp/geomag/Catalogue_magnetograms.pdf, link on the page http://www.wdcb.ru/stp/geomag/magnetograms.html).

The MySQL database "Geomagnetism" is created for geomagnetic data http://stp.wdcb.ru/. The database is created to modernize the technology of storing large data sets and change the interface of access to them.

The database will include data from observatories in Russia and CIS countries: the series of average hourly, minute, annual values of the Earth's magnetic field elements and a unique collection of K-index values, from 1957 to the present. Data In one of three formats - WDC, IAGA2002, CSV - can be obtained from the database. The user interface includes data selection with different search criteria and provides the user with metadata. Currently, the hourly mean and minute absolute values are loaded to the database and can be retrieved.

Metadata compiled according to ISO 19115 for datasets are stored in a separate directory. From the webpages there are links to the corresponding section of the metadata catalog http://www.wdcb.ru/metadata/stp/. We plan to make a searching system for the metadata catalog when the funds are available.

The Center is in the process of assigning the digital object identifier (DOI) registered in the Crossref system to datasets and databases and presents data with DOI which provides the ability to cite data in scientific publications http://www.wdcb.ru/stp/documents/DOI_manual.ru.pdf.

At the first stage of the process, DOIs are assigned to individual data, preferably from the Geomagnetic Variations and the Solar Activity sections (for example http://www.wdcb.ru/stp/solar/solar_proton_events.html). Central repository http://esdb.wdcb.ru/ and landing pages for registered datasets have been created.

WDC for STP is entered in the Register of research data repositories re3data.org

(http://doi.org/10.17616/R3DC7Z) which facilitates access to data resources stored in the Center.

Reviewer 1:

Accept

Reviewer 2:

Accept

XIV. Data reuse

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.

In the WDC for STP metadata standards ISO19115 and DIF (Directory Interchange Format) are used to create metadata. We are in the process of creating a metadata catalog on our server http://www.wdcb.ru/metadata/stp/. We plan to make a searching system for the metadata catalog when the funds are available.

These standards are most suitable for Earth science data, in particular for spatially distributed data. They contain a set of attributes giving information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital data.

DIF standard was used for inclusion of geophysical data in Global Change Master Directory (GCMD) https://gcmd.nasa.gov/. GCMD metadata records are now stored in the Common Metadata Repository (CMR) and available in the CEOS International Directory Network (IDN) - a Gateway to the world of Earth Science data. Standards ISO19115 and DIF are compatible.

Metadata is available on landing pages for datasets with DOI (Central repository http://esdb.wdcb.ru/). (Examples of landing pages: http://esdb.wdcb.ru/doi/2016/mag-variations-database.html, http://esdb.wdcb.ru/doi/2018/esdb-sad-fe-01.html). In case of significant updates of the dataset, a new landing page will appears with its own DOI and details of the data change, and a new metadata file will be created.

To ensure permanent data understandability and easier reuse, data can be accompanied by a 'readme' file or corresponding contextual metadata, as well as format descriptions.

The repository provides data in the formats preferred by the users, usually WDC, IAGA2002, PDF, TIFF formats. PDF, Word document, TXT (ASCII) formats are used for documents.

The Center is working on the transfer of data from the WDC format to the IAGA2002 format. Due to database "Geomagnetism" the user has the opportunity to obtain data in one of three formats: WDC, IAGA2002 and CSV.

For long-term data storage, the original data format and formats in which the data is distributed are used.

If there is a need to migrate data to a new hierarchy, the data in the old form remains available to users. For example, it is still possible to access the geomagnetic data of each observatory as a separate file in WDC format via an FTP server.

Reviewer 1:

Accept

Reviewer 2:

Accept

TECHNOLOGY

XV. Technical infrastructure

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.

WDC for STP has all necessary software required for all types of work with data. Free and open licensed software products or software distributed free of charge MySQL, PHP, Leaflet, OpenOffice (for DataBase), MicrosoftOffice, PaperPort SE14 (software for data processing) are used in the Center. Furthermore, the Center has its own programs (PHP and Turbo Basic language are used for writing the program code) for the analysis, adjustments and modification of data sets.

External devices (HDDs hard disk drives) and network storage (NAS) Synology RS1219+ with 12TB of storage capacity and the increasing capacity by 8 times are used for the purposes of permanent archive storage. Backup and security options are carried out regularly.

In addition, backup copies of the WDC for STP's archive datasets are located in the robotic library (robotic IBM system Storage 3500 Tape Library) of the Research Institute of the Hydrometeorological Information - World Data Center (RIHMI-WDC) with extended period of storage.

Datasets and databases for online remote access are located on a server running the Linux operating system. The user can get data through a download from FTP server or website.

The technical infrastructure is maintained and administered by the IT staff of GC RAS. All documentation for the software is available and stored by the system administrator. He also keeps an inventory of software and provide for effective control of software licenses. The web server is connected by a fiber-optic communication channel to the southern public network of Moscow (with a node in the Moscow State University) with a bandwidth up to 1 Gbit/s. It works twenty four hours a day, seven days a week.

The IT service performs all the functions necessary to ensure the activities of the WDC for STP on the basis of the order of the GC RAS administration – an internal document defining the scope of tasks of this division (IT service), and official instructions for the IT service employees.

In the event of a failure, the disaster recovery plan is executed. It includes a list of actions required to recover data, indicating the most important recovery objects that include WDC for STP data. Backed up copies of data and files are used. They are obtained with the necessary frequency on a schedule in triplicate and stored elsewhere.

Reviewer 1:

Accept

Reviewer 2:

Accept

XVI. Security

Compliance Level: 4

Reviewer 1:

4

Reviewer 2:

4

R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

The Center assesses risks and creates a consistent security system. The repository is taking appropriate measures to counter the threats.

Original data and their backups are stored in different storages. An access to the storage is restricted and under control. Full backup (complete copy of all data) is carried out monthly and quarterly. Every week an incremental backup is done.

For quick restoration of a server, in emergency, the virtualization technology is used. The server works on a virtual machine, which is backed-up once a week. In case of a failure the server can be restored quickly.

For additional protection each member of WDC' staff (further "user") is supplied by the unique identifier with the unique password. If identifiers of users in system are accessible to the system administrator, then passwords of users are stored in the special codes and are known only to the users. Properly authorized persons or processes operating from their name acquire the right to read, write, create and delete the information. At each reference of the user to programs or the data, an admissibility of the given reference (a coherence in action of the specific user with the list of the actions permission for him) is checked.

Network access is made in the following way:

- Authentication, authorization, and access control for input registration data.

- Admission to certain working conditions according to the rules, prescribed for every user.

- Passwords are encrypted at application level, to conceal them from users, protocol HTTPS is supported, used for cryptographically protected channel for exchange of data.

- There is a firewall, which prevents any attempts to get unauthorized access to data.

- The system is monitored around-the-clock.

- The system is protected from harmful programs, including viruses, with the help of corresponding Anti-virus programs, which are linked to firewall, and spam-filters.

The building, where the WDC for STP is located, is guarded 24 hours a day, the entrance to the building is granted according to passes and documents, all received and returned keys are recorded in a journal, the fire and security alarm systems are installed in every part of the building.

Reviewer 1:

Accept

Reviewer 2:

Accept

APPLICANT FEEDBACK

Comments/feedback

The application form in our opinion is too rich and intensive document.

When completing the application form, some requirements caused difficulties. From our point of view, the names of the requirements do not always agree well with the requirements for which information should be provided.

For example, requirement R12 "Archiving takes place according to defined workflows from ingest to dissemination". Questions arise about the following topics: "Appraisal and selection of data" and "The types of data managed and any impact on workflow".