Russian Journal of Earth Sciences

Special Issue: "Data Science, Geoinformatics and Systems Analysis in Geosciences"

Problems of Processing of Raw Data from Magnetic Observatories: Will Modern Mathematical Technologies Help?

S. Yu. Khomutov[1]

[1]Institute of Cosmophysical Research and Radio Wave Propagation, FEB RAS, Paratunka, Kamchatka, Russia
* **Correspondence to:** Sergey Y. Khomutov, khomutov@ikir.ru

Abstract: Various aspects of the measurements and processing of raw magnetic data obtained at observatories are considered. It is noteworthy that the processing can be executed through simple mathematical methods and algorithms at almost all stages. Nevertheless, there are a number of tasks, for example, related to the mass recognition of noise in raw data and the need to fill in gaps, for the effective solution of which it is necessary to involve more powerful mathematical technologies.

Keywords: magnetic observatory, raw data, noise, modern mathematical methods.

Citation: Khomutov, S. Yu. (2025), Problems of Processing of Raw Data from Magnetic Observatories: Will Modern Mathematical Technologies Help?, *Russian Journal of Earth Sciences*, 25, ES2022, EDN: VTMATG, https://doi.org/10.2205/2025ES000994

1. Introduction

There are several problems that magnetic observatories currently face in their daily practice:

- Modern magnetometers allow measurements to be performed with greater frequency and greater sensitivity. This leads both to a simple increase in the volume of data received by orders of magnitude and to an expansion of the semantic content of this data, which was previously unavailable;
- The enhancement of magnetometer parameters and the increase in man-made load on magnetic measurements lead to the fact that noise is recorded, which previously had not been dealt with;
- Previously gained experience of the observatory staff becomes irrelevant.

On the other hand, the rapid growth of the volume and content of information in modern society leads to the development of new mathematical technologies aimed at solving emerging problems. In one way or another, these technologies can and are beginning to be used in magnetic monitoring at observatories.

Some steps of the work of standard INTERMAGNET observatory will be considered here, and an assessment of the need and possibility of using modern mathematical methods in this work will be given.

2. Absolute Observations and the Calculation of the Baselines Values for Variational Magnetometers

Absolute magnetic observations are an essential requirement for INTERMAGNET observatories [*INTERMAGNET*, 2020; *Jankowski et al.*, 1996]. As standard, they include manual observations of magnetic declination D and inclination I, the results of which are usually manually recorded in the protocol. After the observations, the observer types the data into a file. This procedure is one of the most common sources of possible errors that cannot always be detected. Modern text recognition methods, in principle, can avoid the manual procedure of typing data into a computer. However, it seems inefficient to use

Research Article

Received: 15 November 2024
Accepted: 15 April 2025
Published: 23 May 2025

this approach in the real practice of observatories since the protocols are handwritten and the records look very complex (a fragment of a typical record is shown in Figure 1). Even modern text recognition methods would have a hard time dealing with this text.

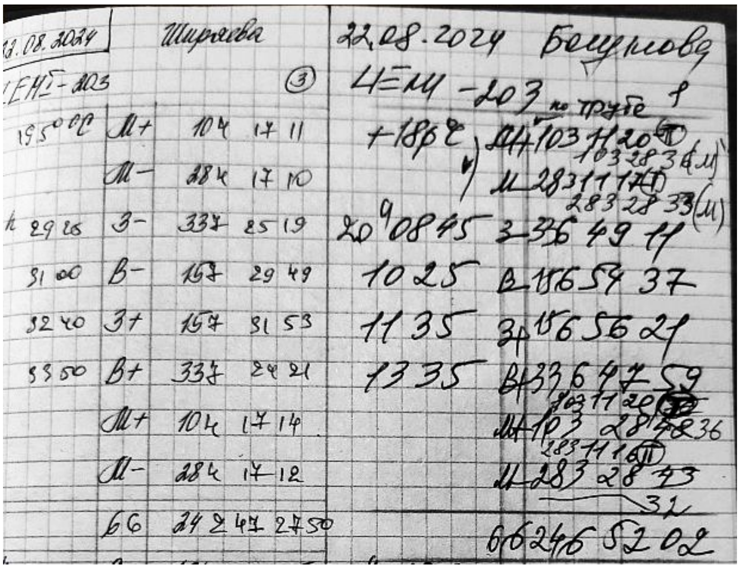

Figure 1. A fragment of the protocol with the absolute observations results (Observatory Paratunka PET).

The obtained baseline values are estimates of certain points in time. According to the absolute observation schedule established at the observatory, such estimates can range from a few values per day to several values per week. At the same time, it is necessary to know the behavior of the baselines as reliably as possible during each variometer measurement. Therefore, it is necessary to approximate the irregularly distributed data on a uniform grid. Various methods of approximation are employed in the practice of observatories, both formal and taking into account the physical features of the baseline's behavior.

To illustrate, you can take the information about the methods used during the preparation of the Definitive data of the INTERMAGNET observatories, which are usually reported in blv-files with baseline values. For example, in 2021, such information was presented in the data sets of 108 observatories:

- The use of polynomial and piecewise polynomial functions was 29%;
- The use of linear and piecewise linear functions was 34%;
- The use of cubic splines, including smoothing ones was 29%;
- The use of filters with various weight functions, visual approximation, etc. was 8%;

As you can see, traditional methods are used, which are often implemented in various standard software packages. This can be explained by the small volume of the processed baseline values, the absence of "hidden" uncertainties, and the effective control of the results by the observatory specialist. However, in extremely difficult cases, such as a long gaps between absolute observations or rapidly and irregularly changing baselines, the methods used remain rather traditional [*Gonsette et al.*, 2017; *Soloviev et al.*, 2018].

3. Daily Processing of Raw Data

The main task of the magnetic observatory is to obtain and provide users with data as complete as possible and, at the same time, as free as possible from any signals of artificial origin (noise). A significant, if not the main, amount of work of a observatory specialist is associated precisely with the analysis of raw magnetic data and the recognition of noise results of hardware failures. The identification problem is often solved if:

- The noise has an amplitude significantly exceeding the amplitude of natural signals;
- The structure (shape) of the signal obviously indicates its artificial origin;

- The signal has a regular repeatability over time and/or if information about its source is available;
- The signal recorded by spatially spaced magnetometers,

and in some other cases. Figure 2 shows an example of a regular noise with an amplitude of up to 4–5 nT in the data from the fluxgate variometer at the Paratunka observatory with deleted points marked with ×. The noise occurs during ionospheric soundings every 15 minutes.

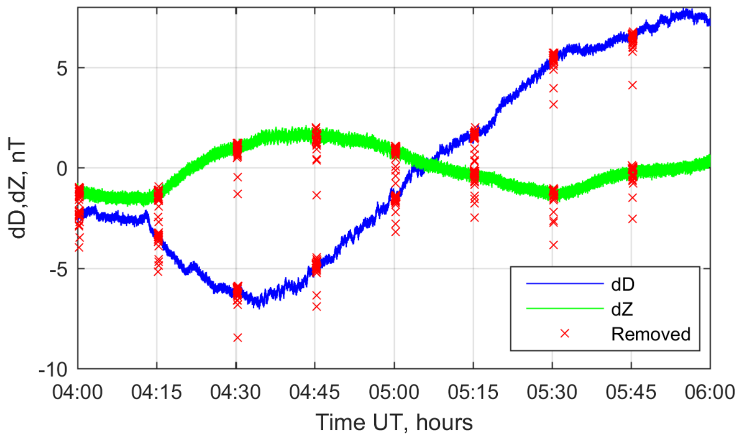

Figure 2. A sample of the records of the variations in the magnetic declination dD and vertical component dZ, in which there is a regular noise from the ionosonde, marked with the symbols ×. This noise was removed automatically.

Simple processing methods become ineffective if, for example, the noise has a small amplitude, morphologically differs little from natural magnetic variations, or there are many of them. Moreover, an artificial signal is more difficult to recognize when online or archived data is being processed. Data processing using modern mathematical technologies that take into account the more detailed and non-obvious characteristics of noise may prove be more effective in this case.

An example of the results of this approach is shown in Figure 3 (see [*Khomutov et al.*, 2017]). The daily recording in Figure 3a shows that in the vertical component Z at the observatory Hyderabad (HYB, CSIR-NGRI, India) there are frequent quasi-spikes with an amplitude of up to 5 nT, which are presumably associated with a branch of the skytrain passing near the observatory. Figure 3c shows a more detailed form of such spike. The result of noise recognition and clearing by the simplest method (by the spike amplitude taking into account its structural features) is shown in Figure 3d. A method for noise recognition based on continuous wavelet transform and threshold functions was also used, the results are shown in Figure 3b, e. It can be seen that a more complex method made it possible to remove spikes with an amplitude of up to 0.5 nT, in contrast to the simple method, which remained the noise with an amplitude of up to 1 nT. This technology has progressed further (see, [*Papsheva et al.*, 2019]), but it remained unclaimed in real practice at magnetic observatories.

In 2009–2012, a few articles were published with algorithms for the recognition of various types of the noise in minute and second magnetic data which were being developed at the Geophysical Center of the Russian Academy of Sciences (see [*Soloviev et al.*, 2009; 2012], [*Sidorov et al.*, 2012]). Algorithms based on the theory of discrete analysis and fuzzy logic methods have demonstrated their efficacy. Unfortunately, there are no details regarding the development of this technology. It is not known if it is actually used at most Russian observatories.

Machine learning methods were used in [*McCuen et al.*, 2023] to identify a noise of various types and origins in the measurement results at observatories in northeastern Canada over several years. To train the system, the expert (manual) processing of annual

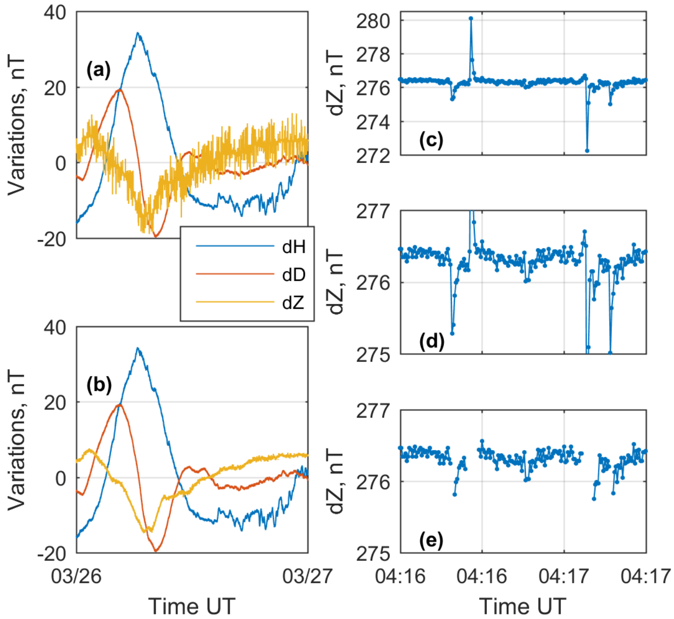

Figure 3. Frequent spikes in the vertical component Z at the Hyderabad Observatory (HYB): (a) Daily records of the variations dH, dD, dZ. (b) as (a), but the recognized spikes are removed. (c) The detailed form of some spikes. (d) Result of recognition and clearing by the simplest method. (e) as (d) but a more complex method was used.

raw data was performed, so a huge volume of magnetic measurement results was analyzed. The article perfectly illustrates the problem that raw data that is not processed promptly significantly increases costs and reduces the quality of the results obtained.

4. Conclusions

The experience of most Russian and many foreign observatories shows that at almost all steps of magnetic measurements, the processing of raw data is quite simple, and when performed every day and directly at observatories, modern high-tech mathematical methods are not required. The most critical points in this process are related to the recognition and correct cleaning of noise in the raw data, especially if there is a lot of noise, archived data is being processed, or the level of qualification of the observatory staff is low. Modern mathematical technologies with elements of artificial intelligence could be relevant.

Acknowledgments. The work was supported by the IKIR FEB RAS State Task (subject registration No. 124012300245-2; Russian Federation). The results presented in this paper rely on data collected at magnetic observatories. I thank the national institutes that support them and INTERMAGNET for promoting high standards of magnetic observatory practice (www.intermagnet.org). I also thank the reviewer, whose comments and suggestions undoubtedly allowed us to correct annoying errors and improve the article as a whole.

References

Gonsette A., Rasson J., Humbled F. In situ vector calibration of magnetic observatories // Geoscientific Instrumentation, Methods and Data Systems. — 2017. — Vol. 6, no. 2. — P. 361–366. — DOI: 10.5194/gi-6-361-2017.

INTERMAGNET Technical Reference Manual, Version 5.1.0 / ed. by B. V. St-Louis. — INTERMAGNET Operations Committee, Executive Council, 2020. — 146 p.

Jankowski J., Sucksdorff C. Guide for magnetic measurements and observatory practice. — Warshaw : International Association of Geomagnetism, Aeronomy, 1996. — 236 p.

Khomutov S. Y., Mandrikova O. V., Budilova E. A., et al. Noise in raw data from magnetic observatories // Geoscientific Instrumentation, Methods and Data Systems. — 2017. — Vol. 6, no. 2. — P. 329–343. — DOI: 10.5194/gi-6-329-2017.

McCuen B. A., Moldwin M. B., Steinmetz E. S., et al. Automated high-frequency geomagnetic disturbance classifier: A machine learning approach to identifying noise while retaining high-frequency components of the geomagnetic field // Journal of Geophysical Research: Space Physics. — 2023. — Vol. 128, no. 2. — P. 1–22. — DOI: 10.1029/2022JA030842.

Papsheva S. Y., Mandrikova O. V., Khomutov S. Y. Method of noise detection in magnetic data // Vestnik KRAUNC. Fiziko-matematicheskie nauki. — 2019. — Vol. 29, no. 4. — P. 87–97. — DOI: 10.26117/2079-6641-2019-29-4-87-97. — (In Russian).

Sidorov R. V., Soloviev A. A., Bogoutdinov S. R. Application of the SP algorithm to the INTERMAGNET magnetograms of the disturbed geomagnetic field // Izvestiya, Physics of the Solid Earth. — 2012. — Vol. 48. — P. 410–414. — DOI: 10.1134/S1069351312040088.

Soloviev A. A., Bogoutdinov S. R., Agayan S. M., et al. Detection of hardware failures at INTERMAGNET observatories: application of artificial intelligence techniques to geomagnetic records study // Russian Journal of Earth Sciences. — 2009. — Vol. 11, no. 2. — ES2006. — DOI: 10.2205/2009ES000387.

Soloviev A. A., Chulliat A., Bogoutdinov S. R., et al. Automated recognition of spikes in 1 Hz data recorded at the Easter Island magnetic observatory // Earth, Planets and Space. — 2012. — Vol. 64. — P. 743–752. — DOI: 10.5047/eps.2012.03.004.

Soloviev A. A., Lesur V., Kudin D. V. On the feasibility of routine baseline improvement in processing of geomagnetic observatory data // Earth, Planets and Space. — 2018. — Vol. 70. — P. 1–14. — DOI: 10.1186/s40623-018-0786-8.