Earth, Planets and Space

FULL PAPER Open Access

Data agreement analysis and correction of comparative geomagnetic vector observations

Zhaobo He[1], Xingxing Hu[1*], Yuntian Teng[1], Xiuxia Zhang[2] and Xiaoyu Shen[1]

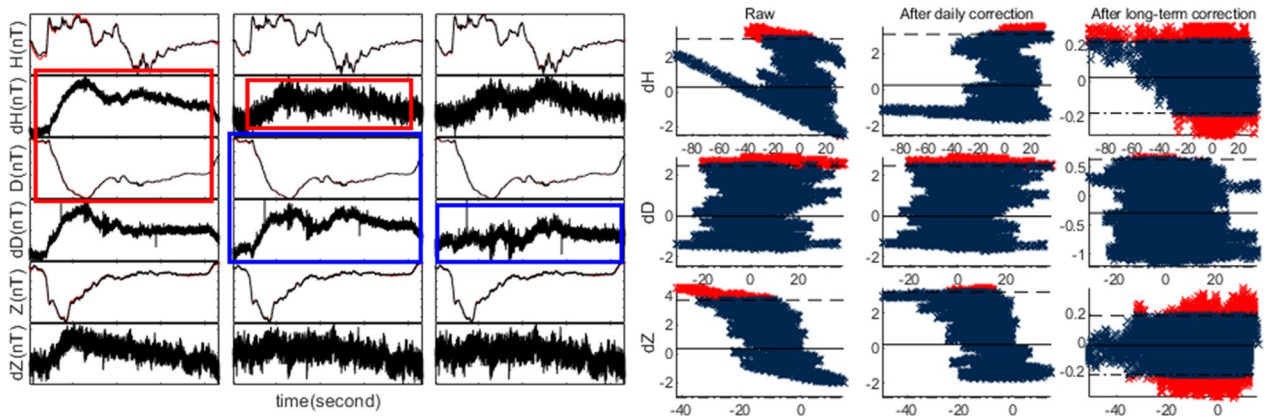
Abstract

Geomagnetism, similar to other areas of geophysics, is an observation-based science. Data agreement between comparative geomagnetic vector observations is one of the most important evaluation criteria for high-quality geomagnetic data. The main influencing factors affecting the agreement between comparative observational data are the attitude angle, scale factor, long-term time drift, and temperature. In this paper, we propose a method based on a genetic algorithm and linear regression to correct for these effects and use the distribution pattern of points in Bland–Altman plots with a 95% confidence interval length to qualitatively and quantitatively evaluate the agreement between the comparative observational data. In Bland–Altman plots with better agreement, that is, with the corrected data, more than 95% of the points are distributed within the 95% confidence interval and there is no obvious pattern in the distribution of the points. Meanwhile, the length of 95% confidence interval decreased significantly after the correction. The method presented here has positive effects on the vector instrumentation detection and would enhance the robustness of geomagnetic observatory by bringing the data quality of the backup variometer data in line with the primary variometer.

Keywords: Geomagnetic vector comparative observations, Data agreement, Genetic Algorithm, Bland–Altman plot

*Correspondence: huxx@cea-igp.ac.cn
[1] Institute of Geophysics, China Earthquake Administration, Beijing 100081, China
Full list of author information is available at the end of the article

Graphical Abstract



Main text

Introduction

Vector observations of the geomagnetic variation field are the primary means used to study internal and external magnetic field sources, such as rapid magnetic variations and magnetospheric currents (Curto et al. 2007; Xu et al. 2015), and the instrument most commonly used for such observations is the fluxgate magnetometer (Jankowski and Sucksdorff 1996). To enhance the operational robustness, some observatories use two sets of fluxgate magnetometers with the same type of probe to enable comparative observations. However, actual comparative observational data can be somewhat divergent, that is, the measurement data from the two sets of instruments do not exactly agree. This leads to a reduction in the credibility of the data recorded by one instrument when the other fails, to the point, where it is impossible to determine whether it is reasonable to use the data from the backup instrument at the time of failure. Furthermore, vector observations of the geomagnetic field are expected to qualitatively reflect the magnetic field variations at the measurement points and, therefore, to invert the relevant physical mechanisms. Morphological differences in the comparative observational data might cause uncertainty on the studies related to physical mechanisms of rapid geomagnetic variations, such as geomagnetic sudden commencements (Araki et al. 2004; Segarra and Curto 2013). Therefore, analyzing and correcting the agreement between the comparative geomagnetic vector observational is a basic but important step for data quality assurance.

Morphological differences in the comparative geomagnetic vector observational data have a variety of causes. For fluxgate magnetometers with the same type of probe, influencing factors that can introduce significant magnetic measurement differences include the attitude angle, scale factor, long-term time drift, and temperature. For daily variations, the attitude angle and the scale factor cause the geomagnetic vector difference to show a characteristic pattern and would bring about 2 nT fluctuations in measurement difference according to different parameter deviations. The effects of the long-term time drift and temperature are more often seen in long-term observations. In China, most geomagnetic observatories have a strict temperature control of 0.04 °C/day in their variation rooms. The variation rooms of the three Chinese geomagnetic observatories selected for this study all meet this requirement. Therefore, when examining the agreement between the daily-variation data, only two aspects—the attitude angle and the scale factor—are analyzed. For long-term observations (greater than or equal to 3 months), the effects of the long-term time drift and temperature cannot be neglected and would bring a significant fluctuation in measurement difference which can exceed even 5 nT.

The traditional index for evaluating the agreement between comparative geomagnetic vector observational data is the Pearson correlation coefficient (Han et al. 2004; Berezin and Tlatov 2020). This correlation coefficient gives a clear agreement judgement when comparing measurement data from two different means of observation (e.g., satellite magnetic observations versus ground-based observations). When the measurement platform is consistent, the observational environment is good, the instrument quality is high, and the correlation coefficient is often very close to 1. However, the geomagnetic vector differences can still be relatively large and show some regularity. This means that the correlation coefficient does not adequately distinguish the degree of agreement between the comparative observational data. This paper proposes a qualitative and quantitative analysis

of the agreement between comparative geomagnetic vector observational data using Bland–Altman (B–A) plots. Disagreement can be visually detected from the shape of the distribution of the points on the B–A plots, and the length of the 95% confidence interval can significantly distinguish the superiority or inferiority of the agreement.

To analyze the influencing factors affecting the agreement between the comparative geomagnetic vector observations and to calculate the corresponding correction parameters, the Lijiang (LIJ), Maguan (MAG), and Yunlong (YUL) geomagnetic observatories in southeastern China, which have two sets of fluxgate magnetometers operating simultaneously, were selected. These three observatories are equipped with two sets of GM4 fluxgate magnetometers with a resolution of 0.01 nT and a sampling rate of 1 Hz. The magnetic variometer GM4 is developed by the Institute of Geophysics, China Earthquake Administration. GM4 has a Φ180 mm × 100 mm probe and linear cores made of permalloy. When GM4 works in compensated operation mode, the dynamic range is ± 2500 nT and the linearity is better than 5‰ (Shen et al. 2021). The selected data range is from January 1, 2020, to July 31, 2021.

Characteristics of and calculation methods for the correction parameters

Attitude angles are important for geomagnetic vector observations. Traditionally, the determination of the azimuth of the fluxgate magnetometer by the staff is made in conjunction with the fluxgate theodolite according to the geographical orientation (Jankowski and Sucksdorff 1996). A high-precision tiltmeter is used to record the tilt of the instrument above the abutment. Alternatively, a suspended fluxgate magnetometer is used to circumvent the effects of tilt. However, in practice, tilt effects are often ignored for observatories in non-permafrost environments. Unattended platforms, such as the SeaFloor ElectroMagnetic Station, use direct attitude measurements by means of tiltmeters and fiber optic gyroscopes (Toh et al. 2006).

Researchers have also tried to correct the attitude between two fluxgate magnetometers using genetic algorithms (Liu et al. 2019). A genetic algorithm is a computational model that simulates the process of natural selection and genetic evolution observed in biological evolution; this is a well-established method for finding the global optimal solution of an objective function (Holland 1992; Weile and Michielssen 1997). Earlier studies have used genetic algorithms to calibrate the orthogonality of fluxgate magnetometer probes (Jiao et al. 2011). Three-component fluxgate magnetometer data naturally contain attitude information, and ideally the difference in the data between the two sets of instruments only arises from the difference in the attitude angles. Accordingly, the genetic algorithm calculates the attitude relationship between two fluxgate magnetometers and has natural advantages such as high measurement accuracy, no interference with the probe, and a minimal use of peripheral instruments.

A genetic algorithm for an optimization search problem usually consists of basic steps such as population initialization, fitness evaluation, selection, recombination, mutation and replacement (Sastry et al. 2005). The number of individuals and their characteristics is usually used to artificially designate population. These characteristics as decision variables are coded to facilitate the computer practice of selection, recombination and mutation steps. In this paper, individuals' characteristics are attitude angles and scale factors. The value range of attitude angle should strictly be [− 180°, 180°], but usually we have a rough judgment of the relative attitude angle of the two sets of instruments. [− 10°, 10°] is sufficient for observatory data. In addition, the range of the scale factor is usually taken as [0, 1.5]. The fitness evaluation should be able to give quantitative indicators to distinguish between good and bad results through the objective function. Selection, recombination, and mutation mimic the principles of nature selection and genetics, giving offspring with different characteristics that are better adapted to the "objective function" to achieve higher fitness. In this paper, universal truncation selection and uniform crossover are chosen as the operators of selection and recombination. To avoid a high possibility of genetic patterns being corrupted and based on the program effect, the crossover probability and the mutation probability are taken as a typical value of 0.7 and 0.047, respectively. For the generality and computational efficiency of genetic algorithms, the choice of operators and parameters is very important. However, for the physical interpretation of specific optimization search problems, more attention should be paid to the objective function which is used for fitness evaluation.

The objective function *Obj* of the genetic algorithm used in this paper is the sum of the absolute values of all elements in the objective matrix ***ObjM*** derived below, as shown in Eq. (1).

$$Obj = \sum_{j=1}^{n} \sum_{i=1}^{3} |D_{i,j}| \tag{1}$$

In Eq. (1), $D_{i,j}$ is the i, j entry of matrix ***ObjM*** and n is the length of the data set. The objective function *Obj* has a global minimum when the decision variables, which are the attitude angle and scale factors in this paper, take appropriate values.

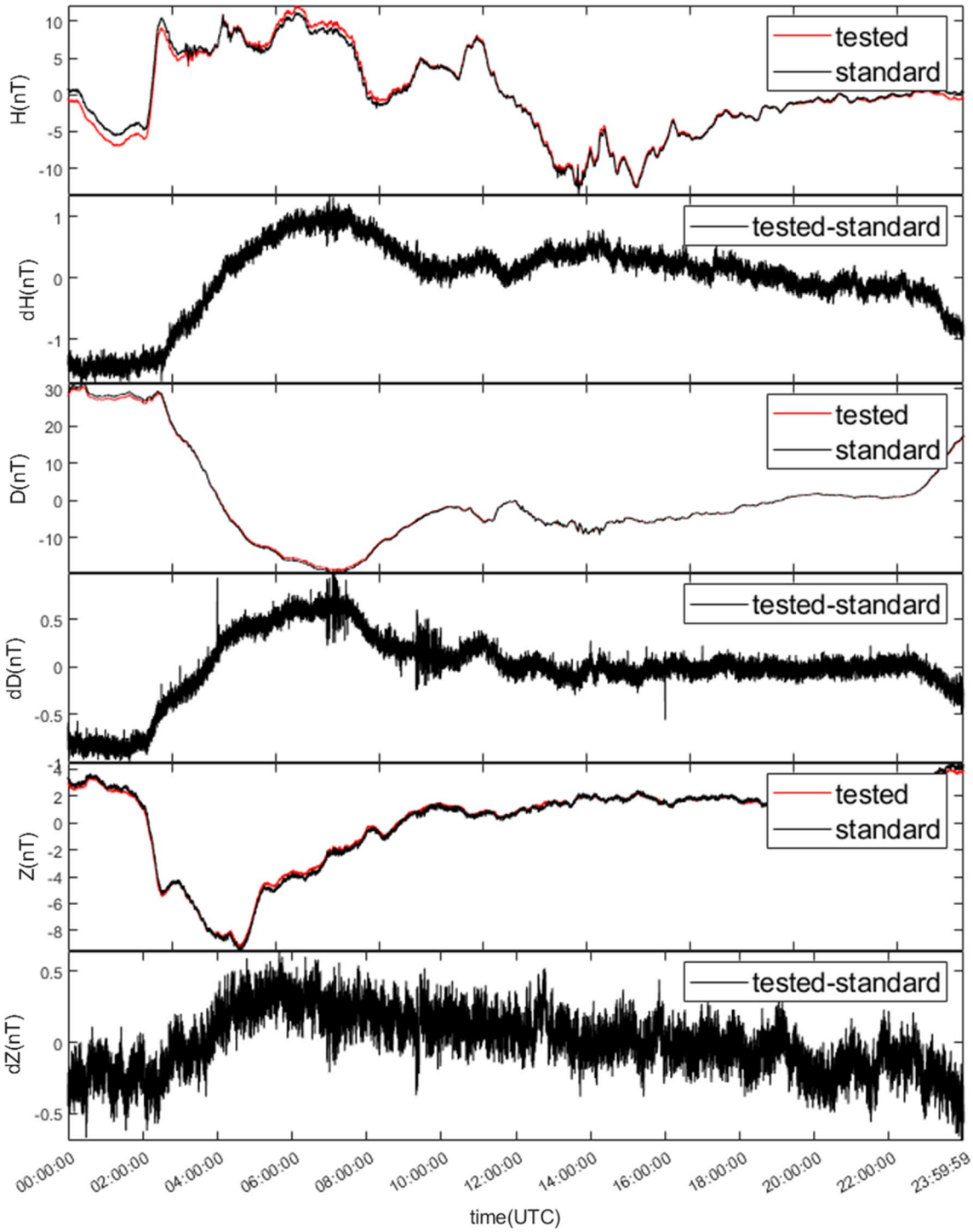


Fig. 1 Influence characteristics of the attitude angle and scale factor in the morphology of comparative geomagnetic observations. The figure shows the daily geomagnetic variation and vector difference on June 17, 2020, at the Yunlong (YUL) observatory. Because of the heading angle deviation, the difference, dH, has a similar morphological change to the D component. Because of the scale factor, the dD and dZ differences are similar to the morphological variations of the D and Z components, respectively

The objective matrix ***ObjM*** is the difference between the value of the tested instrument data normalized to the standard instrument coordinate system and the value of the standard instrument data, as shown in Eq. (2).

$$\begin{aligned} \boldsymbol{ObjM} &= \boldsymbol{Data}'_{\text{standard}} - \boldsymbol{Data}_{\text{standard}} \\ &= \boldsymbol{T}^{-1}\boldsymbol{Data}_{\text{test}} - \boldsymbol{Data}_{\text{standard}} \end{aligned} \tag{2}$$

In Eq. (2), the subscripts standard and test represent the data in the standard coordinate system and in the tested coordinate system, respectively, and the superscript ' represents data converted from another coordinate system. $\text{Data}_{\text{test}}$ and $\text{Data}_{\text{standard}}$ denote the comparative observational data of the H D Z components of the magnetic induction intensity obtained from the observatories. Both sets of data are only performed the demeaning process before entering the genetic algorithm analysis. The calculation results of the vector relative observational data collected from the three observatories and the vector absolute observational data collected from the Lijiang experiment show that the demeaning process in the daily correction does not affect the calculation of the scale factor and can significantly improve the accuracy of the attitude angle calculation. The rotation matrix $\boldsymbol{T}$ is obtained by multiplying the rotation matrices represented by the tilt angles Roll, Pitch, and the heading angle Yaw in sequence during the conversion from the standard coordinate system to the tested coordinate system, as shown in Eq. (3).

$$\begin{aligned} \boldsymbol{T} = \boldsymbol{T}_{\text{Roll}}\boldsymbol{T}_{\text{Pitch}}\boldsymbol{T}_{\text{Yaw}} = &\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\text{Roll}) & \sin(\text{Roll}) \\ 0 & -\sin(\text{Roll}) & \cos(\text{Roll}) \end{bmatrix} \\ \cdot &\begin{bmatrix} \cos(\text{Pitch}) & 0 & -\sin(\text{Pitch}) \\ 0 & 1 & 0 \\ \sin(\text{Pitch}) & 0 & \cos(\text{Pitch}) \end{bmatrix} \cdot \begin{bmatrix} \cos(\text{Yaw}) & \sin(\text{Yaw}) & 0 \\ -\sin(\text{Yaw}) & \cos(\text{Yaw}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{3}$$

This order cannot be changed, because the attitude angles correspond to the rotation angles only when the matrix is multiplied in this order.

The presence of deviations in the attitude angles between the two fluxgate magnetometers results in a geomagnetic vector difference in one direction, reflecting a morphological change in the other direction due to projection (Wang et al. 2017). Taking the deviation of the heading angle Yaw as an example, as shown in Fig. 1, the difference dH of the H component reflects the morphological variation of the D component.

Even for the same probe, the coefficient used to convert the voltage to the magnetic induction intensity during the instrument commissioning process is not the same. In this paper, we refer to the parameter that results in measurement differences between the two instruments due to differences in voltage-magnetic conversion coefficient as the scale factor. Therefore, the scale factor matrix also needs to be added to the objective matrix, as in Eq. (4):

$$\begin{aligned} \boldsymbol{ObjM} &= \boldsymbol{Data}'_{\text{standard}} - \boldsymbol{Data}_{\text{standard}} \\ &= \boldsymbol{Sf}^{-1}\boldsymbol{T}^{-1}\boldsymbol{Data}_{\text{test}} - \boldsymbol{Data}_{\text{standard}} \end{aligned} \tag{4}$$

Here, the matrix ***Sf*** is the scale factor matrix, which is a diagonal array with its diagonal elements being the scale factors of the corresponding components. Even in the ideal case, where there is no relative attitude angle between the two sets of instruments, the voltage-magnetic conversion coefficient is still not the same. Therefor scale factor naturally exists and acts on the tested instrument data first, as in Eq. (5), and the attitude rotation matrix and the scale factor matrix acting on the tested instrument data are not interchangeable.

$$\boldsymbol{Data}_{\text{test}} = \boldsymbol{T} \cdot \boldsymbol{Sf} \cdot \boldsymbol{Data}'_{\text{standard}} \tag{5}$$

Note that the attitude angle and scale factor parameters are for the instrument being tested relative to the standard instrument. The positive direction of the angle is a counterclockwise rotation around the rotation axis, and each scale factor consists of the tested instrument data divided by the standard instrument data.

The presence of the scale factor between the two instruments results in a geomagnetic vector difference in one direction, reflecting a morphological change in this same direction. As shown in Fig. 1, the D component difference, dD, exhibits its own morphological variation, that is, it exhibits the morphological variation of the D component.

In the literature (Liu et al. 2019), after selecting the appropriate genetic algorithm parameters, the tested instrument data are first obtained by an angular rotation

Table 1 Calculated results of the comparative attitude angle observation experiment

No.	Roll (°)	Pitch (°)	Yaw (°)
1	− 1.139	0.674	30.129
2	− 0.866	1.136	30.075
3	− 0.978	0.202	29.953
4	− 1.122	1.737	29.942
5	− 1.001	1.091	29.932
6	− 1.450	1.376	29.922
7	− 1.063	1.565	29.909
Average	− 1.088	1.112	29.980
Standard deviation	0.171	0.491	0.079

of the standard instrument data and then a genetic algorithm is used to calculate the attitude angle. The difference between the corrected tested instrument data and the standard instrument data is up to 10^{-4} nT (the maximum absolute value of the difference). We can make slight improvements by running each attitude angle calculation 10 times; then, after excluding results that are more than one standard deviation from the mean, the mean value can be used as the attitude angle solution to further constrain the convergence of the genetic algorithm solution. In this way, the maximum absolute difference between the tested instrument data after the attitude angle correction and the standard instrument data can reach 10^{-5} nT.

Actual observations are often not so ideal. During the period of June 2–July 15, 2021, we conducted related experiments at the LIJ. Two sets of fluxgate magnetometers of the same type were used to make comparative observations. The standard instrument and the tested instrument were spaced 7-m apart, and both were placed on a stone pier in the variation room, with a deviation of 30° in the heading angle between the two. The 7 days with the smallest standard deviation for the results calculated by the genetic algorithm were taken, and the results are shown in Table 1.

The calculated heading angle of 29.980° is very close to 30° with a small standard deviation of 0.079°, which indicates that the method used here is able to calculate large existing angles between two fluxgate magnetometers with high accuracy. However, smaller angles, where the standard deviation of the calculated angles over multiple days is greater than the mean value, need to be considered separately.

In the comparative geomagnetic vector observations, the static attitudes of the two fluxgate magnetometers often differ to some extent. Furthermore, the vector difference between the tested instrument data and the standard instrument data after the attitude angle correction is stabilized within approximately 0.5 nT. This requires that the standard deviation of the calculated angle be less than 0.057° (e.g., for a residual magnetic field of 100 nT). However, this requirement is not always satisfied as a result of the quality of the data or disturbances in the observational environment. Therefore, when designing the correction process, we chose, as the criterion to judge the correct calculation of the attitude angles, the uncertainty of all three angles of the multi-day calculation results to be less than 0.1° or the uncertainty of the calculation results of significantly large attitude angles (greater than 1°) to be less than 0.057°.

As for the scale factor, the calculated uncertainty (expressed as the standard deviation of the multi-day scale factors) is small. It is approximately 0.002 for observatories with a good observational environment, resulting in an uncertainty of less than 0.1 nT for the magnetic field. In this paper, after determining the relative attitude angles of the two instruments, we calculate the scale factor for each day in 3 consecutive months and perform linear regression to obtain the base scale factor (intercept) and the long-term time drift (slope). Instrumental scientists and engineers assume that the long-term time drift of a fluxgate magnetometer is linear (Gordon and Brown 1972; Esper 2020); such behavior is characterized by tiny fluctuations in the short-term observations and non-negligible and linear variations in the long-term observations. If the long-term time drifts of the two instruments are different, there will be a linear change in the scale factor between the instruments over time. This is a relative relationship and does not specify whether the drift comes from the tested instrument or the standard instrument or both. However, for the correction of the long-term observational agreement, it is sufficient to assume that the drift comes from the tested instrument.

Temperature has an important effect on fluxgate magnetometers (Primdahl 1979). New fluxgate magnetometers have been able to achieve a thermal drift of less than 0.1 nT/°C in the laboratory (Korepanov and Marusenkov 2012). However, the effect of temperature on fluxgate magnetometers is still very important and non-negligible in long-term observations. Even though the temperature difference between two sets of instruments in the same variation room is nearly constant (the mean value of the temperature difference between the two sets of instruments at LIJ from January 1, 2020, to March 31, 2020, was 0.9672 °C, and the standard deviation was 0.0863 °C), the measurement difference caused by the fixed temperature difference is not constant, which means that the temperature change affects the measurement difference between the two sets of instruments. The top section of Fig. 2 shows the Z component difference, dZ, and temperature variation of the two sets of instrumental data from January to March 2020 at LIJ after the attitude angle, scale factor, and long-term time drift corrections. The dZ pattern, showing a decrease followed by an increase, is very similar to the temperature variation. This relationship is approximately linear, as seen in the scatter plot of dZ versus temperature (middle section of Fig. 2). The temperature, rather than the temperature difference, has a linear effect on the vector difference between the two sets of instruments. This can be interpreted as the difference in the temperature coefficients between the two sets of instruments leading to differences in the measurements at different temperature points (the red line at the bottom of Fig. 2), even though the temperature

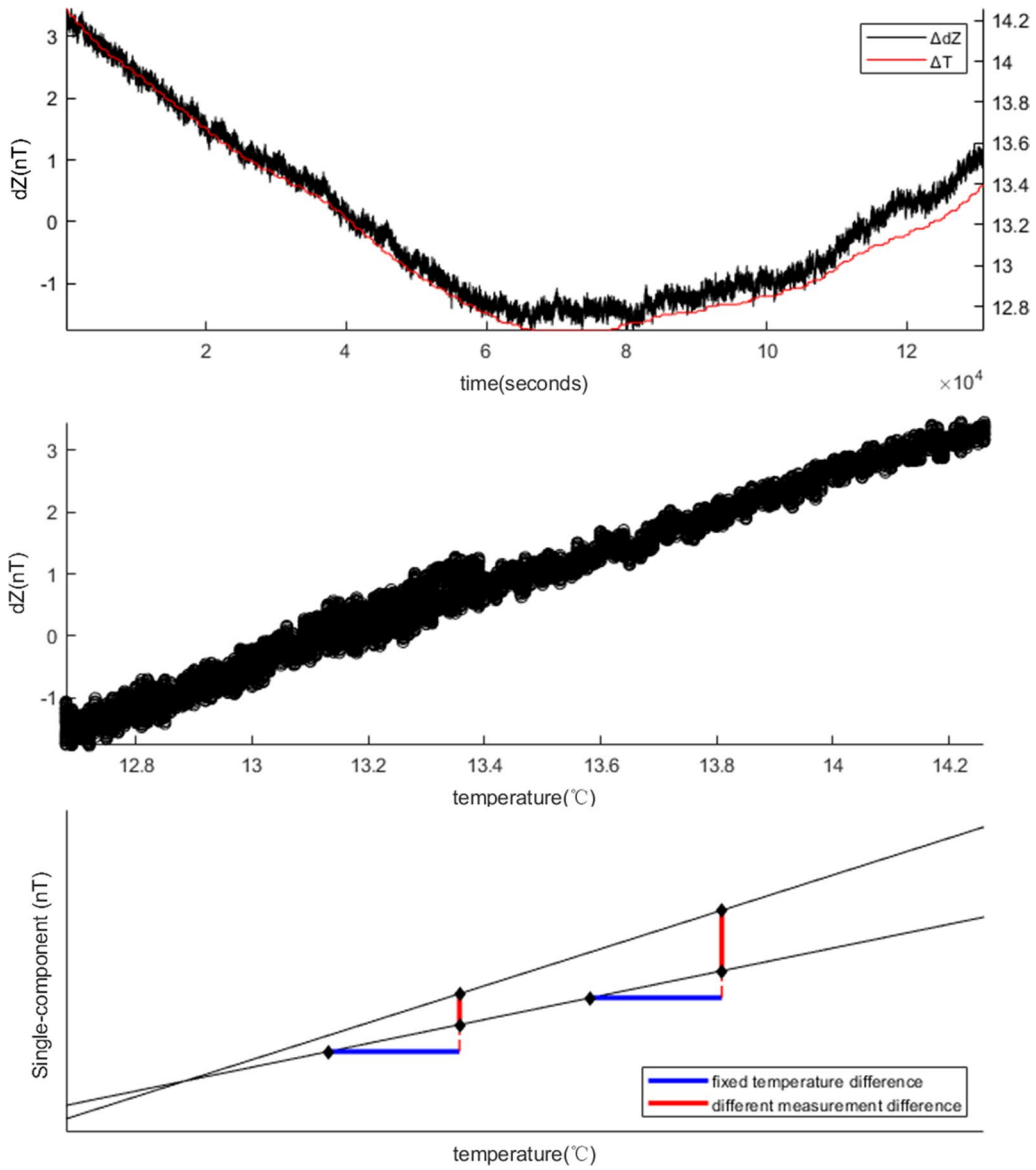


Fig. 2 Influence of temperature on comparative geomagnetic vector observations. The relative temperature coefficient describes the linear effect of the temperature on the vector difference between the comparative observations, mainly resulting from the different temperature coefficients of the two sets of instruments

difference is always constant (the blue line at the bottom of Fig. 2).

Parameter estimation and data corrections

The morphological differences and the data disagreement in the comparative geomagnetic vector observations are primarily due to the attitude angle, scale factor, long-term drift, and relative temperature coefficients. The calculations of the attitude angle and the scale factor are based on the genetic algorithm. The long-term time drift and relative temperature coefficients are then obtained via linear regression. The long-term correction calculation flow considering the above four parameters is shown in Fig. 3.

The calculation results for the attitude angle and the scale factor based on the genetic algorithm are shown in Table 2.

Of the nine attitude angles shown in Table 2, only the three attitude angles of YUL and the heading angle of MAG are available, because their mean values are large and their standard deviations are small. In particular, the computed mean values of the heading angles of the YUL

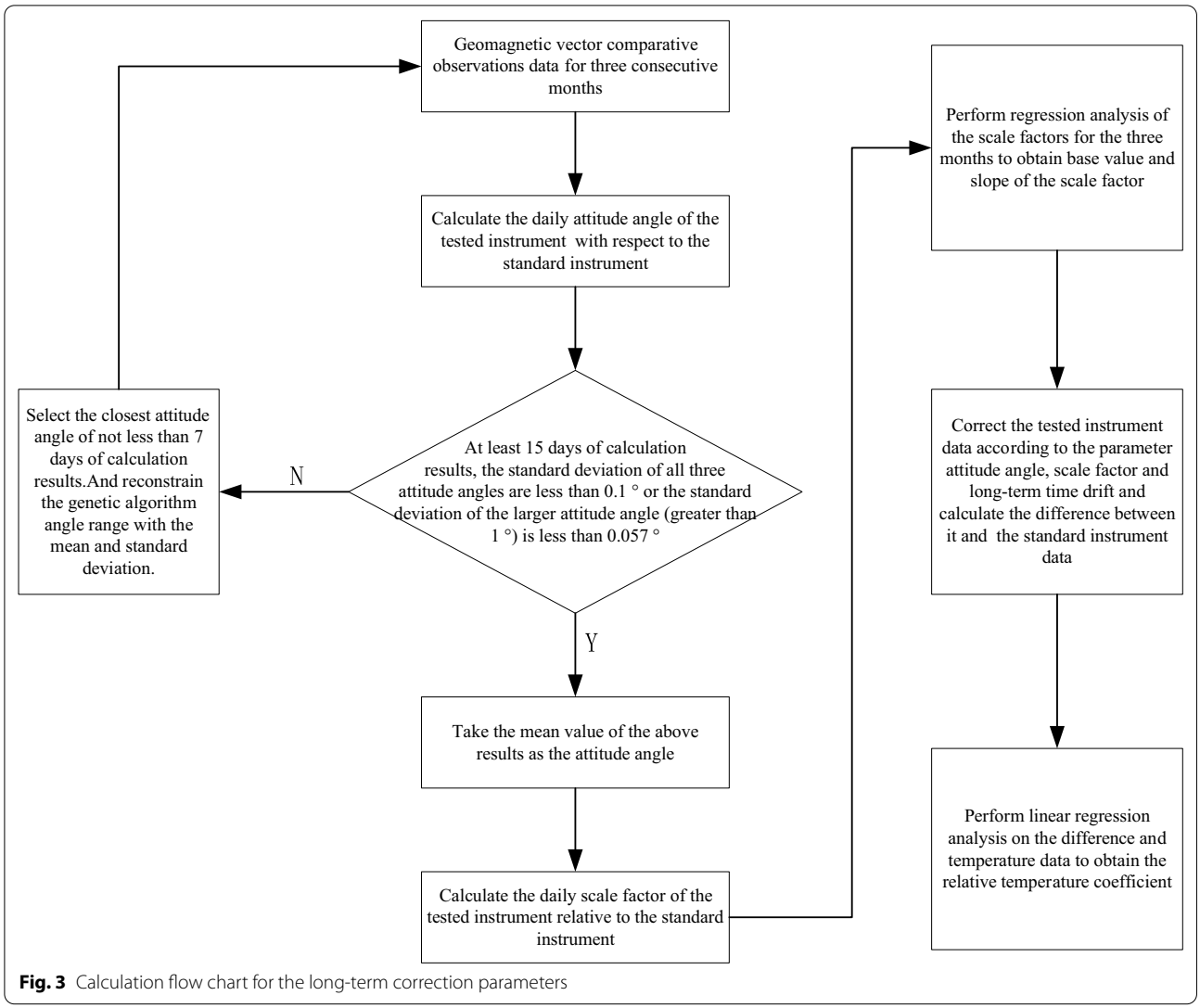


Fig. 3 Calculation flow chart for the long-term correction parameters

and MAG exceed 1°, while their standard deviations are less than 0.057°. The remaining five attitude angles have average values close to 0 and standard deviations greater than the average. In the actual correction procedures, it was assumed that there was no angular deviation on the axis between the two sets of instruments.

All three observatories have large heading angle deviations between the two sets of instruments, while only the YUL has a large tilt angle deviation, which is mainly related to the horizontal calibration and orientation of the instruments.

A common method of orienting the magnetometer to the magnetic field coordinate system is to turn the magnetometer carefully so that the uncompensated D-component shows zero value in undisturbed field after level calibration (Jankowski and Sucksdorff 1996). Orienting the tested and standard instruments at different times obviously brings some deviation in heading angle deviation, which is in good agreement with the large deviation in heading angle at all three observatories. The main reason for the large tilt angle deviation at YUL compared with the other two observatories is considered to be that there is no vacant marble pillar in the variation room at YUL so that the tested non-suspended magnetometer GM4 is calibrated at the ground level and thus some horizontal error occurred.

The calculation of the scale factor was performed after the attitude correction. The results of 3 consecutive months of calculations were linearly regressed after removing significantly erroneous data to obtain the intercept and slope, which were used as the scale factor and long-term time drift correction parameters, respectively.

The parameters used to correct the single-day data are the attitude angle and the scale factor. For the sake of

Table 2 Calculation results for the attitude angle and the scale factor

YUL 2020/05–2020/07	Attitude angle (°)		
	Roll (ave/std)	Pitch (ave/std)	Yaw (ave/std)
	0.487/0.087	0.421/0.097	−2.341/0.074
	Scale factor		
	H	D	Z
	1.008585	0.980915	0.994477
LIJ 2020/01–2020/03	Attitude angle (°)		
	Roll (ave/std)	Pitch (ave/std)	Yaw (ave/std)
	0.053/0.105	−0.018/0.148	−0.067/0.117
	Scale factor		
	H	D	Z
	0.965424	1.002734	0.958902
MAG 2021/01–2021/03	Attitude angle (°)		
	Roll (ave/std)	Pitch (ave/std)	Yaw (ave/std)
	−0.093/0.298	−0.147/0.681	−1.100/0.051
	Scale factor		
	H	D	Z
	0.992251	1.002943	1.001010

brevity, the corrections for attitude angle and scale factor are called daily corrections and their parameters are calculated by the genetic algorithm using 1-day data. The attitude angle, scale factor, long-term time drift, and relative temperature coefficient correction is called the long-term correction whose parameters are obtained by the calculation flow shown in Fig. 3. The daily correction differs from the long-term correction not only in the correction parameters but also in the demeaning operation of the data preprocessing. Because the geomagnetic vector observations are concerned with the variation of the geomagnetic field and the geomagnetic vector baseline is determined by other methods at the observatory, the mean value of each component needs to be subtracted from the raw vector data for the single-day correction; meanwhile, the long-term correction subtracts the mean value of the magnetic field corresponding to the length of the time series.

The daily variation comparison curve of YUL on June 17, 2020 shown in Fig. 4, is taken as an example of the parameter calculation and the correction effect of the daily correction. As shown in the red rectangle in Fig. 4a, the morphology of dH is clearly similar to that of the D component, while the correlation with the H component itself is not obvious. This is a typical feature of morphological disagreements due to a heading angle deviation, which agrees well with the calculated result: a larger heading angle deviation between the two instruments (the Yaw angle at YUL is −2.341°). After just the attitude angle correction, as shown in the red rectangle in

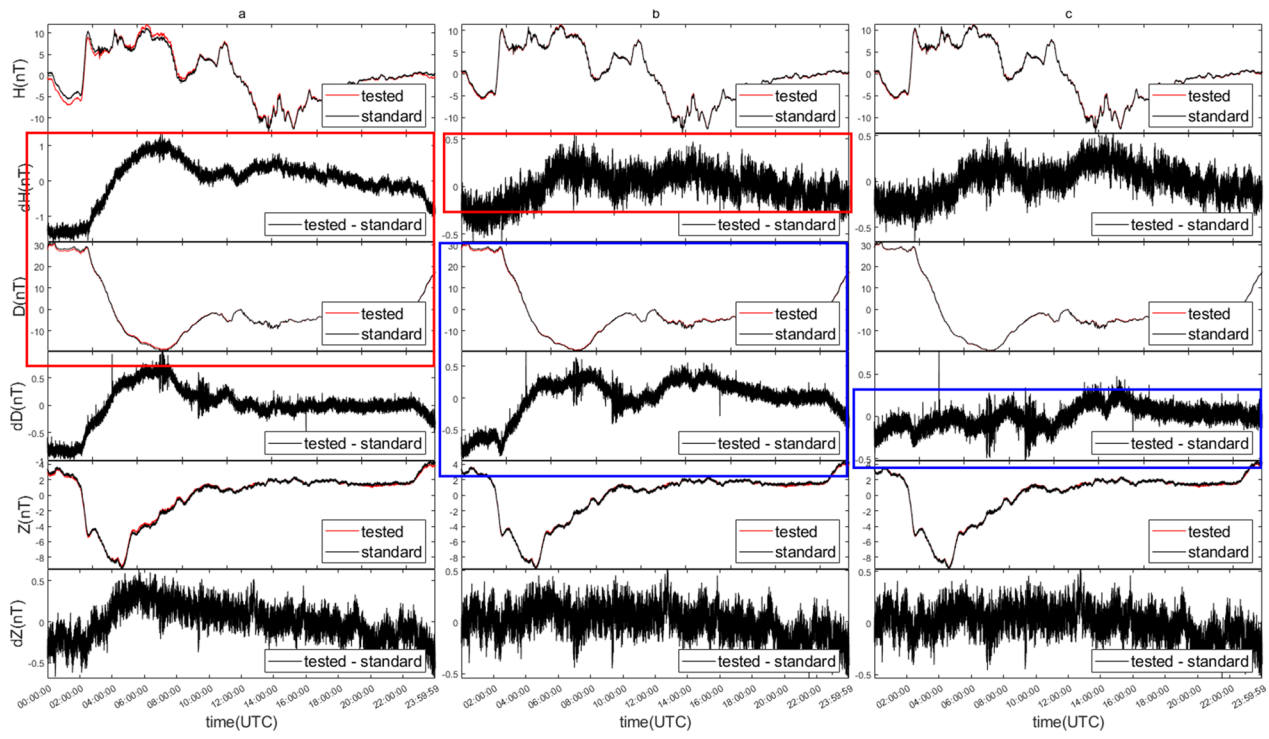


Fig. 4 Daily correction effect. The figure shows the daily geomagnetic variations and the vector differences on June 17, 2020, at YUL **a** before corrections (see explanations in Fig. 1) and **b** after attitude angle corrections only, **c** after both attitude angle and scale factor corrections

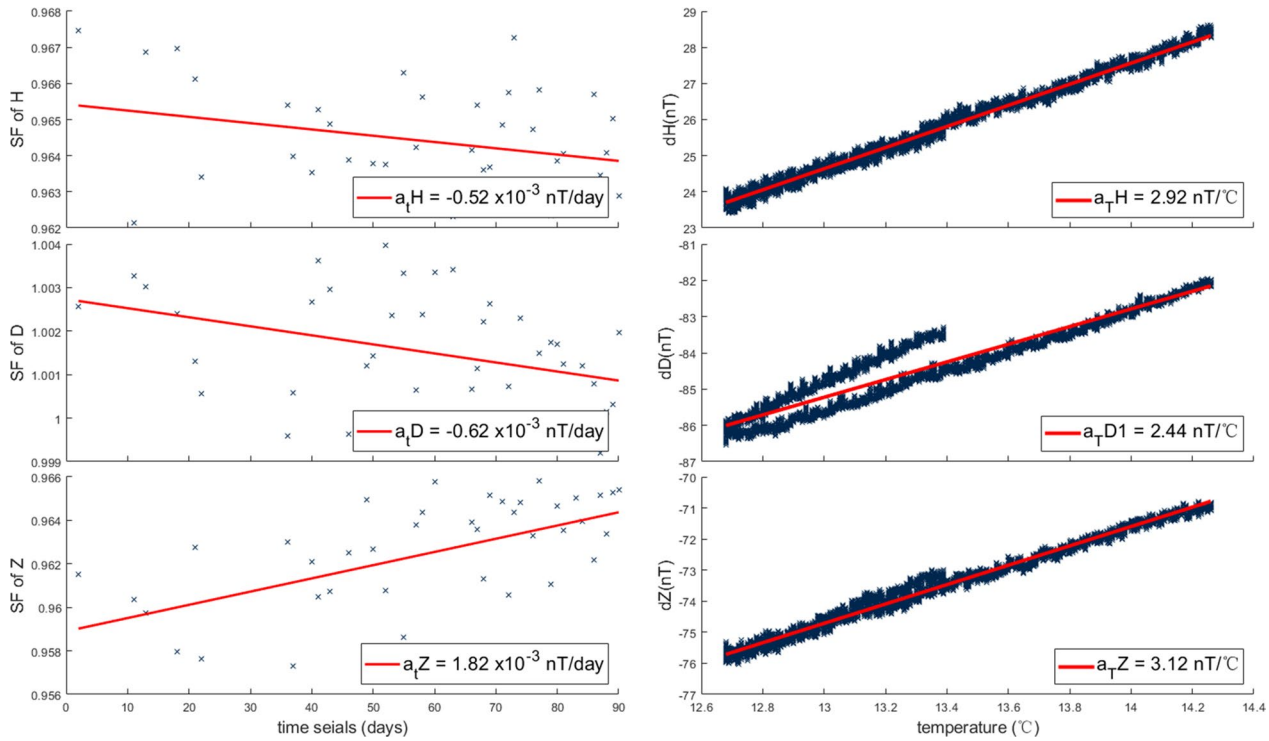


Fig. 5 Linear regression of the long-term time drift (left) and the relative temperature coefficient (right)

Fig. 4b, the morphological correlation between dH and D is weakened; however, the morphological correlation between dD and D is still strong which is highlighted in the blue rectangle in Fig. 4b. As shown in the blue rectangle in Fig. 4c, the correlation between dD and D is weakened after performing the attitude angle and scale factor corrections. The maximum value of the geomagnetic vector difference is reduced from approximately 1.3 nT with the original data to approximately 0.5 nT after the daily correction.

Taking LIJ with its complete temperature measurements and small environmental disturbance as an example, a linear regression analysis of the long-term time drift and relative temperature coefficient over 3 months is shown in Fig. 5. Note that the long-term time drift of the instrument is given in units of nT/day and is obtained by multiplying the slope of the scale factor by the amplitude of the daily variation (in the case of 30 nT). The residual magnetic field strength is not multiplied by the slope of the scale factor here, because when the instrument is misoriented or other factors cause the residual magnetic field strength of a component to be large, this will cause the calculated value of the long-term time drift to be too large. The effect of the long-term time drift varies from instrument to instrument and is generally not significant. However, the effect of temperature is more than was expected. Even though the D component appears to have two fitted straight lines, which are discussed below, the 3-month continuous temperature variation has a good linear relationship with the vector difference and a large slope of approximately 3 nT/°C, which is important for studies of the data agreement between long-term comparative observations.

The difference between the comparative observational data after the long-term correction is shown in Fig. 6. It can be seen that, after the daily correction (Fig. 6b), the daily variation of the geomagnetic vector difference is significantly reduced and its trend is consistent with that of the temperature. After temperature correction (Fig. 6c blue line), the dH and dZ variations are no longer temperature-dependent. After the long-term correction (Fig. 6c black line), there is no clear temporal pattern in the variation of dH and dZ, showing flatter straight trends. In addition, the black line overlaid on the blue line in Fig. 6c is slightly closer to the zero horizontal line due to the long-term time drift correction. Note that, in Fig. 6b, dD deviates significantly from the temperature change after approximately 41 days (the thick red line) but its trend is still the same as that of the temperature, which leads to an increase in dD in the second half of Fig. 6c, instead of continuing as a straight line. This is believed to reflect the case in which a temperature inflection point causes

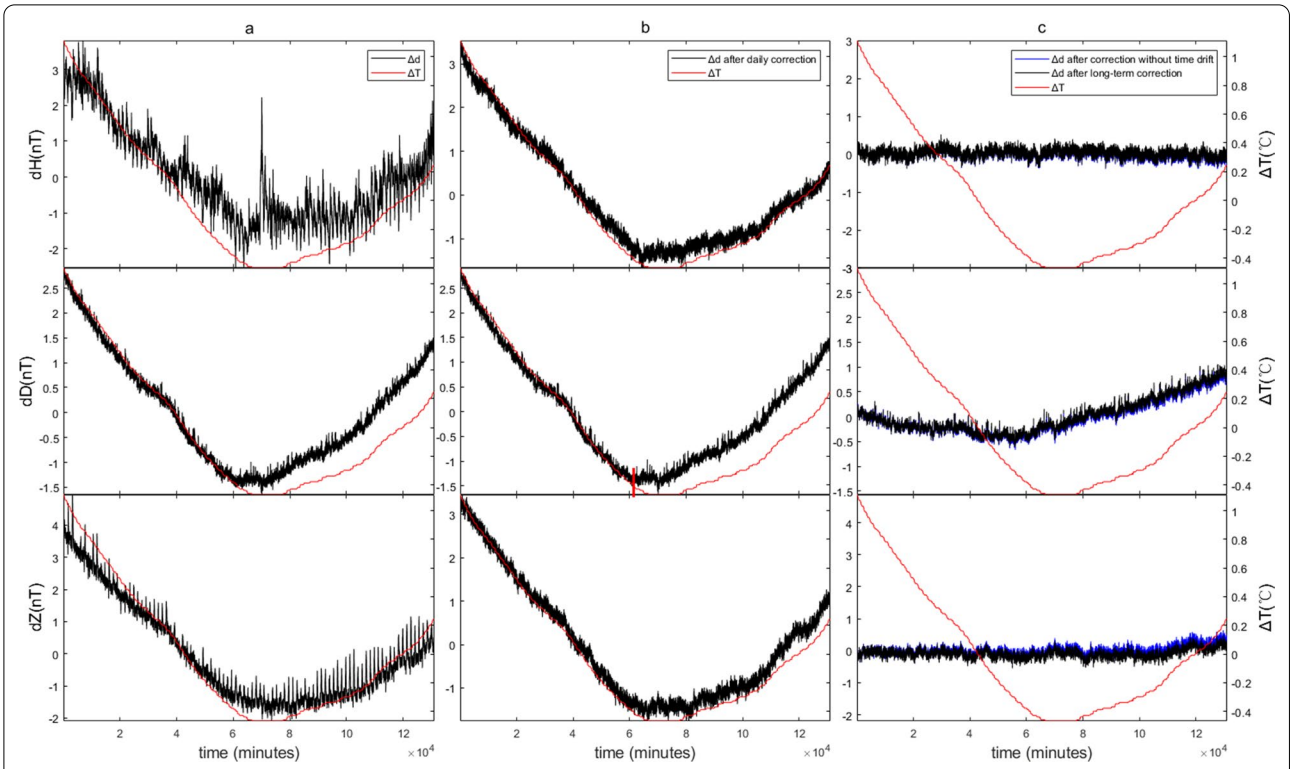


Fig. 6 Long-term-correction effect. **a** Geomagnetic vector difference during the period of January–March 2020 at the Lijiang (LIJ) observatory before and after the long-term correction. **b** The geomagnetic vector difference no longer has a significant daily period after the daily correction. **c** The difference curve tends to level off after the long-term time drift and relative temperature coefficient correction; compared with (**b**), it no longer has significant temperature and time characteristics. In **b**, after the thick red line, dD is separated from the temperature profile but their trends remain consistent, which is thought to indicate a change in the relative temperature coefficient around the temperature inflection point

a change in the trend, resulting in a change in the relative temperature coefficient of the probe pair on a certain component, as explained in detail in the discussion. From a quantitative point of view, the maximum value of the geomagnetic vector difference is reduced from approximately 3 nT to approximately 0.5 nT after the long-term correction.

B–A plots: a more appropriate method for data agreement evaluations of comparative geomagnetic vector observations

Traditionally, the data agreement between comparative geomagnetic vector observations is often expressed by the Pearson correlation coefficient. In the case, where the comparison is between two different observation platforms or two sets of instruments on the same platform that are not precisely calibrated, the correlation coefficient can appropriately distinguish the degree of agreement in the trend of the comparative observational data. However, the resolution of ground-based observations made by currently available fluxgate magnetometers has increased significantly. With the precise orientation adjustments made by observatory staff and the strict temperature control available in geomagnetic variation rooms, the correlation coefficients of the comparative observations are often so high that they cannot adequately describe data disagreements.

Take LIJ as an example. Its correlation coefficient prior to the daily correction is generally higher than 0.9995, and the difference between the correlation coefficients before and after the correction is on the order of 10^{-15}. Prior to the long-term correction, the correlation coefficients of all three components of the comparative observational data for 3 consecutive months were higher than 0.995 and the difference between the correlation coefficients before and after the correction is on the order of 10^{-11}. This indicates that correlation coefficients are not sufficient to reflect disagreements caused by the effects of the attitude angle, scale factor, long-term time drift, and temperature of the comparative observational data and that the correlation coefficient difference is not a good description of the effect of the data agreement

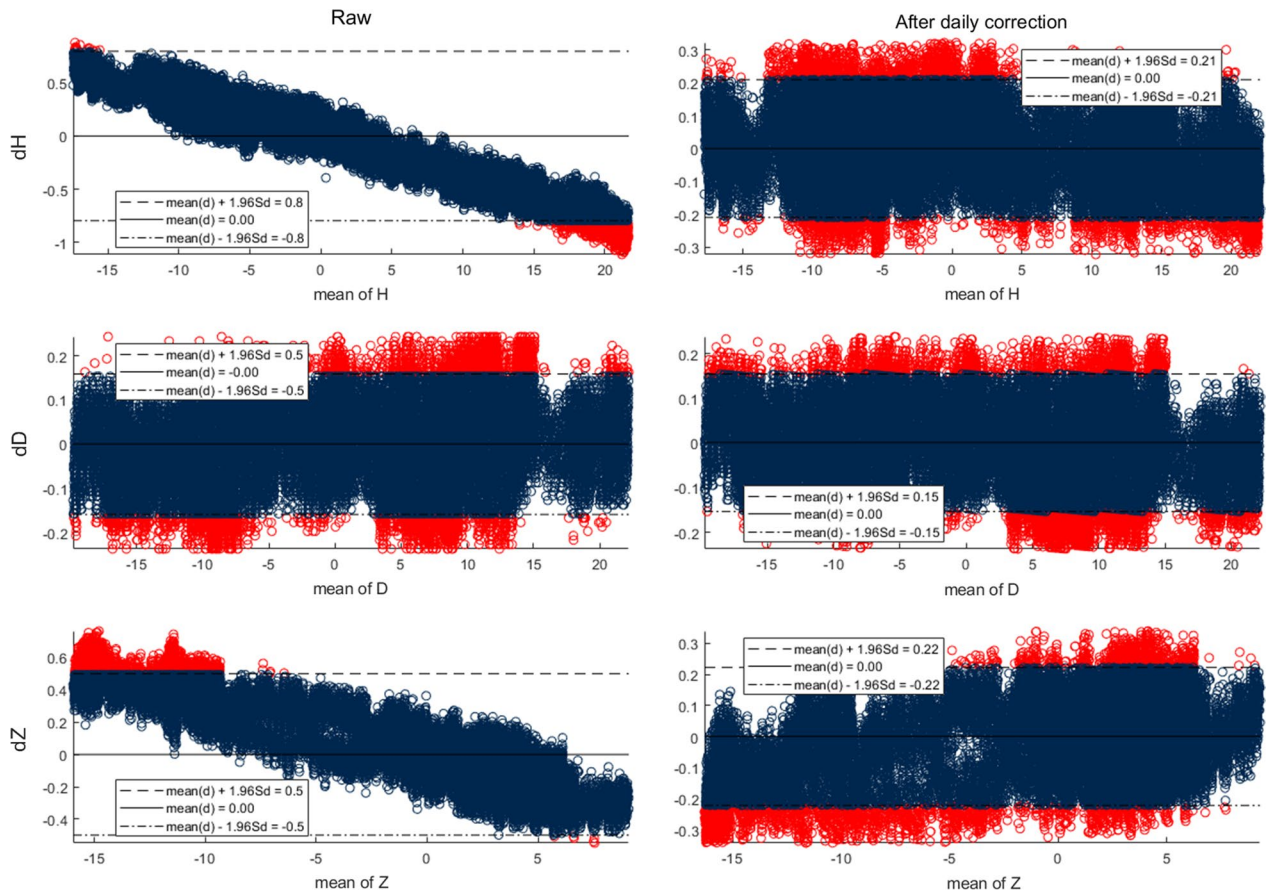


Fig. 7 Bland–Altman (B–A) plots of the comparative geomagnetic vector observations on March 30, 2020, at LIJ

corrections. It has been argued (Giavarina 2015) that correlation studies are inappropriate to assess the agreement between comparative observational data.

B–A plots are often used as a statistical method to analyze the agreement between two quantitative measurements. This method is very popular in the field of analytical chemistry and statistical medicine and has been popularized by J. Martin Bland (Bland and Altman 1986) and Douglas G. Altman (Altman and Bland 1983). The horizontal and vertical axes of a B–A plot represent the mean and difference of two sets of data, respectively. The mean of the difference $\pm$ 1.96 times the standard deviation of the difference is the 95% confidence interval. If the two sets of data are in good agreement, then there are sufficient points distributed within the 95% confidence interval and there is no significant pattern in the distribution of the points. The length of the confidence interval is also relatively small and allows a judgement with respect to whether the agreement between the two data sets meets the criteria according to the specific requirements of the comparative observations. This paper proposes a visual analysis of the agreement between comparative geomagnetic vector observations using B–A plots, which can identify, to some extent, the reasons for the disagreement between the comparative observational data and can visualize the effect of the parameter corrections. Furthermore, the numerical length of the 95% confidence interval in B–A plots enables a quantitative evaluation of the agreement between the comparative observational data, which can be combined with the actual needs of the comparative geomagnetic vector observations to determine the data availability and substitutability.

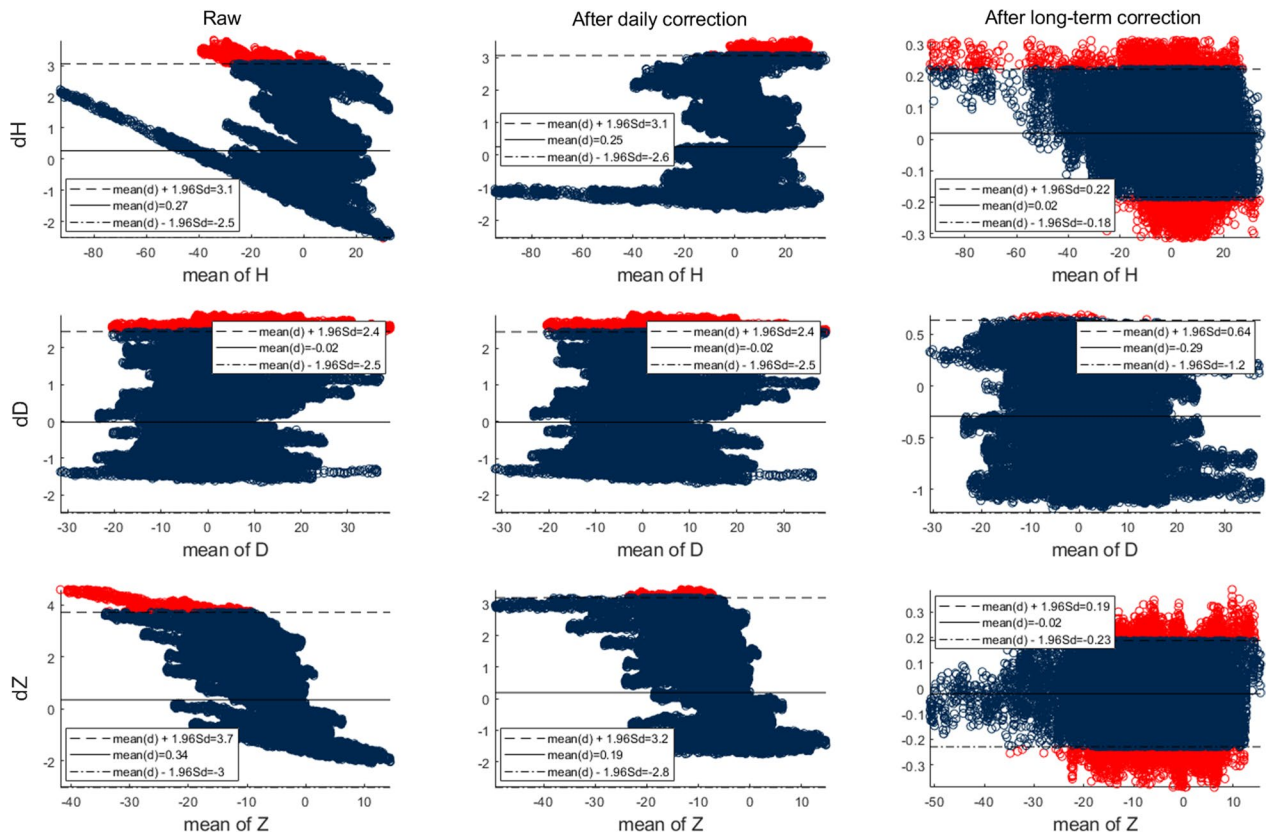


Fig. 8 B–A plots of the comparative geomagnetic vector observations during the period of January–March 2020 at LIJ

Taking LIJ as an example, the B–A plots of the single-day and long-term comparative observational data are shown in Figs. 7 and 8, respectively.

More than 95% of the points in the B–A plots before and after the daily correction of the single-day data are distributed within the 95% confidence interval. This indicates that the comparative observational data of the two instruments are very similar with or without the correction and that their general variation trends are the same. However, there is a significant skew in the distributions of the points in the B–A diagram of the H and Z components that disappears after the daily correction. This indicates that the geomagnetic vector difference in the original data is influenced by the attitude angle and the scale factor and that the agreement is significantly improved after correcting the corresponding parameters. Meanwhile, the values of the 95% confidence interval decreased significantly before and after the correction, for example, from [− 0.8, 0.8] to [− 0.21, 0.21] for the H component, and the interval length was reduced from 1.6 nT to 0.42 nT. The 95% confidence interval lengths for the D and Z components reached 0.3 nT and 0.44 nT after the correction, respectively. This means that, for the three components of the geomagnetic daily variation after the daily correction, the absolute value of the difference between the vast majority of the comparative observations is less than 0.22 nT and the fluctuation of the difference is less than 0.44 nT.

The B–A plots of the long-term comparative observations are similar. The raw data, the daily corrected data, and the long-term-corrected data all satisfy the condition in which more than 95% of the points are distributed within the confidence interval. However, the particular characteristics of the point distributions weakened sequentially. Again, the absolute value of the 95%

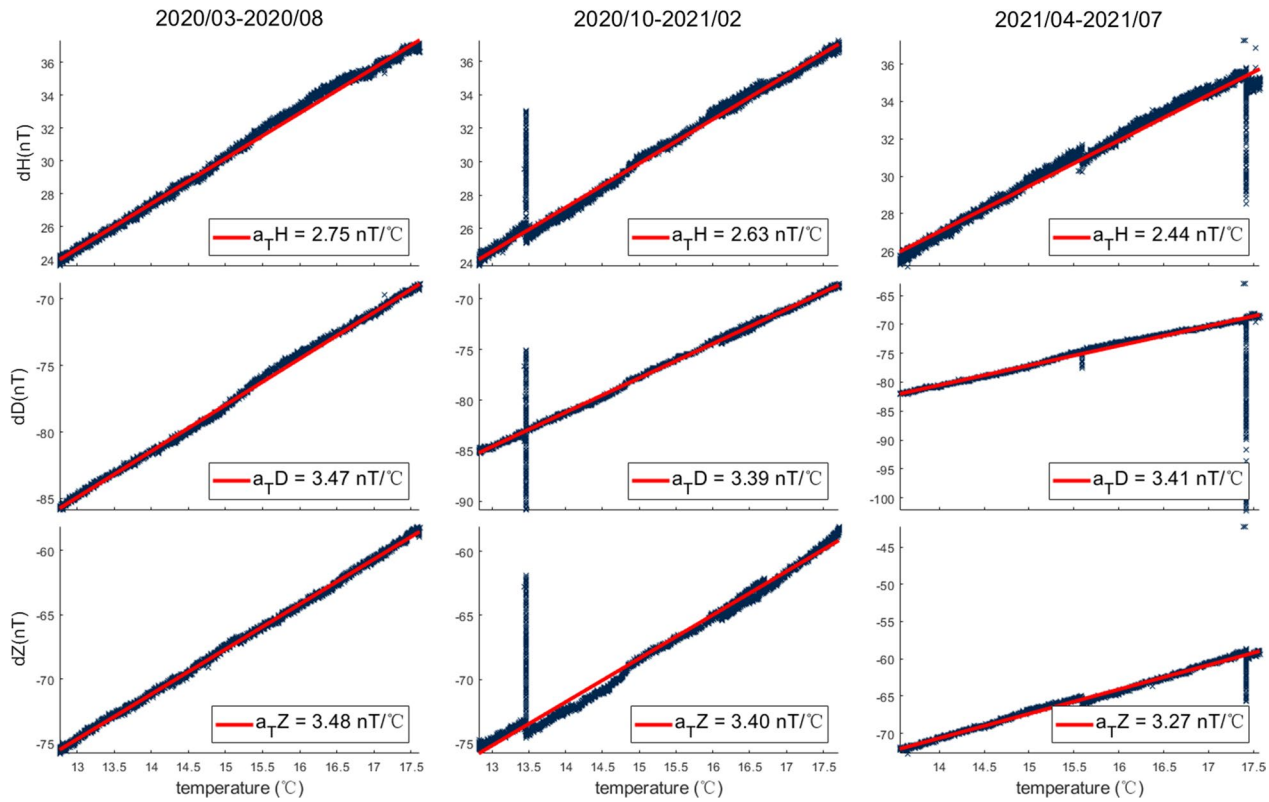


Fig. 9 Linear regression of the relative temperature coefficients before and after the temperature inflection points

confidence interval decreased significantly after the long-term correction. The interval lengths of the H, D, and Z components decreased from 5.6 nT, 4.9 nT, and 6.7 nT to 0.4 nT, 1.84 nT, and 0.42 nT, respectively. The larger length of the 95% confidence interval after the long-term correction of the D component is related to the temperature inflection point, which affects the relative temperature coefficient as described above.

Discussion

(1) In Fig. 6b, the D component difference, dD (black line), does not match the temperature (red line) in the second half of the time period well. Even though both trends remain the same, the variation curve of the geomagnetic vector difference deviates from that of the temperature. This deviation causes the long-term-corrected dD in Fig. 6c to climb in the second half of the panel. The change in the relative temperature coefficient can be visually described by a linear regression plot of the geomagnetic vector difference versus the temperature. As shown in Fig. 5, there are two well-fitted regression lines in the scatter plot of dD versus the temperature. Combined with the actual data, it can be determined that the relative temperature coefficient of the probe pair of the D component changes after the arrival of the temperature inflection point. Consequently, we believe that the calculation and correction of the relative temperature coefficient can be roughly determined by the timing of the temperature inflection point.

No geomagnetic vector comparison observations were performed at LIJ prior to January 2020, and therefore, we divided the long-term comparative observational data (from March 1, 2020, to July 31, 2021) into three time periods based on the approximate temperature inflection points that occurred in September 2020 and March 2021; the time periods are from March 2020 to August 2020 (warming), from October 2020 to February 2021 (cooling), and from April 2021 to July 2021(warming). The linear regressions of the relative temperature coefficients over the three time periods and the long-term-corrected B–A plots are shown in Figs. 9 and 10, respectively.

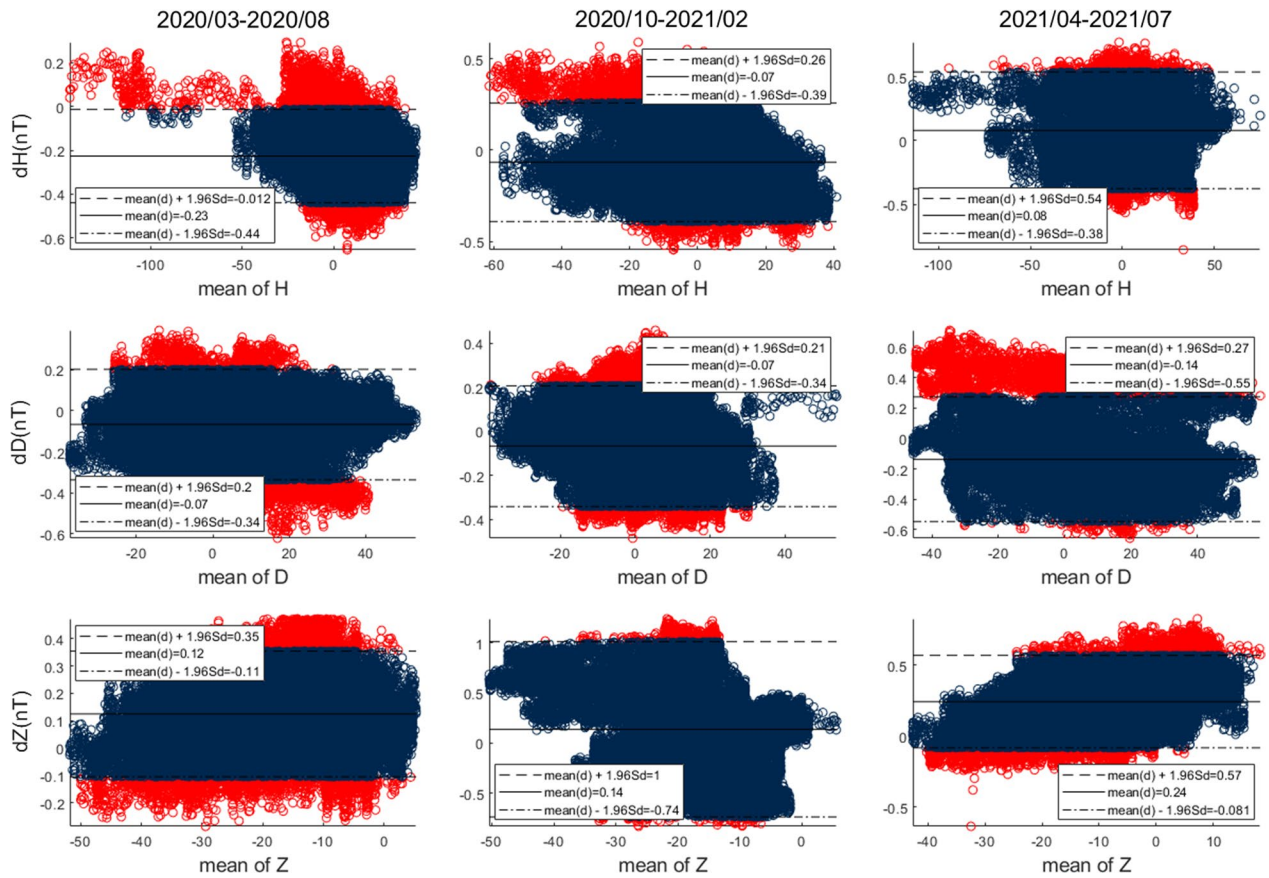


Fig. 10 B–A plots of three long-term observations

Figure 9 indicates that the relative temperature coefficients of the probe pairs for the different components may change before and after the temperature inflection points. The relative temperature coefficients for the different time periods can be obtained from a linear regression analysis of the geomagnetic vector difference and the temperature. Figure 10 shows that the agreement between the comparative observational data after the long-term correction is good in all three time periods, with no significant influence of the attitude angle or the scale factor. Moreover, the length of the 95% confidence interval is significantly reduced compared with that prior to the correction.

(2) The effects of the four parameters on the agreement between the comparative geomagnetic vector observations show different characteristics in the B–A plots. As shown in the top section of Fig. 11, the attitude angle parameter causes the distribution of the points to take the form of a "connected domain" in a B–A diagram, that is, the white block surrounded by the data points. The scale factor, as shown in the bottom section of Fig. 11, causes the points to be distributed diagonally. As shown in the top section of Fig. 12, the long-term time drift parameter causes the mean value of the geomagnetic vector difference in the B–A diagram to deviate somewhat from the zero point. The relative temperature coefficient parameter, as shown in the bottom section of Fig. 12, results in a horizontal streak-like distribution of points and an overall larger deviation to one side of the mean value of the difference. In general, before the correction, the measured values of the two fluxgate magnetometers always include the contributions of these influencing factors, making the difference between the measured values of the comparative geomagnetic vector observations related to the mean of the measured values, resulting in points on the B–A plot that do not conform to a normal distribution. This is key to the B–A plot clearly and qualitatively describing the agreement between the comparative geomagnetic vector observation data. Furthermore, additional comparative observational experiments

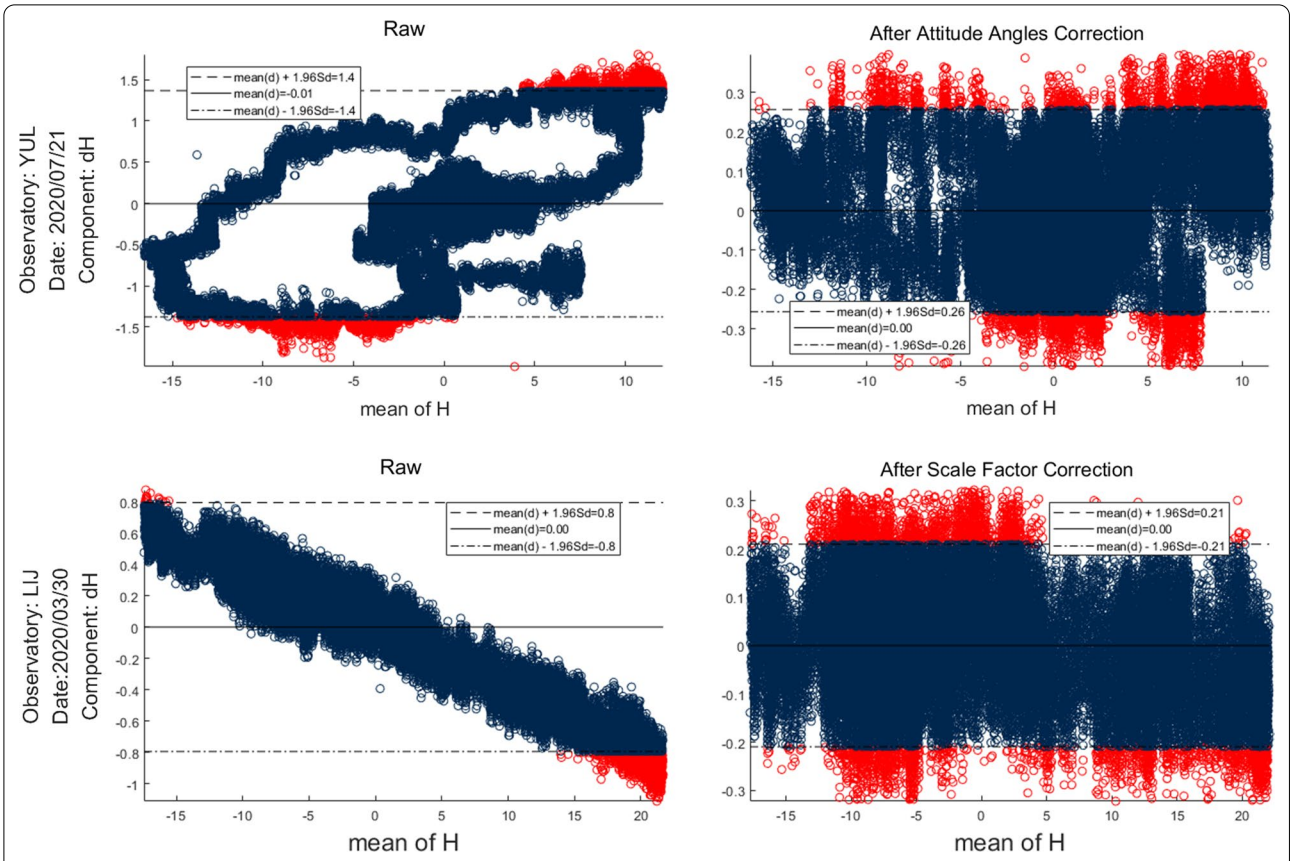


Fig. 11 Characteristics of the attitude angle and the scale factor in a B–A plot

are needed to explain, in detail, the reasons for the unique distributions in the B–A plots caused by these correction parameters or influencing factors.

(3) The correction parameters for the agreement between the comparative geomagnetic vector observations are not constant, especially the attitude angle and the relative temperature coefficient. The relative temperature coefficient was discussed previously. Even though most observatory instruments are set up on bedrock or marble piers, the long-term accumulation of slow changes or seismic activity can cause slight changes in the attitude angle. Therefore, it is necessary to determine whether the points distributions of the B–A plots of daily and long-term observations show an unusual pattern and the correction parameters need to be recalculated every 3 or 4 months.

Summary

In the variation room of a geomagnetic observatory with a good environment, the attitude angle and the scale factor are the main influencing factors that affect the agreement between the single-day comparative observational data. For long-term observations, the long-term time drift and the temperature also need to be included. In this paper, we analyzed the characteristics of these influencing factors using geomagnetic vector variation plots and calculated the corresponding correction parameters based on a genetic algorithm and linear regression analysis. The effect before and after the attitude angle and scale factor corrections was analyzed for YUL, which has a large heading angle deviation. Meanwhile, the effect before and after the long-term time drift and relative temperature factor corrections was analyzed for LIJ, which has complete temperature

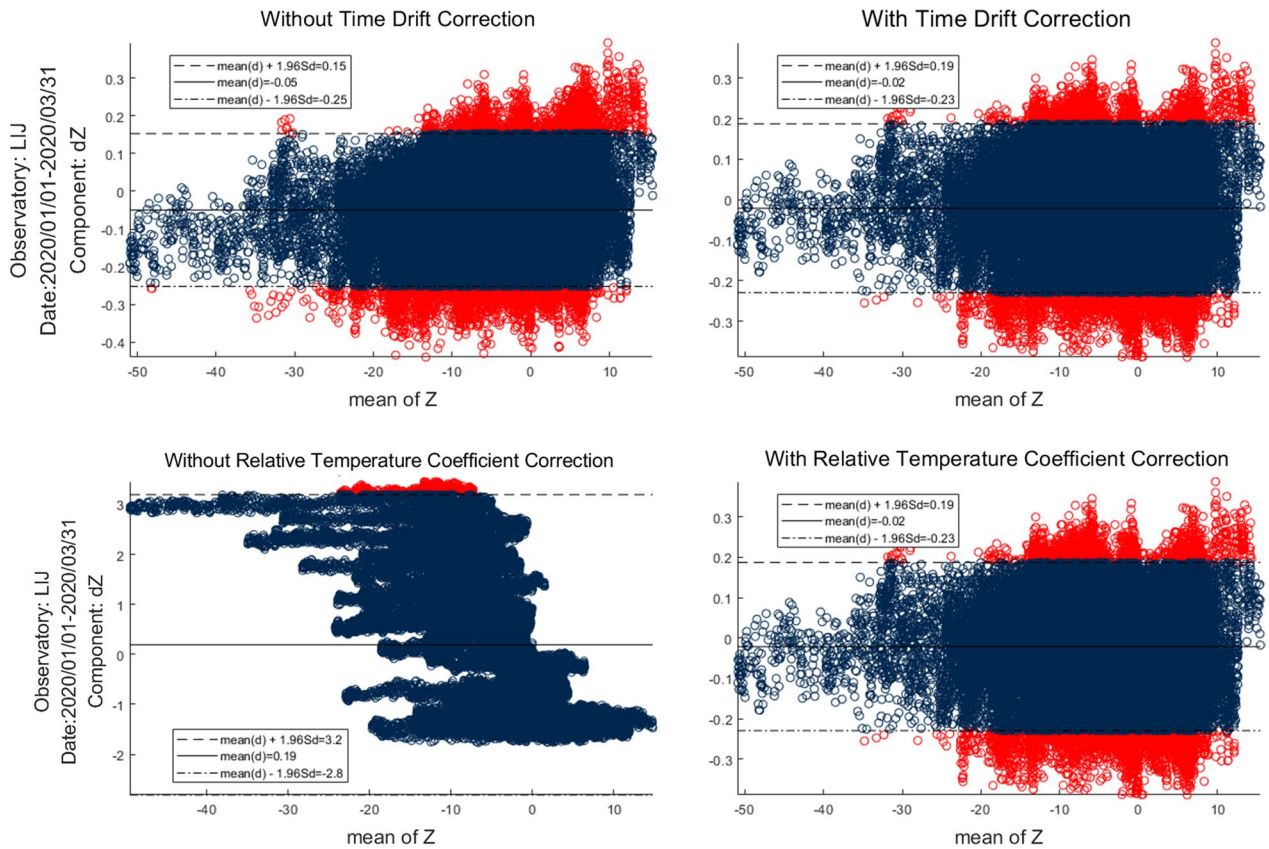


Fig. 12 Characteristics of the long-term time drift and the relative temperature coefficients in a B–A plot

records. A B–A plot enables a qualitative and quantitative evaluation of the data agreement between comparative geomagnetic vector observations. For comparative observational data with good agreement, that usually is, with the corrected data, more than 95% of the points in the B–A plot are distributed within the 95% confidence interval and without significant patterns. Meanwhile, the length of 95% confidence interval decreased significantly after the correction. We found that the relative temperature coefficient changes before and after temperature curve inflection points. Furthermore, we discussed the special distribution characteristics of most of the points in a B–A plot under each influencing factor.

Abbreviations

LIJ: Lijiang Geomagnetic Observatory; MAG: Maguan Geomagnetic Observatory; YUL: Yunlong Geomagnetic Observatory.

Acknowledgements

We thank the Geomagnetic Network of China for the availability of geomagnetic data. We also thank the Lijiang Geomagnetic Observatory for helping to carry out attitude angle experiments. We also thank the handling editor and the two anonymous reviewers for their suggestions and comments, which greatly improved the manuscript.

Authors' contributions

XXH developed a miniaturized fluxgate magnetometer for the large attitude angle comparative observation experiment conducted at LIJ. YTT determined the research direction of the paper and participated in the discussion of the B–A plots to evaluate the data agreement between the comparative observations. XXZ helped analyze the different influencing factors in the long-term comparative observational data. XYS helped analyze the calculation method for the correction parameters. ZBH performed the comparative observation experiment at LIJ, analyzed the different influencing factors and corresponding correction parameters in the geomagnetic variation plots and B–A plots, respectively, and was a major contributor to the writing of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by National Key Research and Development Program of China (2018YFC1503803), Earthquake Science Spark Program of China Earthquake Administration (XH20021), and the Special Fund of the Institute of Geophysics, China Earthquake Administration (DQJB19B22).

Availability of data and materials

The data sets analyzed during the current study are available from the corresponding author on reasonable request. The generic algorithm program and some example data are available in [HeZhaoBo/GAs] repository on GitHub, [http://github.com/HeZhaoBo/GAs]. The geomagnetic data are provided by Geomagnetic Network of China.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details
[1]Institute of Geophysics, China Earthquake Administration, Beijing 100081, China. [2]Earthquake Administration of Jiangsu Province, Nanjing 210014, China.

Received: 16 October 2021 Accepted: 20 January 2022
Published online: 11 February 2022

References

Altman DG, Bland JM (1983) Measurement in medicine: the analysis of method comparison studies. Statistician 32(3):307–317. https://doi.org/10.2307/2987937
Araki T, Takeuchi T, Araki Y (2004) Rise time of geomagnetic sudden commencements —Statistical analysis of ground geomagnetic data—. Earth Planets Space 56:289–293. https://doi.org/10.1186/BF03353411
Berezin IA, Tlatov AG (2020) Comparative analysis of terrestrial and satellite observations of photospheric magnetic field in an appendix to simulation of parameters of coronal holes and solar wind. Geomagn Aeron 60:872–875. https://doi.org/10.1134/S0016793220070051
Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327(8476):307–310. https://doi.org/10.1016/S0140-6736(86)90837-8
Curto JJ, Cardús JO, Alberca LF, Blanch E (2007) Milestones of the IAGA international service of rapid magnetic variations and its contribution to geomagnetic field knowledge. Earth Planets Space 59:463–471. https://doi.org/10.1186/BF03352708
Esper A, Dufay B, Saez S, Dolabdjian C (2020) Theoretical and experimental investigation of temperature-compensated off-diagonal GMI magnetometer and its long-term Stability. IEEE Sens J 20(16):9046–9055. https://doi.org/10.1109/JSEN.2020.2988939
Giavarina D (2015) Understanding bland altman analysis. Biochemia Medica 25(2):141–151. https://doi.org/10.11613/BM.2015.015
Gordon D, Brown R (1972) Recent advances in fluxgate magnetometry. IEEE Trans Magn 8(1):76–82. https://doi.org/10.1109/TMAG.1972.1067268
Han DS, Iyemori T, Nose M, McCreadie H, Gao Y, Yang F, Yamashita S, Stauning P (2004) A comparative analysis of low-latitude Pi2 pulsations observed by Ørsted and ground stations. J Geophys Res 109:A10209. https://doi.org/10.1029/2004JA010576
Holland JH (1992) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT press, Cambridge
Jankowski J, Sucksdorff C (1996) IAGA Guide for Magnetic Measurements and Observatory Practice. IAGA, WARSAW
Jiao BG, Gu W, Zhang SY (2011) Error calibration of three-component fluxgate sensor's nonorthogonality. Modern Electronics Technique 34(13):123–126. https://doi.org/10.3969/j.issn.1004-373X.2011.13.036
Korepanov V, Marusenkov A (2012) Fluxgate magnetometers design peculiarities. Surv Geophys 33:1059–1079. https://doi.org/10.1007/s10712-012-9197-8
Liu C, Teng YT, Wang XM, Wu Q, Zhang XT (2019) Consistency correction for the observed data of fluxgate magnetometer based on genetic algorithms. Progr Geophys (in Chinese) 34(2):0751–0756. https://doi.org/10.6038/pg2019CC0024
Primdahl P (1979) The fluxgate magnetometer. J Phys e: Sci Instrum 12:241. https://doi.org/10.1088/0022-3735/12/4/001
Sastry K, Goldberg D, Kendall G (2005) Genetic Algorithms. In: Burke E K, Kendall G (eds) Search Methodologies. Springer, Boston, MA. https://doi.org/10.1007/0-387-28356-0_4
Segarra A, Curto JJ (2013) Automatic detection of sudden commencements using neural networks. Earth Planets Space 65:791–797. https://doi.org/10.5047/eps.2012.12.011
Shen X, Teng Y, Hu X (2021) Design of a low-cost small-size fluxgate sensor. Sensors 21(19):6598. https://doi.org/10.3390/s21196598
Toh H, Hamano Y, Ichiki M (2006) Long-term seafloor geomagnetic station in the northwest Pacific: A possible candidate for a seafloor geomagnetic observatory. Earth Planets Space 58:697–705. https://doi.org/10.1186/BF03351970
Wang XM, Teng YT, Tan J, Wu Q, Wang Z (2017) Influencing factors of tri-axial fluxgate sensor horizontal and orientation on the accuracy of geomagnetic variation records. Acta Seismol Sin 39(3):429–435
Weile DS, Michielssen E (1997) Genetic algorithm optimization applied to electromagnetics: a review. IEEE Trans Antennas Propag 45(3):343–353. https://doi.org/10.1109/8.558650
Xu D, Chen H, Gao M (2015) Observed geomagnetic induction effect on Dst-related magnetic observations under different disturbance intensities of the magnetospheric ring current. Earth Planets Space 67:15. https://doi.org/10.1186/s40623-015-0189-z

Publisher's Note