General comments

The paper evaluates accuracy performance of a commercial anisotropic magnetoresistive magnetometer. This kind of the instrument are going to be used in two low orbit CubeSat satellites intended for scientific observations of the radio emission of the Earth's aurora. These aspects discussed in the article are within the scope of GI.

The paper attempts to expand existing approaches to calibrating such magnetometer parameters as zero offsets, scale factors, angular (orientation) errors, and to estimate the linear temperature coefficients of these parameters. However, the conclusions made in the article are not fully supported by the presented results. The authors do not give arguments why the chosen method of calibration and data processing was used. From the calibration results, the authors conclude that “lack of characteristic calibration data can cause overfitting due to degeneracy of the fit to the available data”. However, it remains unclear why additional experiments were not carried out to complete the data set to the desired level. The proposed calibration method applied for processing ground-based data is planned to be used in the future to calibrate the magnetometer in orbit, but the paper does not provide arguments to support such an intention. In particular, the proposed method is not an attitude-independent calibration, since it relies on the knowledge of three components of the calibration signal. However, it is not clear how real magnetic field vectors derived from the global magnetic models will be defined in orthogonal, spacecraft-fixed coordinates.

The authors present a complete and accurate description of the experiments and calculations performed. Additional materials contain code and data for reproducing calculations.

The authors based their research on the analysis of a number of related studies, which is reflected in the corresponding bibliography. However, a deeper analysis of vector magnetometer calibration methods would be useful. The title clearly reflects the contents of the paper. The abstract provides a concise and complete summary. The overall presentation is well structured, but some aspects have to be clarified. The language is fluent and precise. In general, mathematical formulas, symbols, and abbreviations are defined and used correctly. The units of some parameters require clarification. Some subsections of the manuscript need clarification, more details are given below in the "Specific Comments" section. The number and quality of references are appropriate, but more references on the calibration of vector magnetometers can be recommended. The amount and quality of supplementary material is appropriate and very helpful.

Specific comments

p. 2, Subsection 1.3

The concept of magnetometer calibration, the conditions of its operation on the satellites AERO and VISTA, and the requirements for measurement accuracy are not entirely clear. For example, from paragraph 1.3 a reader may conclude that a magnetometer calibrated on the ground should provide a measurement accuracy of 100 nT. On the other hand, from the information in sections 3.2, 3.3 and 4 (*“In orbit, AERO and VISTA will gather calibration data at low latitudes using a global magnetic map as a reference source. The regression parameters will be used to achieve the desired accuracy in the science gathering region near Earth's aurora.”*) it is unequivocally stated that the instrument is supposed to be calibrated in orbit.

First, under what conditions (temperature range and field measurement range) is it necessary to ensure a measurement accuracy of 100 nT? The value of 100 nT itself refers to the modulus of the field or to each component, 100 nT - is this the maximum or root mean square value?

Secondly, if in-orbit calibration by the proposed method is supposed, how will this be done? The method compares the readings of the components of the tested and the reference magnetometer, and not the field module, as is done, for example, in the attitude-independent calibration method in the paper by Archer et al. In addition to the position in orbit and, accordingly, the values of the field components according to IGRF or other global magnetic model, it is necessary to know the orientation of the spacecraft in order to recalculate the magnetic field components in the satellite coordinate system. After all, there will not be a reference magnetometer there, right?

p. 3, Subsection 1.4
"The work by Archer et al. fits calibration coefficients for gain, offset, and angular position using on-orbit magnetometer data and the IGRF as a reference."

In fact, Archer et al. estimated also temperature coefficients of gain and offset, at least for the outboard magnetometer.

p. 3, Subsection 2.1
"A 3D printed mechanical mount for the EDU constrains the DUT in space at about 1 cm distance to the reference magnetometer (a Meda FVM400)."

Is there no mutual interference of the sensors with such a small distance between them?

p. 4, Subsection 2.3
"Cross-axis coupling"

The cross-axis effect in magnetic sensors is "a change in sensitivity based on the applied in the transverse, or cross-axis, direction". (https://aerospace.honeywell.com/content/dam/aerobt/en/documents/learn/products/sensors/application-notes/AN205_Magnetic_Sensor_Cross-Axis_Effect.pdf).
The response of the sensor is non-linear in respect to the applied transverse signal. The calibration model (Eq. 1) does not take into account this non-linear effect, in my opinion.
The parameter "cross-axis coupling" may be misunderstood by a reader as the cross-axis effect. It would be useful to clarify the meaning of this parameter.

p. 8, Subsection 2.3.4
"Given that both magnetometers reported similar hysteresis effects, the source of the hysteresis is likely magnetization of material near both magnetometers and not an effect inherent to either magnetometer alone."

This is an important issue, if there is an object near magnetometers, which disturbs the calibration field. Would not such an object introduce distortions into the results of other tests?

p. 10, Section 3
"This work extends the calibration equation reported in work by Archer et al. (2015) (Eq. 1) by including parameters for linear drift of all gain and offset parameters with temperature."

The calibration equation presented in Archer et al. (2015) differs significantly from the one proposed in this paper. First, Archer et al. used the reference signal as an independent variables (or predictors) and the magnetometer data as a dependent variables (or response variables). The present paper uses the instrument magnetic and temperature data as predictors and the reference magnetometer data as response variables. In some cases, errors in the independent variables (predictors) can lead to biased parameter estimates.
Second, Archer et al. evaluated the set of calibration parameters that minimizes the square difference of the field magnitude. This makes their calibration method attitude-independent – the

knowledge of the satellite orientation is not necessary. In the present work, three subsets of calibration parameters are evaluated, each minimizing the difference of one of the three magnetic field components.
And finally, Archer et al. also included in their model temperature dependence of the magnetometer gains and zero offsets: "Therefore, we subsequently applied a temperature-dependent calibration to the science mode data to account for the large temperature drift during this interval. This was achieved by modifying the attitude-independent procedure, requiring a linear relationship of the offsets and gains with the temperature measured by the thermistor at each time, e.g. $O_x(t) = c_x T(t) + d_x$, where $O_x(t)$ is now a time-varying magnetometer offset, $T(t)$ is the temperature measured by the thermistor and c_x and d_x are the constants estimated through the iterative calibration procedure."

p. 11, Subsection 3.1

"The K_{Syy} term is also anomalously large at -0.052 as compared to less than magnitude 0.01 for all other sensitivity terms. ... This shows a that lack of characteristic calibration data can cause overfitting due to degeneracy of the fit to the available data."

Does it mean that the estimated value (-0.052) of the K_{Syy} term is wrong? If yes, why the data set was not expanded till the level sufficient to properly estimate all calibration parameters? For example, temperature experiments can be done so that the Y probe and then the Z probe are parallel to the axis of the calibration coils. If the proposed calibration method is so sensitive to the nature of the data, it is necessary to perform some preliminary operations in order to assess the applicability of the method. By the way, Archer et al. estimated how uniform the data covers of the attitude sphere in order to reliably extract calibration parameters.
If a magnetometer actually has a scale factor temperature coefficient of 5%/°C, then it is unlikely that such an instrument can measure a 50,000 nT magnetic field with an accuracy of 100 nT (0.2% of full scale in another formulation).

p. 12. Subsection 3.1

"The offset values themselves and their own temperature dependence are similar to the linear fit values from Section 2.3."

It is not clear. First of all, the offset estimations in Tables 1 and 7 have opposite signs. Mentioning that adding the offset values from Table 7 to B_{meas} should fix the offset values shown in Table 1 would be helpful. Secondly, reducing $\overline{O}$ and $\overline{K}_O$ values from Table 7 to a temperature of 25 °C we obtain:
Ox=-1.21+25·0.036=-0.346 µT
Oy=-0.0071+25·(-0.0009)=-0.0935 µT
Oz=4.323+25·(-0.1607)=0.3055 µT
On the other hand, the offset of the Z component from Table 1 is -0.467 µT, so -0.467 + 0.3055 = -0.1615 µT, which exceeds the required limit of 100 nT.
It seems that the value of the temperature coefficient K_{Oz} is also anomalously high. Is this a reliable estimate or a side effect of the insufficient dataset used for calibration?

p. 13, Section 4

"This experiment has simultaneously validated the magnetometer design and calibration method for use on the AERO-VISTA mission"

In my opinion, the results presented in the manuscript are not sufficient to draw the above conclusions.

The magnetometer calibration parameters had not been validated by applying them to other datasets.
There are doubts about the accuracy of estimating the calibration parameters K_{Syy}, O_z, and K_{Oz}. The applicability of the calibration method to processing on-orbit data is not clear without understanding how the reference magnetic field values derived from global magnetic models will be transformed to the satellite coordinate frame.

Technical corrections

p. 1, Abstract
"...multivariate non-linear regression using a 27 parameter measurement equation"
p. 10, Section 3
"...9-element model..."

I counted only 24 calibration parameters estimated by the method of multivariate non-linear regression. They are 9 elements of the matrix **S**, 9 elements of the matrix $\mathbf{K_S}$, 3 elements of the vector $\overline{O}$ and 3 elements of the vector $\overline{K}_O$. 9+9+3+3=24. A subset of 8 calibration parameters is evaluated at each of the three executions of the MATLAB fitnlm function.

p. 7, Subsection 2.3.3
"The measured fields over temperature are reported in Figure 3. The linear fit to the X-axis data pictured derives a linear temperature coefficient of 4.37 nT per degree C."
In my perception of Figure 3, X component was drifting at -4600 nT over a temperature range of 34 Celsius, so the linear temperature coefficient is -135 nT/°C.

p. 8, Figure 3
The temperature varies in the range 33 – 73 K (Kelvin) , whereas in the subsection 2.3.3 we read "the DUT was heated to about 65 Celsius and allowed to cool to steady state—approximately a 30 degree Celsius temperature range...".

p. 11, Table 7
"Derived regression coefficients. Units of °C and μT"

In my opinion, if units of B_{act} and B_{meas} is μT, then **S** has to be dimensionless, $\mathbf{K_S}$ – 1/°C, O – μT, K_O – μT/°C, RMSE – μT. Thus, only terms O and RMSE have unit of μT, units of all other coefficients are neither °C, nor μT. If units of B_{act} is μT and B_{meas} is dimensionless, then **S** has to be μT, $\mathbf{K_S}$ – μT/°C, O – μT, K_O – μT/°C, RMSE – μT. Thus, none of the coefficients has a unit of °C.