On-orbit Calibration of Magnetometer Using Stochastic Gradient Descent

Joshua Abbey, Samuel Boland
UNSW Canberra Space
Building 16-208, Northcott Drive, Campbell ACT 2612; (02) 6268 8937
j.abbey@unsw.edu.au

ABSTRACT

Magnetometers are a key component of small satellite attitude determination and control systems (ADCS). They are typically calibrated on the ground during the spacecraft Assembly, Integration, and Testing (AIT) program.

UNSW Canberra Space have observed magnetometer calibration changes during the AIT program, with the most likely cause being spacecraft component magnetisation. Vibration testing has been noted to result in significant changes to magneto-inductive magnetic sensor calibration biases, and to a lesser extent, scale gains.

Ground-based calibration can be time consuming, manpower intensive and susceptible to human error. For larger spacecraft production runs, it is desirable to reduce or eliminate the time required to conduct calibration.

This paper outlines the use of stochastic gradient descent as a way of calibrating magnetometers on-orbit.

The method was tested using the Buccaneer Risk Mitigation Mission satellite developed and operated jointly by Australia's Defence Science and Technology Group and University of New South Wales Canberra Space. The newly tuned calibration parameters were successfully tested on-orbit. The results are compared with the values generated during ground-based calibration.

The method reduced error by approximately 20% during a six-orbit test period. The updated parameters result in an angular change in the indicated magnetic direction of up to 7.2 deg.

INTRODUCTION

Magnetometers are a key component of small satellite attitude determination and control systems (ADCS). They are typically calibrated on the ground during the spacecraft Assembly, Integration, and Testing (AIT) program[1].

UNSW Canberra Space have observed magnetometer calibration changes during the AIT program, with the most likely cause being spacecraft component magnetisation. Vibration testing has been noted to result in significant changes to magneto-inductive magnetic sensor calibration biases, and to a lesser extent, scale gains. Both the mechanical vibration and the high magnetic field from the vibration table are expected to induce changes in the calibration offsets.

Ground-based calibration can be time consuming, manpower intensive and susceptible to human error. For larger spacecraft production runs, it is desirable to reduce or eliminate the time required to conduct calibration.

On-orbit calibration of magnetometers is of interest to UNSW Canberra Space (UCS) to reduce AIT workload and to avoid issues that occur due to changes in magnetometer calibration. Typically, on-orbit calibration is conducted using non-linear least-squares or other numerical error minimization methods[2].

There is an increasing interest in the application of machine learning methods to spacecraft attitude determination and control solutions[3,4]. The application of simple machine learning methods offers a starting point for the development of more complex applications of machine learning to spacecraft attitude determination and control.

Additionally, for spacecraft with deployable hardware, it can be infeasible to calibrate magnetometers in flight configuration; requiring on-orbit calibration methods. Primary magnetometers are commonly mounted on a deployable boom arm. Launch Services Provider (LSP) do not always provide authorisation to deploy and restow these boom arm after acceptance vibration testing.

This paper outlines UCS application of stochastic gradient descent to an in-orbit magnetometer to improve magnetometer calibration, and UCS intention to develop these methods for future missions.

ON-ORBIT CALIBRATION METHODS

Numerical methods are commonly used for the calibration of magnetometers and other attitude determination sensors on spacecraft in-orbit. Such methods include the TWOSTEP algorithm[5], the differential value approach[6] and Kalman Filters[7].

Gradient Descent (GD) is a machine learning technique via which neural networks learn. The Buccaneer Risk

Mitigation Mission (BRMM) is UCS's first satellite developed in conjunction with DSTG. Magnetic field data from BRMM was used to demonstrate the on-orbit calibration process using GD.

The calibration gains and biases were trained and validated against seven hundred and fifty valid BRMM data points and then tested against unseen data. Several different calibration sets were produced, with the best performing set loaded to the spacecraft to test against the existing gains and biases on orbit.

The gradient descent calibration values were compared with the existing calibration values which were generated using analytical methods, and ground generated non-linear least-squares generated calibration values.

Input Data

The input data for BRMM on orbit calibration was comprised of telemetry packets from 10 May 2018 to 27 Aug 2018.

The telemetry packets included measured magnetic field strength in the three spacecraft axes and spacecraft position (Earth-Centred Inertial (ECI) J2000[8]).

For BRMM on-orbit calibration, the available data, after the quality filtering detailed below, was allocated into training, validation and test data.

Training data was used exclusively for training the gains and biases, validation data was used alongside training to measure how well the training was generalizing. Test data was used to evaluate final gains and biases against one another.

750 valid data points from 10 May 2018 to 3 Aug 2018 were divided into a training data set size of 675 and a validation data set size of 75. 269 data points from 4 Aug 2018 to 27 Aug 2018 were used as the test data for BRMM.

Data transformations

In order to calibrate using the BRMM data, several data transformations had to be performed.

Since magnetic field data was already calibrated in the BRMM datasets, the onboard data had to be uncalibrated using the existing analytical calibration gains and biases (see Table 3) and then converted to engineering units.

These changes were made to simulate the application of the process to an uncalibrated spacecraft magnetometer.

The World Magnetic Model (WMM) values for total magnetic field strength were used as the reference values when training on each of the on-orbit data points. As such, the process is independent of spacecraft attitude.

To obtain the WMM values of magnetic field, BRMM position was converted from ECI to Earth-Centred Earth Fixed (ECEF) and then to Latitude, Longitude, Altitude (LLA). The WMM-predicted magnetic field strength was obtained for each on-orbit measurement.

The difference in the total magnetic field strength as measured by BRMM and the WMM was used as the loss which GD aimed to minimize.

Data Filtering

From the BRMM datasets it was observed that many of the measured magnetic field values were significantly different to the WMM’s value to the point that they were beyond the software limits of calibration, and suspected outliers. These values, greater than 25% delta from the WMM value were not used to train or validate calibration gains and biases. A portion of these points were confirmed as invalid, such as being from flight portions where the on-spacecraft orbit propagator epoch was incorrect resulting in invalid WMM references.

The magnetometer was also mounted close to the primary payload power regulator (a known design issue). This was a known source of magnetic noise. Data points where the payload current was not equal to zero were removed from the dataset.

Magnetometer Calibration

BRMM used the following magnetometer calibration equation:

$$\boldsymbol{B}_C = \boldsymbol{g} * (\boldsymbol{B}_M + \boldsymbol{b}) \tag{1}$$

Where $\mathbf{B}_C$ = Calibrated Magnetic Field Strength; $\mathbf{g}$ = Calibration Gains; $\mathbf{B}_M$ = Measured Magnetic Field Strength; and $\mathbf{b}$ = Calibration Biases

The above formula was expanded in order to match the more general form used in GD: y = mx + b.

$$\boldsymbol{B}_C = \boldsymbol{g}\boldsymbol{B}_M + \boldsymbol{b}_{NN} \tag{2}$$

Where $\mathbf{b}_{NN} = \mathbf{g}*\mathbf{b}$, the GD-form calibration biases.

Gradient Descent

The above formula fits the pattern of a basic neural network (NN):

$$Input*\boldsymbol{w} + \boldsymbol{b} = Output \tag{3}$$

Where $\boldsymbol{w}$ = the weights matrix and $\boldsymbol{b}$ = the bias matrix.

Gradient descent relies on calculating the difference, or loss, between the output vector of the network (the predicted value), and a reference value. This loss is then minimized by back propagating the loss gradient through the network, resulting in small changes to the gain and bias matrices.

The WMM output included the total field strength as well as the individual components of the vector. Since the coordinate systems of BRMM and the WMM were not congruent, the total field strength was used as the point of comparison. This total field strength was calculated using the root-sum-square of the X,Y and Z components of the magnetic field strength vector.

The loss function chosen for BRMM GD was:

$$L = 0.5 * (S_{WMM} - S_C)^2 \quad (4)$$

Where L = loss, S_{WMM} = WMM RSS Magnetic Field Strength and S_C = Calibrated RSS Magnetic Field Strength.

This loss function produces a neat loss gradient:

$$\Delta L = S_{WMM} - S_C \quad (5)$$

To back propagate this loss through the network, the loss gradients with respect to the calibration parameters were calculated using the chain rule and can be seen below:

$$dL/d\boldsymbol{g} = \Delta L(\boldsymbol{B}_C / S_C)\boldsymbol{B}_M \quad (6)$$

$$dL/d\boldsymbol{b} = \Delta L(\boldsymbol{B}_C / S_C) \quad (7)$$

These three by one gradient vectors, calculated for each training data point, were then multiplied by the learning rates and subtracted from the gains and biases. This process tuned the gains and biases step by step, minimising the difference between predicted magnetic field strength value and the WMM reference value.

Stochastic Mini-batches

Two possible issues with gradient descent are that it can be susceptible to over-fitting and also to remaining in a local, rather than the global, minimum. Regularisation is a common way of combating these issues.

One method of achieving regularization in a NN is to implement mini-batch gradient descent or stochastic gradient descent.

Mini-batch gradient descent takes the average gradients across a batch of data points from the training set and applies this average gradient, rather than individual updates, to the gains and biases. This averages out the gradient steps applied to the gains and biases.

Stochastic gradient descent adds to this by randomly choosing the data allocated to batches rather than moving through the data sequentially. This averaging and shuffling of data balances out the influence of any one data point.

Stochastic, mini-batch gradient descent (SGD) was used for training the BRMM calibration parameters.

Hyper Parameter Optimisation

There are several hyper parameters which need to be tuned to optimise SGD, namely: the learning rates and batch size. Both were tuned for BRMM calibration, with the tuned parameters shown in Table 1. These hyper parameters were tuned by varying each as an independent variable until the minimum average loss across the training set was found.

Table 1: Optimised SGD Hyper Parameters

Hyper Parameter	**BRMM**
Gains Learning Rate	1.0e-12 and 1.0e-11*
Biases Learning Rate	1.0e-02
Batch size	25 and 1*

Training

Training presented a few irregularities during early method development.

Two of the trained gains and bias sets used the existing analytical calibration values as starting values during SGD training. This is opposed to the standard approach of initializing the gains at 1 and the biases at 0. This approach demonstrated the ability of SGD to optimize existing calibration values on orbit.

The optimised gains learning rate for this non-standard approached was 1.0e-11 with a batch size of 1 performing best (strict stochastic gradient descent).

Initially, training ran for 10 epochs, or passes through the entire data set, whereas later testing ran for 100 epochs. In order to optimize over the 100 epochs, only when the validation loss was reduced, an indication of overall improvement, were the gains and biases saved. This prevented over-fitting after the 100 epochs of training.

Non-linear Least Squares Comparison

A nonlinear least squares solution was generated using existing MatLab functions to compare the SGD solution with more traditional numerical methods.

RESULTS

Test Data Results

As can be seen in Table 2 below, both sets of SGD calibration values reduced the average difference between the predicted and measured magnetic field strength ('error') over the test set by 50%. The B1SB calibration parameters (batch size of 1, starting with the existing analytical calibration values), performed slightly better than the B25S0 (batch size of 25 with standard starting values) parameters.

Table 2: Test Data Results

Calibration Values	Average Error (nT)
Default Gains and Biases	20,256
Existing Analytical Calibration Values	2191.4
Batches of 1, BRMM Start (B1SB)	906.7
Batches of 25, Default Start (B25S0)	998.4
Nonlinear Least-Squares	1150.0

New calibration values

The calibration values shown in Table 3 show some significant differences and a few similarities between the sets.

The Bx and By values are all very similar across the three sets of calibration parameters. Bz changes moderately between the two SGD sets and is vastly different in the existing calibration set.

The magnitude ordering of gains for both SGD sets is the same but the B25S0 values are more polemical. This is understandable given that B1SB had a smaller learning rate and began at an already partial optimized position. The existing analytical calibration values are still a similar magnitude to the other sets but are on average lower and have different magnitude ordering.

The NLLS calibration values are very similar to those produced by SGD. They appear to be halfway between the two SGD sets except for Bz, which could account for the performance difference below.

This all suggests that there are several local minima at which these values are sitting in the vector space. It is possible that none of these values are sitting in the true global minima or that they are all circling a large global minimum.

Table 3: New Calibration Values

Calibration Value	B25S0	B1SB	Initial BRMM Values	Nonlinear Least-Squares
Gx	0.8064	0.8303	0.84	0.8151
Gy	0.8950	0.8690	0.79	0.8836
Gz	0.8759	0.8608	0.81	0.8649
Bx (nT)	34282.7	34669.3	35766.7	34616.5
By (nT)	-6004.3	-5924.3	-5601.3	-5968.1
Bz (nT)	-599.3	-853.4	1037.5	-133.0

On-ground Testing

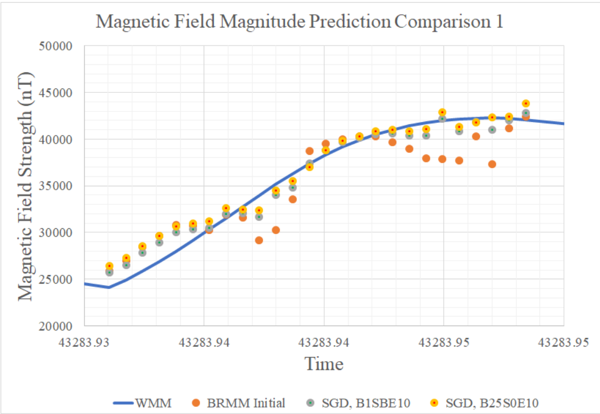

Figure 1: On-ground magnetic field prediction 1

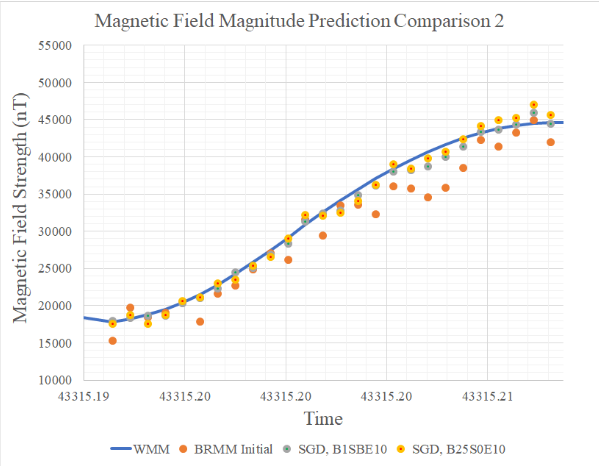

Figure 2: On-ground Magnetic Field Prediction 2

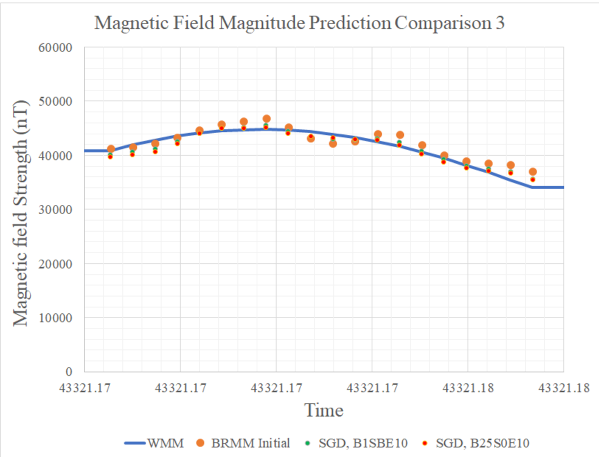

Figure 3: On-ground Magnetic Field Prediction 3

Figures 1-3 suggest that the SGD calibration values are generally more accurate compared to the existing analytical calibration values. The analytical values appear more sporadic in Figures 1-2 but much more stable in Figure 3. It appears that the SGD calibration values are 'damped' further and less prone to larger, inconsistent inaccuracies.

Both SGD calibration values outperform the NLLS by 13% and 20% respectively. The maximum errors produced by both methods are similar though, as shown in Figures 5 and 6, but SGD produces significantly more low-error calibrations.

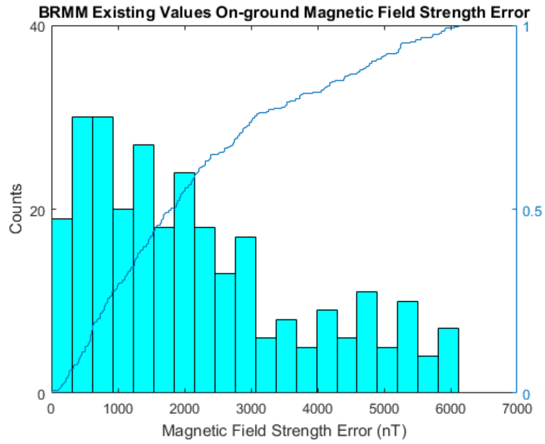

Figure 4: On-ground initial calibration values magnetic field strength error

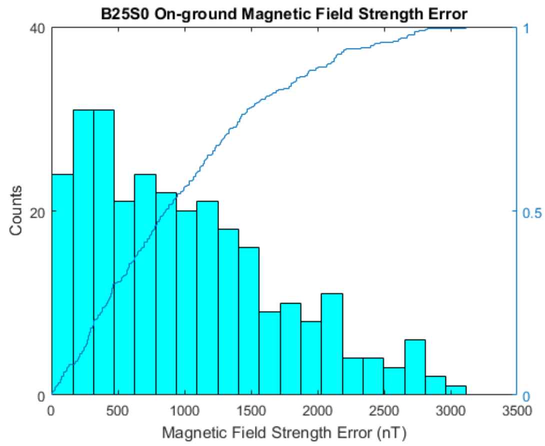

Figure 5: On-ground B1SB calibration values magnetic field strength error

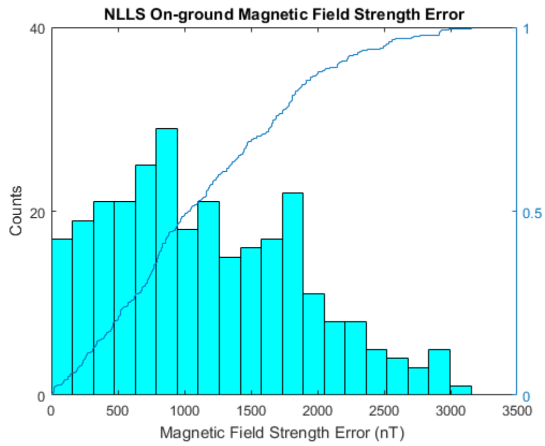

Figure 6: On-ground B25S0 calibration values magnetic field strength error

Figures 4-6 demonstrate the reduction in magnetic field strength error with the SGD calibration values. B25S0 and B1SB performed very similarly.

The SGD values produced much more consistent magnetic field strength predictions compared to the existing analytical calibration values.

On-orbit Testing

The calibration values from B25S0 were chosen for testing on orbit given they were generated from standard initial calibration values and were significantly different from the initial BRMM values. The parameters were tested on BRMM on 13/14 Sep UTC 2018. The attitude determination system was run in test mode (no actuation) for three orbits, and then the B25S0 calibration parameters were commanded and observed for three orbits. The results follow.

Table 4: On-orbit Calibration Values Comparison

Calibration Set	Magnetic Field Strength Average Error (nT)	Sun-Mag Angle Average Error (Deg)
Initial BRMM Values	1025	10.17
B25S0	827	7.12

Overall, the new SGD-learned values performed better than the existing analytical calibration values on orbit. Both calibration sets performed better on orbit than they did in on-ground testing against on-orbit data. It is likely that the small on-orbit test size had an impact on these results given the on-ground test set data was measured over a month.

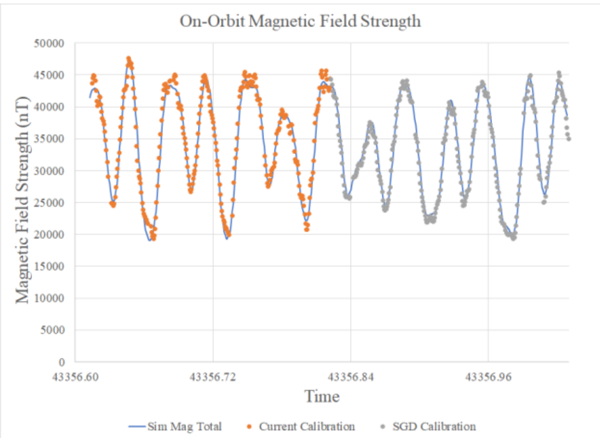

Figure 7: On-orbit magnetic field strength measurement comparison

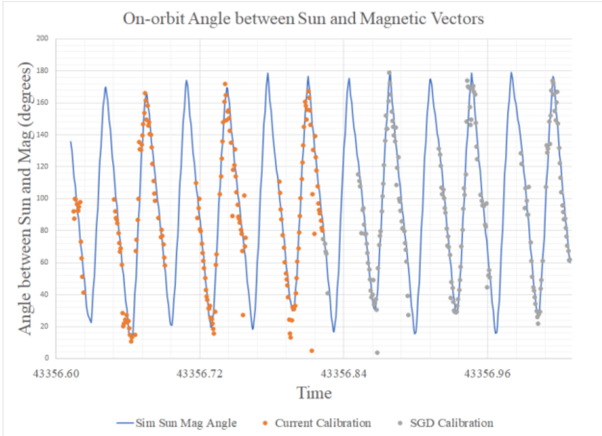

Figure 8: On-orbit measured angle between sun and magnetic vector comparison

The SGD calibration values show less oscillation around the simulated magnetic field strength in Figure 7 than the existing analytical calibration values, particularly at the peaks and troughs. This suggests that the SGD values handle significant changes in the magnetic field better than the existing analytical values do. There was only a small observable improvement in the sun and magnetic field vector angle as seen in Figure 8. Again, the SGD values perform slightly better at peaks and troughs.

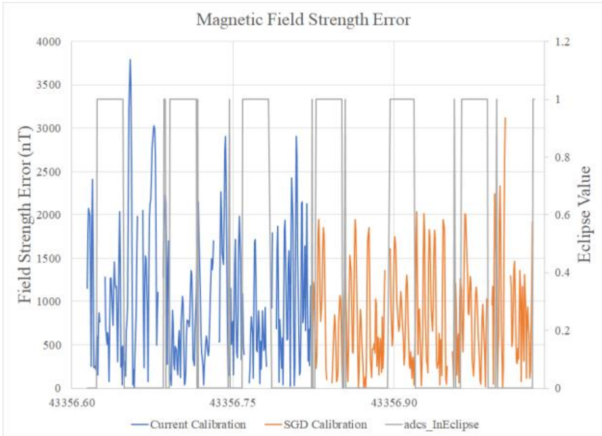

Figure 9: On-orbit magnetic field strength error comparison

The magnetic field strength error looks like it marginally improved with the SGD values. The existing calibration values perform better during eclipse than the SGD values. While out of eclipse both values perform worse, with the SGD values performing the better of the two.

It is possible that the existing calibration values perform better in the absence of magnetic field noise given that the SGD values have been trained with potentially noisy data and so have learned to compensate for it. This may account for the performance of the existing values in eclipse, there are fewer systems in use and no solar cell power compared to their performance out of eclipse. The BRMM spacecraft was not designed to reduce unbalanced current loops.

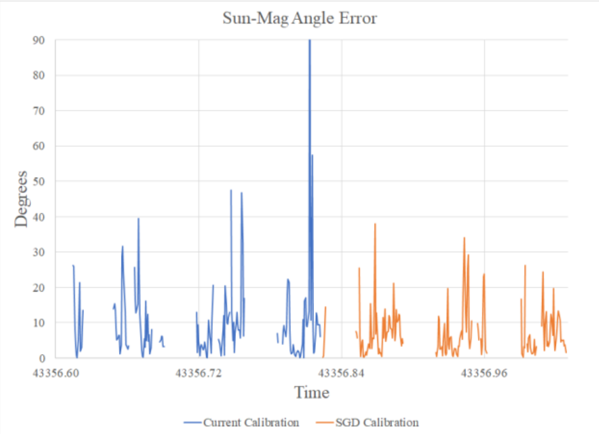

Figure 10: On-orbit angle between sun and magnetic vector error comparison

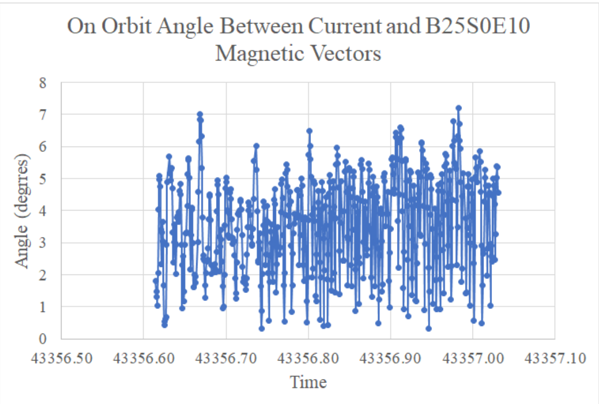

Figure 11: On-orbit angular difference between existing and SGD magnetic vectors

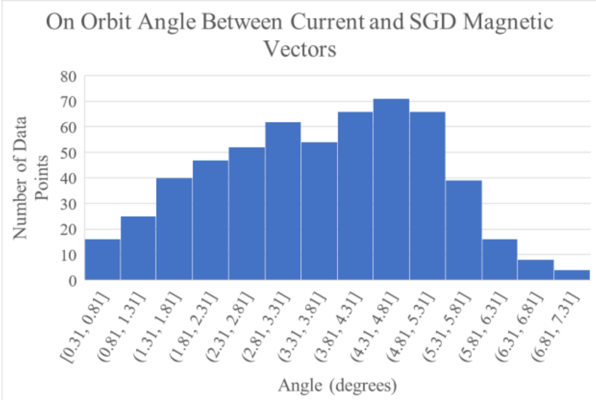

Figure 11: A histogram of the on-orbit angular difference between existing and SGD magnetic vectors

Figures 11 and 12 show an angular change of up to 7.2 degrees across the six orbits between the existing and SGD calibration produced magnetic vectors. The mean difference is 3.18 degrees with a standard deviation of 1.49

This change in magnetic vector and the decrease in magnetic field error suggests that the on-orbit SGD calibration resulted in significant improvement to the attitude accuracy of BRMM.

DISCUSSION

The implementation of SGD to calibrate a magnetometer on-orbit can be considered successful based on the improvements over the existing calibration values in both magnetic field strength accuracy and in sun magnetic field vector angle accuracy.

However, both the quantity and quality of data that was used to train the SGD values could have been improved.

Although the SGD calibration values on BRMM seemed to reach convergence after the ten epochs, the data set is relatively small once false readings were taken out – 1019 total data points.

One improvement in the methodology of training would have been to take the entire data set and split the data into training, testing and validation data sets randomly. The results obtained in on-ground testing are possibly skewed because the testing set came exclusively from the data from August 4 to 27.

Future SGD calibration should seek to maximise the amount of data but also how varied this data is and randomly assign this data into the appropriate data sets.

Nonlinear Least Squares performed marginally worse than expected compared to SGD. It is currently used by UCS to calibrate magnetometers on-ground. It is likely that SGD outperformed NLLS on this data set given the size and variation of the data set. NLLS would be expected to perform better compared to SGD with a smaller, less varied data set

However, given that the NLLS calibration values were not tested on-orbit, to complete a fair comparison between the two methods, on orbit testing would be required.

FUTURE WORK

SGD has been successfully demonstrated as a means by which to calibrate a magnetometer on-orbit. Given that NLLS calibration produced similar but greater losses compared to SGD calibration in on ground testing, on orbit testing of NLLS is required to determine conclusively that SGD is superior for on-orbit magnetometer calibration.

Testing both NLLS and SGD calibration values on-orbit over many orbits should be able to determine more clearly which of the two methods is more suitable for on orbit magnetometer calibration.

In order to get better training results in the future, the data acquired should be as representative as possible of all the states that the spacecraft is likely to experience. This should include coordinate, attitude and system-state variation.

UNSW Canberra Space also has plans to attempt applying on-orbit calibration to other on-board sensors such as coarse sun sensors and earth horizon sensors.

Building further on this desire is to move towards the on-orbit calibration of multiple sensors in concert, given that the readings of one attitude sensor directly impacts the effectiveness of the on-orbit calibration of another sensor.

ACKNOWLEDGMENTS

UNSW Canberra Space acknowledges the excellent work and support of Defence Science and Technology Group, the collaborating partner on Buccaneer Risk Mitigation Mission

REFERENCES

1. Renaudin, V., M.H. Afzal and G. Lachapelle, "Complete Triaxis Magnetometer Calibration in the Magnetic Domain," Journal of Sensors, vol. 2010, Article ID 967245, 2010
2. Pang, H., J. Li, D. Chen, M. Pan, S. Luo, Q. Zhang, F. Luo, "Calibration of Three-axis Fluxgate Magnetometers with Nonlinear Least Square Method," Measurement, vol. 46, No. 4, May 2013
3. Nassar, B., W. Hussein, M. Medhat, " Supervised Learning Algorithms for Spacecraft Attitude Determination and Control System Health Monitoring," IEEE Aerospace and Electronic Systems Magazine, vol. 32, No. 4, April 2017
4. Ma, Z., Y. Wang, Y. Yang, Z. Wang, L. Tang, S. Ackland, "Reinforcement Learning-Based Satellite Attitude Stabilisation Method for Non-Cooperative Target Capturing," Sensors, vol.18, No. 12, December 2018
5. De Amorim, J. and L.S Martins-Filho, "Experimental Magnetometer Calibration for Nanosatellites' Navigation System," Journal of Aerospace Technology and Management, vol.8, No.1, January-March 2016.
6. Zhang, Z., J. Xiong and J. Jin, "On-orbit Real Time Magnetometer Bias Determination for Micro-Satellites Without Attitude Information," Chinese Journal of Aeronautics, vol.28, No.5, October 2015.
7. Hajiyev, C., "Orbital Calibration of Microsatellite Magnetometers Using a Linear Kalman Filter," Measurement Techniques, vol.58, No.9, December 2015
8. Sanz Subirana, J., J.M Juan Zornoza and M. Hernandez-Pajares, "Conventional Celestial Reference System," 2010, https://gssc.esa.int/navipedia/index.php/Conventional_Celestial_Reference_System