Using an artificial neural network to improve the orientation accuracy of unmanned aerial vehicles

Ammar Assad, Sergey Serikov, Nader Al Bitar

***Abstract*— The importance of Magnetometer measurement in all the fields that are related to moving bodies including UAV, they are used for calculation the attitude of moving body, they are also can integrated with other sensors to calculate accurate position moving body. According to that, there is an urgent need to get accurate data from magnetometer, the raw data can be used directly or can be processed by a calibration process. In this article, an assessment of using neural network (NN) in calibration magnetometer is conducted, the NN model is a sequential model that consists of many layers, the dataset contains raw measurements and calibrated data. The model achieved learning accuracy 99.38 % and loss 1.411 μT; the loss metric is mean square error. The dataset was split into three parts, train, test and validation. Train and validation parts are used in the training process. The trained model is tested on the test part, on the whole dataset and on noisy dataset by adding some noise on raw data and test the efficiency of NN model. The results showed high efficiency of NN model in calculating calibrated output directly from raw measurements. A structure of NN is also discussed. To validate the proposed method in the field of UAV orientation, heading angle by using magnetometer data is calculated for raw, calibrated, output of proposed method and another calibration method on the same dataset.**

***Keywords*— Magnetometer, Calibration, Neural Network, Sequential model.**

I. Introduction

In the field of scientific and technological advancement, there is an urgent need of precision and accuracy. This holds in many fields that need measurements from sensors, since sensors are the most important tools in any real systems and in the same time, they are most susceptible to noise. In the field of magnetometry, where the measurement of magnetic fields plays an important role in various applications that range from geophysics and environmental monitoring to medical diagnostics and aerospace navigation. The accuracy of magnetometer measurements directly influences the reliability and efficacy of the data obtained when used in real application and calculations. This shows the fundamental aspect of magnetometry calibration [1].

Magnetometers are sensors that are used to measure magnetic field in the location where they are placed. A magnetic field's strength, direction, or relative change at a specific area can all be measured using different types of magnetometers. Calibration, the process of adjusting and tuning sensors outcomes, the main aim of magnetometer calibration is to ensure the accuracy and reliability of magnetometer data. In this exploration, it should be explained briefly that the outputs of magnetometer can be affected by many conditions. Calibration can be also defined as a process of converting raw data of measurements of magnetometer into calibrated data [2].

According to the fact that the magnetic components on the aircraft itself can interfere with the desired magnetic measurements that are required to navigate. The measurements of magnetometers contain magnetic signals from both the (desired) earth field and (undesired) aircraft field. According to this inference, it is difficult to separate the two signals. Many studies used linear methods to clean the magnetic signal, uses a single pair of scalar and vector magnetometers [3].

There are basically two reasons why raw measurements will not be accurately equal the true value of magnetic field [4]:

- Errors in the magnetometer triad: These errors contain some components
  - Non-orthogonality of the magnetometer axes. This is defined by matrix (3×3) C_{no}
  - Existing of zero bias: this holds that magnetometer measures a non-zero magnetic field even when magnetic field is zero, this is defined by B_z.
  - There is a difference in sensitivity between the three axes of magnetometer, this is represented by matrix (3×3) C_s; it is diagonal matrix.
  - Noise in the magnetometer measurements. This in most cases is assumed to be Gaussian noise; this is defined by N_G.
- Existing of magnetic disturbances: magnetometer does not measure only the local magnetic field, but also an additional magnetic field component. In general, magnetic disturbances are not stationary and constant especially when existing iron bodies near the magnetometer, rigidly installation of sensor might decrease disturbances. The existing of iron body will negatively affect the measurements because of hard and soft iron. Hard iron effects are due to the permanent magnetization of the magnetic material and lead to a constant additional magnetic field. The vector representing these effects is denoted as B_{hi}. Soft iron effects are due to magnetization of the material as a result of an external magnetic field and

will therefore depend on the orientation of the material with respect to the external field. It can change both the magnitude and the orientation of the measured magnetic field. This is defined as matrix (3×3) C_{si}.

After all, the equation between raw measurements and true values is

$$M_{c(3\times1)} = C_s * C_{no} * (C_{si} * M_{3\times1} + B_{hi}) + B_z + N_G$$
$$M_{c(3\times1)} = C_s * C_{no} * C_{si} * M_{3\times1} + C_s * C_{no} * B_{hi} + B_z + N_G \quad (1)$$

Where $M_{3\times1}$ is 3 axes measured magnetic fields, $M_{c(3\times1)}$ is 3 axes calibrated magnetic fields. The Eq. 1 shows that the true representation contains linear disturbance of magnetometer. The effect of temperature changes is not shown in this equation, this needs experimental calculations; this effect might be nonlinear.

The traditional equation of calibration is basically as follows

$$M_{c(3\times3)} = A_{3\times1} * M_{3\times1} + B_{3\times1} \quad (2)$$

Where A and B are calibration parameters. To obtain the calibration parameters A and B, a set of measurements of the magnetic field M along with corresponding known values of the magnetic field M_c should be collected. This forms a set of equations that can be solved to find the calibration parameters.

The calibration process involves solving for A and B in the equation Eq. 1. Depending on the number of measurements and the precision of calibration process, it may be used different techniques such as least squares regression to find the best-fitting calibration parameters. Here are some common methods for magnetometer calibration [5]:

Linear Regression: Using linear regression techniques to find the best-fitting calibration matrix A and bias vector B that minimize the difference between measured and true magnetic field values.

Ellipsoid Fitting: Model the calibration as an ellipsoid fitting problem. This is particularly useful when the magnetic field measurements are subject to distortions.

Least Squares Optimization: Using optimization techniques such as least squares to minimize the difference between the measured and true magnetic field values.

Geomagnetic Field Modeling: Using mathematical models of the Earth's magnetic field to calibrate magnetometer readings. This is common in navigation applications.

Manufacturing Calibration: Perform factory calibration by characterizing each sensor's individual characteristics during the manufacturing process. This information can be used to create compensation parameters.

Magnetometers has an important role in orientation estimation. Their significance concludes in ability to measure information about the Earth's magnetic field, which enables the estimation of an object's orientation in the space. Also, it can be integrated into sensor fusion algorithms with other sensors such as accelerometers and gyroscopes; magnetometers improve the accuracy and robustness of orientation estimation. The orientation information serves to get a global reference frame, allowing for precise heading estimation. This is especially valuable in navigation systems. Moreover, accurate magnetometers measurements have a key role in correcting drift in gyroscope measurements. In environments where GPS signals are weak or denied, such as indoors or underwater, magnetometers become indispensable for maintaining accurate orientation estimates. Beyond navigation, magnetometers have now used in features like automatic screen rotation in smartphones and tablets, also in attitude control of unmanned aerial vehicles.

The main aim of the article is to reduce the complexity of calibrating the magnetometer and increase reliability of their outputs, thus will directly lead to increase the accuracy of determining the orientation angles of the UAV. The research will depend on using of a neural network to correct errors or raw outputs of magnetometers and analysis the structure of NN to find best structure.

II. Related Works

Authors in [6] investigated calibration of the magnetometers of non-dedicated satellites using neural networks. The transformation is encapsulated in a workflow named Macaw (Magnetometer Calibration Workflow). They used raw magnetometer measurements, housekeeping data, and telemetry data. They merged data for one month. Macaw contains many layers that do many stages of preprocessing the main aim is calibration to accurately calculate the earth magnetic field monthly or daily at specific location.

In [7], authors conducted yaw/heading estimation method and focus how to achieve high accuracy of estimation through a Machine Learning (ML) approach, particularly when the calibration motion range of the vehicle/device is limited. They used Random Forest (RF) algorithm and employ it after training to address magnetometer uncertainty. They had reference from the Pan Tilt Unit-C46 (PTU-C46) with precise positioning serves as a reference heading value for labeling magnetic features in the learning model. Their main approach facilitates estimation yaw regarding challenging conditions to overcome many challenges that are related to susceptibility of magnetometer to hard and soft iron in the surrounding environment. RF model used 5 features: 3 axes magnetic fields (M_x, M_y, M_z), the sphere radius, and the ratio between M_x and M_y and the output is Yaw angle. The experiment was carried out using a low-cost platform equipped with Micro-Electro-Mechanical System (MEMS) sensors as gyroscope, accelerometer, and magnetometer. They applied sensor fusion approach to track the yaw value after the level calibration despite various error conduction. The RF model accomplishes a superior result with more stability and more minor error. Under iron disturbance or calibration absence, the ML model still maintains the good tracking command with maximum Mean Square Error of about 0.3◦.

Authors in [8] implemented neural networks for calibration of magnetometers by using additional magnetometer and other flight data. Their approach has been shown to outperform the state-of-the-art model when only in-cabin data is used. On held out testing data, magnetic signal errors of less than 6.5 nT and navigation position errors of less than 45 meter are consistently achieved

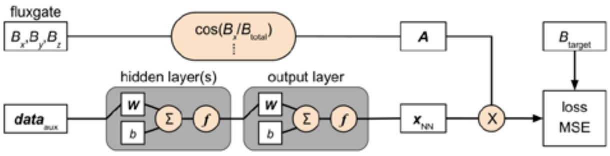

Figure 1: Approach in [8]

Data aux: raw measurements and other data (using the Tolles-Lawson model [9], which is a linear model that uses measurements from a vector magnetometer to remove aircraft magnetic field contributions to the scalar magnetometer measurements, which are used for navigation), the output of NN is multiplied by Matrix A from calculations of Tolles-Lawson model, this needs another magnetometer (fluxgate) to calculate the matrix A.
The main difference between study [8] and our approach is that the proposed approach needs after training only the measured values from the sensor. In addition, the main significance of proposed approach may lie in choosing the optimal neural network architecture for calibration and correcting the error of the magnetometer.

III. Methods

A. Neural network

A neural network can be defined as function approximator, it is composed of many interconnected layers of nonlinear functions. This architectural enables neural networks to effectively model and approximate complex, nonlinear relationships within data. Composition of these nonlinear transformations allow neural network to represent mathematical functions with ability to adapt and learn from diverse datasets for a wide range of applications in machine learning and artificial intelligence such as pattern recognition, classification, regression, and decision-making.

Connected nodes have weight and biases; during a training phase, the weights of connections between neurons are adjusted to process information through the network. They are adept at deriving complicated patterns from data and forecasting outcomes using those patterns [10].
Sequential model is used in this research, sequential model is a type of neural network architecture that is defined using a linear stack of layers, which is implemented in various deep learning frameworks, including TensorFlow and Keras [11]. The Figure 2 shows the overall process. The data is measured by the board that contains FXOS8700 3-Axis accelerometer and magnetometer, and the FXAS21002 3-axis gyroscope (Adafruit Precision NXP 9-DOF Breakout Board - FXOS8700 + FXAS21002) [1]

The using of neural network has advantage that the calibration parameters are not only A and B as in Eq. 2; they are many parameters of network weights and biases, the next equation shows the output of first layer neurons;
Suppose M_j is vector of input of layer j, L_j is length of vector M_j, A_n is vector of weights of neuron n, f_n is activation function of neuron n. B_n is bias on neuron n. N_j is number of neurons in each layer

$$M_{j+1(n)} = f_n\left(\sum_{i=0}^{L_j} A_{n_i} * M_{j_i} + B_n\right)$$

$$\begin{pmatrix} M_{j+1(1)} \\ \vdots \\ M_{j+1(N_j)} \end{pmatrix} = \begin{pmatrix} f_1\left(\sum_{i=0}^{L_j} A_{1_i} * M_{j_i} + B_1\right) \\ \vdots \\ f_{N_j}\left(\sum_{i=0}^{L_j} A_{N_{j_i}} * M_{j_i} + B_{N_j}\right) \end{pmatrix} \tag{3}$$

M_0 is the raw measurements of magnetometer. By comparing Eq. 3 and Eq. 2, it can be concluded that NN with its many layers has many parameters that would make calibration more robust.

B. Dataset

The dataset was available on GitHub, Figure 4 shows raw and calibrated values of magnetometer readings. The main steps to use NN as approach:

- Get Magnetometer measurements in location where earth magnetic field is known. As the size of data increases, the accuracy increases only up to a certain limit, after specific limit the increase of dataset will be useless and it should reconsider other ways of increasing accuracy such as structure of neural network; however, the most important thing is to measure in different positions (as suggestion: measurements in all possible combinations of North, South, East and West will be sufficient). The collected data at least should contain these combinations for accurate calibrations.
- Train the model
- Using Trained model from previous step

In the case of perfect calibration, a magnetometer measures the local magnetic field and its measurements will therefore lie on a sphere with a radius equal to the local magnetic field and centered in 0. Figure 4 shows raw data and calibrated data from the dataset (when plotting X and Z coordinates of data)

[1] https://www.adafruit.com/product/3463

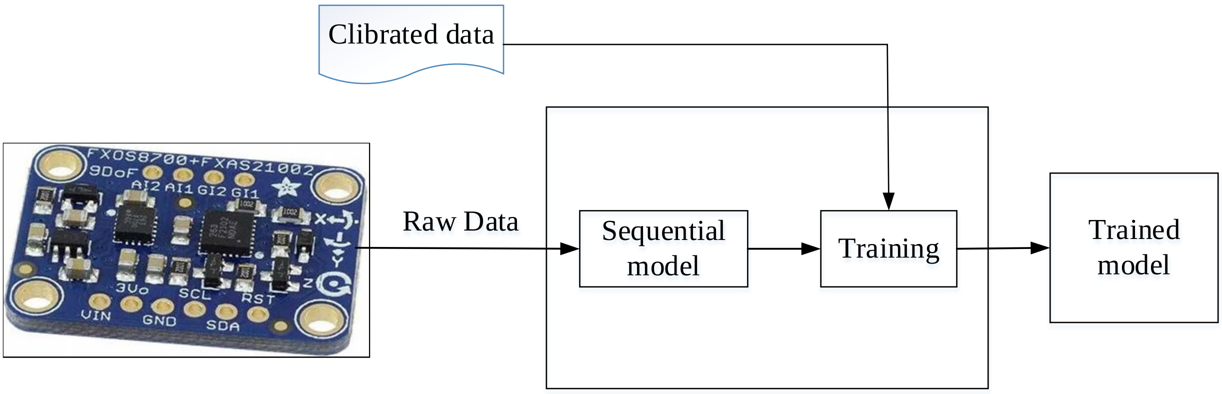

Figure 2: Overall all approach

The evaluation process is shown in Figure 3

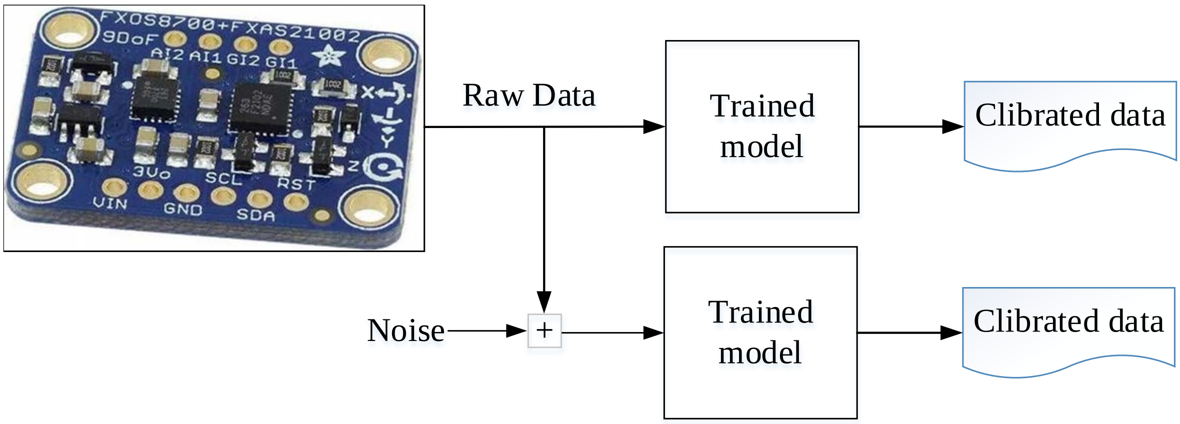

Figure 3: Evaluation process

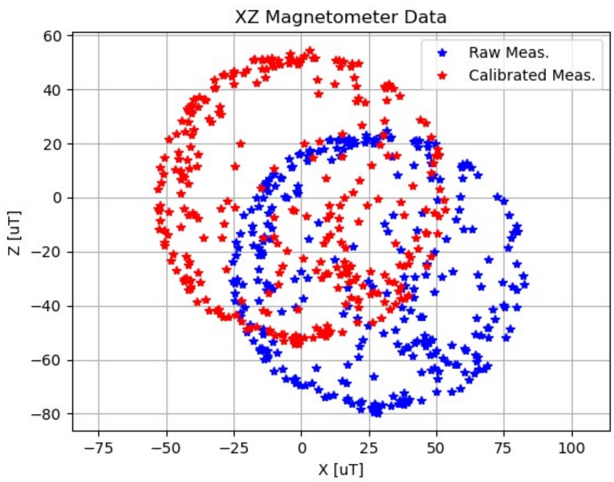

Figure 4: Magnetometer (raw and calibrated) from dataset

IV. Implementation and Results

A. Training

The implementation was accomplished using python programming language. Sequential model and Dense layers were used from Keras library. The first sequential model that was tested contains:

- First dense layer with 128 units (neurons), activation function is Relu: Rectified linear unit
- Second layer with 64 units (Activation function is Relu)
- Third layer with 32 units (Activation function is Relu)
- Forth layer (the last) with 3 units (Activation function is linear); the problem here is not classification or clustering problem, it is regression problem and that needs linear activation function in last layer and loos function of type regression such as mean square error.

Dataset was split into three parts: test part (124 samples) train part (160 samples) and validation part (40 samples). The process of splitting is shown below in Figure 5

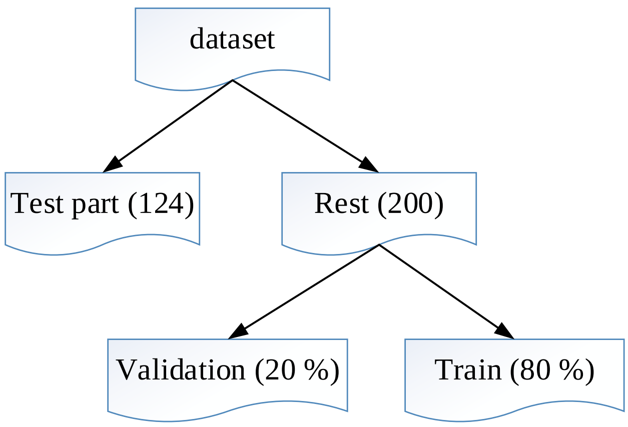

Figure 5: Splitting of dataset

The neural network's parameters are trained using the training set, and hyperparameter tuning is greatly aided by the validation set. Hyperparameters are configuration options, like the number of needed hidden layers or number of neurons in each layer, that affect learning but are not learned during training.

The test set is used to assess the neural network's capacity for generalization after the model has been trained and hyperparameters have been adjusted using the validation set. The final accuracy is 99.38%. this accuracy is how model is trained, it is just index if the model is trained or not and how well it is trained, the effective metric is the loss value of model because the proposed model is regression model, the final loss is 1.411 μT. Figure 6 shows loss and accuracy of the model during training and validation processes.

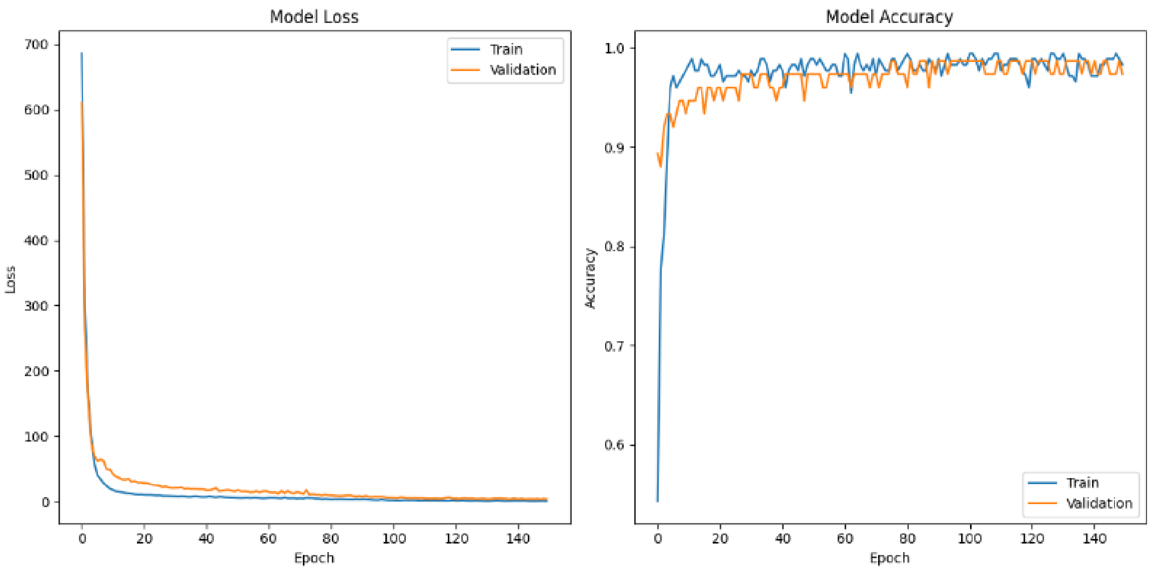

Figure 6: Loss (in μT) and validation of training and validation process

As it can be seen, after 20 epochs, the model is trained well since the loss decreased rapidly. This is due to that neural networks can automatically learn relevant features from the data. If there are complex patterns or interactions between different components of the magnetic field, the neural network can potentially discover and exploit them. The batch size in training process was set to *8*. The dataset size is 324. The figures shows that the system is will fitted (no over or under fitting) [12].

When batch size is 16 the accuracy reached 98.1 % and the validation loss decreased to 34.05 with 150 epochs. Then batch size was then decreased to 8, this value resulted in best accuracy of model training. It should be noted that this accuracy is not train accuracy, it is the validation accuracy.

B. Structure analysis

Many experiments have been conducted to get best structure, back to calibration parameters; matrix $A_{3\times3}$ and $B_{3\times1}$, there are 12 variables, at least we need 12 equations to calculate them, but in the proposed method, the model not just is solving these 12 parameters but all linear and nonlinear parameters.

Number of Layers: According to dataset size and the complexity of the conducted problem, 4 layers are sufficient; since more complex tasks or larger datasets may benefit from deeper architectures. 4 layers will avoid unnecessarily deep network, since that deep network may lead to overfitting or increased computational cost without significant performance improvement.

Number of neurons: the number of neurons in each layer is chosen to be decreasing from one layer to another starting with (128, 64, 32).

Batch size: according to best number of neurons and best number of layers, complexity of model is calculated as follows:

Table 1: Model specifications

Layer	Number of parameters
First	$128 * 3 + 128 = 512$
Second	$64 * 128 + 64 = 8256$
Third	$32 * 64 + 32 = 2080$
Forth	$3 * 32 + 3 = 99$

The total number of parameters are: 10947, the number of datapoints is in training 160 datapoint, the full complexity is $O(Ba * 10947 * 160) = O(Ba * 1.8 * 10^6)$, batch size should be {2, 4, 8, 16, 32, …, 2^r; r is positive integer number} 32 or 16 will increase too much the complexity of calculation, 4 is small since only 4 datapoint is handled together, 8 is best one. The best experiment is when the model contains 4 layers, and (128, 64, 32) neurons and batch size is 8. The batch size 16 resulted in 3.2830 μT loss; batch size 8 resulted in 1.411 μT loss value

C. Validation

First validation on the test part dataset itself and make a comparison with calibrated data, Figures below show the results (each figure shows two axes together at the same time) the calibration process.

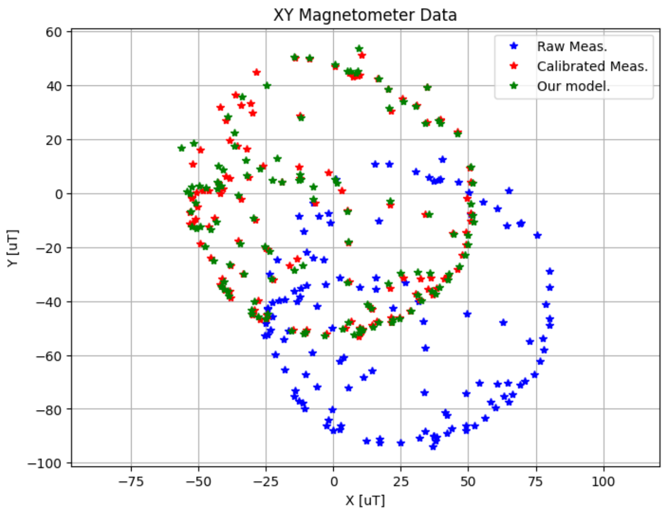

Figure 7: XY Magnetometer data (test part) (Raw, calibrated, our model)

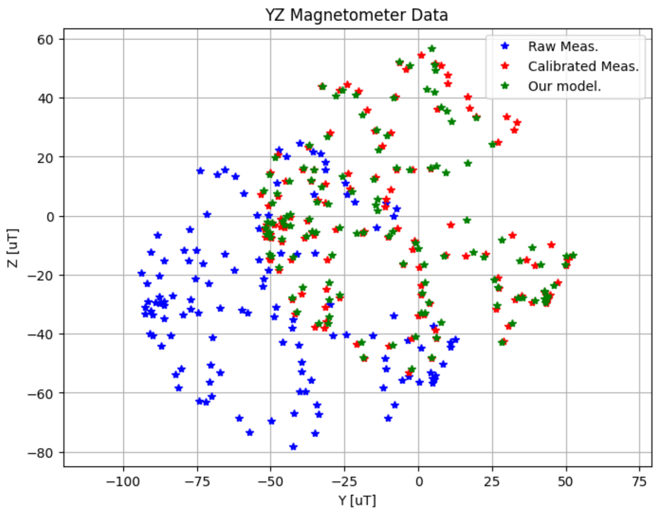

Figure 8: YZ Magnetometer data (test part) (Raw, calibrated, our model)

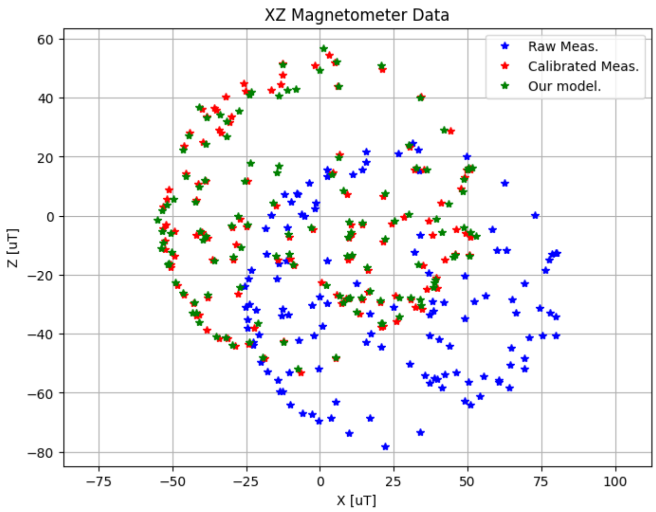

Figure 9: XZ Magnetometer data (test part) (Raw, calibrated, our model)

Figure 10, Figure 11, Figure 12, Figure 13 show the results when testing trained model on the full dataset (324 samples)

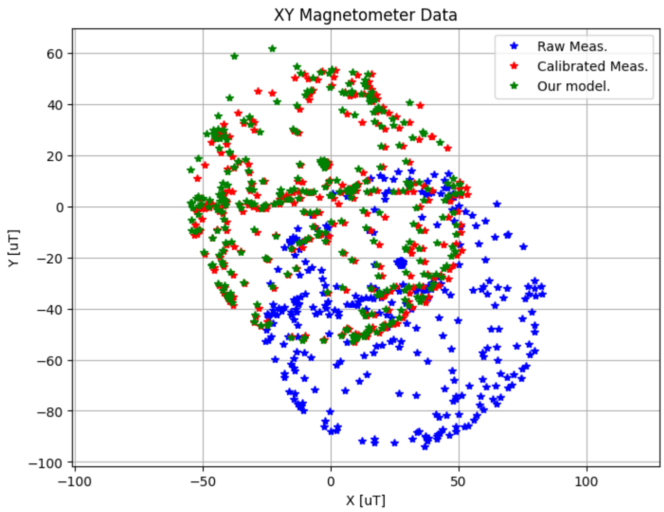

Figure 10: XY Magnetometer data (Raw, calibrated, our model's output)

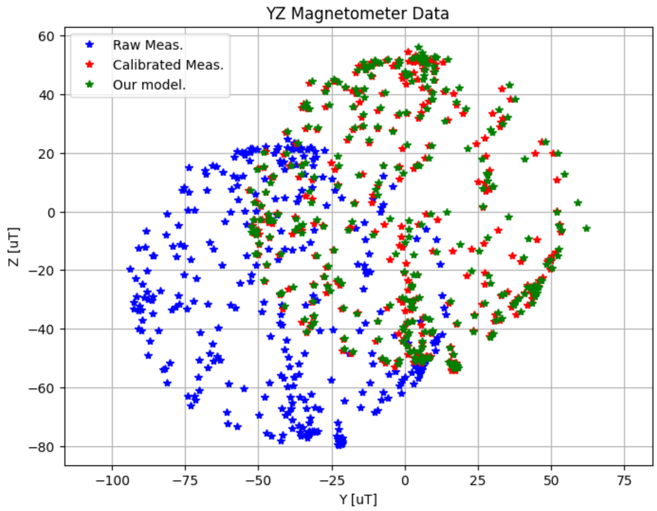

Figure 11: YZ Magnetometer data (Raw, calibrated, our model's output)

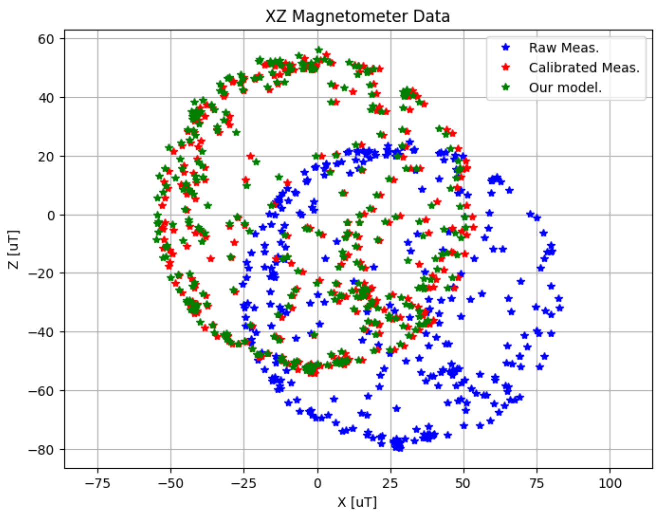

Figure 12: XZ Magnetometer data (Raw, calibrated, our model's output)

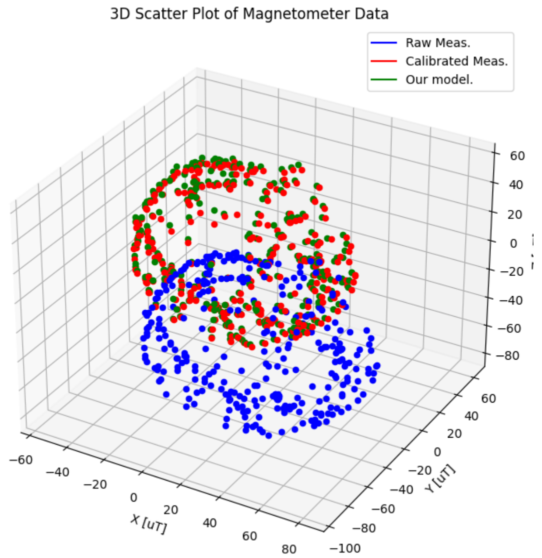

Figure 13: 3D scatter of magnetometer data (Raw, calibrated, our model's output)

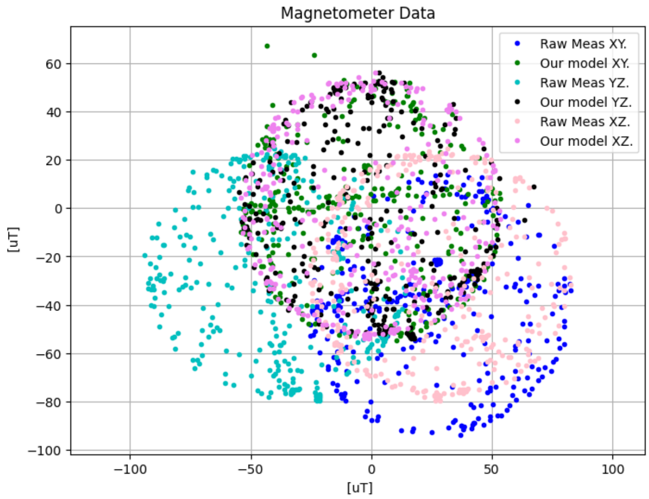

Figure 14: Raw data and our approach output (All axis)

Second validation: adding noise to raw data and test the model, the adding noise is Gaussian white noise, with noise factor as in Eq. 4. This is test for robustness of trained neural network. This is applied just on the first layer

$$M_{1(n)} = f_n\left(\sum_{i=0}^{L_j} A_{n_i} * (M_{0_i} + No_i) + B_n\right) \quad (4)$$

$$No = noise\ factor * Noise$$

Figure 15 shows the calculation of heading angle of raw data, true calibrated data and the proposed model,

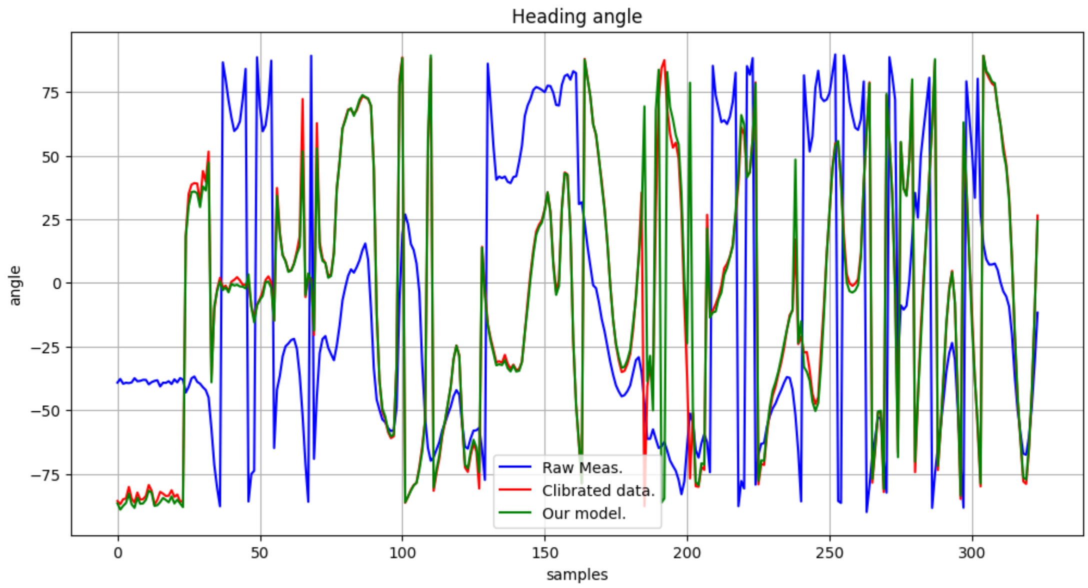

Figure 15: Heading angle

Ellipsoid method was applied to adopted dataset to validate proposed method with other methods. the calibration Ellipsoid results were used to calculate heading and the results are shown below.

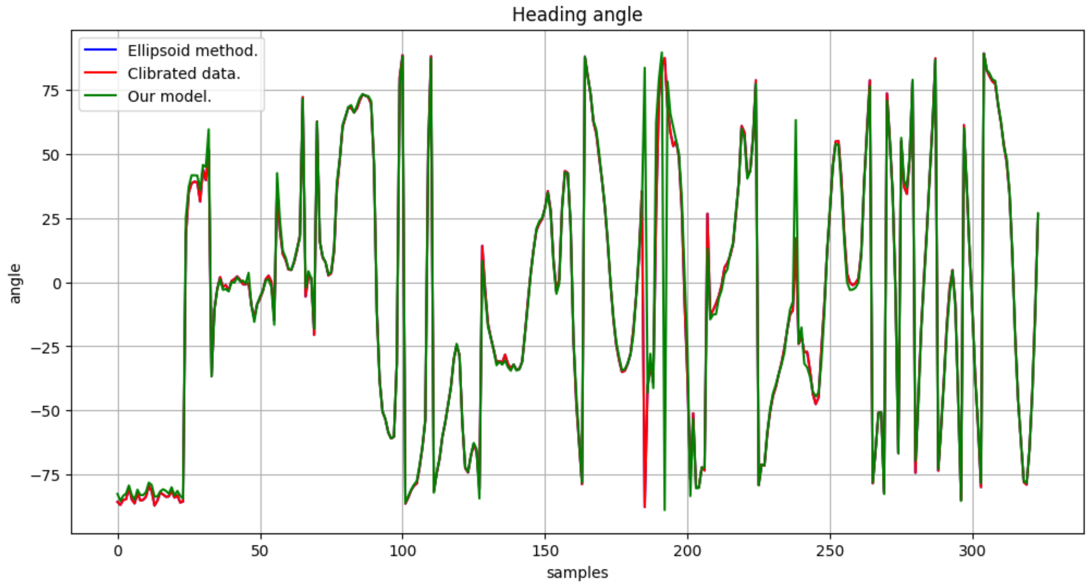

Figure 16: proposed method vs Ellipsoid method

The results showed that the proposed method is efficient as Ellipsoid in calibration raw magnetometer data.

V. Discussion

In general, adaptability of neural networks to different environmental conditions; this model can be used only with the magnetometer sensor that raw data was obtained from it. If indeed a calibration needs to be recalculated, the initial values for weight and biases can be got from previous calibration. Neural networks are capable of capturing non-linear relationships between the measured values and calibrated ones. This is especially beneficial if the calibration process involves complex transformations, linear and nonlinear disturbances, the powerful of nonlinearity of NN can be shown in many factors:

- Using of nonlinear activation functions, such as the rectified linear unit (ReLU).
- Multiple layers allow learning hierarchical representations of data. Each layer has ability to capture different levels of data, thus leads model to represent complex patterns and features in the data.
- NN has ability to handle cases of high complexity data.

If the calibrated data is available, then the calibrations parameters can be calculated and used, but NN have an advantage of more generalizations. Since in normal calibration, the measured data will be applied in (Eq. 1) and get calculated calibrated data; it can be described as static calculation or static calibration and if there is wrong value on specific axe it will affect the calibrated data directly. On another hand, approach of NN when it is applied, the results will go through many calculations (in hidden layers) these will extract features of input with many weights and biases and in case of wrong values on specific axe, this wrong value will be processed by many weights and many biases and this might resolve it (Many weights and biases not only one matrix and only one vector of bias).

VI. Conclusion

In conclusion, this study highlights the importance of magnetometer measurements in applications involving moving bodies. The need for accurate magnetometer data is evident, and this can be achieved through the using of raw data or by handled it by a calibration process. This article addresses the calibration of magnetometer data using a Neural Network (NN) model. The employed NN model follows a sequential architecture with multiple layers. NN model achieves a high accuracy rate of 99.38%. Furthermore, the trained NN model is tested on the entire dataset, as well as on a noisy dataset generated by adding noise to the raw data. The outcomes confirm the high efficiency of the NN model in directly calculating calibrated output from raw measurements. The findings of this study show the ability of NN to provide accurate results in the calibration process.

Availability of data: the data is existed as raw data and true calibration parameters

(https://github.com/michaelwro/mag-cal-example)

Acknowledgment

I would like to express my sincere gratitude to my supervisor, **Serikov Sergey Anatilievich**, for his invaluable guidance and unwavering support throughout the research process. His expertise, insightful feedback, and encouragement have been instrumental in shaping this article. I am truly fortunate to have had the opportunity to work under their mentorship.

References

[1] Bennett, J. S., Vyhnalek, B. E., Greenall, H., Bridge, E. M., Gotardo, F., Forstner, S., ... & Bowen, W. P. (2021). Precision magnetometers for aerospace applications: A review. Sensors, 21(16), 5568.

[2] Kok, M., & Schön, T. B. (2016). Magnetometer calibration using inertial sensors. IEEE Sensors Journal, 16(14), 5679-5689.

[3] Ousaloo, H. S., Sharifi, G. H., Mahdian, J., & Nodeh, M. T. (2017). Complete calibration of three-axis strapdown magnetometer in mounting frame. IEEE Sensors Journal, 17(23), 7886-7893.
[4] Kok, M., Hol, J. D., Schön, T. B., Gustafsson, F., & Luinge, H. (2012, July). Calibration of a magnetometer in combination with inertial sensors. In 2012 15th International Conference on Information Fusion (pp. 787-793). IEEE.
[5] Soken, H. E. (2018). A survey of calibration algorithms for small satellite magnetometers. Measurement, 122, 417-423.
[6] Bader, J., Styp-Rekowski, K., Doehler, L., Becker, S., & Kao, O. (2022, November). Macaw: The machine learning magnetometer calibration workflow. In 2022 IEEE International Conference on Data Mining Workshops (ICDMW) (pp. 1095-1101). IEEE.
[7] Hoang, M. L., & Pietrosanto, A. (2022). Yaw/Heading optimization by Machine learning model based on MEMS magnetometer under harsh conditions. Measurement, 193, 111013.
[8] Gnadt, A. (2022). Machine Learning-Enhanced Magnetic Calibration for Airborne Magnetic Anomaly Navigation. In AIAA SciTech 2022 Forum (p. 1760).
[9] Han, Q., Dou, Z., Tong, X., Peng, X., & Guo, H. (2017). A modified Tolles–Lawson model robust to the errors of the three-axis strapdown magnetometer. IEEE Geoscience and Remote Sensing Letters, 14(3), 334-338.
[10] Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. Heliyon, 4(11).
[11] Fraccaro, M., Sønderby, S. K., Paquet, U., & Winther, O. (2016). Sequential neural models with stochastic layers. Advances in neural information processing systems, 29.
[12] Li, H., Zhang, Z., & Liu, Z. (2017). Application of artificial neural networks for catalysis: a review. Catalysts, 7(10), 306.

Ammar Assad: Postgraduate student at St. Petersburg State University of Aerospace Instrumentation, Department of Electromechanics and Robotics.
Email: ammar.assad225@gmail.com
ORCID: https://orcid.org/0000-0001-9148-6752

Serikov Sergey Anatolievich, Associate Professor, Doctor of Technical Sciences, St. Petersburg State University of Aerospace Instrumentation, Professor, Department of Electromechanics and Robotics,
Email: srkv@inbox.ru
ORCID: https://orcid.org/0000-0002-3668-2459

Nader Al Bitar, Department of Electronic & Mechanical Systems, Higher Institute for Applied Sciences and Technology (HIAST), Damascus, Syria
Email: nader.albitar@hiast.edu.sy
ORCID: https://orcid.org/0000-0002-2505-3219