Flux Gate Magnetometer Data Predictive Correction Using LSTM Recurrent Neural Networks

E. Hurtado [1], C. De La Jara [2], O. Veliz [2], R. Rojas [2]

[1] Universidad Nacional de Ingeniería, Lima, Peru

[1] Radio Observatorio de Jicamarca - Instituto Geofísico del Perú, Lima, Peru

Abstract

Fluxgate magnetometers play an important role in ionospheric research, providing valuable data for understanding Earth's magnetic field variations influenced by processes occurring within the ionosphere, such as geomagnetic storms, solar flares, and ionospheric disturbances. However, these instruments are susceptible to errors induced by magnetic interference; presence of ferromagnetic objects in the vicinity of the sensor; sensor tampering, among others. In this study, we present a novel approach utilizing Long Short-Term Memory (LSTM) recurrent neural networks to predict and correct erroneous data points in magnetometer recordings. Our model uses the temporal dependencies inherent in magnetometer data to accurately identify and rectify anomalies caused by external disturbances. By training on historical datasets with known errors, the model learns to predict the correct magnetic field values corresponding to erroneous readings.

1. Introduction

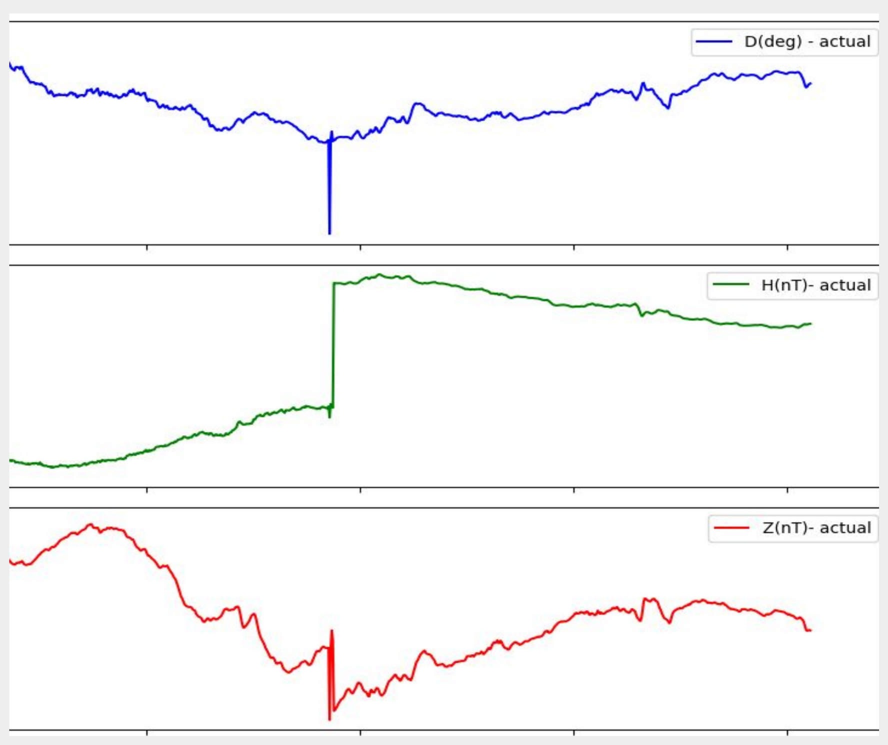


Figure 1. Examples of errors in magnetometer data

Data with errors: Magnetometer logs often present errors. Correcting them is essential to ensure the accuracy, reliability, and consistency of the data, which is fundamental to perform meaningful studies.

Outliers, baseline jumps, and data gaps are some of the most common errors in magnetometer data. These are caused by a variety of reasons like environmental interference, physical impacts, vibration, human activity, sensor or electronics malfunction, power interruptions, among others

Magnetometer data as a time series: Magnetometer readings are captured continuously over time, providing a sequence of measurements. Each reading is dependent on the previous ones, creating a natural order that is crucial for time-series analysis. Each reading is associated with a specific time, creating a chronological record.

PIURA MAGROJ-04 <263> 1 Min. Reported data

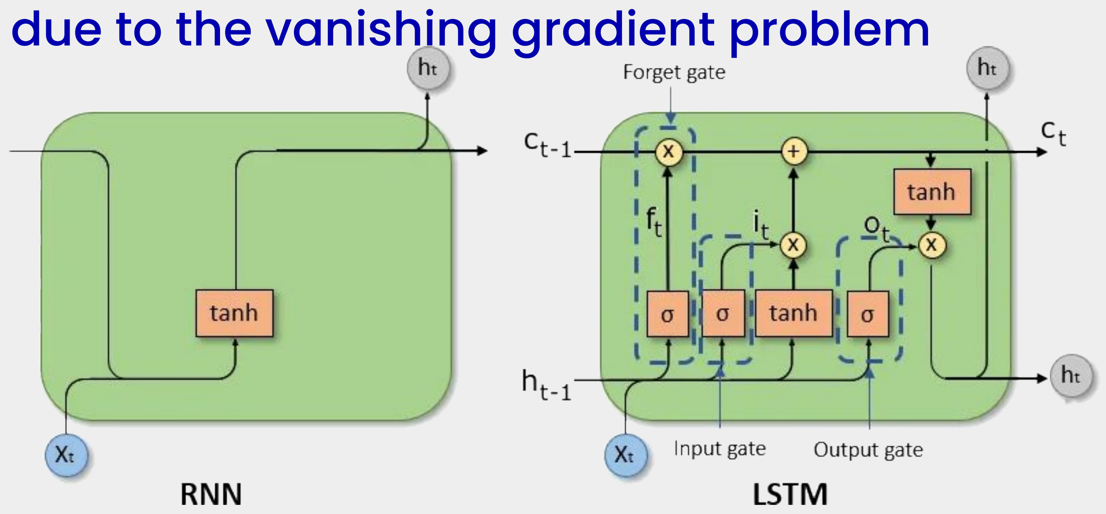
DD	MM	YYYY	hh	mm	D(deg)	H(nT)	Z(nT)	I(deg)	F(nT)
19	09	2016	00	00	-1.3265	26573.9	6007.9	12.7393	27244.6
19	09	2016	00	01	-1.3265	26573.8	6007.8	12.7392	27244.5
19	09	2016	00	02	-1.3265	26574.6	6007.8	12.7388	27245.2
19	09	2016	00	03	-1.3265	26574.5	6007.7	12.7388	27245.1
19	09	2016	00	04	-1.3265	26574.8	6007.7	12.7386	27245.4
19	09	2016	00	05	-1.3265	26572.9	6007.7	12.7395	27243.6
19	09	2016	00	06	-1.3265	26574.0	6007.7	12.7389	27244.6
19	09	2016	00	07	-1.3264	26574.1	6007.7	12.7388	27244.7
19	09	2016	00	08	-1.3265	26573.0	6007.6	12.7392	27243.7

Figure 2. Sequential nature of magnetometer data

2. Recurrent Neural Networks

Recurrent neural networks (RNN), particularly Long Short Term Memory (LSTM) networks, are well suited for predicting and correcting sequential data due to their ability to capture temporal dependencies.

RNNs have loops allowing information to be passed from one step to the next in the sequence. This makes them capable of handling sequential data. For traditional RNNs is difficult to learn long range dependencies due to the vanishing gradient problem



$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f)$$
$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i)$$
$$o_t=\sigma(W_o\cdot[h_{t-1},x_t]+b_o)$$
$$h_t=o_t*\tanh(C_t)$$

Figure 3. RNN and LSTM network models. Forget, input and output gate in LSTMs are used to control the flow of information, effectively managing both short-term and long-term memory

LSTMs are a type of RNN designed to handle the vanishing gradient problem. They have a more complex architecture: input gates, output gates, and forget gates, which regulate the flow of information. LSTMs learn patterns in the sequential data and can remember relevant information from previous time steps and use it to make predictions about future time steps.

3. Data correction with RNN LSTM

An LSTM model was trained using 500 magnetograms from the LISN network database. Mean Squared Error was used as loss function and Adaptive Moment Estimation as optimization function. The network was trained to recognize and correct anomalies by learning the normal patterns and identifying deviations from them.

10 consecutive values are used to predict the next one, if the predicted value is different from the actual value, an anomaly was detected.

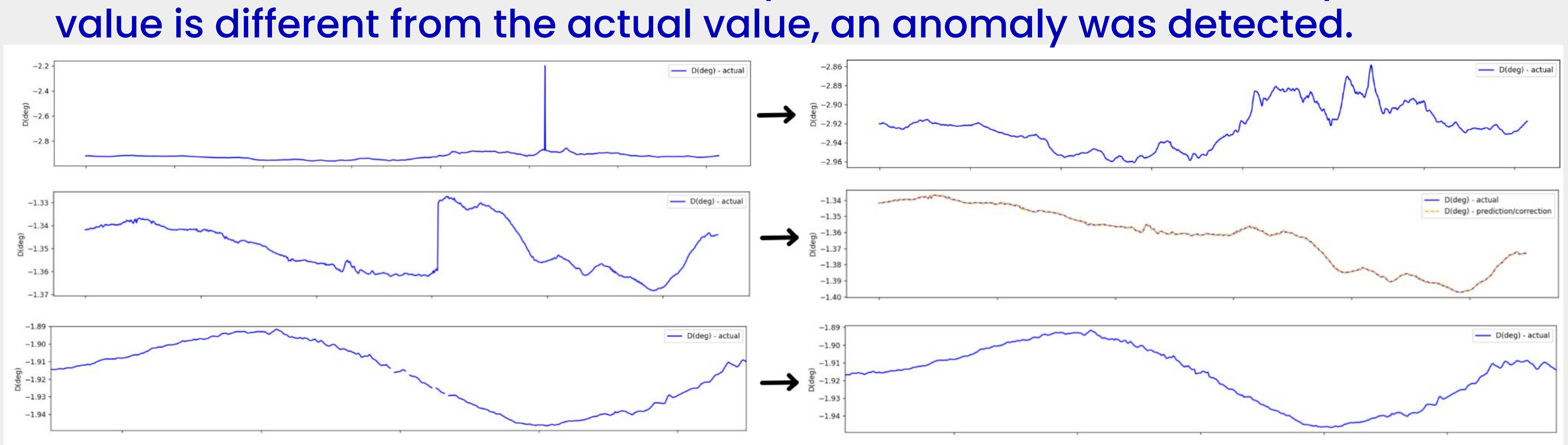


Figure 4. Data with anomalies (left) corrected using the LSTM model

Average difference between predicted and actual data was less than 0.1% indicating the high accuracy of the model.

5. Conclusions

- The LSTM model is capable of detecting and correcting anomalies in magnetometer data.
- A prediction error of less than 0.1% indicates high accuracy and precision of the model, suggesting that it has effectively learned the underlying patterns in the data and is capable of making highly accurate predictions.
- A system to correct magnetometer data in real time is going to be implemented in the LISN server, ensuring that the data is consistently accurate and reliable, and providing more precise insights into geomagnetic activities, contributing to better understanding and forecasting of ionospheric phenomena.

Presented at CEDAR 2024, San Diego, CA, USA

email: cdelajara@igp.gob.pe