EGU2020-10385
https://doi.org/10.5194/egusphere-egu2020-10385
EGU General Assembly 2020


# Using AutoRegressive Integrated Moving Average and Gaussian Processes with LSTM neural networks to predict discrete geomagnetic signals

**Laurentiu Asimopolos**[1], Alexandru Stanciu[2], Natalia-Silvia Asimopolos[1], Bogdan Balea[1], Andreea Dinu[2], and Adrian-Aristide Asimopolos[3]

[1]Geological Institute of Romania, Surlari Geomagnetic Observatory, BUCHAREST, Romania (asimopolos@gmail.com)
[2]National Institute for Research and Development in Informatics BUCHAREST, Romania (alex@ici.ro)
[3]Polytechnic University of Bucharest, Faculty of Transportation, BUCHAREST, Romania (adrian.asimopolos@gmail.com)

In this paper, we present the results obtained for the geomagnetic data acquired at the Surlari Observatory, located about 30 Km North of Bucharest - Romania. The observatory database contains records from the last seven solar cycles, with different sampling rates.

We used AR, MA, ARMA and ARIMA (AutoRegressive Integrated Moving Average) type models for time series forecasting and phenomenological extrapolation. ARIMA model is a generalization of an autoregressive moving average (ARMA) model, fitted to time series data to predict future points in the series

We made spectral analysis using Fourier Transform, that gives us a relevant picture of the frequency spectrum of the signal component, but without locating it in time, while the wavelet analysis provides us with information regarding the time of occurrence of these frequencies.

Wavelet allows local analysis of magnetic field components through variable frequency windows. Windows with longer time intervals allow us to extract low-frequency information, medium-sized intervals of different sizes lead to medium-frequency information extraction, and very narrow windows highlight the high-frequencies or details of the analysed signals.

We extend the study of geomagnetic data analysis and predictive modelling by implementing a Long Short-Term Memory (LSTM) recurrent neural network that is capable of modelling long-term dependencies and is suitable for time series forecasting. This method includes a Gaussian process (GP) model in order to obtain probabilistic forecasts based on the LSTM outputs.

The evaluation of the proposed hybrid model is conducted using the Receiver Operating Characteristic (ROC) Curve that provides a probabilistic forecast of geomagnetic storm events.

In addition, reliability diagrams are provided in order to support the analysis of the probabilistic forecasting models.

The implementation of the solution for predicting certain geomagnetic parameters is implemented in the MATLAB language, using the Toolbox Deep Learning Toolbox, which provides a framework for the design and implementation of deep learning models.

Also, in addition to using the MATLAB environment, the solution can be accessed, modified, or improved in the Jupyter Notebook computing environment.