XGBoost Algorithm for Estimating Equatorial Plasmaspheric Mass Density Using ULF Wave Measurements

Raffaello Foldes, Alfredo Del Corpo, Ermanno Pietropaolo, and Massimo Vellante
University of L’Aquila, Physical and Chemical Sciences, Italy (raffaello.foldes@graduate.univaq.it)

Monitoring the plasmasphere is an important task to achieve in the Space Weather context. A consolidated technique consists of remotely inferring the equatorial plasma mass density in the inner magnetosphere using Field Line Resonance (FLR) frequency estimates derived from Ultra-Low Frequency (ULF) measurements. FLR frequencies can be obtained via cross-phase analysis of magnetic signals recorded from pairs of latitude separated stations. In the last years, machine learning (ML) has been successfully applied in Space Weather, but this is the first attempt to estimate FLR frequencies with these techniques. EXtreme Gradient Boosting (XGB) is a recent ensemble-based algorithm that is resulted in being highly efficient in regression/classification competitions and many real-world applications. Here we employ XGB for identifying FLR frequencies by using measurements of the European quasi-Meridional Magnetometer Array (EMMA). Our algorithm takes as input the 30-min cross-phase spectra of magnetic signals and returns the FLR frequency as output; we evaluated the algorithm performance on four different station pairs from L=2.4 to L=5.5. Results show that XGB algorithm can be a robust and accurate method to achieve this goal. Its performances slightly decrease with increasing latitude and tend to deteriorate during nighttime. However, at high latitudes, the error increases during highly disturbed geomagnetic conditions such as the storm’s main phase. Finally, we compare the equatorial plasmaspheric mass density obtained by XGB estimates with the density profiles by Del Corpo et al. (2019) for a case study, the geomagnetic storm of the 1st June 2013. Our approach may represent a prominent space weather tool included in an automatic monitoring system of the plasmasphere. This work represents only a preliminary step in this direction; applying this technique on a more extensive data set and on more pairs of stations is straightforward and necessary to create more robust and accurate models.