Annales
Geophysicae
Discussions

EGU
Open Access

Interactive comment on "Automatic detection of the Earth Bow Shock and Magnetopause from in-situ data with machine learning" *by* Gautier Nguyen et al.

Gautier Nguyen et al.

gautier.nguyen@lpp.polytechnique.fr

Received and published: 29 November 2019

We believe that there is a big misunderstandings about our answer. Please, let us try to clarify the situation.

The referee seems to think that we train a first model with very few points and *automatically* add the positive predictions to the training set of another, etc. We agree with the referee that this methodology would obviously be wrong, as explained in its comments.

This is **not** what is done.

All labelled points added to the dataset are selected by visual inspection for all

spacecraft. The predictions of intermediate trainings are **just used as a visual suggestion** for labelling the data, to guide the eye. We **manually** build the new training/test set and train a brand new model with our new visually made labels.

The visual guidance the intermediate prediction provides when we label additional data only serves to speed up the data browsing. Indeed, if not perfect, the intermediate predictions are not stupidly wrong either and quite better than random. This prediction is just used to zoom in the data intervals of interest and manually select points we identify as belonging to a given class. Labelling is done also by looking at the spectrograms, as presented in the paper, while this data is not included in the training dataset.

The results presented in the paper are produced with a totally independent, unique, model, trained with the dataset obtained after the whole visual labelling process.

Furthermore, an important part of the paper is dedicated to a **massive prediction on unseen data**. This prediction is presented and is very good. Contrary to many (almost all) studies, we made all codes and predictions available, which can confirm the prediction is good on unseen data. This could not be the case with an algorithm that would be based on only 6 hours of data and successively automatically trained on raw positive inputs. We will clarify this explanation in the revised paper.