Automatic detection of the electron density from the WHISPER instrument onboard CLUSTER II

E. De Leon [(1)], N. Gilet [(1)] , X. Vallières [(1)], L. Bucciantini [(1)] , P. Henri [(1)] , J.L. Rauch [(1)],

[(1)] CNRS-LPC2E, Orléans, France, +33 (0)2 38 25 78 25, emmanuel.de-leon@cnrs-orleans.fr

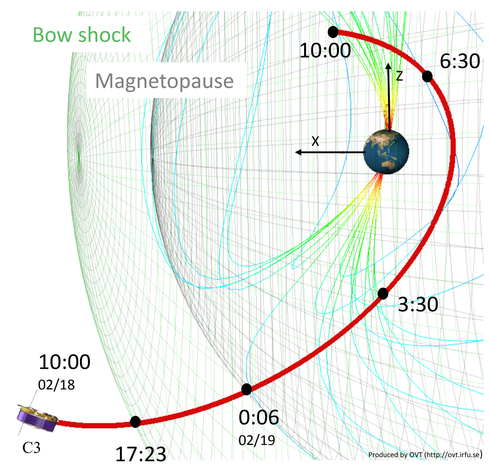
CONTEXT

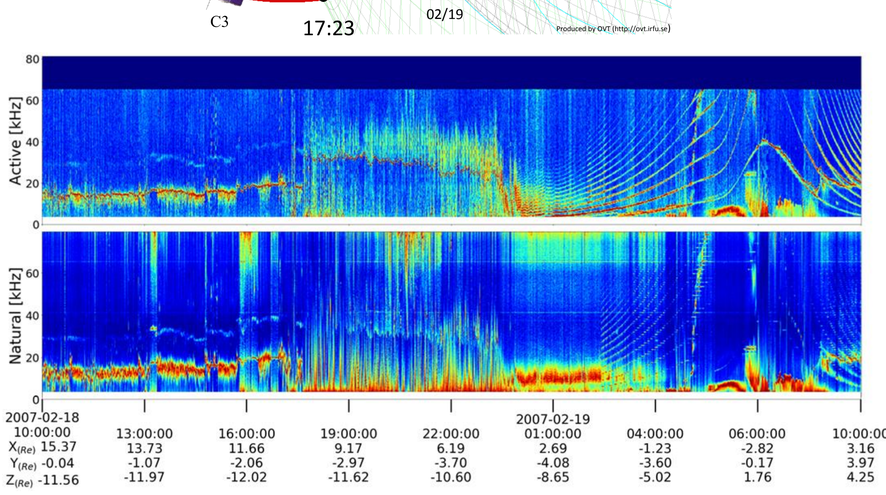
The ESA CLUSTER mission has been successfully monitoring the plasma bulk properties of the Earth magnetosphere for more than 20 years. The payload is identical for each of the four spacecraft, consisting of 11 experiments dedicated to the measurements of the surrounding plasma. As part of the Wave Experiment Consortium, the WHISPER instrument is a relaxation sounder that monitors the in situ space plasma bulk properties such as the thermal electron density. It can be operated in two modes:

- Active (sounding) mode : an active spectrum is built on-board after local excitation of the surrounding plasma over the 3.5–82 kHz frequency range.
- Natural mode: the transmitter is switched off and the instrument becomes a natural wave receiver delivering spectrograms over the 2–80 kHz frequency range. The plasma bulk properties can be extracted, under certain conditions, from the active and natural electric field spectrogram.

In particular, the thermal electron density, one of the key datasets provided by the WHISPER experiment, is used for scientific studies and as the main driver of the calibration of particle instruments and is routinely delivered to the Cluster Science Archive.

The extraction of the electron density, directly related to the frequency of plasma signatures (Fpe), requires a careful analysis of active and/or natural spectra and can be difficult in some plasma regions. Indeed, the location of plasma frequency in the electric field spectrograms strongly depends on the plasma conditions such as the magnetic field or spacecraft position. The thermal electron density has been extracted by various ad-hoc algorithms developed throughout the mission lifetime, from fully manual to semi-automatic, but always requiring time-consuming manual intervention.





OBJECTIVES

In order to improve data analysis, we developed a tool to an automatically analyze the WHISPER spectra without need of manual inputs. Our final objective is a fully automatic electron density production on a subset of magnetospheric regions, i.e., solarwind (SW) and magnetosheath (MS), in order to deliver data faster, reduce manpower and subjectivity due to manual intervention. For this purpose, we have implemented automatic routines based on neural networks. The main constraint is to be able to determine the plasma frequency only with the WHISPER instrument data. The priority is to derive electron density in active mode.

1. The automatic electron density determination in active mode, performed in two steps :

- Identify the magnetosphere regions with neuronal networks classification (ANN, Multi-Layer Pecreptron)
- Determine the electron density with a Neuronal Network adapted to the region (GRU, Recurent Neuronal Network)

2. The developement of operational pipeline using these methods to reduce human intervention

3. Improve and extend the thermal electron density determination coverage by applying the neuronal networks only to the natural spectra.

1. Automatic electron density determination with Neuronal Networks

Several model architectures have been tested. Inputs for training are Active and Natural normalized spectra (in dB) from a substantial set of data, randomly chosen during the mission lifetime, with :

- Corresponding region labels derived manually and with other data products from CLUSTER
- Electron density values, obtained from ad-hoc and manual determination.

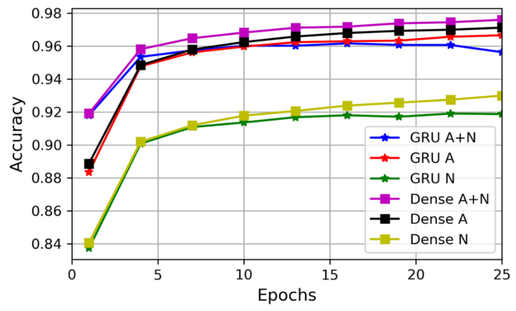
- **Region Classification**

  The output for region determination is the plasma region (MS, SW or Other).
  The models were trained with ~400K spectra and tested on ~30K spectra.
  Best accuracy is obtained with a 4 Layer Multi-Layer Perceptron with up to **97.5**% accuracy

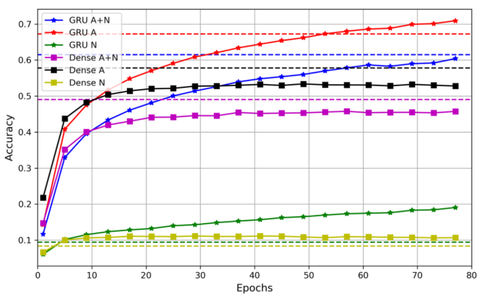
- **Electron density determination from Active spectra**

  The output of these models is the Fpe index inside the spectrogram.
  The models were trained with ~150k spectra and tested on ~16K spectra.
  Best accuracy is obtained with a model is a GRU (recurrent neuronal network) , it has an accuracy going up to **97%** for the MS and SW region, with 1 frequency bin (+/- 0.165kHz) bin tolerance.

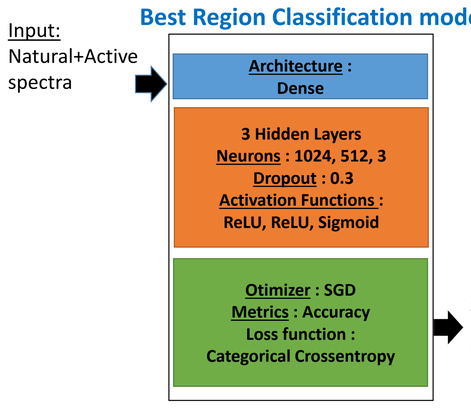
Region Classification models accuracy



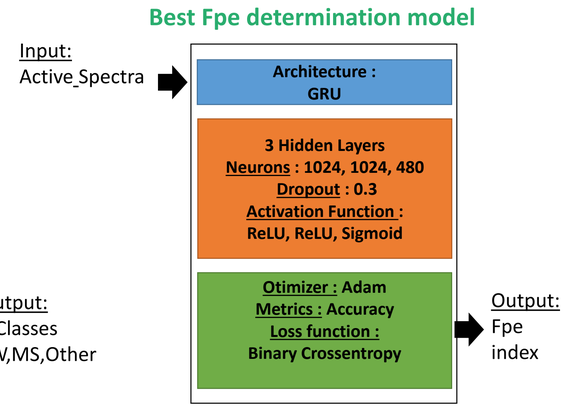
Fpe determination models accuracy



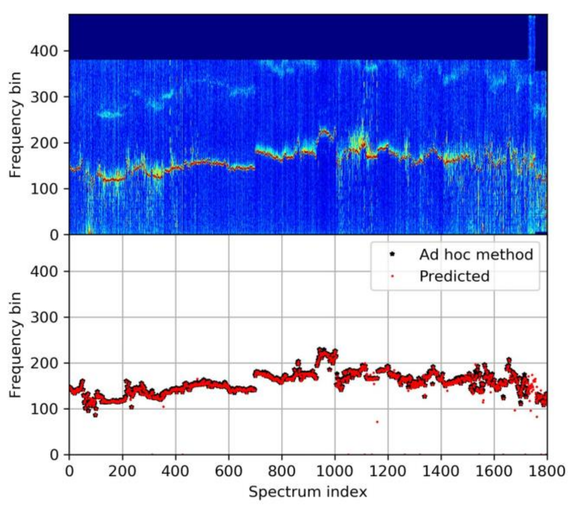
Best Region Classification model



Best Fpe determination model



Region Classification and Fpe Determination in the Solar Wind



2. Operational Pipeline : JedAI

The « Judging Electron Density by Artificial Intelligence » pipeline allows to plot, visualize and validate the model outputs. It is developed with the main objective of reducing the tedious and time-consuming electron density manual determination or ad-hoc methods selection. The pipeline is operational in the solar wind and the magnetosheath regions. It has been designed as a modular tool capable of adding new models for other regions and updating current models.

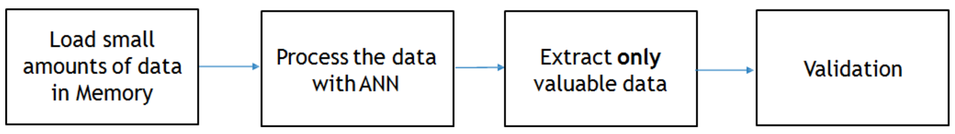
The pipeline works on successive time intervals based on the region classification output, where severe constraints have been imposed to ensure data quality and avoid ambiguous detection :

- A Thresohd on the predicted region probability
- A Threshold density estimate probability
- Time continuity

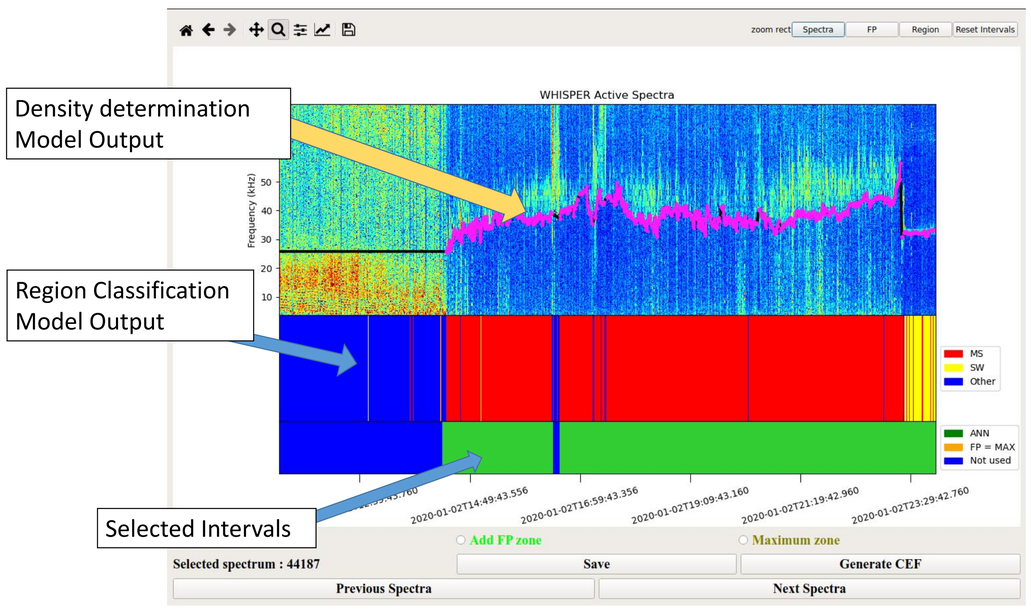
This can lead up to a maximum of 25% loss in the number of electron densities given by neural networks as compared to ad hoc methods. On the other hand, the processing of one month of data for the four spacecraft now requires approximatively 4 h of manual validation whereas the ad hoc methods required around 40 h. The output of JedAI is the electron density which is ready to deliver to the Cluster Science Archive (CSA) as a high level WHISPER data product.

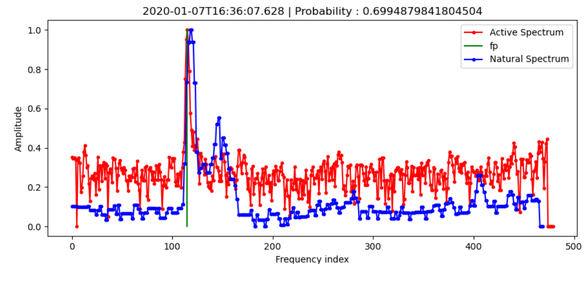
3. Electron density determination from Natural spectra (work in progress)

A specific dense neuronal network model is created for each region using only natural spectra as input.The output of these models is a subset of spectra corresponding to the signature of the selected plasma region. As the data volume for natural spectra is much higher than in active mode, the plasma frequency is extracted only from the subset by a full automatic ad-hoc algorithm. Each model is trained with ~600K spectra and tested on ~100K spectra. The accuracy scores are promising in some regions (up to ~**97**% in MS and SW, ~**80**% in Cusp/Tail and ~**70**% in Plasmasphere).



Fp JedAI validation interface





CONCLUSION AND FUTURE WORK

- We have developed a new density determination pipeline which is capable of reducing human intervention, now limited to validation, up to 10 times compared to the previous methods. This pipeline is now operational and is routinely used to produce the Whisper densities delivered to CSA.
- Current work is focused on improving the naural spectra analysis and on improving usability, ergonomics and validation process of the pipeline.
- The long term goal is to study if the automatic determination methods could be applied for future data processing of electric field experiments on-board BepiColombo (Hermean environment) and JUICE (Jovian environment). Indeed, this study is a proof of concept for an operational automatic analysis of massive datasets potentially provided by future missions or constellations.