SM31D-3532: Recognizing Patterns in EMIC Waves Spectrograms Using Machine Learning

AGU 100 | FALL MEETING Washington, D.C. | 10-14 Dec 2018

C. Medeiros[1], V. M. Souza[1], L. E. A. Vieira[1], D. G. Sibeck[2], L. R. Alves[1], L. A. Da Silva[1] , J. P. Marchezi[1,2], P. R. Jauer[1], M. Rockenbach[1], A. Dal Lago[1], O. Mendes[1] and C. Kletzing[3]

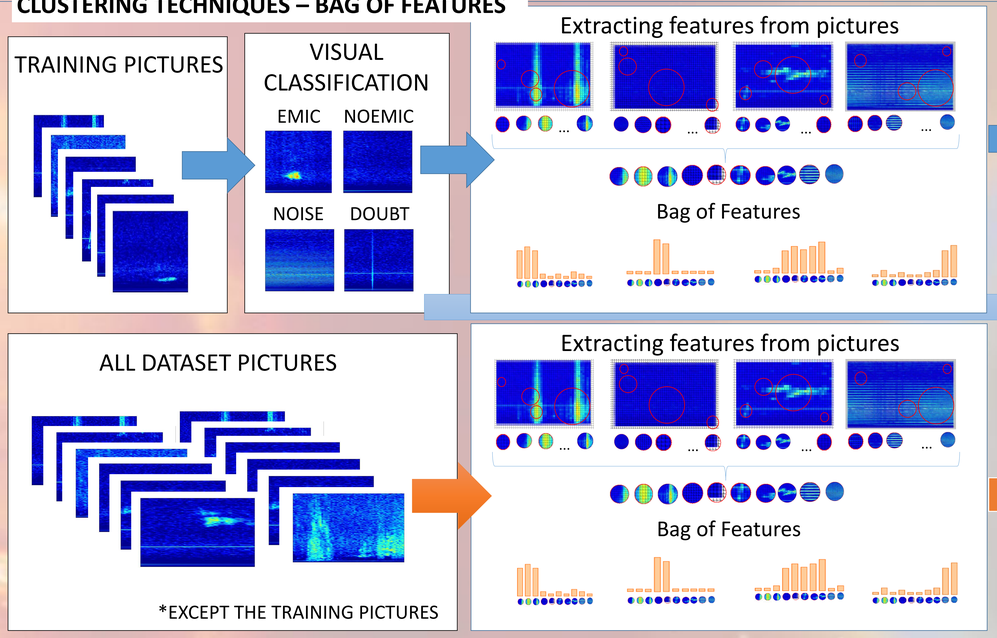
[1]National Institute for Space Research – INPE São José dos Campos, São Paulo, Brazil, [2]Goddard Space Flight Center – NASA, Greenbelt, Maryland, USA, [3] Department of Physics & Astronomy University of Iowa, Iowa City, USA

claudia.medeiros@inpe.br

ABSTRACT

The EMFISIS instrument on board the twin Van Allen Probes has been measuring high resolution magnetic field data during the last 6 years. Spectrograms can be obtained from such measurements which can evidence the presence of electromagnetic ion cyclotron (EMIC) waves. Several studies have been showing the relevance of EMIC waves on the pitch angle scattering of energetic particles into the loss cone, hence contributing to the net loss of relativistic electrons from the outer radiation belt to the upper atmosphere. The huge amount of data collected thus far provides us with the opportunity to use a data clustering technique based on a neural network referred to as Bag-of-Features (BoF). When applied to images of magnetic field spectrograms in the frequency range of EMIC waves, the BoF allows us to distinguish several patterns in these spectrograms expediting their analysis which in turn can be relevant to describe physical aspects of EMIC waves. Specifically, the BoF technique is able to detect and classify distinct patterns in the spectrograms which are identified as signatures of EMIC wave packets. Each spectrogram image provided as input to the BoF corresponds to the windowed Fourier transform of an one hour interval of EMFISIS's magnetic field data. Our dataset spans the September 2012 to December 2016 period, where only in situ data acquired at geocentric distances larger than or equal to 3 Earth radii were selected. Preliminary results revealed that the clustering technique employed here successfully detected, in an automated way, different EMIC wave's signatures like propagation in Hydrogen, Helium and Oxygen bands, as well as, fairly monochromatic mode waves, and further details that will be discussed.

CLUSTERING TECHNIQUES – BAG OF FEATURES



PROCESSING DATASET

- High resolution (64 Hz sampling rate) magnetic field data provided by the EMFISIS instrument on board both Van Allen Probe were selected for the interval between September 2012 to December 2016.
- Data acquired above 3 Earth radii were binned according to the figure below and their spectrograms were calculated using FFT transforms and printed as pictures like the ones shown in Figure 2. It provided us with 66,665 pictures.

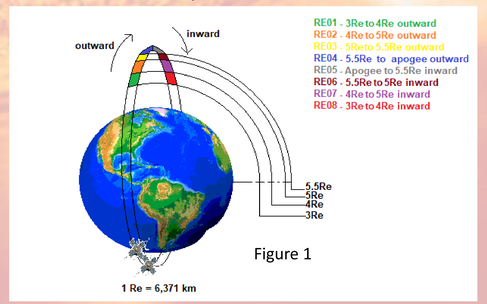


Figure 1

- All pictures from 2012 were visually classified into 4 classes, namely:

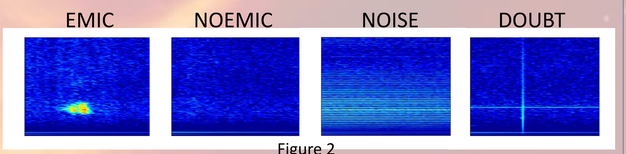


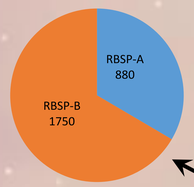
Figure 2

- To train the network we used 248 (from 2012) pictures (62 per class).
- After training the network with pre-classified features we obtain the Bag-of-Features which will be used as parameters to classify our dataset
- All 66,417 images from both Van Allen Probes containing the magnetic field spectrograms calculated, were used as inputs to be classified.
- The classification process takes nearly 5,000 sec to classify this huge amount of data using MATLAB R2017 with Imaging Processing, Statistics and Machine Learning Toolboxes without using Parallel Computing Toolbox™ which could accelerate this execution.
- The classification algorithm distributed all input images according to their "scores" , and then assigned each image to one of the four classes mentioned above (EMIC, NOEMIC, NOISE, DOUBT).
- We can then obtain the statistics about EMIC wave occurrence from the subset of all input images that fell within the "EMIC" class.
- For the current study, we keep some informations about the wave's date of occurrence on the filename and we can track the original data from it.

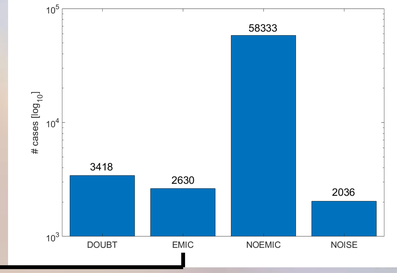
FIRST RESULTS

Accuracy results for the training dataset:

EMIC	NOEMIC	NOISE	DOUBT
92%	82%	98%	78%



BoF classification for the whole dataset excluding training dataset



ONGOING SITUATION

- Analyze all EMIC wave events and their properties
- Correlate all events with the interplanetary medium and local radiation belts properties
- Verify the classification algorithm's accuracy for the whole dataset
- Compare classification results with different techniques

SOME CONSIDERATIONS

- Care is needed when defining EMIC waves criteria so to avoid misclassification
- **It is only a data mining process.**

FUTURE WORKS

- Distinguish different EMIC patterns automatically.
- Apply the same methodology to different waves.
- Improve classification algorithm's execution time.
- Add new parameters to the classification process.

REFERENCES

- Wang, D., et al. (2015), Statistical characteristics of EMIC waves: Van Allen Probe observations, J. Geophys. Res. Space Physics, 120, 4400–4408, doi:10.1002/2015JA021089.
- https://www.mathworks.com/help/vision/examples/image-category-classification-using-bag-of-features.html

ACKNOWLEDGEMENTS

We acknowledge the NASA Van Allen Probes and Craig Kletzing (University of Iowa) for use of data. All RBSP-ECT data are publicly available at the Web site http://www.RBSP-ect.lanl.gov/.