WHISTLER WAVES DETECTION USING MACHINE LEARNING

Othniel Konan (kylekojey@gmail.com) & Dr Amit Mishra (amit.mishra@uct.ac.za), Department of Electrical Engineering, University of Cape Town
Dr Stefan Lotz (slotz@sansa.org.za), Science Research and Application, South African National Space Agency

BACKGROUND

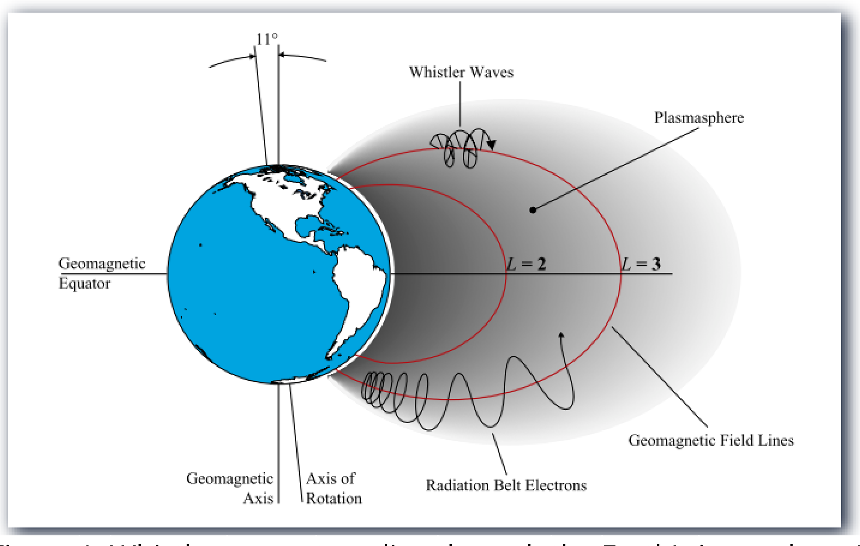


Figure 1: Whistler waves traveling through the Earth's ionosphere [1].

The propagation time delay of these waves is dependent on the plasma density along the propagation path. **This enables the use of whistler wave observations for characterising the plasmasphere in terms of particle number and energy density.**

DATA

The data provided for the research is collected and labeled using the Automatic Whistler Detector (AWD) designed by Janos L. et. al. [2].

Each sample has a length of 4 to 12 seconds depending of the number of whistler detected and the time interval between them.

For this research, **2195 labeled samples collected at Marion island in 2013** are used. The samples are partitioned into **1471 training samples** and **725 testing samples.**

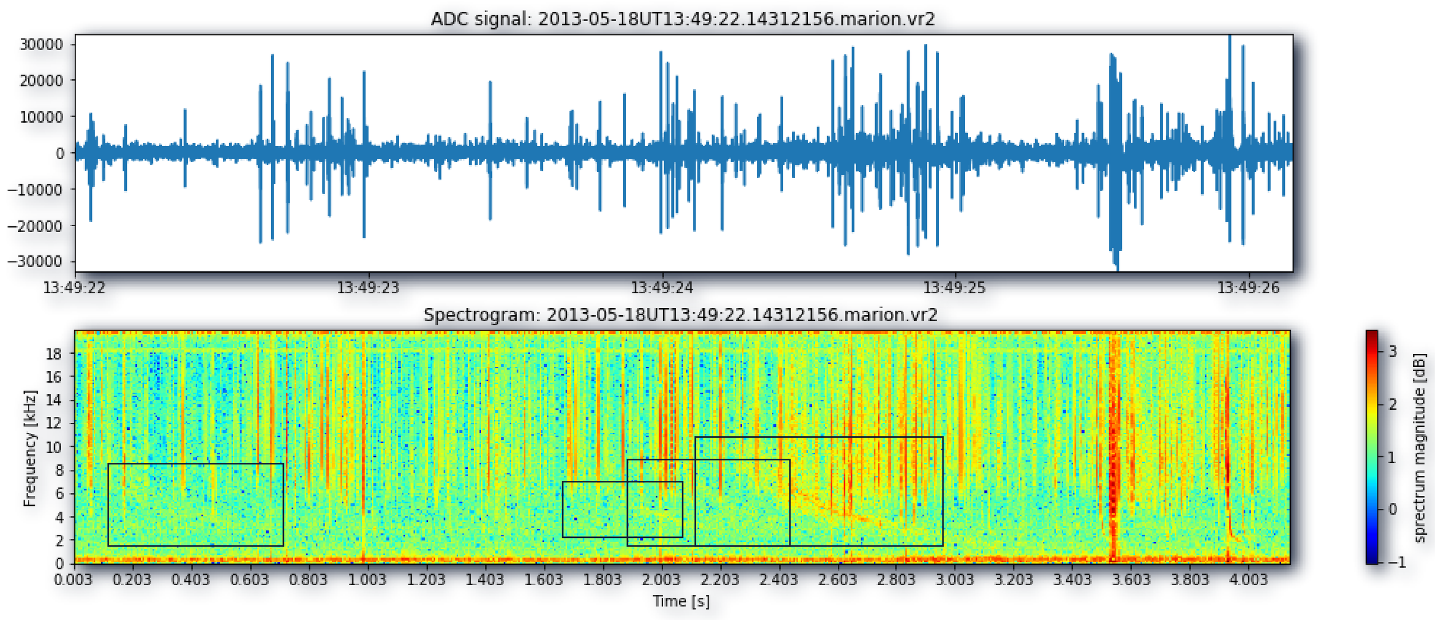


Figure 2: (a) Raw data collected by a Automatic Whistler Detector (AWD) [2] node at Marion island showing strong presence of spherics of magnitude above 20000. (b) Time-frequency representation of the data showing the 5 kHz positions of the whistlers at 0.3, 2, 2.2, and 2.6 s.

PRE-PROCESSING

Z-score Transform

$X_z = Z\{X\} = (X-\mu)/\sigma$

$\mu_{Xz} = 0, \ \sigma_{Xz} = 1$

Apply on each axis:

$z_{ij} = Z_i\{ Z_j \{x_{ij}\} \}$

Figure 3: Whistler (seen at 5 kHz) and spheric magnitude respectively located at 2.5 s and 2 s. (a) Whistler is -0.5 dB below the spheric (b) After pre-processing, the whistler is +1.5 dB above the spheric.

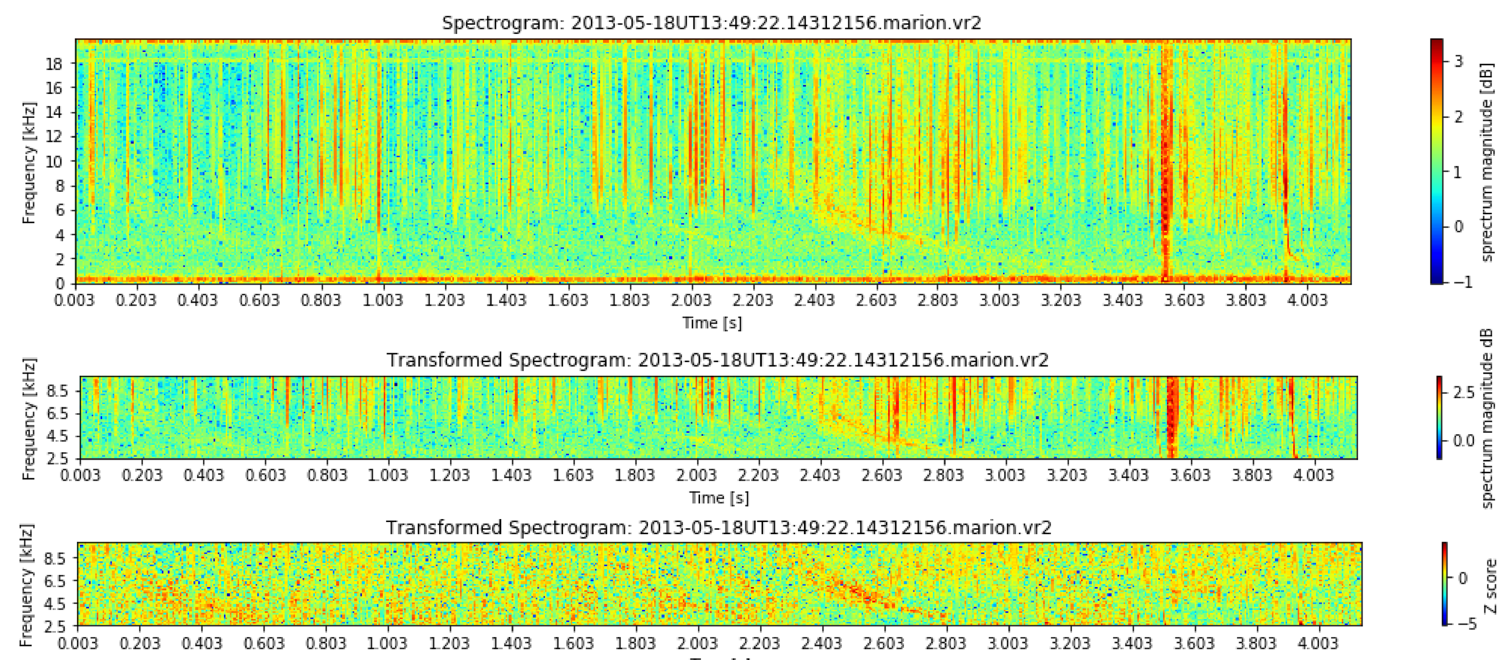
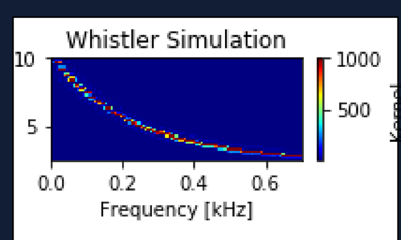


Figure 4: (a) Unprocessed spectrogram. (b) Trimmed spectrogram to remove part of the signal which do not contain the whistlers. (c) Application of the Z-score transform to reduce the noise level.

METHOD

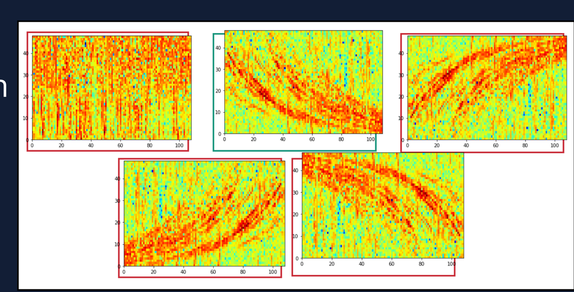
Correlation with Simulated Whistler and Adaptive Threshold (MF)



Whistlers represented as **exponential decay of 0.7s** from **2.5-10 kHz**

Detection using a Convolutional Neural Network (ML)

We use the output provided by AWD to create **cuts of 0.7 s from 2.5-10 kHz range**. We train a CNN with a total of **13417 samples** (using data augmentation) with **20% for cross-validation**. The CNN is made of **4 2-D convolutional layers, a Max pooling layer, 3 Dense layers and few Dropout layers**. The classifier obtained has a **accuracy of 95%** and a **loss of 0.1** after the 25th epoch.



RESULTS

PERFORMANCE ON THIS SAMPLE

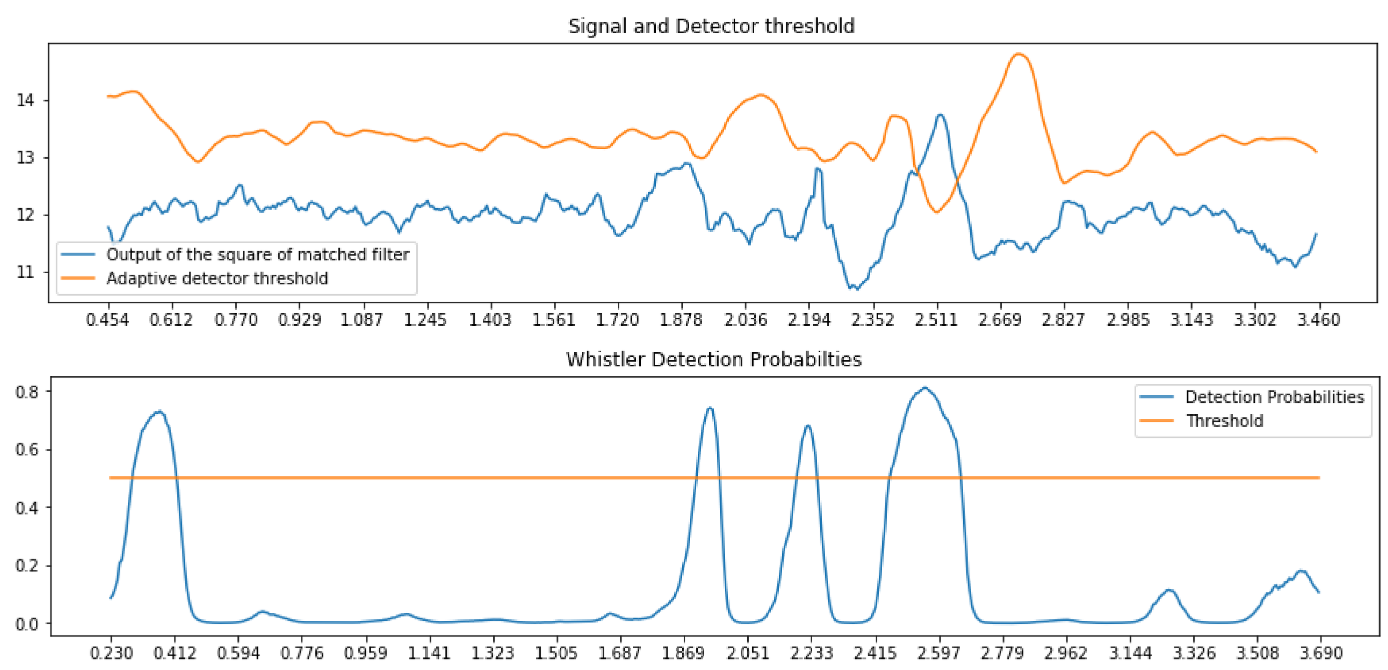


Figure 5: (a) Detection level with adaptive threshold for Correlation method.
(b) Probabilities of whistler location with a 50% threshold.

- AWD detection: 2 whistlers at **2.43 s and 2.56 s**
- MF detection: 1 whistler at **2.5 s**
- ML detection: 4 whistlers at **0.35 s, 1.87 s, 2.21 s, 2.56 s**

OVERAL PERFORMANCE

Table 1: Overall performance of the matched filtering (MF) method and the machine learning (ML) method on the data.

	TRAINING DATA		TEST DATA	
	MF	ML	MF	ML
PRECISION	79.9%	69.7%	80.7%	70.5%
RECALL	72.9%	82.5%	71.5%	81.2%
F1-SCORE	76.2%	75.6%	75.8%	75.5%
FALSE ALARM RATE	20.1%	30.3%	19.3%	29.5%
MISDETECTION	27.1%	17.5%	28.5%	18.8%

- **AWD** has a **misdetection rate of 10%** and a **false alarm rate of 20-50%** [2]. Since AWD is used as the ground truth, its error propagates through metric of evaluation of the matched filtering method and the machine learning model.
- **MF detection** has a **misdetection rate of 28.5%** and a **false alarm rate of 20.1%**.
- **ML detection** has a **misdetection rate of 18.8%** and a **false alarm rate of 29.5%**. The false alarm are questionable since the AWD missed some events.

CONCLUSION & CHALLENGES

- The data available for the research is limited. Only 2195 samples, that is 3 hours of collected data are used to create the model above.
- The ground truth is not available to us, as a result, the labels provided by the AWD algorithm is used as a ground truth.
- The Matched Filtering approach using an adaptive threshold has many tweaking parameters and finding an optimal set is difficult.
- The Machine learning model perform well in practice and can detect whistler difficult to see with the human eye and not detected by the AWD algorithm.

ACKNOWLEDGEMENTS

- The Mandela-Rhodes Foundation
- ML-Helio

REFERENCES

[1] W. C. Burgess, "Lighting-induced coupling if the radiation belts to geomagnetically conjugate ionosphere ...", May 1993
[2] J. Lichtenberger et. al. "Automatic Whistler Detector and Analyzer system: ..." Journal of Geophysical Research: Space Physics, 2008