



Clustering and data reduction algorithm applied to spectra measured with a multi-static HF sounding system in Peru

V. Vásquez¹, M. Milla², K. Kuyeng¹, J. Gonzales³ and D. Scipión¹.

¹Jicamarca Radio Observatory, Instituto Geofísico del Perú, Lima, Perú.

²Sección Electricidad y Electrónica, Pontificia Universidad Católica del Perú, Lima, Perú.

³Facultad de Ingeniería Electrónica y Eléctrica, Universidad Nacional Mayor de San Marcos, Lima, Perú.

Abstract

The HF radar is a network of HF radio beacons and receivers for ionospheric sounding, and it has been operating in Peru since 2016. The purpose of this instrument is to measure the group delay, Doppler shift, power, and other parameters in order to estimate the regional plasma density as a function of space and time. This information is crucial for forecasting the occurrence of Spread-F. To improve the quality of the spectral data, two changes were made. The first one involved transmitting frequencies separated by 3.3 Hz in each transmission station, which allowed to spectrally separate and identify the signals coming from a given station, displacing the cross-talk in frequency but not eliminating it. Therefore, the second change was made, which corresponds to the development of an algorithm that extracts only the signals of interest from the measured spectrum. It is capable of detecting clusters of data in the spectra classifying them as coherent echoes, while noise sectors are discarded. The algorithm procedure and the comparisons of the spectra and final data are shown in this work.

1. Introduction

The HF radar is a multi-static continuous wave radar, and it was composed of three transmitting and six receiving stations, deployed around the center coast and the Andes region of Peru, operated continuously until December 2022. Currently, the radar maintains its three transmitting stations but only two receiving stations.

Since 2016, the three transmitters stations used the same two frequencies: 2.72 MHz and 3.64 MHz modulated with a different pseudorandom code, however, interference of the other transmitters were observed in the obtained spectrum, affecting the signals closer to DC as shown in Figure 2. To solve this problem, multi-frequency transmission separated by 3.3 Hz was implemented in April 2022. And then a clustering algorithm was developed to classify coherent echoes and obtain noise-free data.

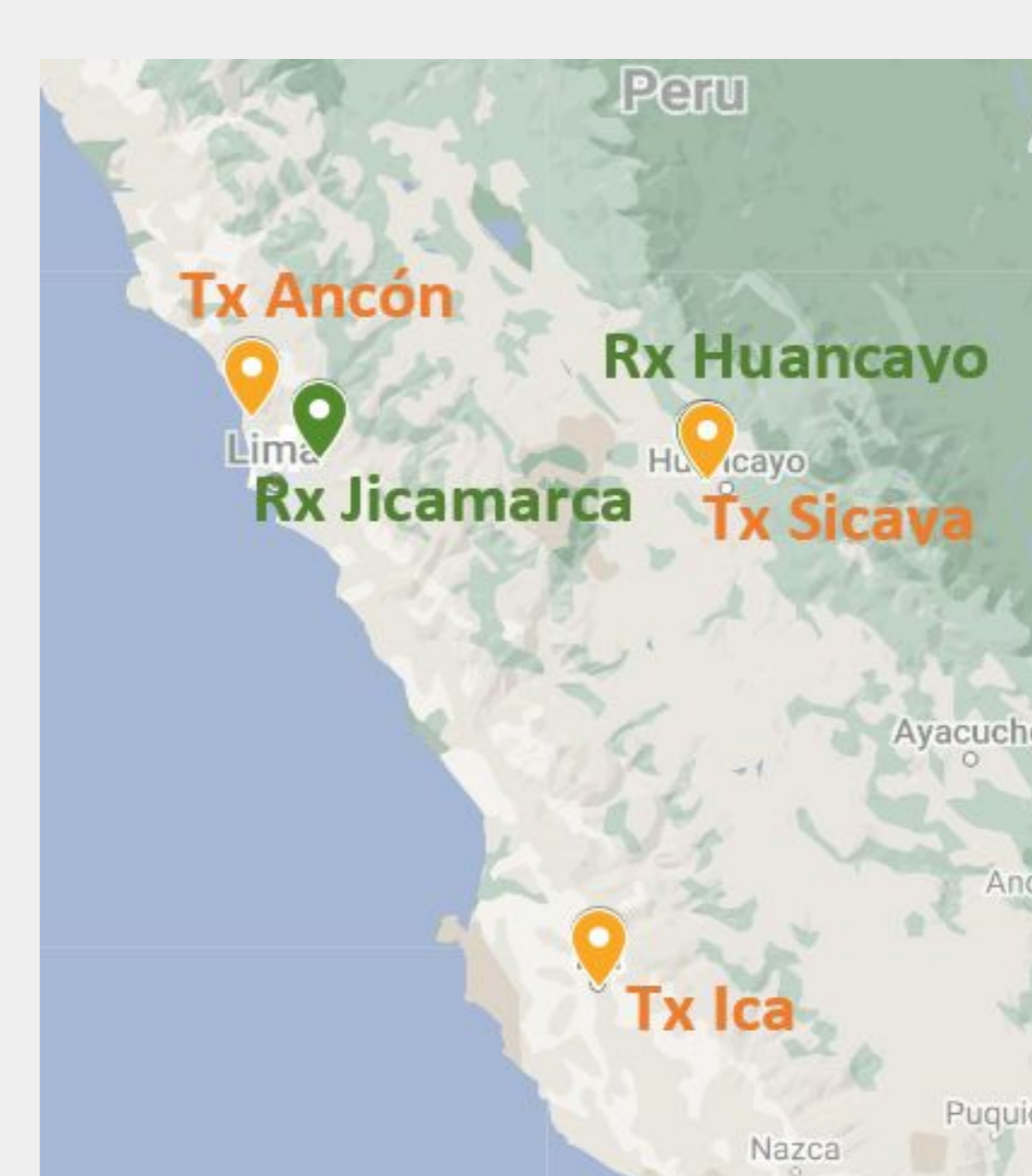


Figure 1. Map of HF beacon stations in Peru.

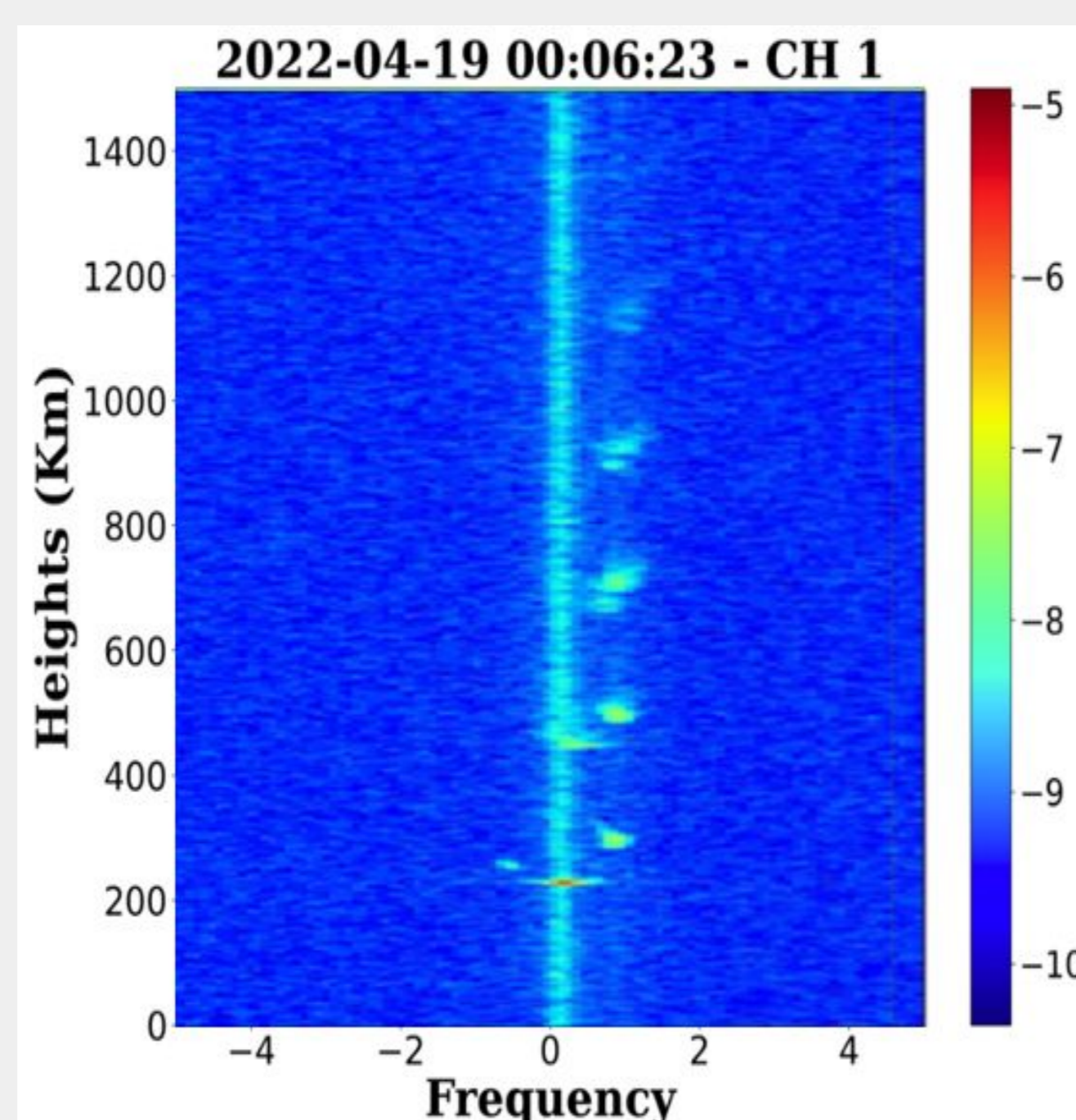


Figure 2. Spectrum Rx: JRO Tx: Ica.

2. Multi-frequencies system in HF radar.

The change in the transmission system was made for the sending of multi-frequency modulated by 3.3 Hz in each transmitter. The frequencies used can be seen in Table 1.

Transmitting stations	2.72 MHz Range	3.64 MHz Range
Ancón (code 0)	2722167.9687500	3649902.3437500
Sicaya (code 1)	2722164.639271800	3649905.673228200
Ica (code 2)	2722171.298228200	3649899.014271800

Table 1. HF Radar Frequencies (Hz).

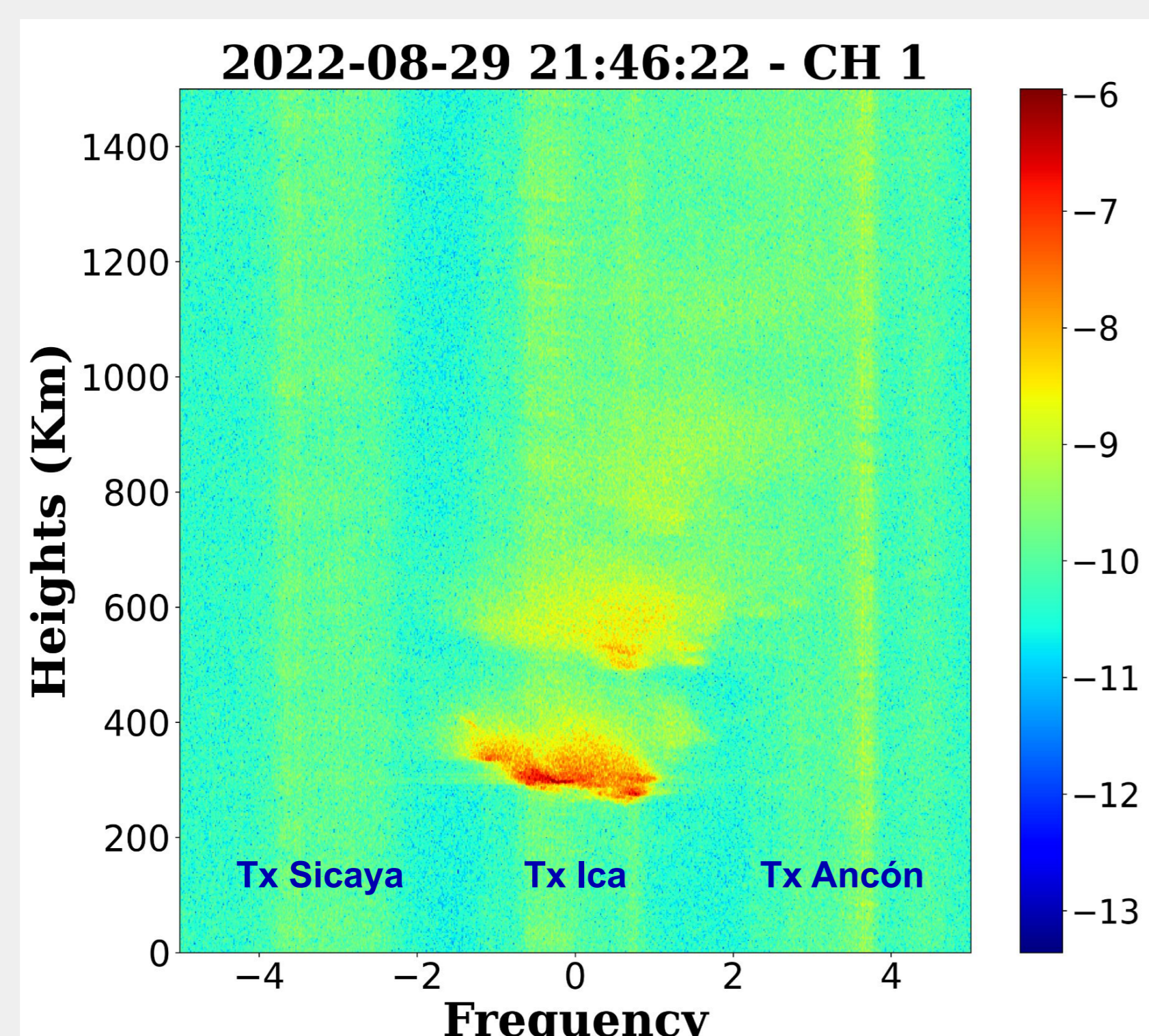


Figure 3. Spectrum Rx: JRO Tx: Ica. Freq: 3.64 MHz.

3. Clustering and Data Reduction algorithm

This algorithm is capable of detecting clusters of data in the spectra, classifying them as the coherent echoes of interest, while tagging the remaining sectors as "noise" that will be discarded.

These are the following steps of the algorithm:

a. A median filter was applied to effectively reduce the interferences as you can see in Figure 4.a..

b. A percentile value is selected to identify regions with the highest spectral power levels, excluding areas without any signals.

c. The DBSCAN algorithm is applied to identify clusters based on density and separate them from noisy areas, enabling the extraction of the desired signal coordinates.

d. A frequency sweep is conducted to obtain the minimum power to replace a fixed noise value, which is used to replace the discarded areas considered as noise. Figure 5.b.

Finally, the coordinates of the filtered signals and the fixed noise value are stored. As a result, interference-free spectrums are obtained and the storage size of each spectrum is significantly reduced.

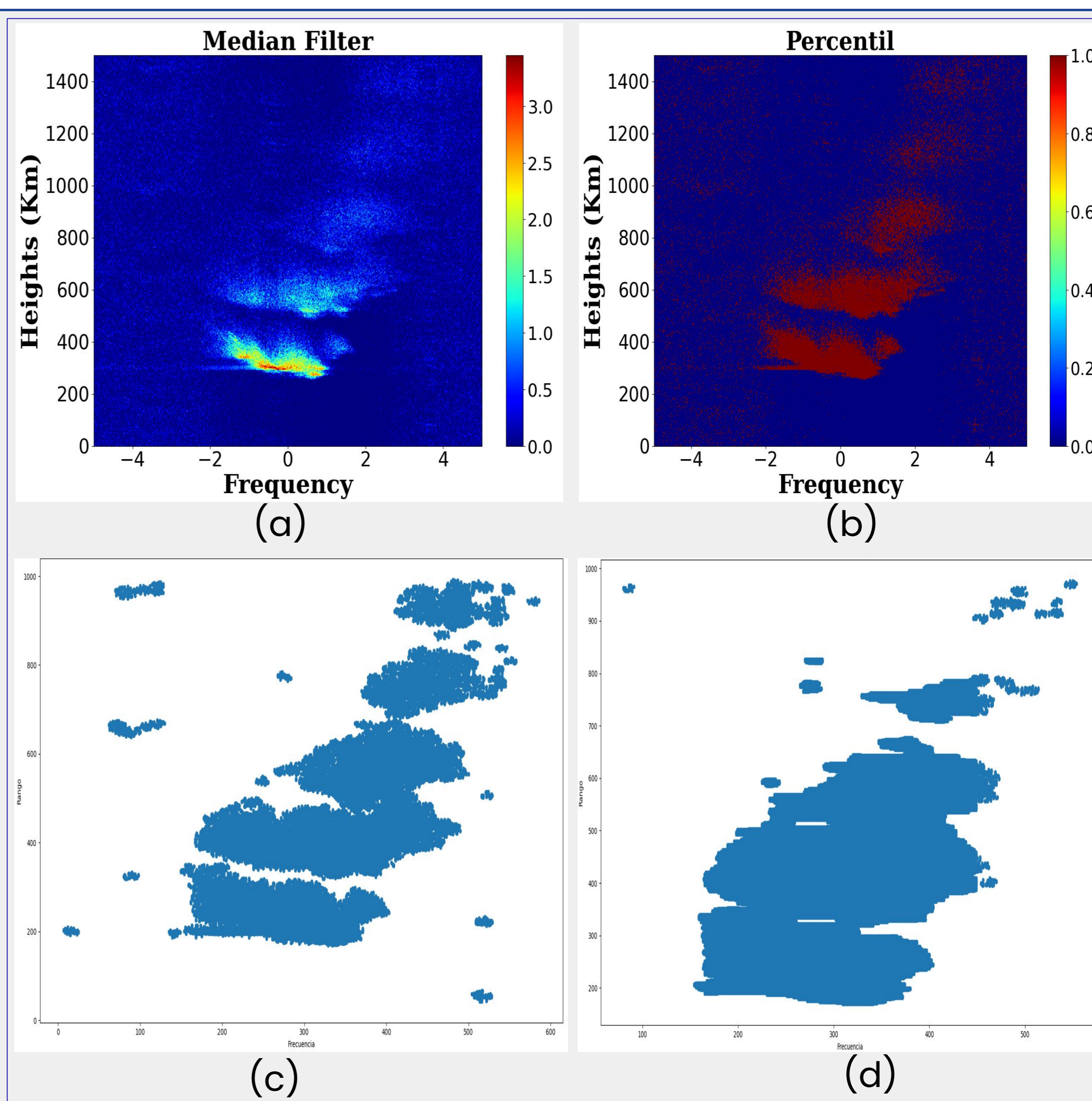


Figure 4. Procedure of Clustering and Reduction algorithm. (a) Median Filter applied to original spectrum. (b) Percentile of highest values of spectral power. (c) DBSCAN Algorithm. (d) Data Fill.

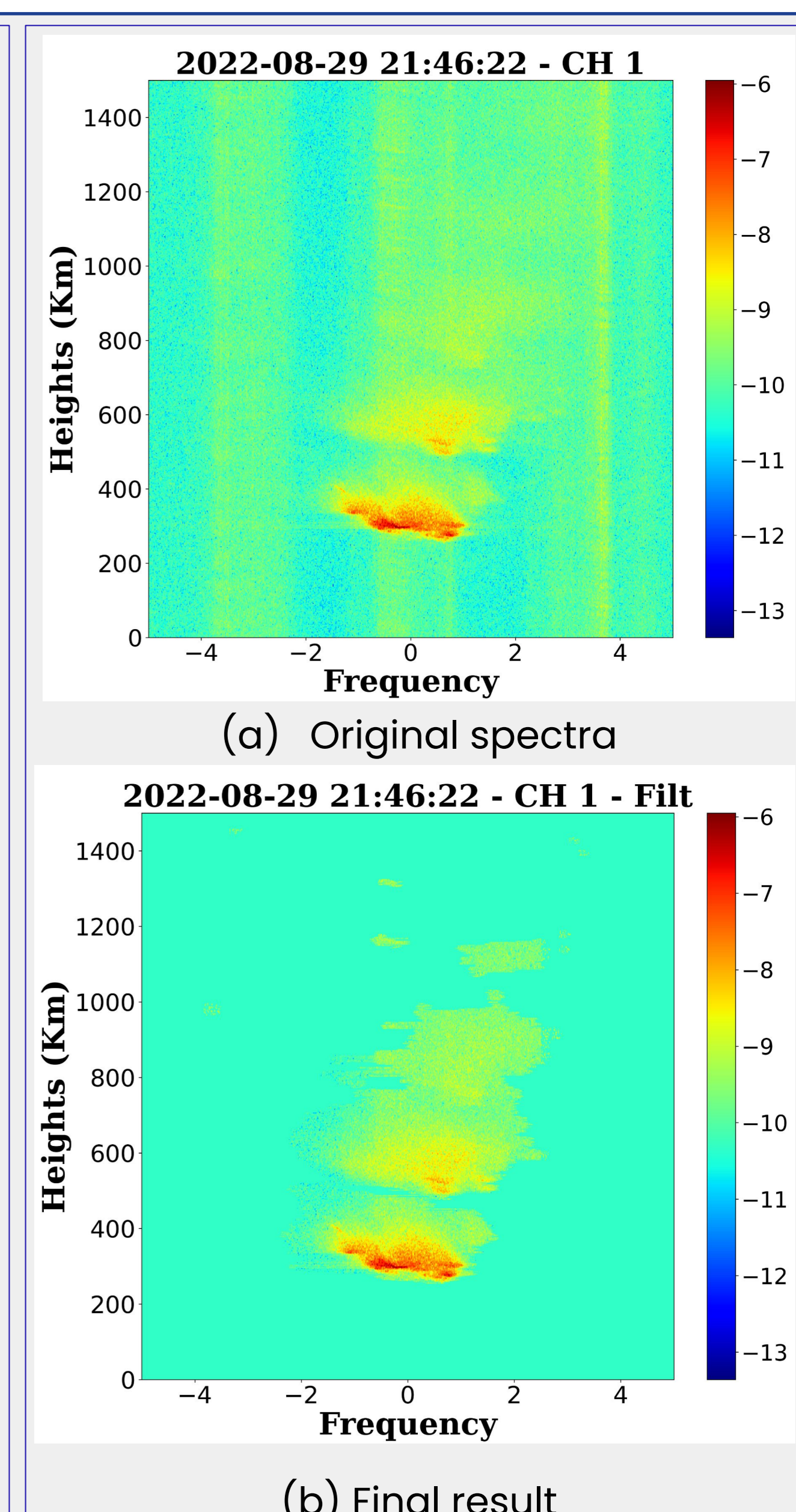


Figure 5. Spectrum a) Original spectra. (b) Final Result.

4. Comparisons of final parameters

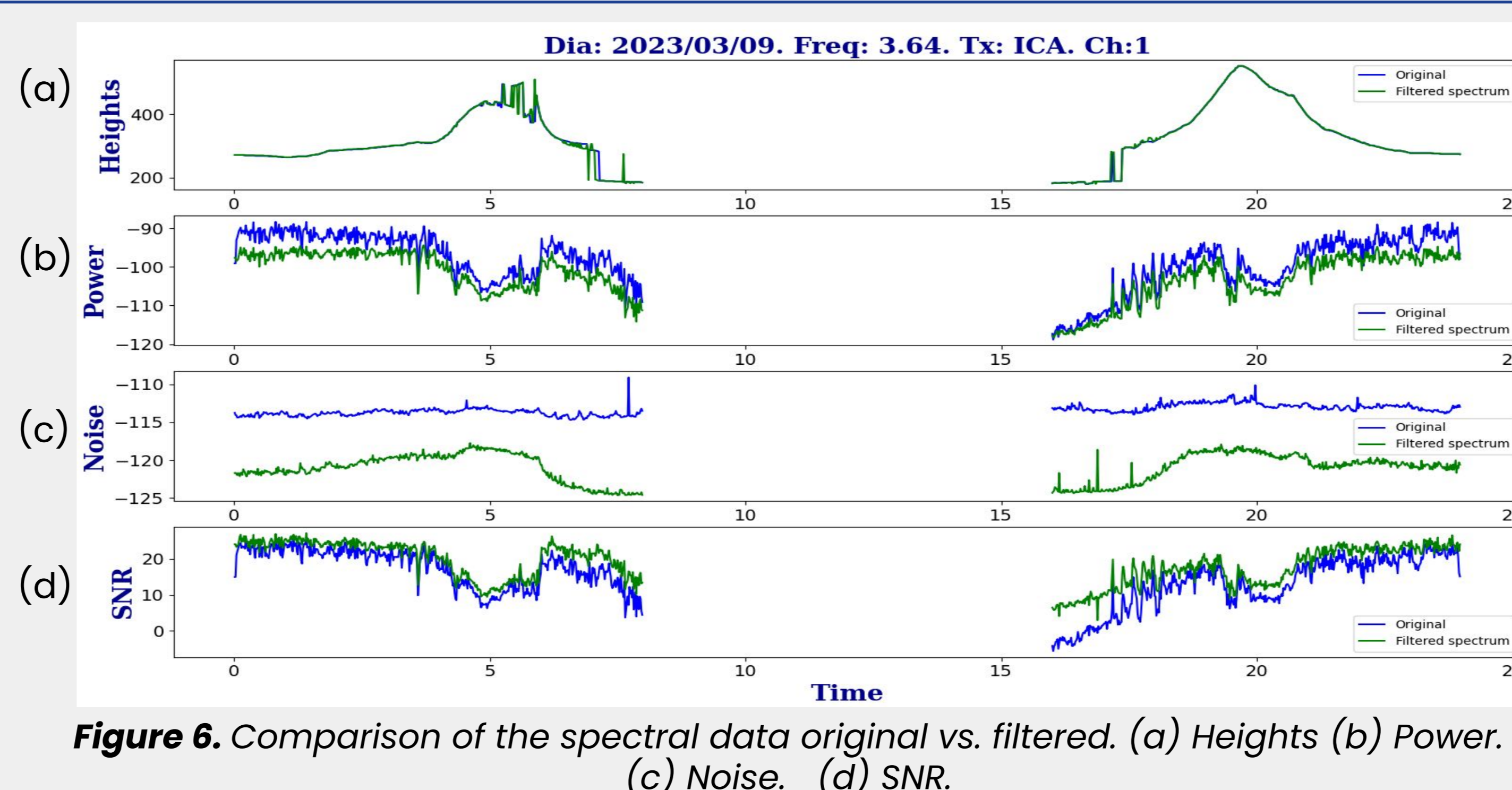


Figure 6. Comparison of the spectral data original vs. filtered. (a) Heights (b) Power. (c) Noise. (d) SNR.

- The spectral data is utilized to determine the altitude of the first ionospheric reflection. Based on this altitude, the Signal-to-Noise Ratio (SNR), power, and Doppler parameters are obtained.
- After obtaining the new noise value without interferences and replacing it in the spectrum of the noisy area (refer to Figure 5.b), the power level slightly decreases (refer to Figure 6.b). This is because the previous power level accounted for the interferences from the other transmitters.
- The noise level is significantly reduced as it is obtained without considering the interferences. (Refer to Figure 5.c)
- These changes result in a slight increase in the Signal-to-Noise Ratio (SNR) levels, as demonstrated in Figure 5.d.

5. Conclusions

- The signals from each transmitter in the spectral data were successfully extracted, free from interference caused by other transmitters and without any cross-talk.
- The developed algorithm effectively reduced the noise level in the HF system.
- As a result, a slight increase in the Signal-to-Noise Ratio (SNR) value was achieved, as the algorithm allows for the extraction of low significant noise values.
- Additionally, the implementation of this algorithm resulted in a significant reduction in the storage size of the spectra, reducing it by approximately 75%.

6. References

[1] Hysell, D. L., Milla, M. A., & Kuyeng, K. (2019). Radio beacon and radar assessment and forecasting of equatorial F region ionospheric stability. *Journal of Geophysical Research: Space Physics*, 124 (11), 9511-9524. <https://doi.org/10.1029/2019JA026991>.

[2] Vierinen, J., Chau, J.L., Pfeffer, N., Clahsen, M., Stober, G. (2016). Coded continuous wave meteor radar. *Atmospheric Measurement Techniques*.

7. Acknowledgments

The Jicamarca Radio Observatory is a facility of the Instituto Geofísico del Perú operated with support from the NSF AGS-1732209 through Cornell University.