Observation of Galactic noise and identification of background sources in RNO-G

Jethro Stoffels[*]**on behalf of the RNO-G collaboration**

[*]*Vrije Universiteit Brussel,*
Pleinlaan 2, Brussels, Belgium

E-mail: jethro.stoffels@vub.be

The Radio Neutrino Observatory in Greenland (RNO-G) aims to detect EeV energy neutrinos. Operating primarily in the 200-700 MHz range, RNO-G is sensitive to broadband signals originating from many sources, including thermal noise, galactic noise, and various types of anthropogenic and environmental backgrounds. Each of these backgrounds will have a unique event signature that can be identified through reconstruction, polarization, and frequency content. During the 2021 and 2022 deployment seasons, 7 of the envisioned 35 stations have been installed and are taking data. This data lends itself well for initial analyses and the characterisation of the classes of noise recorded by RNO-G. In this contribution, I will discuss the current status of background identification with RNO-G. In particular, I will present our efforts to detect noise from our Galaxy, which might be used as a standard candle for future calibration. I will also discuss our efforts to develop new types of event classification strategies, including anomaly detection and other machine learning tools.

1. Introduction

The Radio Neutrino Observatory in Greenland (RNO-G) is an array of detector stations located at Summit Station in Greenland which aims to detect a neutrino flux in the EeV energy range [1]. Initial deployment began in the summer of 2021 and so far 7 of the envisioned 35 stations have been installed. This means that for every station up to three summers of data has been taken, which lends itself well to a study of the typical backgrounds that are encountered. Inspecting the frequency content of surface antennas shown in Figure 1, one sees that the downwards facing surface antennas predominantly see the thermal noise which was also measured in the lab. The upwards facing surface antennas have an additional low frequency contribution, which might be indicative of a galactic noise contribution, and superimposed onto this some anthropogenic single frequency peaks. Of these noise classes, galactic noise will be discussed first.
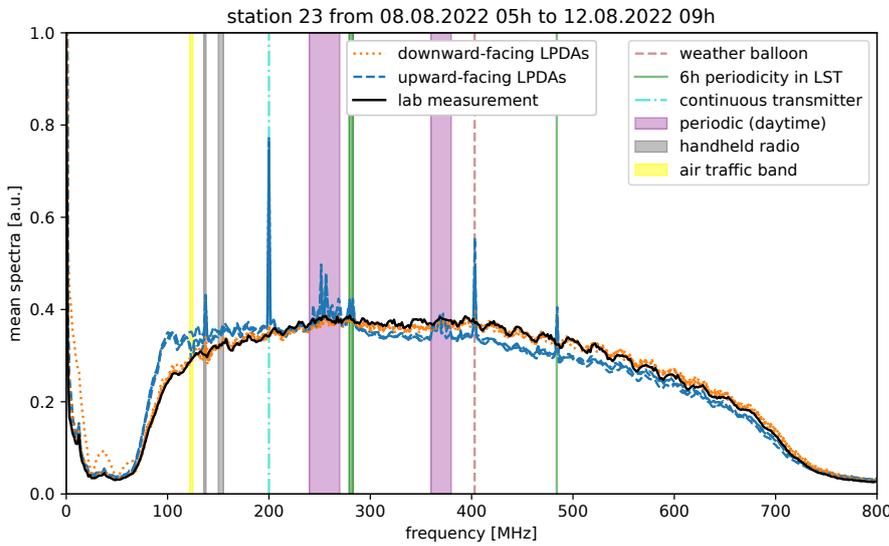

Figure 1: Frequency spectrum as measured by an in situ upwards facing antenna (blue), in situ downwards facing antenna (orange) and the noise measured in the lab (black). [2]

2. Galactic noise

In order to study Galactic noise, one has to make use of transit curves [3]. These plot the voltage variations (V_{RMS}) value of timetraces as a function of Local Sidereal Time (LST), where the LST time scheme is used as this will average out any kind of signal which does not have a periodicity of 1 local sidereal day while signals with a periodicity of one local sidereal day, such as the Milky Way, become more prominent. As can be seen in Figure 2, the Galactic center is not visible from the detector location at Summit Station. As a result, the galactic contribution will be weakened but still present since a significant part of the Milky Way remains visible. From Figure 2 one can also infer that the Galactic contribution will be strongest from 18hr to 20hr LST. For this study, 2 months of 2022 summer data of three upwards facing surface antennas was used.
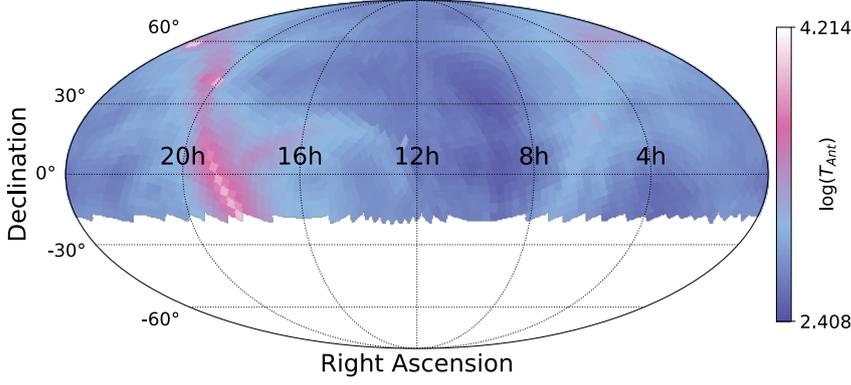
Figure 2: The radio sky model at 110MHz generated by pygdsm [6] and used for simulating the galactic noise signal, represented in Equatorial coordinates. The radio sky map has been pixellised in the same manner as for the simulation and clipped with the RNO-G field of view

In order to construct these transit curves, two months of 2022 summer data was taken and for each timetrace the V_{RMS} value is in practice computed as the standard deviation of the recorded voltage via the following definition:

$$V_{RMS} = \sqrt{\frac{1}{N} \sum_i (V_i - \langle V \rangle)^2}. \tag{1}$$

Additionally, the time of occurence of each timetrace is converted to LST such that the VRMS values can be binned by LST time.

Datacleaning is performed by first filtering the Galactic dominant region, which is up to 110MHz and subsequently performing two consecutive 3σ quality cuts. In order to validate the transit curves obtained from data, the background noise has been simulated with the NuRadioMC software package [5]. In this simulation, the background is assumed to only consists of instrumental noise and galactic noise. The V_{RMS} value of the instrumental noise is given by

$$V_{RMS} = \sqrt{k_b T R \Delta v}, \tag{2}$$

where k_b is the Boltzmann constant, T is the thermal noise temperature, Δv the considered bandwidth and R the resistivity of the system [4]. The Galactic noise is computed by reading in the radio sky map [7, 8] using the pygdsm python package [6] and dividing this into smaller pixels. For each pixel the antenna noise temperature is known, from which its contribution to the antenna is calculated. Summation of the contributions of all the pixels then yields the total galactic noise. The thermal noise temperature is left as a fit parameter of the simulation; which allows the transit curve to shift up or down while retaining its shape.

The data transit curves for all channels of all stations have been calculated. In Figure 3, some of these results can be seen for station 23. The transit curves for the upwards facing antennas are more prominent, have less variation per bin and are generally at higher V_{RMS} values compared to the transit curves for the downwards facing antennas. This qualitative trend is consistent with expectation.

A comparison of the data and simulated transit curves for antenna 16 of station 23 can be seen in Figure 4. There is a good agreement between the data and simulation except between 18hr and 21hr, the cause of which is still under investigation.
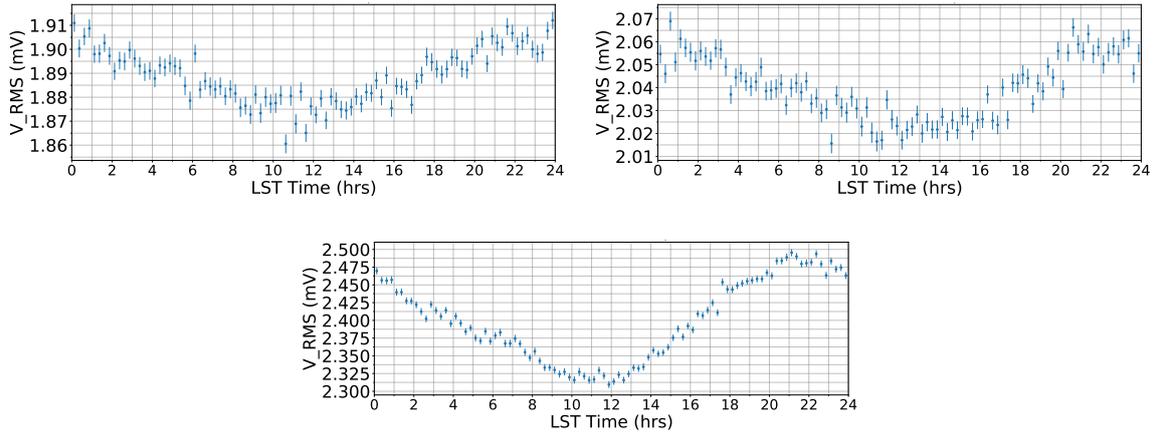
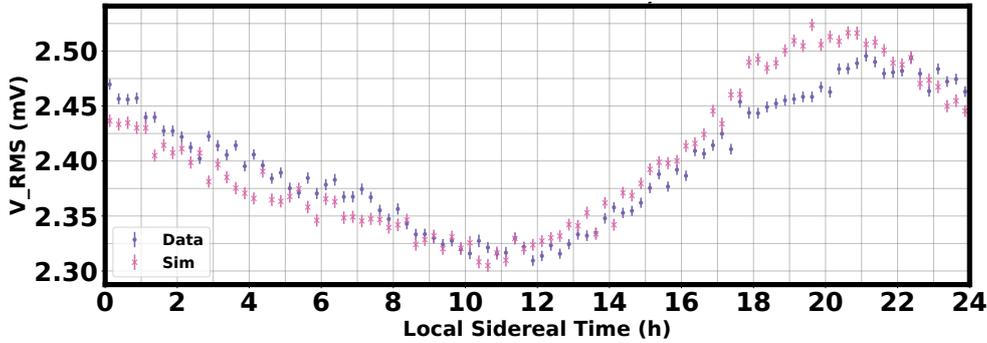**Figure 3:** Transit curves constructed from data of station 23, channels 15 (top left), 16 (bottom) and 17 (top right).

Figure 4: Comparison between data and simulations transit curves for station 23, channel 16

3. Anthropogenic noise

Antropogenic noise encompasses any kind of noise which is man-made. At Summit Station, a myriad of anthropogenic noise sources are present such as weather balloons (which are launched at specific hours during the day), airplanes, general station activity (communication equipment, snow mobiles, heavy machinery and so forth) and the wind turbines that have been installed at some RNO-G stations. Additionally, a single frequency signal at 200MHz is permanently visible in some stations, the cause of which is still under investigation.

The commercial airplanes flying over Greenland emit a transient signal which is strong enough to trigger the detector. Subsequently these events can be reconstructed to yield the arrival direction. As can be seen in Figure 5, the reconstructed directions agree with the flight path. These events are avoidable in any case as the flight plans are known beforehand.
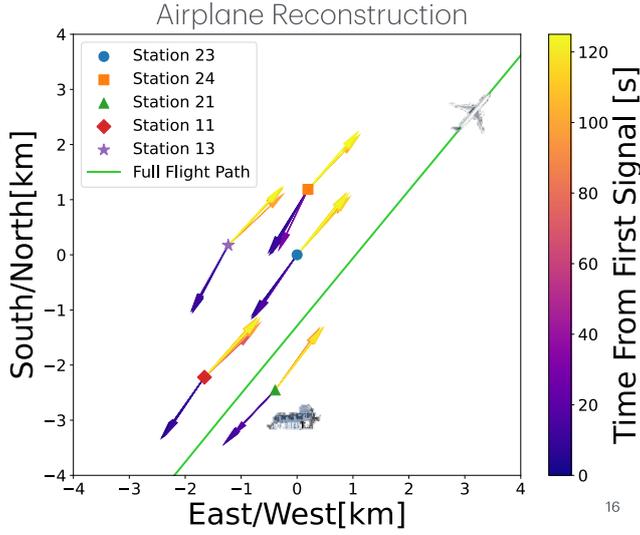
Figure 5: A top-down view of the current station array with the reconstructed flight directions shown as arrows. The colors of the arrows indicate its time with respect to the first signal. The green line represents the actual flight path. Credits: Z. Meyers, private communication

The weather balloon emits a transient single frequency signal and records its position at any time. These measurements can be used as a reconstruction test, the results of which are shown in Figure 6. The reconstructed zenith angle follows the same trend as the zenth angle computed from the weather balloon GPS measurements. Note that some frequency harmonics can be seen repeating this pattern. Additionally, a measure of the index of refraction of the ice can be extracted by investigating the difference in arrival time of this signal for antennas on the same string [9].

Figure 6: Reconstruction of the zenith angle of the weather balloon by two stations. The colormap indicates the zenith angles reconstructed from data and the red line the actual zenith angle. [2]
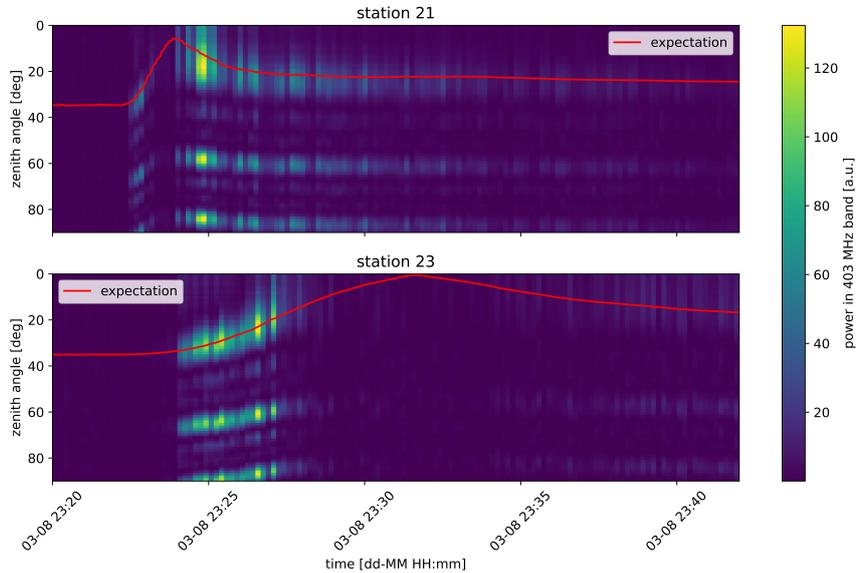
4. Environmental noise

The main type of environmental noise measured at Summit Station is noise resulting from high wind periods. As can be seen in Figure 7, once the wind speed at Summit Station exceeds 10m/s, the trigger rate of the detector increases significantly. This is thought to be due to the triboelectric effect,

which describes electrostatic discharges due to force being applied onto a physical discontinuity [10]. In the case of RNO-G this can be caused by snow hitting the solar panels. Reconstruction of these wind events can be seen in Figure 8. For station 21, which is closest to the main building of Summit Station, these events reconstruct to this location. For the other stations, these events reconstruct to the general direction of the solar panels, further reinforcing the perception that the triboelectric effect is the cause of these events.
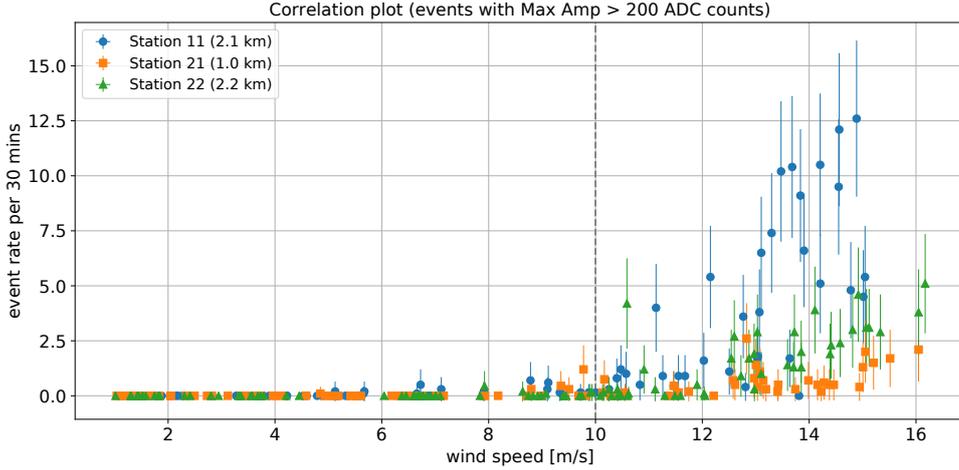
Figure 7: Detector trigger rate as a function of wind speed for different stations. [10]

5. Background mitigation software efforts

In order to eliminate these backgrounds, several approaches have been investigated, two of which will be discussed here. The first effort is an anomaly detection code which trains a convolutional autoencoder-decoder network onto spectrograms of pure background [11]. This makes the network very efficient at recognising pure background, or equivalently, whether an event is not pure background dubbed anomalous. Once a neural network has been trained, it can operate fast and computationally inexpensive. As a result, it might be utilised in an online manner.
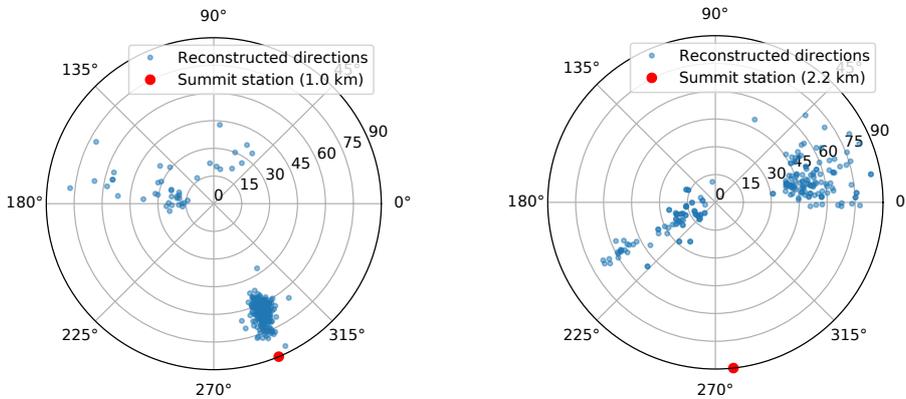
Figure 8: Reconstruction of wind events for station 21 (left) and station 22 (right). [10]

The second effort is a data classification method which consists of a two step process [12]. At first, it transforms the data onto a latent space with a specific set of parameters. For a specific choice of transformation and parameters, noise events coming from same origins or with same signatures will cluster together. In the second step, a clustering algorithm is employed to tag these groups of noise, effectively classifying the data in groups. The results of this method can be seen in Figure 9.
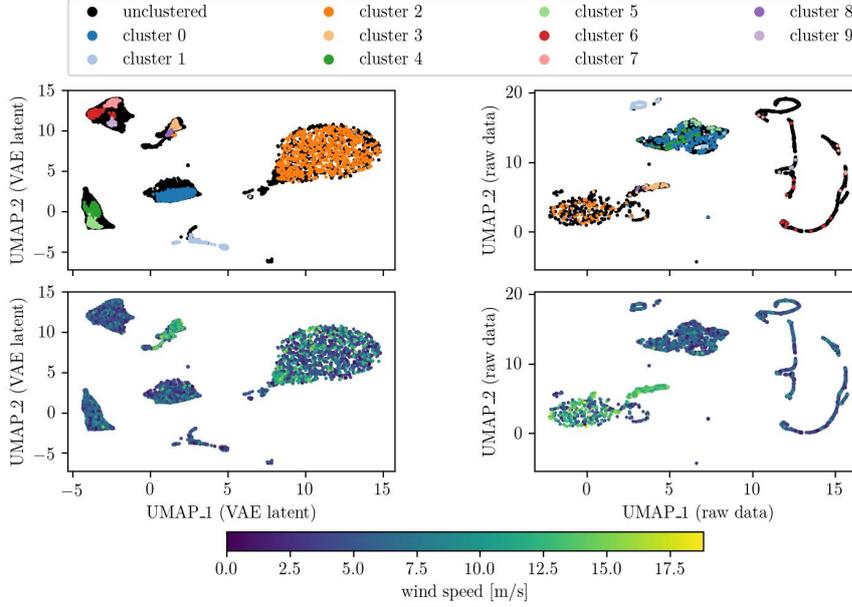

Figure 9: Noise events from station 21 plotted in latent space (left column) and as raw data (raw column). Overlaid are the tagged clusters (top) and the measured wind speed (bottom). [12]

6. Summary

Using the first three summers of data taken by the installed RNO-G stations, an initial search for backgrounds is performed. The most prevalent backgrounds have been shown to consist of instrumental thermal, galactic, anthropogenic and environmental noise. While each of these noise backgrounds pose an obstacle, some aid in calibrating the detector. In order to mitigate these backgrounds altogether, several approaches have been explored such as anomaly detection and data classification methods.

References

[1] RNO-G Collaboration, J. A. Aguilar et al. JINST 16 no. 03, (2021) P03025. [Erratum: JINST 18, E03001 (2023)].

[2] RNO-G Collaboration, J. A. Aguilar et al. PoS ICRC2023, (2023) 1043..

[3] Escudie, A. Ecole nationale supérieure Mines-Télécom Atlantique no. 2019IMTA0145, (2019).

[4] David Staelin . *Receivers, antennas, and signals*. Massachusetts Institute of Technology, course number: 6.661. DOI: N/A . arXiv:N/A.

[5] Glaser, C. et al. The European Physical Journal C 80 no. 02, (2020).

[6] Price, Danny C. The European Astrophysics Source Code Library, (2016).

[7] de Oliveira-Costa et al. Monthly Notices of the Royal Astronomical Society 388 no. 01, (2008) 247-260.

[8] Zheng, H. et al. Monthly Notices of the Royal Astronomical Society 464 no. 03, (2016) 3486-3497.

[9] RNO-G Collaboration, Oeyen, B. et al. PoS ICRC2023, (2023) 1042.

[10] RNO-G Collaboration, J. A. Aguilar et al. Astroparticle Physics 145, (2023) 102790..

[11] RNO-G Collaboration, Meyers, Z. S. et al. PoS ICRC2023, (2023) 1142..

[12] RNO-G Collaboration, Glüsenkamp, Thorsten et al. PoS ICRC2023, (2023) 1056..

Acknowledgements