The Radio Meteor Zoo: a citizen science project

Stijn Calders[1], Cis Verbeeck[2], Hervé Lamy[1] and Antonio Martínez Picar[2]

[1] **Royal Belgian Institute for Space Aeronomy**
`Stijn.Calders@aeronomy.be`

[2] **Royal Observatory of Belgium**

Scientists from the BRAMS radio meteor network have started a citizen science project called Radio Meteor Zoo in collaboration with Zooniverse in order to identify meteor reflections in BRAMS spectrograms. First, a small-scale version of the Radio Meteor Zoo was carried out with a sample of meteor identifications in 12 spectrograms by 35 volunteers. Results are presented here and allowed us to define a method that reliably detects meteor reflections based on the identifications by the volunteers. It turns out that, if each spectrogram is inspected by 10 volunteers, hit and false detection percentages of 95% respectively 6% are expected. The Radio Meteor Zoo is online at https://www.zooniverse.org/projects/zooniverse/radio-meteor-zoo. Citizen scientists are kindly invited to inspect spectrograms.

1 Introduction

The BRAMS (Belgian RAdio Meteor Stations) network consists of ~30 receiving stations spread all over the Belgian territory and a single radio transmitter installed at the Geophysical Centre of the Royal Meteorological Institute (RMI) in Dourbes (Calders et al., 2014; Lamy et al., 2015). This radio transmitter emits a sine wave with circular polarization at a frequency of 49.97 MHz and with a constant power of 150 W. At each receiving station, the signal is sampled with a frequency of 5512 Hz, providing a bandwidth of ~2.5 kHz. Data are saved as WAV (sound) files every 5 minutes. BRAMS data are usually presented as spectrograms, which provide the frequency content of the signal as a function of time. Spectrograms are built from raw data using the FFT on 16384 samples and with an overlap of 90%. Only 200 Hz of the whole bandwidth, centered on the direct signal of the transmitter, are usually shown as the majority of the meteor echoes appear there. Spectrograms are very useful because the spectral signatures of meteor echoes are very different from those due to "spurious" signals such as e.g. reflections on airplanes or broad-band interferences.

Each BRAMS receiving station is recording continuously, producing each day 288 WAV files and detecting ~ 1500–2000 meteors. This huge amount of data requires the use of automatic detection algorithms. Several attempts were made to identify meteor reflections either in raw data or in spectrograms by using automatic detection algorithms, with varying degrees of success as discussed in detail in (Calders and Lamy, 2014; Lamy et al., 2015). The automatic detection of overdense radio meteor echoes in particular remains a difficult task due to the various and complex shapes they produce in spectrograms. This problem is particularly striking during meteor showers where these types of meteor echoes are observed abundantly. In this paper, a new strategy for the identification of meteor reflections in the spectrograms is explored. Instead of detecting meteor reflections automatically by means of software, we propose to rely on the best detector which is the (trained) human eye. This is a well-established method nowadays in observational science, known as crowdsourcing or citizen science (Lintott, 2008).

2 The Radio Meteor Zoo

The authors have started collaboration with the scientists at Zooniverse to use their platform to host a project called *Radio Meteor Zoo*.
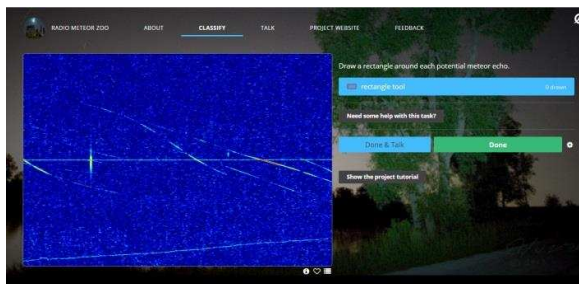

Figure 1 – The Radio Meteor Zoo website.

In order to be able to analyze the *Radio Meteor Zoo* contributions, we must be able to answer the two following important questions:

- What is the minimum number of volunteers we need to inspect a given spectrogram such that we can statistically be confident in the results?
- In a given spectrogram, how can we accurately derive the number and position of meteor echoes based on individual contributions?

In order to answer these questions, a test was performed with 12 spectrograms and 35 users.

3 A small-scale version of the Radio Meteor Zoo

Description of the test data set
We used 12 spectrograms from the BRAMS receiving station in Ottignies obtained on 15 March 2015 between 0^h and 1^h UT. The authors carefully inspected the spectrograms together. In total 120 meteor reflections

were identified in the 12 spectrograms. These detections represent the reference dataset.

Composition of the test group

The test group consisted of 35 volunteers with a strong physics background, and most of them are interested in meteor research but not necessarily familiar with radio observations.

Meteor identification interface

Both the authors and the test group used the same interactive web tool to inspect the spectrograms and to identify meteor reflections. This tool is accessible online[1]. With this tool a user can draw a rectangle around each feature in the spectrogram that he considers to be a meteor. Once the user has identified all meteor reflections in the spectrogram, he can navigate to the next spectrogram. He can also navigate back to check his identifications in a previous spectrogram.

The coordinates of the rectangles, both in pixels coordinates and in frequency/time coordinates, are saved in a comma-separated values (CSV) file. A CSV file was created for each user.

Training

The volunteers of the test group were asked to read first a tutorial. This tutorial explains what a spectrogram looks like and provides examples of typical signatures of meteor reflections and common distortions (like reflections on airplanes or broad-band interferences). Finally the tutorial explains what is expected from the volunteer: drawing rectangles around potential meteor echoes and how to do it correctly.
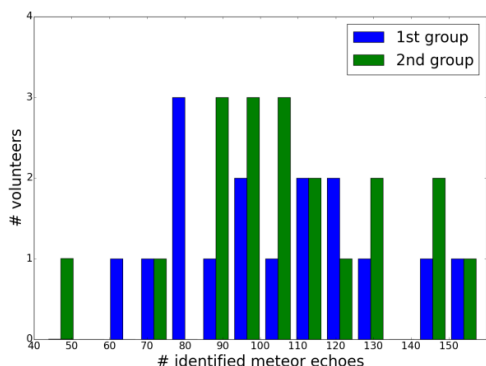

Figure 2 – Total number of meteor echoes identified by the different volunteers in the 12 spectrograms.

From the results of a first group of volunteers (16 people in the test group of 35), the authors learned that the median volunteer identified 99 meteor reflections (median absolute deviation MAD=19). This is far less than the 120 meteors that the authors had identified. The difference between the counts from the test group and the reference dataset was mostly due to the faintest meteor echoes. Therefore the tutorial was updated asking the users to draw a rectangle even when they have a doubt about a faint meteor. After all, it is easier to filter out a

false detection than to retrieve a missed meteor detection. The median volunteer of the second group identified 107 meteor reflections in the spectrograms (MAD=17). A histogram of the individual counts is given in *Figure 2*.

Challenges

It was soon realized that there is a large spread on the number of meteors identified by the 35 persons. For instance, one volunteer identified 7 meteor reflections in the spectrogram in *Figure 3*, while another volunteer identified 17 meteor reflections, and the reference dataset yielded 15 meteor reflections.
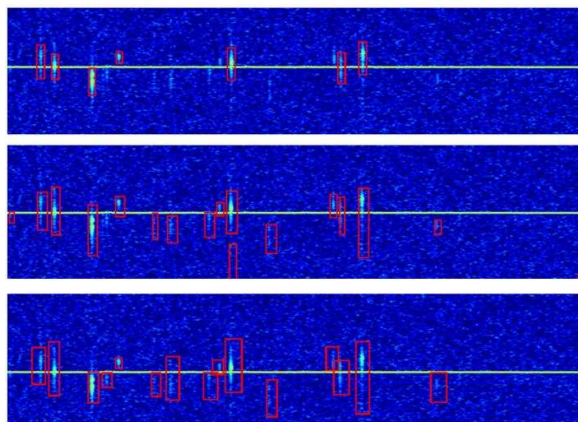

Figure 3 – Meteor reflection identifications by different volunteers in the same spectrogram. Top: volunteer 1 identifies 7 meteor reflections; middle: volunteer 2 identifies 17 meteor reflections; bottom: the reference detection identifies 15 meteor reflections.

4 How to interpret the *Radio Meteor Zoo* identifications?

First let us try to answer the second question from *Section 2*: in a given spectrogram, how can we accurately derive the number and position of meteor echoes based on individual contributions? In order to investigate this closer, we performed for every of the 12 spectrograms the following analysis for all values of *i* between 1 and 35.

- Create a binary version of the spectrogram called $image_0$ which has pixel value 1 for pixels identified as a meteor pixel in the reference spectrogram, and pixel value 0 for all other pixels.
- Create a binary version of the spectrogram called $image_1$ which has pixel value 1 for pixels which were identified as a meteor pixel by at least *i* volunteers, and pixel value 0 for all other pixels.
- Calculate the number $D(i)$ of pixels for which $image_0$ and $image_1$ have different pixel values.

This allowed us to determine the value of *i* which minimizes $D(i)$.

It turns out that the number $D(i)$ of pixels where "at least *i* volunteers" and the reference spectrogram disagree, has a minimum at $i_{optimal} = 12$ volunteers. This means that in order to best reconstruct the reference spectrogram, we should consider as meteor pixels those pixels that have been identified as meteor pixels by at

least 12 of the 35 volunteers. The corresponding spectrogram is called the optimal identification spectrogram.

Now let us try to answer the first question of *Section 2*: what is the minimum number of volunteers we need to inspect a given spectrogram? Indeed, in the *Radio Meteor Zoo* project, it would be better to have a number of volunteers n that have to inspect a single spectrogram well below 35. For example, if 2000 spectrograms have to be investigated (corresponding to approximately one week of data for one receiving station), that would already amount to 70000 individual inspections for 35 users. So we repeated the analysis above for each number of volunteers n between 1 and 35. For every n we randomly selected 1000 combinations of n volunteers out of 35 to have a significant number of simulations without making it too CPU intensive.

Figure 4 provides for every n the optimal number of volunteers $i_{optimal}(n)$ which minimizes D(i). For instance, $i_{optimal}(35) = 12$ for $n = 35$, as was explained before. As expected, $i_{optimal}(n)$ increases with n.

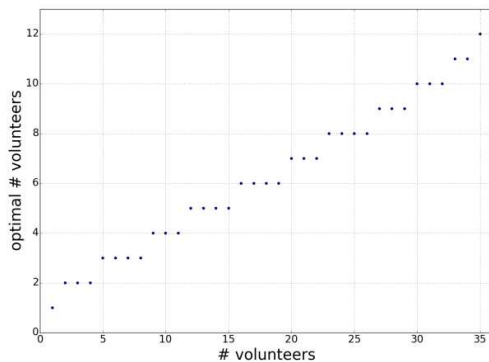*Figure 4* – For every number of volunteers n on the x-axis, the y-axis shows the optimal number of volunteers $i_{optimal}(n)$ that minimizes D(i).
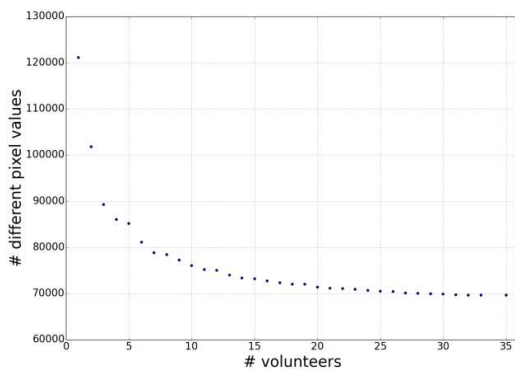

Figure 5 – The number of pixels D(i) for which the optimal identification spectrogram for n volunteers and the reference spectrogram have different pixel values, as a function of number n of volunteers considered.

For every value of n (number of volunteers), we can now derive an optimal identification spectrogram of meteor reflections by considering a pixel as a meteor pixel if it is identified as such by at least $i_{optimal}(n)$ volunteers. In

Figure 5, the number of pixels D(i) where the optimal identification spectrogram disagrees with the reference spectrogram is plotted as a function of number of volunteers n. Note that each spectrogram contains 595 × 864 pixels in the 200 Hz range shown to the volunteers. For 12 spectrograms the total amount of pixels is therefore larger than 6×10^6. The values of D(i) shown in *Figure 5* represents thus maximum ~2% of the total number of pixels in the worst case. This curve varies smoothly as a function of n, and of course has a minimal (best) value at $n = 35$.

Figure 5 allows us to select a value for the number of volunteers n which is much smaller than 35 but yet still delivers accurate meteor reflection identifications. For the *Radio Meteor Zoo*, we selected $n = 10$ volunteers per spectrogram. Indeed, using 10 volunteers instead of 35, corresponds only to an increase of 9% of D(i), i.e. 9% more pixels with different values in the optimal identification spectrogram and in the reference spectrogram.

5   Results

When we apply the identification method described above on the spectrogram from *Figure 6* (with number of volunteers $n = 12$), the same 15 meteors are identified as in the reference spectrogram.
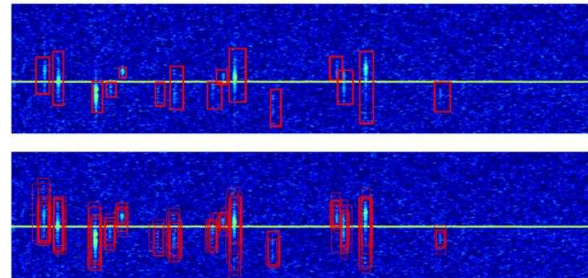

Figure 6 – Comparison of the reference meteor spectrogram (top) and optimal identification spectrogram by the method described above with 12 volunteers (bottom). The same meteors were identified in both cases.

Until now, we have only considered meteor pixels instead of meteor reflections as a whole. To which extent does meteor reflection identification by the above method correspond to the meteor echoes in the reference dataset? In order to investigate this, we have applied a minimum bounding box algorithm to group meteor pixels into individual meteor echoes.

For every value of n between 1 and 35, 1000 random combinations of $i_{optimal}(n)$ out of n volunteers are considered. For each combination, the number of hits (meteor reflections identified by both the reference and the proposed method) and the number of false detections (meteor reflections identified by the proposed method but not by the reference) are calculated.

In *Figure 7*, the medians of the percentage of hits and false detections over these 1000 iterations are plotted as a function of the number of volunteers n. For the *Radio*

Meteor Zoo, we will employ 10 volunteers per spectrogram, which amounts to a median percentage of hits of 95% and a median percentage of false detections of 6%. The median percentage of false detections is quite low because we have very few airplane echoes at night.
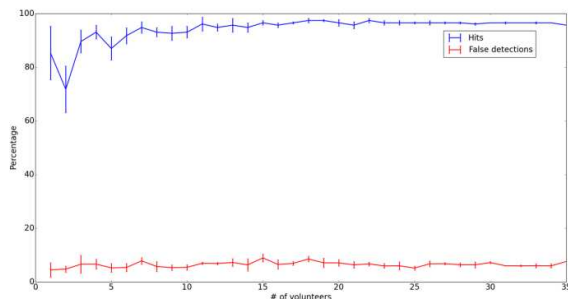

Figure 7 – Median percentage of hits and false detections of the proposed method as compared to the reference detection, as a function of number of volunteers *n*.

6 Discussion

Employing the meteor reflection identifications of 12 spectrograms by 35 volunteers, we were able to define a statistical method to identify meteor reflections. Based on $n = 10$ volunteers inspecting each spectrogram, a median percentage of hits of 95% and a median percentage of false detections of 6% is obtained. Note that post-processing (e.g. looking at the power profile) can be invoked to analyze and reject false detections after detection.

Since the population of *Radio Meteor Zoo* volunteers may differ systematically from the population of 35 volunteers in the test (in particular with regard to their physics background), a similar test will be performed in order to validate our approach with the *Radio Meteor Zoo* volunteers.

These manual identifications will prove to be extremely useful during meteor showers because these contain many complex overdense meteor echoes. They will also be of great use to calibrate and test the pre-existing and potential new automatic detection algorithms.

We are ready to start analyzing *Radio Meteor Zoo* data. It is accessible via the following URL: https://www.zooniverse.org/projects/zooniverse/radio-meteor-zoo. We kindly invite all readers to help us by identifying meteor reflections and promoting this website!

Acknowledgments

References

Calders S., Lamy H., Gamby E. and Ranvier S. (2014). "Recent developments in the BRAMS project". In Gyssens M., Roggemans P. and Zoladek P., editors, *Proceedings of the International Meteor Conference*, Poznan, Poland, 22-25 August 2013. IMO, pages 170–172.

Calders S. and Lamy H. (2014). "Automatic detection of meteors in BRAMS data". In Rault J.-L. and Roggemans P., editors, *Proceedings of the International Meteor Conference*, Giron, France, 18–21 September 2014. IMO, pages 194–196.

Lamy H., Anciaux M., Ranvier S., Calders S., Gamby E., Martinez Picar A. and Verbeeck C. (2015). "Recent advances in the BRAMS network". In Rault J.-L. and Roggemans P., editors, *Proceedings of the International Meteor Conference*, Mistelbach, Austria, 27-30 August 2015. IMO, pages 171–175.

Lintott C. J., Schawinski K., Slosar A., Land K., Bamford S., Thomas D., Raddick M. J., Nichol R. C., Szalay A., Andreescu D., Murray P. and Vandenberg J. (2008). "Galaxy Zoo: Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey". *Monthly Notices of the Royal Astronomical Society*, **389**, 1179–1189.