



Technologie de l'Informatique

Avenue du Ciseau 15

1348 Ottignies

---

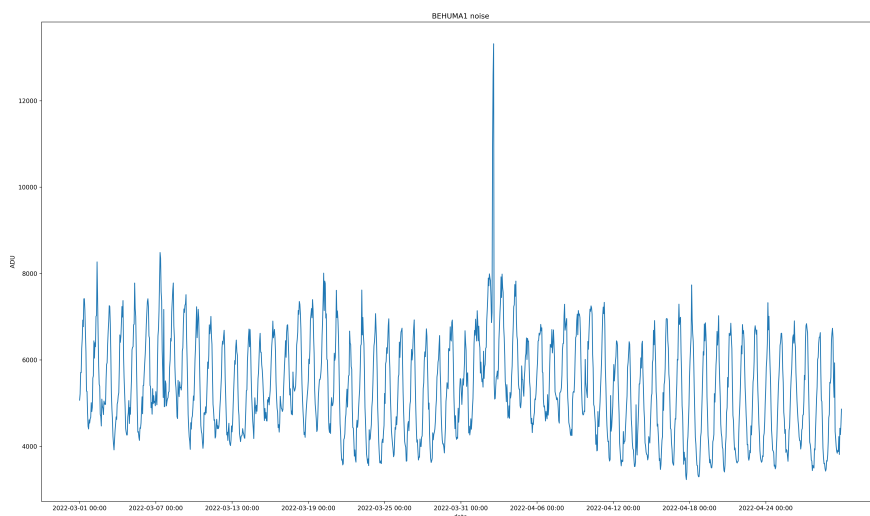
## Monitoring des données BRAMS et détection automatique des échos de météore

---

*Miguel Antoons*

**Rapporteur**

*Monsieur Arnaud Dewulf*



*2021-2022*

## Remerciements

*Je tiens tout d'abord à remercier toutes les personnes qui m'ont aidé à réaliser ce projet de fin d'études.*

*En commençant par mon professeur rapporteur, à qui j'ai pu poser mes questions en cas de besoin et qui s'est assuré que tout se passe bien tout au long du projet.*

*Ensuite, je voudrai remercier Mr Hervé Lamy pour avoir proposé ce sujet de fin d'études, mais également pour m'avoir expliqué, de façon claire et précise, toutes les notions qui nécessitaient des explications.*

*Je tiens également à remercier Mrs Antoine Calegaro et Michel Anciaux, qui m'ont guidé quand c'était nécessaire et qui m'ont conseillé durant le projet de fin d'études.*

*Enfin, je voudrais exprimer ma reconnaissance envers toutes les personnes qui m'ont conseillé sur, et ont relu ce rapport de projet de fin d'études.*

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Le projet BRAMS . . . . .	3
1.1.1	Le Réseau BRAMS . . . . .	3
1.1.2	Le Format WAV . . . . .	3
1.1.3	La Réflexion Spéculaire . . . . .	6
1.1.4	Les Données BRAMS . . . . .	7
1.1.5	L'Archive BRAMS . . . . .	10
1.2	Problématique liée à l'étude des données BRAMS . . . . .	12
<b>2</b>	<b>Méthodologie</b>	<b>13</b>
<b>3</b>	<b>Technologies utilisées</b>	<b>14</b>
3.1	Python . . . . .	14
3.1.1	Librairies Utilisées . . . . .	14
3.2	MariaDB . . . . .	16
<b>4</b>	<b>Lecture des Fichiers BRAMS</b>	<b>17</b>
4.1	Fonctionnement du Module . . . . .	17
4.1.1	La méthode FFT . . . . .	19
<b>5</b>	<b>Monitoring des Données BRAMS</b>	<b>20</b>
5.1	Fonctionnement du Programme . . . . .	20
5.2	Résultats . . . . .	22
<b>6</b>	<b>Détection des Météores</b>	<b>23</b>
6.1	Analyse . . . . .	23
6.2	Fonctionnement du Programme . . . . .	25
<b>7</b>	<b>Pistes d'Améliorations</b>	<b>28</b>
<b>8</b>	<b>Conclusion</b>	<b>29</b>

# 1 Introduction

## 1.1 Le projet BRAMS

Lancé en 2010 par Monsieur Hervé Lamy à l'Institut Royal d'Aéronomie Spatiale de Belgique, le projet BRAMS (Belgian RAdio Meteor Stations) à pour but de collecter et stocker des données relatives à des objets rentrant dans l'atmosphère belge et, plus précisément, des météores. Ces données pourront ensuite être analysés afin de retrouver la trajectoire, la vitesse ou encore la masse d'un ou de plusieurs météores.

### 1.1.1 Le Réseau BRAMS

Afin d'accomplir ce but, le projet BRAMS dispose d'un réseau de stations nommé le réseau BRAMS. Ce réseau est composé d'un ensemble de quarante-deux stations émettrices, situées en Belgique ou dans les pays avoisinants, et d'un émetteur dédié situé à Dourbes, dans le sud de la Belgique.

Il existe 2 types de station de réception :

- Il y a d'abord les stations avec des récepteurs Icom IC-R75, que j'appellerai les récepteurs ICOM dans ce document. Ce sont les premières stations réceptrices mises en service pour le réseau BRAMS. Actuellement, ces stations ne sont plus utilisées suite à l'arrêt de la commercialisation du récepteur Icom IC-R75. De plus, une variation de température pouvait causer une légère déstabilisation en fréquence.
- C'est alors que les stations utilisant le récepteur RSP 2 ont été développés. Ces stations n'éliminaient pas seulement en grande partie les problèmes des anciennes stations, mais sont également plus compacts et plus faciles à installer. Une image des stations RSP 2 peut être trouvé à la figure 1.

### 1.1.2 Le Format WAV

Le format de fichier WAV (Waveform Audio File) est un format destiné au stockage de signal audio développé par Microsoft et IBM. Il est construit conformément au RIFF (Ressource Interchange File Format) ce qui veut dire que le fichier est organisé en blocs de données, aussi nommés des "chunks" ou "data chunks".

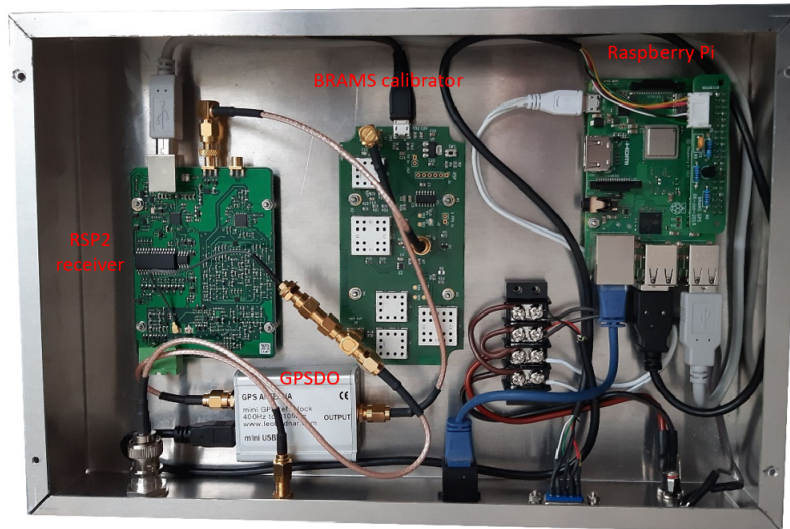


Figure 1: Station RSP 2 du réseau BRAMS

Chaque bloc de données dispose d'un ID codé sur 4 octets, qui représente souvent un mot de 4 lettres. Suit ensuite le champ "ChunkSize" indiquant la taille des données à venir dans l'ensemble du fichier, codé également sur 4 octets. Cette valeur exclut donc les champs "ID" et "ChunkSize" ainsi que la taille de tous les éventuels blocs de données venant avant le bloc courant. Après ces 2 champs, on est libre de rajouter le nombre de champs que l'on souhaite, avec la taille en octets que l'on souhaite.

Pour un fichier WAV, on retrouve typiquement 3 blocs de données. On retrouve premièrement le bloc "RIFF" qui contient la taille de l'entièreté du fichier moins 8 octets (4 octets du champ "ID" et 4 octets du champ "Chunk-Size") suivi du champ indiquant le format du fichier. Dans le cas du fichier WAV, ce dernier contient les quatre lettres "WAVE".

Le deuxième bloc, nommé le bloc FMT<sup>1</sup>, contenu dans un fichier WAV ordinaire contient toutes les données techniques des données. On y trouve notamment la fréquence d'échantillonnage, le nombre de pistes audio ou encore le nombre d'octets par secondes.

---

<sup>1</sup>FMT pour format

Nom du Bloc	Nom du Champ	Position (en octets)	Taille (en octets)
<b>Bloc RIFF</b>	ID	0	4
	ChunkSize	4	4
	Format	8	4
<b>Bloc FMT</b> contenant les données techniques du fichier WAV.	ID	12	4
	ChunkSize	16	4
	AudioFormat	20	2
	NumChannels	22	2
	SampleRate	24	4
	ByteRate	28	4
	BlockAlign	32	2
	BitsPerSample	34	2
<b>Bloc de Données</b>	ID	36	4
	ChunkSize	40	4
	Data	44	*

Figure 2: Structure d'un fichier WAV ordinaire

Vient enfin le bloc principal du fichier : le bloc de données. C'est dans ce bloc que se trouvent les données audio brutes. La taille de ce bloc varie bien-entendu selon les caractéristiques techniques des données audio (longueur, fréquence d'échantillonnage, nombre de pistes, etc.).

Une chose à noter à propos des fichiers WAV est que leur structure offre une grande flexibilité. En effet, elle permet non-seulement de lire et interpréter les fichiers facilement, mais également de rajouter des blocs de données personnalisés sans corrompre les autres données.

La figure 2 montre la structure d'un fichier WAV ordinaire.

### 1.1.3 La Réflexion Spéculaire

Pour comprendre le fonctionnement du réseau BRAMS, la première chose à savoir est : lorsqu'un météore passe dans l'atmosphère, il laisse derrière lui une trainée ionisée. Cette trainée a la propriété de réfléchir les ondes radio. Dans la plus grande partie des cas cette réflexion se fait en un seul point le long de la trajectoire, phénomène qui s'appelle la réflexion spéculaire (illustré à la figure 3). La position de ce point, nommé le point de réflexion spéculaire, dépend notamment du positionnement de l'émetteur, du positionnement du récepteur et de la trajectoire du météore.

Pour exploiter le phénomène de la réflexion spéculaire et réussir à détecter les météores à l'aide d'ondes radios, le réseau BRAMS fonctionne de la façon suivante :

1. Un émetteur, situé à Dourbes, transmet de façon continue un signal à une fréquence de 49.97 MHz. Ce signal est émis en direction du ciel et peut être réfléchi sur des trainées ionisées dans le sillage des météores.
2. Lorsque le signal est réfléchi, il peut être détecté par une ou plusieurs stations réceptrices faisant partie du réseau BRAMS. Un signal calibre est alors additionné au signal venant du ciel. Ce signal est injecté à 49.9705 MHz, c'est-à-dire 500 Hz plus haut que le signal direct. Il dispose d'une amplitude fixe et sert de référence d'amplitude pour le reste du signal. La station réceptrice décale ensuite le signal direct de 49.97 MHz vers une fréquence de 1 kHz.
3. Ensuite, la station réceptrice enregistre l'ensemble du signal dans un fichier audio de type WAV. Le signal est échantillonné à une fréquence de 5512,5 Hz pour les stations ICOM et 6048 Hz pour les stations RSP2. Si tout se passe bien, un fichier est généré toutes les cinq minutes et chaque fichier devrait commencer et terminer à un temps prédéfini (par exemple : 16 h 00 à 16 h 05, 16 h 05 à 16 h 10, etc.).
4. Tous les fichiers générés par les stations réceptrices sont envoyés à intervalle régulier aux serveurs utilisés pour le projet BRAMS par internet.
5. Une fois sur le serveur, les fichiers WAV sont archivés, manipulés et étudiés par les scientifiques du projet BRAMS afin d'en extraire les informations utiles.

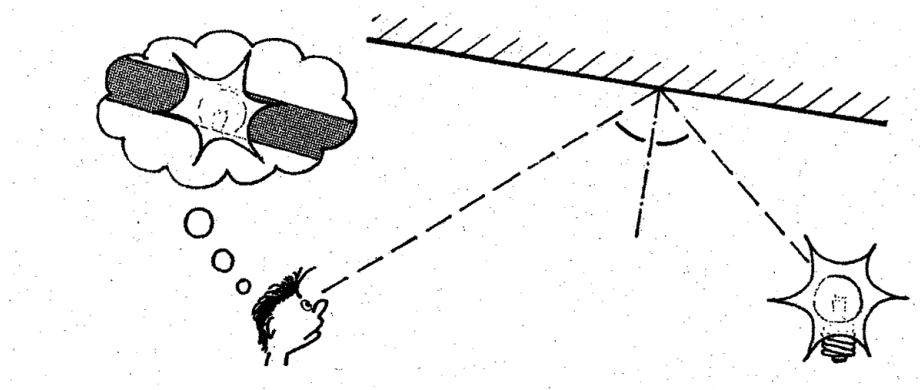


Figure 3: Représentation de la réflexion spéculaire.

#### 1.1.4 Les Données BRAMS

Un fichier WAV venant d'une station réceptrice contient donc le signal capté par l'antenne ensemble avec un signal calibreur. Il est composé d'une seule piste audio, échantillonnée à une fréquence de 5512.5 Hz ou 6048 Hz. Cette fréquence permet d'enregistrer des données dans une bande de fréquences allant jusqu'à 2756.25 Hz ou 3024 Hz (ou la moitié de la fréquence d'échantillonnage), selon le théorème de Nyquist. Sachant que les échos de météores apparaissent typiquement dans une bande de fréquence de 100 Hz autour du signal de l'émetteur à 1000 Hz, cette bande de fréquence couvre l'ensemble des signaux utiles à l'étude des météores.

À chaque fichier WAV est rajouté un bloc de données (data chunk) conçu pour faciliter l'étude des données. Dans ce bloc, on retrouve quelques informations relatives à la station de réception, la station émettrice, le signal GPS<sup>2</sup> ou encore, la fréquence d'échantillonnage. Ce bloc de données se trouve entre le bloc FMT et le bloc de données.

Ces fichiers audio, comme montrée à la figure 4, sont très difficiles à interpréter. Ils sont composés de bruits et de parasites sur l'entièreté de leur bande de fréquence. Par contre, lorsque l'on calcule le spectrogramme du fichier WAV, les données deviennent beaucoup plus simples à lire.

Un spectrogramme est une représentation différente des données contenues

---

<sup>2</sup>Global Positioning System



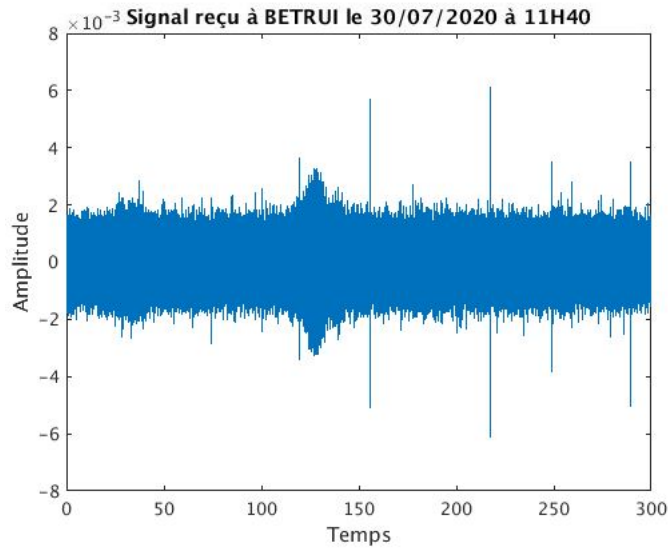


Figure 4: Données brutes venant d'un fichier audio d'une station réceptrice.

dans un fichier audio. Au lieu de représenter la puissance sur l'ordonnée et le temps sur l'abscisse, un spectrogramme représente la répartition des fréquences au cours du temps. Il contient donc trois axes :

- Sur l'abscisse on retrouve toujours le temps.
- Sur l'ordonnée on retrouve la fréquence.
- Sur la cote, on retrouve la puissance. Cet axe est souvent représenté par un code couleur.

Pour générer un spectrogramme, il faut calculer un ensemble de Transformées de Fourier. La Transformée de Fourier permet de représenter la répartition de puissance entre les fréquences contenues dans un signal ou une partie d'un signal temporel. Elle produit donc ce qu'on appelle le spectre du signal. Elle se calcule sur un nombre quelconque d'échantillons qui se suivent dans un signal temporel. Dans l'informatique, elle est souvent calculée à l'aide de la FFT<sup>3</sup>. C'est un algorithme qui permet de calculer une Transformée de Fourier plus rapidement, mais qui exige un nombre d'échantillons  $2^n$ .

---

<sup>3</sup>Faste Fourier transform

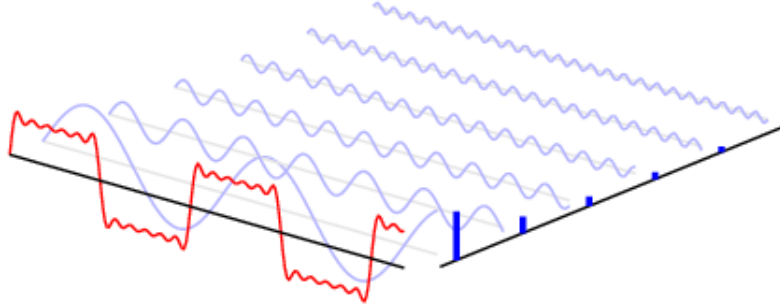


Figure 5: Représentation visuelle de la Transformée de Fourier.

Sur la figure 5, on peut voir en rouge un signal temporel et le spectre de ce même signal en bleu. Les signaux en mauve claire sont les signaux à une fréquence, dont le signal rouge est composé.

Dans le cas des spectrogrammes pour le projet BRAMS, les FFT contenus dans un spectrogramme sont calculés sur 16384 échantillons. Si on suppose qu'un fichier WAV venant d'une station réceptrice dure en moyenne trois-cents secondes (ou cinq minutes), un spectrogramme est composé d'environ 101 spectres pour les stations ICOM et 111 pour les stations RSP2. Ce nombre est obtenu avec la formule ci-dessous, où  $F_s$  est la fréquence d'échantillonnage,  $T$  la durée en secondes du signal et  $nfft$  le nombre d'échantillons utilisés pour générer un spectre.

$$\frac{F_s * T}{nfft}$$

Le spectrogramme généré aura une résolution fréquentielle, donnée par la formule  $F_s/nfft$ , de 0.34 Hz pour les stations ICOM et 0.37 Hz pour les stations RSP2. Sa résolution temporelle, donnée par la formule  $nfft/F_s$ , est de 2.97 secondes pour les stations ICOM et de 2.7 secondes pour les stations RSP2. Un exemple de spectrogramme venant d'une station de réception est affiché à la figure 6.

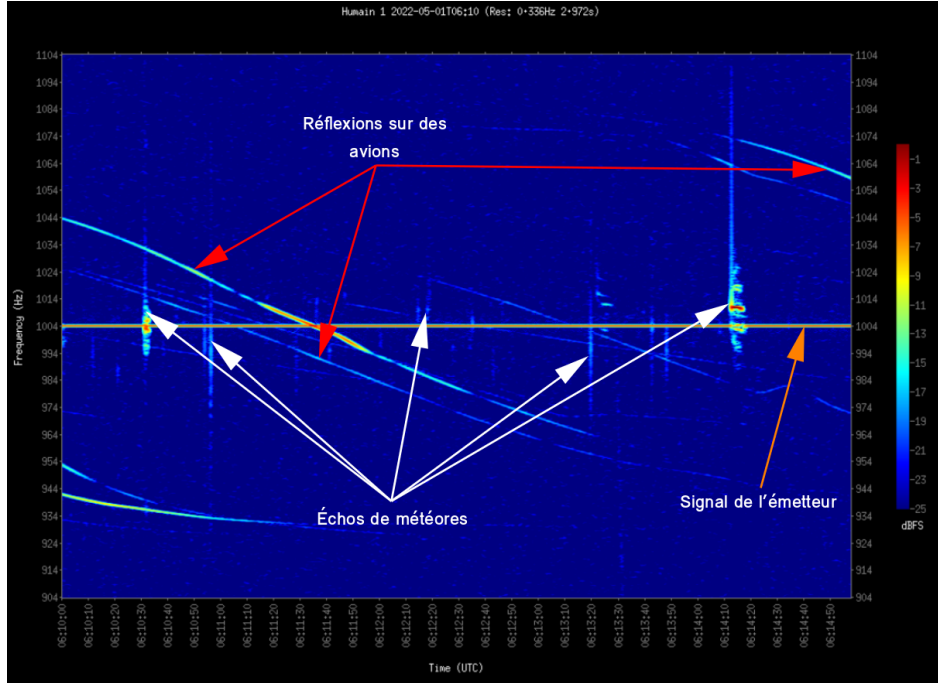


Figure 6: Exemple d'un spectrogramme du projet BRAMS.

### 1.1.5 L'Archive BRAMS

Quand les fichiers venant des stations de réception sont archivés, ils sont placés dans une structure de répertoires spécifique. Cette structure dépend de la date de début d'enregistrement des données contenues dans le fichier.

Dans la première couche de l'archive, on retrouve un répertoire par station BRAMS. Chaque nom de répertoire est le code de location de la station dont elle contient les données. Ensuite, on retrouve un nombre de répertoires, chacun contenant les fichiers d'une année spécifique. Le nom de ces répertoires est donc l'année des données qu'elles contiennent. Si on rentre dans le répertoire d'une année, on se retrouve dans la troisième couche de l'archive BRAMS. Ici, les données sont encore une fois séparées selon le mois d'enregistrement du fichier. Finalement, dans la quatrième couche de l'archive, on retrouve un répertoire par jour d'enregistrement d'un fichier.

Dans chaque répertoire représentant un jour spécifique, on retrouve les don-

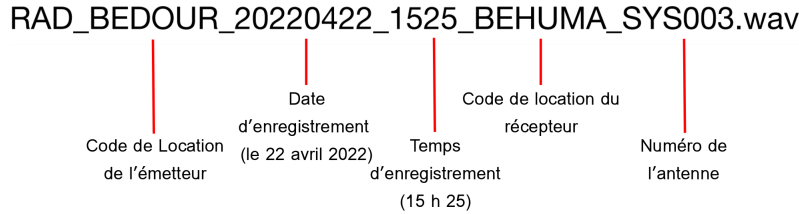


Figure 7: Exemple de nom de fichier BRAMS.

nées BRAMS organisés par heure dans des fichiers de format TAR. Ce format permet de regrouper plusieurs fichiers dans un même fichier, qui prend alors le format TAR. Le nom des fichiers TAR et les fichiers WAV qui se trouvent à l'intérieur des fichiers TAR suivent également une structure fixe. Il est composé de six informations ordonnées de la manière suivante :

1. Les trois lettres **RAD**.
2. Le code de la station émettrice. Comme il n'y a qu'une station réceptrice à présent, la valeur sera toujours **BEDOUR**.
3. La date du début de l'enregistrement du fichier. Elle respecte le format **YYYYMMDD**.
4. Le temps du début de l'enregistrement du fichier, précis à la minute. Elle respecte le format **HHMM**.
5. Le code de location de la station émettrice. Un code de location est composé de deux informations. Les deux premiers caractères représentent le code du pays où se trouve la station (e.g. BE, FR, ...) et les quatre lettres restantes représentent typiquement les quatre premières lettres de la ville où est situé la station (e.g. HUMA pour Humain, UCCL pour Uccle, ...).
6. L'alias de la station. Elle est composée des trois caractères 'SYS', suivie du numéro de l'antenne qu'a capté les données. Le numéro de l'antenne est toujours écrit sur trois chiffres (e.g. 001, 002, ...).

Toutes ses valeurs sont séparées par un tiret bas. Un exemple de nom de fichier est affichée à la figure 7.

## 1.2 Problématique liée à l'étude des données BRAMS

Afin de pouvoir trouver la trajectoire d'un météore dans l'atmosphère, il est nécessaire de rechercher par quelles stations réceptrices le météore est détecté. Ceci est une procédure qui prend un temps non négligeable pour les scientifiques du projet. En effet, toutes les stations de réception ne détectent pas tous les météores. De plus, certaines données peuvent être corrompues ou encore, non-existantes.

De plus, savoir, ou au moins avoir une idée, si un météore détecté par une station réceptrice est le même qu'un météore détecté sur une autre station réceptrice serait une grosse économie de temps.

C'est pour cette raison que les membres du projet BRAMS souhaitent disposer d'un programme qui remplit deux fonctions. D'une part, elle doit être capable de vérifier si un fichier contient des données non-corrompues et utilisables. D'autre part, elle devra, sur base d'un écho de météore d'une station de réception, retrouver les échos de ce même météore sur des fichiers d'autres stations de réception.

## 2 Méthodologie

Afin d'arriver à un résultat final de qualité et qui est conforme aux requis des membres du projet BRAMS, il est important d'utiliser une bonne méthodologie.

La première chose que j'ai fait est : comprendre quoi exactement le programme devrait faire. Pour ce faire, plusieurs réunions ont eu place avant la réalisation du travail. Durant ces réunions, j'ai pu poser mes questions et demander des explications sur les concepts à connaître pour pouvoir réaliser le travail. Ayant reçu des documents qui expliquent de façon claire et précise le fonctionnement du réseau BRAMS, j'ai pu me préparer avant de m'attaquer à l'analyse et la réalisation du travail.

Durant la période d'analyse et de réalisation du TFE, j'ai pu régulièrement demander validation quant à la direction que je prenais pour mon travail. Je présentais régulièrement mes réalisations aux scientifiques du projet BRAMS afin d'avoir un feedback régulier et de pouvoir perfectionner mon programme un maximum.

Dans le but d'être plus efficace lors de l'écriture des fonctionnalités pour le programme, je décrivais avant ce que cette fonctionnalité devait faire. Ensuite, je la divisais en tâches techniques afin de pouvoir m'organiser plus facilement. Chaque tâche technique contenait et expliquait les étapes que le programme devrait exécuter pour accomplir cette tâche technique. Les étapes étaient décrites textuellement, par pseudo-code, par schéma ou encore, par le mélange de deux ou trois des moyens cités.

Durant l'entièreté du projet de fin d'études, j'ai tenté de garder un rythme de travail régulier. Je me suis organisé de telle façon de pouvoir travailler une moyenne de 23 heures par semaine. Une grande partie de ces heures se sont déroulés lorsque je rentrais de mon stage ou pendant le week-end.

## 3 Technologies utilisées

### 3.1 Python

Le programme réalisé est entièrement écrit en Python. Ce choix a été pris premièrement pour les librairies performantes et open-source qu’offre Python. En effet, des librairies comme numpy, scipy ou encore matplotlib ont été très utiles pour arriver à un résultat performant et fonctionnelle. Ces librairies offrent beaucoup de flexibilité et offrent une haute performance grâce à leur écriture en C, C++ et même en Fortran.

Un autre avantage du Python est sa portabilité. Puisque c’est un langage interprété, largement répandu, quasiment tous les systèmes d’exploitation le supportent. Ceci a facilité aussi bien la phase de développement que la phase de testing du projet de fin d’études.

Un des désavantages souvent évoqués pour Python, est que le langage est lent pour des gros traitement de données. Cependant, dans le cas de ce travail, ceci ne fut pas un problème et les librairies utilisées étaient largement assez rapide. Ceci est dû, entre autres, grâce à la nature open-source de ces librairies ainsi que leur âge. Suite à ces deux facteurs, des milliers de personnes travaillent sur ces librairies depuis plusieurs années dans le but d’optimiser le plus possible ses fonctions et méthodes.

#### 3.1.1 Librairies Utilisées

Dans cette section sont listées toutes les librairies Python utilisés pour la réalisation du travail de fin d’études.

##### **Argparse**

Argparse est une librairie Python qui permet de facilement créer une interface ergonomique en ligne de commande. On spécifie les paramètres nécessaires et optionnels pour le programme qu’on crée, et lorsqu’on lance le programme, Argparse pourra retrouver les paramètres entrés par l’utilisateur. Il indiquera également une erreur si un paramètre obligatoire manque, ou qu’un type de paramètre n’est pas respecté. Argparse est une librairie installée par défaut avec Python.

##### **Datetime**

La librairie Datetime offre une suite de fonctions et de classes permettant la manipulation et la conversion de temps et de dates. Cette librairie fait

partie des modules standards de Python et ne doit donc pas être installé.

### **GeoPy**

GeoPy est une librairie Python qui facilite la recherche de coordonnées d'adresses, de villes ou encore de lieux connus. Elle permet également de trouver la distance entre deux lieux en donnant les coordonnées de ces endroits.

### **Math**

Le module math offre toute une série de fonctions mathématique de base. Elle est installée par défaut avec Python.

### **Matplotlib**

La librairie Python Matplotlib permet de générer des graphiques et de les enregistrer en tant qu'image.

### **NumPy**

NumPy est une librairie Python utilisée lorsqu'on travaille avec des listes très larges. Il contient de nombreuses fonctions pour l'algèbre linéaire, les Transformées de Fourier et les matrices. Par rapport aux listes classiques de Python, les listes NumPy (aussi appelés des ndarray) sont beaucoup plus rapides.

Ceci est dû à trois raisons principales. La première est que les listes NumPy sont stockés à un endroit continu dans la mémoire RAM<sup>4</sup>, ce qui permet un accès et une manipulation plus efficace des données. Ensuite, comme NumPy est une librairie open-source, il est continuellement adapté à des nouvelles architectures de processeurs. Finalement, une grande partie de la librairie NumPy est écrit dans des langages de bas niveau tels que le C, le C++ ou encore le Fortran. Ceci résulte à un code beaucoup plus proche du langage machine et qui est donc plus optimisé.

### **Os**

La librairie os de Python contient une suite de fonctions permettant d'interagir avec le système d'exploitation. Cette librairie vient par défaut avec une installation de Python, ce qui évite de devoir l'installer lorsque nécessaire.

### **SciPy**

Scipy est une librairie qui est basé sur la librairie NumPy. Elle partage donc

---

<sup>4</sup>Random Access Memory



de nombreuses fonctions et méthodes avec cette dernière et fonctionne également avec des `ndarray`. Cependant, cette librairie ajoute des fonctions pour la science des données. De plus, Scipy optimise des fonctions déjà présentes dans la librairie NumPy. Ceci permet de construire un programme qui demande moins de ressources machine et qui est donc plus rapide.

### **Simplejson**

Simplejson est une librairie Python offrant la possibilité de facilement écrire et lire des fichiers JSON<sup>5</sup>. Elle vient par défaut avec Python sous le nom `'json'`. Cependant, installer Simplejson séparément ajoute l'avantage d'avoir des mises à jour plus régulières et de disposer des dernières fonctionnalités plus rapidement.

### **Tarfile**

Le module Python Tarfile permet de lire et écrire des fichiers du format TAR. Cette librairie est incluse par défaut avec Python et ne doit donc pas être installée séparément.

### **Tqdm**

Tqdm est une librairie Python permettant d'afficher des barres de chargements en ligne de commande. Elle permet également de donner une estimation de temps de chargement restant.

## **3.2 MariaDB**

Afin de sauvegarder les données produites par le programme, une base de données est requise. Les données seront enregistrées dans la base de données existante du projet BRAMS. Le type de base de données du projet BRAMS est MariaDB.

---

<sup>5</sup>JavaScript Object Notation

## 4 Lecture des Fichiers BRAMS

Avant de commencer à développer des outils pour la détection des météores ou pour l'analyse des fichiers BRAMS, il a fallu développer un logiciel permettant de lire ces fichiers. Ce logiciel sera implémenté dans un module Python contenant une classe nommée `BramsWavFile`. Ceci veut dire qu'il pourra facilement être utilisé par d'autres fichiers Python et qu'il doit donc être flexible. Dans cette section, le fonctionnement de ce module est expliqué en détail.

### 4.1 Fonctionnement du Module

Le contenu principal du module est donc la classe `BramsWavFile`. Cette classe permet de lire et de récupérer des informations sur les fichiers WAV du réseau BRAMS de manière efficace. Elle est basée sur un code écrit par Michel Anciaux, membre du projet BRAMS. Cependant, elle a fortement été retravaillée afin de l'adapter aux besoins du programme et de l'optimiser. La classe prend six paramètres en entrée :

- Le paramètre **`date_time`** indique le temps et la date du fichier recherché. Ce paramètre est obligatoire et doit être de type `datetime`, type qui est défini dans le module `datetime`.
- Le paramètre **`station`** indique le code de location de la station recherchée (e.g. BEHUMA, BEHAAC, BEUCCL). Les deux premières lettres de ce code représentent le code pays (e.g. BE, FR). Typiquement, les quatre lettres suivantes sont les quatre premières lettres de la ville où se trouve la station. Ce paramètre est obligatoire et doit être de type `str` (chaîne de caractères).
- Le paramètre **`alias`** indique l'antenne, se trouvant à la location spécifiée, dont on veut récupérer un fichier. En effet, certaines stations disposent de plusieurs antennes. Les antennes sont spécifiées à l'aide de six caractères : les trois premiers contiennent les lettres 'SYS' et les trois suivants contiennent le numéro de l'antenne (e.g. 001, 002, ...). La valeur par défaut de ce paramètre est 'SYS001', c'est donc un paramètre optionnel. Le type du paramètre est `str` (chaîne de caractères).
- Le paramètre **`is_wav`** est un booléen indiquant à la classe si le fichier recherché se trouve dans un fichier de format TAR (dans quel cas la

valeur devrait être 'faux') ou pas (dans quel cas la valeur devrait être 'vrai'). Ce paramètre est optionnel et prend la valeur 'faux' par défaut.

- Le paramètre **parent\_directory** indique le répertoire parent ou il faudra commencer la recherche du fichier demandé. C'est un paramètre optionnel de type str (chaîne de caractère) qui prend le répertoire de l'archive BRAMS comme valeur par défaut.
- Le paramètre **from\_archive** est un booléen indiquant si le fichier recherché se trouve dans un répertoire respectant la structure de l'archive BRAMS (dans quel cas la valeur est 'vrai') ou non (dans quel cas la valeur est 'faux'). Elle prend la valeur 'vrai' par défaut et est donc optionnel.

Lorsque la classe est initialisée, elle vérifie d'abord si le fichier recherché existe. Si le paramètre 'from\_archive' contient la valeur 'vrai', la classe rajoute, au chemin spécifié par le paramètre 'parent\_directory', les dossiers nécessaires pour que le chemin respecte la structure de l'archive BRAMS. Ensuite, le programme essaye de lister les contenus du dossier. Dans le cas où le dossier n'est pas trouvé, l'initialisation de la classe est interrompue et l'exception 'DirectoryNotFoundError' est levée.

La classe recherche ensuite pour le bon fichier dans le contenu du répertoire en comparant la date et le temps dans le nom du fichier avec la date et le temps reçu comme paramètre. Quand les fichiers WAV sont contenus dans une archive TAR, l'étape précédente est répétée pour trouver le bon fichier dans l'archive. Si aucun fichier correspondant à la date et la station reçue en paramètres est trouvé, l'initialisation de la classe est interrompue et l'exception 'BramsError' est levée. Une fois le bon fichier trouvé, toutes les données brutes sont lues et enregistrées dans un buffer (variable).

La prochaine étape est la séparation des différents blocs de données (ou chunk) contenus dans le fichier WAV. Pour ça, la classe parcourt le buffer contenant les données du fichier WAV. Il utilise un pointeur de lecture indiquant la position en octets du prochain bloc de données dans le buffer. Ce pointeur est calculé à partir du champ 'subChunkSize' indiquant la longueur du bloc de données courant.

Lorsque le programme retrouve un bloc de données avec un champ ID correspondant à 'fmt', 'BRA1' ou 'data', il extrait toutes ces données et les place dans un ndarray. Ceci est fait à l'aide de la fonction 'frombuffer' venant de

la librairie NumPy. Cette fonction offre la possibilité de convertir, de façon très rapide, des données vers un ndarray. Une des contraintes pour utiliser cette fonction est l'indication du type de données qu'on veut placer dans un ndarray.

Les données contenues dans la classe seront la fréquence d'échantillonnage, extrait du bloc de données 'BRA1', et la piste audio du fichier WAV, extrait du bloc de données 'data'.

#### **4.1.1 La méthode FFT**

La classe BramsWavFile dispose d'une méthode publique. Cette méthode se nomme 'FFT' et permet, comme son nom l'indique, de récupérer la Transformée de Fourier de l'entièreté de la piste audio du fichier WAV. Ceci est fait à l'aide de la fonction 'rfft' de la librairie Scipy, multipliée par une fenêtre de Hann. La fonction 'rfft' existe également dans la librairie NumPy, cependant dans ce cas de figure la librairie SciPy permettait de calculer la Transformée de Fourier plus rapidement. Enfin, la fonction FFT retourne la Transformée de Fourier normalisée, son abscisse et sa résolution fréquentielle.

## 5 Monitoring des Données BRAMS

Comme indiqué précédemment, les fichiers ne contiennent pas toujours des données utilisables. Pour des diverses raisons, il peut parfois arriver que le bruit est trop fort pour pouvoir détecter des météores. Ou encore, lorsque les stations ICOM arrivent en fin de vie, la puissance des données détectées décroît de telle sorte qu'on détecte du bruit seulement. Bref, de nombreux problèmes peuvent survenir avec les données. Pour optimiser la détection de météores, les données corrompus doivent être détectés avant celle-ci. C'est pourquoi, un outil permettant le monitoring des données a été développé.

Cet outil se base sur deux éléments dans les données BRAMS indiquant si un fichier est utile ou non. Le premier est le bruit. En effet, lorsque le bruit augmente ou diminue beaucoup d'un coup ou graduellement sans revenir à un niveau normal, il est fort probable qu'il y ait un problème avec la station réceptrice ou son environnement.

Le deuxième élément permettant de détecter un problème est le signal du calibre. Le signal calibre devrait toujours être présent dans un fichier produit par une station réceptrice. Typiquement elle devrait se trouver autour de la fréquence 1500 Hz. Cependant, avec les anciennes stations, elle peut varier jusqu'à 250 Hz dépendant de la chaleur. Dans le cas où elle n'est pas aux alentours de 1500 Hz, la probabilité qu'il y ait une erreur avec la station est grande.

### 5.1 Fonctionnement du Programme

Le programme permet donc de mesurer le niveau du bruit et le signal calibre d'un ou de plusieurs fichiers WAV du projet BRAMS. Il fait ceci en calculant la densité spectrale de puissance (dsp) de ceux-ci. La densité spectrale de puissance montre la distribution des puissances à travers les différentes fréquences contenues dans le signal de base. Ceci permet de comparer ces différentes valeurs et de remarquer des variations non typiques.

Lorsqu'on lance le programme, trois paramètres peuvent être ajoutés :

1. Le paramètre **START DATE** indique la date à partir de laquelle on veut mesurer la dsp des fichiers. Par défaut, elle prendra la date d'un jour avant le lancement du programme.

2. Le paramètre **END DATE** indique la date avant laquelle on veut mesurer la dsp des fichiers. Dans le cas où le paramètre 'START DATE' n'est pas donné, elle prend la valeur de la date de lancement du programme.
3. Le paramètre **STATIONS** indique les stations pour lesquelles on veut mesurer la dsp des fichiers. Si ce paramètre n'est pas donné, le programme mesurera la dsp des fichiers de toutes les stations.

Le programme calculera donc la dsp des fichiers de tous les fichiers générés entre **START DATE** et **END DATE** venant des stations **STATIONS**.

La première va donc, pour chaque fichier demandé, performer une suite d'étapes. Premièrement, elle va chercher le fichier à l'aide de la classe 'BramsWavFile'. Si une erreur se produit lors de la recherche du fichier, on considère que le fichier n'existe pas et le programme passe au fichier suivant. Lorsqu'on a le fichier, on peut calculer la dsp du bruit et du signal calibreur.

Dans le cas du bruit, on prend toujours la moyenne de la dsp entre 800 Hz et 900 Hz. Entre ces deux fréquences on retrouve normalement uniquement du bruit. Dans les cas où il se trouverait qu'il y ait un signal autre que le bruit, il ne devrait pas avoir grand impact sur la dsp puisque cette dernière se calcule sur 100 Hz, sur l'entièreté d'un fichier.

Pour le signal calibreur, on ne suffit pas de calculer la moyenne de la dsp à 1500 Hz. En effet, la fréquence du calibreur peut varier jusqu'à 250 Hz, autour de la fréquence de 1500 Hz, d'un fichier à l'autre. Comme le signal calibreur est le seul signal se trouvant entre 1350 Hz et 1750 Hz, il suffit de rechercher la fréquence, entre ces deux valeurs, avec la valeur de dsp maximale.

Enfin, le programme enregistre toutes les données de dsp dans la base de données BRAMS existante. Chaque valeur de dsp est alors liée à son fichier. Si l'option **-p** (**-plot**) a été donné au lancement du programme, le programme génère également un graphique représentant l'évolution de la dsp au cours du temps.

## **5.2 Résultats**

Lors des tests, quelques résultats ont déjà montrés des informations intéressantes. Un bon exemple est l'antenne quatre de la station à Humain entre

## 6 Détection des Météores

Finalement, on arrive à la détection des météores. Ce programme effectue deux tâches principales. La première tâche consiste à, à partir des coordonnées d'un météore donné, détecté sur une station donnée, trouver tous les échos qui viennent potentiellement du même météore. Ensuite, la deuxième tâche du programme est de donner l'utilisateur le plus d'informations possibles pour confirmer qu'un écho correspond à l'écho de météore reçu en entrée ou non.

### 6.1 Analyse

C'est ici la première étape qui pose la plus grande difficulté. En effet, la détection de météores consiste à retrouver des objets d'une forme spécifique dans une image. Cette image est, dans ce cas si, un spectrogramme calculé à partir d'un fichier WAV venant d'une station de réception BRAMS. De plus, on retrouve sur ce spectrogramme un nombre d'autres parasites que le programme doit pouvoir différencier d'un vrai météore tel que les échos d'avions qui passent, ou encore le signal direct de l'émetteur.

Afin de faciliter la recherche d'échos de météores dans un spectrogramme, l'amplification de ces échos peut faciliter la tâche. Un écho de météore est, dans un spectrogramme, caractérisé par un signal qui s'étale sur minimum 10 Hz et qui dure typiquement entre une et trois secondes. Ceci est possible en performant une convolution sur le spectrogramme. Une convolution est une opération mathématique qui prend deux fonctions en entrée et qui produit un nouveau signal. Dans ce cas si, le premier signal en entrée sera le spectrogramme, ou une partie du spectrogramme, et le deuxième signal est le filtre. Afin d'arriver à un bon résultat, il a fallu tester plusieurs filtres. Le filtre offrant les meilleurs résultats est illustré à la figure N. Elle permet de faire deux choses :

- Premièrement, elle amplifie les éléments qui font plus de 10 Hz. En effet, pour avoir un résultat d'une valeur maximale il faut que le haut et le bas du filtre soient multipliés par des grandes valeurs.
- Deuxièmement, elle permet de diminuer l'intensité des éléments qui ont une longue forme horizontale (typiquement des échos d'avions).

Suit alors la suppression du signal direct de l'émetteur. En effet, pour détecter les échos de météores, le signal direct nous complique la tâche.



Pour l'éliminer, le programme recherche d'abord sa fréquence entre 800 Hz et 1200 Hz, en prenant la valeur maximale sur l'entièreté de la durée du fichier, entre ces deux fréquences. S'il n'est pas retrouvé, on considère qu'il n'est pas détecté (ce qui est le cas sur certaines stations réceptrices). Par contre, dans le cas où il est détecté, pour chaque seconde, on remplace la valeur du signal direct par la moyenne des valeurs se trouvant autour de la fréquence du signal direct.

Ensuite, il faut trouver une méthode permettant de séparer, dans le spectrogramme, les objets du bruit. La première méthode testée se basait sur un seuil fixe. Toutes les valeurs qui se trouvaient en dessous de ce seuil étaient considérées comme du bruit et leur valeur était mise à zéro. Bien que cette méthode fonctionnait bien sur un fichier spécifique, elle posait problème si on voulait la lancer sur plusieurs fichiers différents. En effet, comme toutes les stations de réception se trouvent dans des environnements différents et disposent parfois même de matériel différent, différent fichier ont des niveaux de bruits différents.

Pour la deuxième méthode, le programme essayait de trouver un seuil qui se base sur la valeur du bruit sur l'entièreté d'un spectrogramme. Pour ce faire, la méthode divisait le spectrogramme, entre les fréquences de 600 Hz et 1400 Hz, en trente zones. Elle mesurait ensuite, à l'aide de la variance de chaque zone, laquelle contenait le moins d'objets possibles. Finalement, la méthode calculait la 95e percentile du bloc contenant le moins d'objet possibles. Toutes les valeurs du spectrogramme se trouvant en dessous de cette valeur seraient considérés comme du bruit et mises à zéro. Malgré que cette méthode produisait des résultats marginalement meilleurs que celle à seuil fixe, elle n'était toujours pas parfaite étant donné qu'elle se basait sur le bruit d'une partie du spectrogramme, or le bruit peut varier au cours du temps.

Finalement, la méthode qui fonctionnait le mieux calculait la 95e percentile pour chaque colonne du spectrogramme. Ensuite, elle remplace toutes les valeurs de cette colonne se trouvant en dessous de la 95e percentile avec zéro. Les résidus de bruits qui ne sont pas encore éliminés sont négligeables puisqu'ils s'étalent sur moins de 10 Hz et ne sont donc pas considérés comme des météores.

Il faut maintenant éliminer tous les objets qui ne sont pas des échos de météores, mais qui s'étalent sur plus de 10 Hz. Ceci n'est pas une tâche facile puisque des parasites peuvent venir se mélanger à des échos de météore formant un objet. La première chose à faire est : nettoyer le spectrogramme des

petits parasites, ou encore de résidus de bruits. Le programme fait cela en labellisant le spectrogramme et en éliminant tout objet s'étalant sur moins de 10 Hz. Ensuite, on récupère tous les objets restants. Pour chacun de ces objets, on vérifie s'il n'est pas composé de plusieurs parasites se superposant sur plus de 10 Hz. Ceci est fait en allant vérifier sur une longueur d'environ treize secondes autour de l'objet détecté, si on retrouve des parasites qui s'étalent sur la même fréquence que cet objet. Lorsque c'est le cas, l'objet n'est pas considéré comme un météore et est ignoré. À partir de ce stade, tous les objets détectés sont considérés comme étant des échos de météores.

Pour la deuxième étape, le programme récolte le plus d'informations possibles à propos des échos pouvant potentiellement correspondre à l'écho de météore reçu en entrée. Parmi ces informations, on retrouve tout d'abord le temps où un écho s'est produit. Ceci est important pour pouvoir situer l'écho si on veut l'étudier en profondeur.

Ensuite, le programme calcule la distance entre la station qui a détecté l'écho reçu en entrée et la station dont un écho pourrait correspondre à celui en entrée. En effet, un météore n'est pas détecté par toutes les stations, et dans le cas où il est détecté par plusieurs stations, ces stations se trouvent souvent proches les uns des autres.

La troisième information donnée indique le nombre d'échos trouvés pour une station, pouvant correspondre à l'écho en entrée. Dans ce cas, il est important d'en informer l'utilisateur, puisque ces échos viennent chacun d'un météore différent et ne peuvent donc pas tous venir de cet écho reçu en entrée.

La dernière information donnée par le programme sont la fréquence minimum et la fréquence maximum des échos détectés. Ces deux fréquences permettront, ensemble avec le temps où l'écho s'est produit, de faciliter la recherche de cet écho pour l'utilisateur.

## 6.2 Fonctionnement du Programme

Le programme prend en entrée le temps et la date de l'écho du météore recherché ainsi que le code de location de la station qui l'a détecté. À partir du temps et de la date reçue en entrée, le programme calcule l'intervalle de temps dans laquelle elle doit chercher des échos pour les autres stations réceptrices. Cet intervalle est toujours de six secondes autour du temps reçu en entrée. Cette valeur a été choisie puisque toute détection en dehors de ces six secondes a très peu de probabilité de venir du même météore que l'écho en entrée. Toutes les actions sur le spectrogramme s'appliqueront

uniquement à l'intérieur de cet intervalle afin d'optimiser le programme.

Le programme va ensuite chercher, pour toutes les stations, les informations des fichiers contenant cet intervalle. Il fait pour ça une requête à la base de données BRAMS. Puis, pour chaque station, le programme calcule la distance entre la station et la location où l'écho de météore reçu en entrée a été détecté.

Quatrièmement, pour tous les fichiers dont on a récolté des informations, on va rechercher les météores dans l'intervalle calculé précédemment. Pour cette étape, le programme fait d'abord appelle à la classe `BramsWavFile`, qui ira chercher le fichier et extraire les données audio. L'étape suivante consiste à générer le spectrogramme à partir de ces données audio. Afin de générer ce spectrogramme, il faut faire appelle à la classe `Spectrogram`. Ceci est une classe spécialement faite pour ce logiciel, qui permet de créer un spectrogramme et de le modifier à l'aide d'une suite de méthodes. Lorsqu'on appelle cette classe, il faut passer comme paramètre les données audio. À partir de ces données, la classe va créer le spectrogramme du fichier avec la fonction `'scipy.signal.spectrogram()'`. La classe gardera une copie originale du spectrogramme et une copie pour les modifications, dont elle supprime, dès l'initialisation, le signal direct de l'émetteur.

Suit alors une convolution du spectrogramme avec le filtre permettant d'amplifier les échos de météores, cette convolution est faite à l'aide de la fonction `'scipy.ndimage.convolve()'`. Ensuite, le filtre supprimant toutes les valeurs en dessous de la 95e percentile sera appliquée. Afin d'éliminer tous les résidus de bruits, le programme labellise le spectrogramme et supprime tout objet s'étalant sur plus de 6 Hz. Dans le but de reconnecter les météores qui auraient été divisés en deux parties différentes, on applique un filtre moyenneur sur le spectrogramme. Puis, le logiciel cherche les coordonnées des échos de météore sur le spectrogramme. Avec ces coordonnées, on pourra alors retrouver la fréquence maximale et la fréquence minimale de l'écho.

Enfin, le programme crée un fichier CSV<sup>6</sup> qui contiendra une ligne par écho de météore détecté. Chaque ligne contiendra :

1. Le code de location de la station où est détecté l'écho.
2. Le numéro de l'antenne qui a détecté l'écho.

---

<sup>6</sup>Comma Separated Values

3. La date et le temps de début d'enregistrement du fichier d'où vient l'écho.
4. Le nombres d'écho comptés dans le même fichier que celui d'où vient l'écho.
5. Le temps, précis à la microseconde, de détection de l'écho.
6. La fréquence minimum de l'écho.
7. La fréquence maximum de l'écho.
8. La distance entre la station où est détecté l'écho et la station de l'écho reçu en entrée.

De cette façon, les données pourront facilement être lues et réutilisées pour étudier ces météores.

## 7 Pistes d'Améliorations

Suite à une réflexion sur les programmes réalisés, j'ai pu trouver quelques pistes d'améliorations qui sont intéressantes pour le futur :

-

## 8 Conclusion