Acoustic pollinator detection to measure the pollination activity

**Abstract**

**Introduction**

Plusieurs méthodes de mesure de l’activité de pollinisation (pan trap, sweep net…)

Mais ces méthodes ont des lacunes => limite les questions qu’on peut poser

Bio acoustique utilisée pour d’autres taxa (chiro, oiseaux, criquet, etc…)

Paragraphe qui annonce ce qu’on va faire dans le papier

**Material & Methods**

**Site description & sunflower fields selection**

We selected 30 sunflower fields from the Long-Term Social-Ecological Research site “Zone Atelier Plaine et Val de Sèvre” located in South-West France. This site was an area of 435 km², on which 87% was cultivated (REF: Bretagnolle et al., 2018). This site was dominated by cereals (Valeur en 2020) and sunflower represents 18% (a verif) of the cultivated area. Sunflower fields were selected according to landscape gradients (coverage of meadows, hedgerows, tillage, semi-natural habitats, and organic farmland) which were calculated within a 1km radius of the field's centroid. It was also ensured that the flowering of these fields was not simultaneous to be able to monitor the entire flowering period of the crop. 11 of these fields were cultivated organically. (voir si ajout période floraison plus précise (date moyenne, taille des parcelles…))

**Recording method**

In each field, we placed an audiomoth(REF), which is a microphone (see …. For the description and parameters), between 10 to 20cm in front of a sunflower plant, randomly selected, situated at approximately 20m of the edge. The audiomoth was covered by a windscreen to reduce the sound of the wind. We made recordings between 6 am to 10 pm. (ajout nombre d’heures d’enrregistrement par parcelle min/max, …)

**Focal**

To calibrate the acoustic method, we needed to know some recordings in which insects had come. For this, we coupled the audio recordings with visual observation. The observer was placed at about 50 centimeters from the plant, in front of the capitulum. For five minutes, every 15 seconds, each insect close to the sunflower capitulum was noted. For each pollinator, its guild (honeybee, Bombus, other wild bees, hoverfly, other Diptera, Coleoptera, other) and its position (on male flowers, on female flowers, around the capitulum) were noted. These observations were repeated 3 times per week before, during, and after the flowering period of the capitulum in each field. During a week, 3 observations were made, one in the morning, one in the middle of the day, and one at the end of the day. We had a total of 291 visual observations coupled with audio recordings.

**A reference database building**

To analyze our recordings, it was necessary to cut them into little recordings of 10 sec thanks to Kaleidoscope(REF) software and to increase the time/frequency ratio. For this, we accelerated all recordings by 10 times. Then, we used Tadarida-D (ref) to detect sound events, which were isolated by frequency and time, on all recordings. To train Tadarida to recognize automatically the sound emitted by pollinators, we build a reference sound database. For this, 12 rounds of visual observation coupled with recordings were used. In these rounds, we knew that pollinators had been seen for some. All the sound events that were detected and that we could recognize by listening to them were labeled with Tadarida-L (REF Github). After this step, we had 14 (a verif) classes of sounds (see the Table .. to see the list of classes and the number of sound events we had labeled per class) and, thanks to Tadarida-C(ref GitHub), a first classifier that we could use to at least roughly detect pollinators in other recordings. To increase our reference sound database, all recordings, located outside the visual observation periods, were classified according to the field in which the plant was located (30 levels), the hour of the recording (16 levels), and the flowering stage of the capitulum (3 levels: not flowering, flowering, deflowering). Within each class, 10% of the recordings were randomly selected. We classified all sound events detected in these recordings based on our first classifier. We obtained for each sound event a confidence score for the "pollinator" class. Higher is this score, the more the "probability" that the recording contains a sound emitted by a pollinator is higher. We defined 6 classes for the confidence score for the pollinator class (1: 0-0.5; 2: >0.5-0.6; 3: >0.6-0.7; 4: >0.7-0.8; 5: >0.8-0.9; 6: >0.9). We conducted stratified random sampling by field and pollinator class confidence score, with an increasing number of samples recordings as the confidence score class increased (mettre le nb d’enregistrements théoriques tiré par classe et le nombre reel en tableau peut-être). All of these recordings were listened to and all sound events were labeled with Tadarida-L. We obtained a second reference sound database with (Nouvelles classes + nb d’évènements sonores par classe en tableau?).

Voir ou rajouter l’ajout de voix humaine

**Building of the classifier**

To find the best classifier to recognize automatically the sound emitted by pollinators during their flight, we tested two different methods of machine learning. The first allows us to build a table containing the acoustic features of all labeled sound events and to build the classification trees necessary for the random forest method used by Tadarida-C on R(ref) to build classifiers. The second is based on picture recognition and used TadaridaDeep(ref) on Python(ref) which is based on a deep learning process to build a classifier. This method allows classifying recordings that have a frequency between 0.5 to 20 kHz and duration greater or equal to 9 ms.

**Classifier testing**

We tested the efficiency of the classifiers on recordings from visual observations coupled with recordings not yet labeled. To be able to make comparable tests between classifiers, we classified all the sound events of these recordings with the random forest method and we randomly selected 100 recordings per score class described above. We listened to all of these recordings and determined if they contained at least one sound made by a pollinator. We counted 94 recordings that contained at least one pollinator sound.

**Results**

Courbe ROC meilleur classif avec AUC (résultats des autres classif en annexes ?)

**Discussion**