Assessing the identification of earthworm categories by non-expert samplers in a citizen science project

Lola Morand[1], Diane Fiala[1], Nathan Lévêque[1], Kevin Hoeffner[1], Sarah Guillocheau[1,2], Daniel Cluzeau[1]

1. University of Rennes, CNRS, ECOBIO [(Ecosystèmes, biodiversité, évolution)] - UMR 6553, Rennes, France
2. Micro-company, MME Guillocheau Sarah, Mauron, France

Context & problematic

There is currently **no global framework allowing to understand the impact of climate, soil condition and anthropological pressure on earthworm communities**. Therefore, the specialists need to obtain data on global scale in order to understand the dynamics of these factors. Besides, since the 2000's, public policies, associated with the growing interest of the public on environmental issues, led to numerous partnerships. Laboratories and citizen organizations combine expertise and working force to bring various interests for both sides :

Pedagogical interests :
- Bring citizens a more accurate **perception of soil ecology**
- Improvement of **awareness in environmental issues**

Citizen science

Research interest :
- Access to data of various and large territories

Is the data quality suitable ?

However, using data from citizen science project raises questions about their reliability while only a few studies exist on that subject. This being the case with **OPAL (Open Air Laboratory),** which, among others, assess the **identification skills** of citizen samplers in Great Britain. The study (1) showed an average of 88% of success rate in the identification of earthworm categories. But is this applicable to other contexts and territories ?
Earthworm identification requires complex skills and knowledges. Various tools are used by scientist to help to empower the non-expert citizen on this matter such as guides or identification keys.

Do non-expert identification data from citizen science project can be used on research question ?

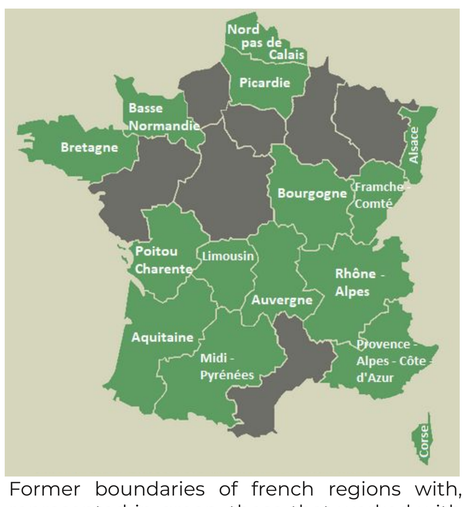
Earthworm citizen science projects in Europe still in operation today :

- **OPAL (Open Air Laboratory)**
- **The OPAL Soil and Earthworm Survey** (is one of five OPAL surveys across England to learn more about the state of the environment)

- **OPVT (Participative Earthworm Observatory)**
- **EcoBioSoil** (collaborative platform dedicated to the study of the biodiversity of soils in temperate environments)

Materials & methods

For this study a part of the SBT-ENI data was used. **SBT-ENI** is a **monitoring survey of the impacts of agricultural practices** on several indicators including earthworms. The **University of Rennes** contributes to this program for earthworms diversity identification.



Former boundaries of french regions with, represented in green, those that worked with the university between 2012 and 2018 (mustard protocol).

15 regions that worked with University of rennes

A group of **samplers** per region
↳ Organised by local authorities

Training & tools
↳ **Training sessions** and identification guides are provided.

Annual sampling since 2012

For a total of **1100 sampling plots**

Year	2012	2013	2014	2015	2016	2017	2018
Number of samples	85	233	145	169	220	117	140

For the boxplot and bar graphs the 1100 sampling sessions were grouped into "one region/one year" groups resulting **in 57 groups**.

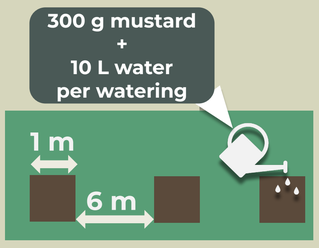
Identification focuses on **ecological categories** but it also goes further for **the anecics by dividing it in two categories**. Thus anecic *Lumbricus* and anecic *Aporrectodea,* are constituting, with epigeics and endogeics, the 4 **earthworm categories** studied.

Knowing their distribution provides informations about their functions on soils. Also, studying the composition of earthworm communities brings an understanding on the vulnerabilities of these categories under the pressure of various factors.

1 Sampling

The samplers collect the earthworms using the **mustard protocol** (adapted from the standardised method ISO 23611-1:2006).

3 squares of 1 m² are sprayed with a solution of mustard diluted in water. Earthworms that come to the surface are then collected.



2 watering for **each square**

Watering

15 minutes of collect between each watering

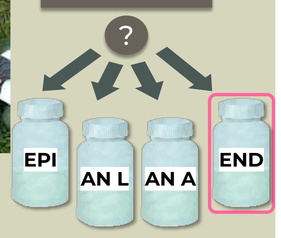
Collecting

2 Field identification (samplers)

The sampler identify the **4 earthworm categories :**

- Epigeic **EPI**
- Anecic *Lumbricus* **AN L**
- Anecic *Aporrectodea* **AN A**
- Endogeic **END**

Sampler identification

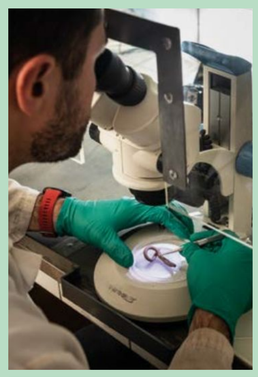


The sampler puts each earthworm identified in the bottle (filled with ethanol 96%) he believes it belongs (EPI, AN L, AN A, END).

After the identification, the samples are sent to the laboratory of the University of Rennes 1.

3 Laboratory identification

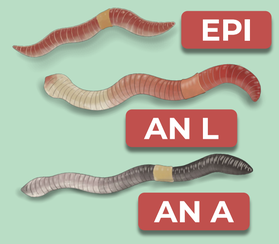
Each individual collected is identified to the species level (Bouché, 1972). The **laboratory technician notes the difference** between his identification and the one made by the sampler.



Sampler identification successful

END

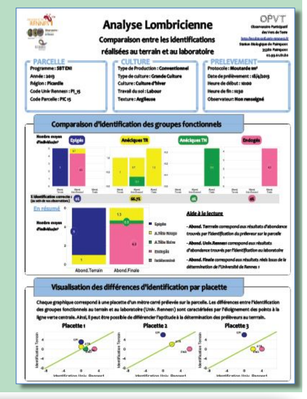
Sampler misidentification



Laboratory technician identify individuals at the lowest taxonomic level possible (sub species, species or genus).

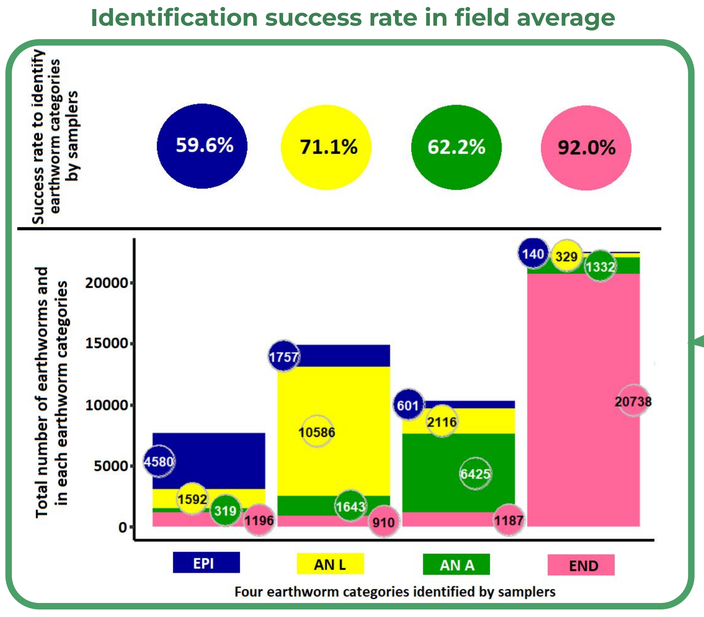
4 Restitution

Restitution documents of the identification success rate in field are made and transmitted to the samplers.



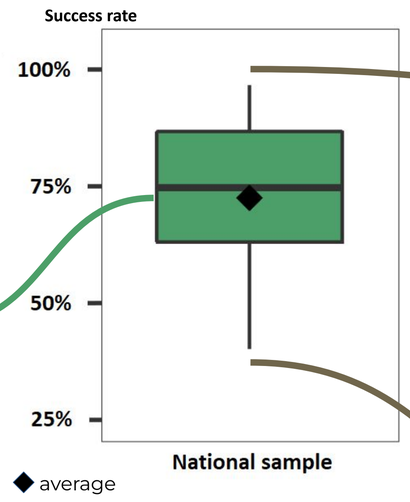
Results & discussion

Identification success rate in field average



Top: success rates in the samplers identification of earthworm categories.
Bottom: differences in identification of earthworm categories between the field and the laboratory.

On average at the national level, the errors are mainly on **epigeics** with only 60% of success, AN A with 62% and AN L with 71%. **These three categories have less obvious distinguishing features than the endogeics** which showed on this study and on OPAL's one the same success rate (92%) (1).



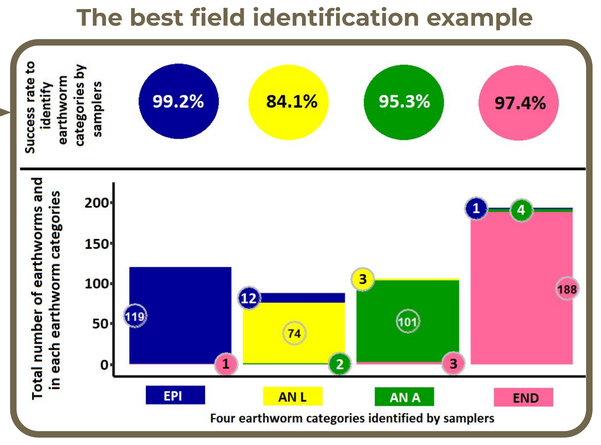
◆ average

Samplers success rate in field identification of earthworm categories (national sample).

Samplers success rates in the identification of earthworm categories (top of the graphs). Differences in identification of earthworm categories between the field and the laboratory (bottom of the graphs).

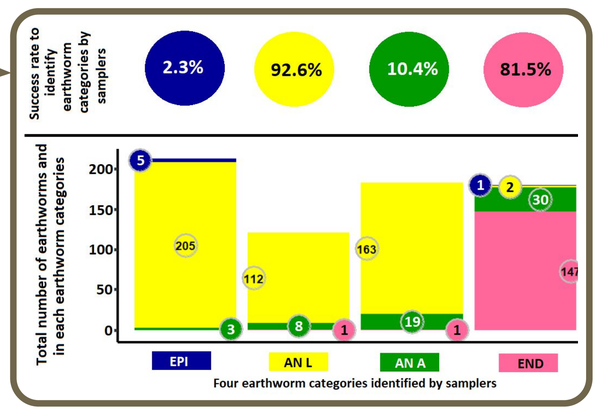
Important heterogeneity : between **40 and 96,5%** of success rate with an **average of 74% while the OPAL study shown an average success rate of 88%.** The difference can be explain by a more precise protocol used by OPAL samplers who identified the individuals at the specie level. This approach is relevant on Britain where only the 12 most common species reflect 93% of the specimens observed (1). However, this can't be applied on the french territory which contains a larger specific diversity.

The best field identification example



In this **best field identification example** there are less than 5% of misidentification for EPI, AN A and END. One of those still visible is that 12 out of 88 (13,6%) **epigeics are mistaken for *Lumbricus* anecics**.
This group was one of the most involved on the project as they communicated regularly and participated in the creation of a booklet.

The worst field identification example



In this **worst field identification example,** thereare a **lot of errors between epigeics and anecics** earthworms mainly. Only 2% of the earthworms identified as epigeics are epigeics. Furthermore, only 10% of *Aporrectodea* anecics are correctly identified.
This example may partly explain the variability showed on the national assessment. Lack of communication or training and frequent turnover between years in sampler team can lead to repeated misidentifications as in this example where AN A and EPI were massively identified as AN L.

Conclusion

Datasets from citizen sciences in a certain context can be used on research. Indeed, the overall data on **the number of earthworms** found is always interesting. Considering the identification of earthworm categories, the abundance of endogeics could be known thanks to citizen science. This observation can be extended to all earthworms categories but only on aeras with a low specific richness.

However, this study shows the **limits in the reliability** of the samplers identification on a national project like SBT-ENI. A consequent variability of success rate and on the nature of the mistakes could be related to the difficulty of organising the training, the communication and the sampling on such a global scale with plural actors involved.

The difficulty in the scientific management of this type of project can be reduced.
More scientific support is needed to limit errors by :
- Improving the communication between research teams and samplers teams during field campaign and encouraging involvement and initiatives.
- Bringing more efforts on regularity of training sessions.
- Focussing trainings on the distinction between anecics and epigeics.

Furthermore, citizen science provides an **important opportunity to raise public awareness about environmental issues.** Indeed, they will become aware of the diversity of earthworm categories and more globally of the biodiversity present in the soil. Thanks to the restitution of the results of the identifications to the laboratory, the public can become aware of the **diversity of earthworm species and the different roles they play in the soil.**

(1) Archer, M., Barraclough, D., Bone, J., Eggleton, P., Head, M., Jones, D. T., & Voulvoulis, N. (2009). The OPAL soil and earthworm survey report. London: Imperial College.