Do correcting methods for biodiversity metrics estimations compensate earthworm sampling methods divergences?

Marc Roucaute[a], Kévin Hoeffner[b], Sarah Guillocheau[b,c], Daniel Cluzeau[b], Guénola Pérès[a]

[a] UMR 1069 SAS, INRAE, Institut Agro Rennes-Angers, Rennes, France
[b] University of Rennes, CNRS, ECOBIO [(Ecosystèmes, biodiversité, évolution)] - UMR 6553, Rennes, France
[c] Micro-company, MME Guillocheau Sarah, 56430 Mauron, France

Context and objectives

Assessing earthworm biodiversity at wide geographical scales relies on meta analysis of existing data [1] and new large scale surveys. But those sources often rely on different sampling methods which may not give comparable results. Previous studies have already explored differences on abundance, biomass or functions of communities sampled with different methods and how they respond to different cultural practices [2] [3]. Here we focus on the ability of two sampling methods based on ISO standard (ISO 23611-1:2018) [4] to characterize biodiversity. We explore if mathematical corrections developed to overcome sampling bias in biodiversity estimations allow for a better comparison between them. We open perspectives for future earthworms biodiversity surveys.

Material and methods

Earthworm sampling

We used data which were recorded during previous projects and we compared i) hand sorting (HS method) of soil monoliths (**fig. 1a**) and ii) chemical extraction (formaldehyde) followed by hand sorting of soil monoliths (CE method) (**fig.1b**). Both methods were applied simultaneously on the same plots, The dataset comes from different studies covering 31 plots distributed on 7 different sites in France (**fig. 2**). They were sampled in 2013, 2014 and 2017. The plots cover different soil uses: grassland, crop field, agroforestry, orchard and forest

Biodiversity estimates

Metrics (according to [5]) : species richness, exponential of Shannon entropy (Shannon diversity) and the inverse Simpson concentration (Simpson diversity), Species richness is sensitive to rare taxa (species found at low abundances) and therefore to sampling effort. Simpson diversity reflects dominance between species. Shannon diversity lies in between, considering both rare species and dominance in the community,

Estimates : observed values ; interpolation/extrapolation for a standard sample size (n=100 worms) and asymptotical extrapolation [6], All were computed using iNEXT package for R [7].

We tested the correlation between the estimates obtained for both sampling techniques and explored the effect of standardisation on this correlation

Figure 1 : sampling methods : **a** Hand sorting (HS) of soil monoliths (number and dimensions); **b** chemical extractant followed by hand sorting of a soil monolith (CE) (number and dimensions)

Site	Year	Soil uses (nb plots)
1[a]	2014	Grassland (1)
2[a]	2014	Crop (1)
3[a]	2014	Crop (1), grassland (1), agroforestry (2), forest (1)
4[a]	2014	Grassland (1)
5[a]	2013	Crop (1), grassland (1), agroforestry (3)
6[a/b]	2014[a]/ 2017[b]	Crop (0[a]/4[b]), grassland (2[a]/8[b])
7[a]	2013	Crop (1), orchard (3)

[a] 6 20 x 20 x 25cm; [b] 4 25 x 25 25cm

Figure 2 : location of the study sites (red stars), year of sampling, soil uses and the corresponding number of plots
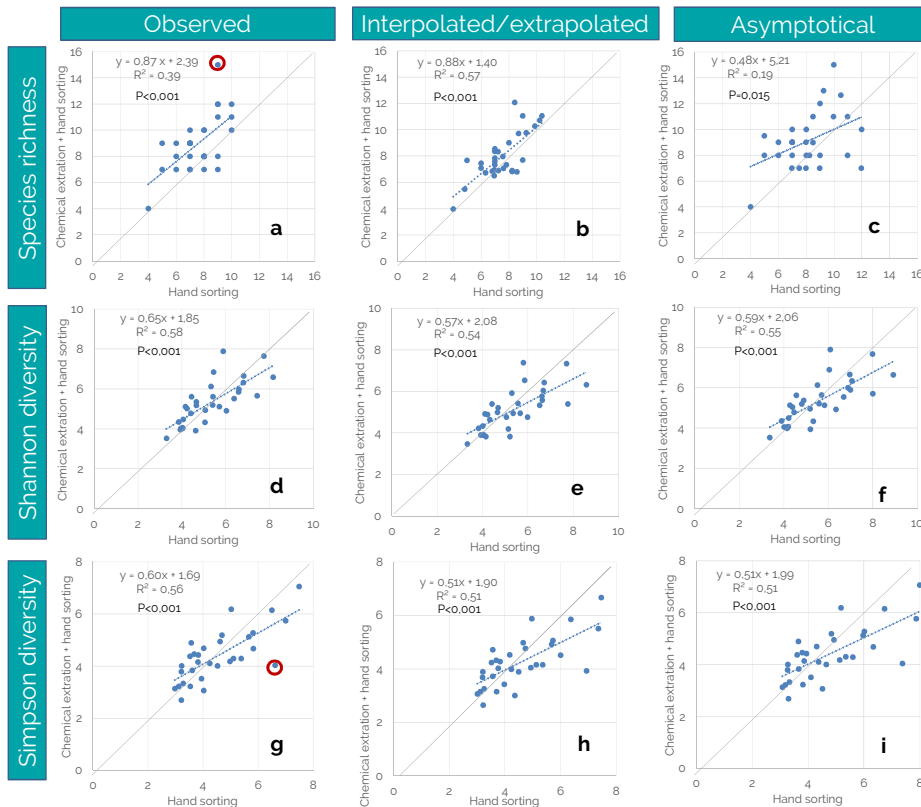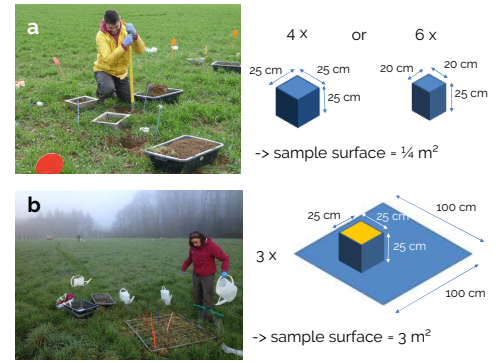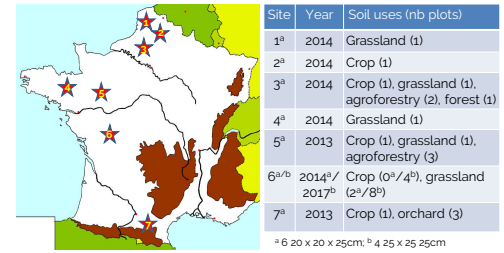
Results and discussion

Figure 3 : Graphs of correlations between species number equivalent estimates obtained from CE or HS methods for species richness (**a,b,c**), Shannon diversity (**d, e, f**) and Simpson diversity (**g, h, i**), on observed values (**a, d, g**), interpolation/extrapolation for a standard sample of 100 worms (**b, e, h**) and asymptotical extrapolation (**c, f, i**). Linear regression (blue lines), its equation, correlation coefficient and significance are presented on each graph. Red circles points extreme outliers.
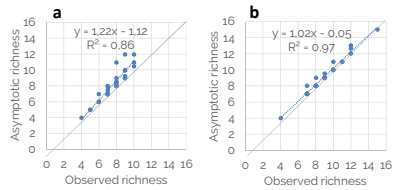

Figure 4 Correlation between observed and asymptotical estimates of richness for HS (a) and CE (b)

HS underestimates species richness compared to CE (**fig. 3a**). HS captures less rare species : In the circled outlier point In **fig. 3a** 7 species were found in CE and not in HS. They represented from 0,01% to 6% of the community.

Sample size standardisation allows to better compare species richness for sites sampled using different protocols but underestimates richness of the richest communities (**fig. 3b**),

Asymptotical estimation of species richness (**fig. 3c**) creates a distortion : low correction for poor sites and a stronger one on rich ones (**fig, 4a**),

Asymptotical estimation of species richness not very different from observed all along the richness gradient for CE (**fig. 4b**),

Shannon and Simpson diversity from the two methods are better correlated for observed values (**fig. 3d and g**) as compared to species richness.

Standardisation and asymptotical estimates of Shannon and Simpson diversity do not improve correlation (**fig.3 e,f,h,i**).

HS presents higher Simpson diversity than CE for the most diversified sites (**fig. 3d to i**). This means a more balanced community structure. In the most outlier point of **fig.3g**, 56 animals were collected in HS and 598 in CE, This reduces the contrast between the proportions of the different species : the proportions of species range from 0.33% to 43% in CE and only from 1.8% to 23% in HS.

Conclusions

Chemical extraction associated with hand sorting (CE) better captures rare species (low abundance) and observed richness value is close to the asymptotical one: CE seems to provide a good estimation of true richness. Our results highlight that merging data issued from different sampling methods may induce bias in data analysis. Standardization to a fix sample size improves the concordance of species richness estimation. Despite the fact that Shannon and Simpson diversity are more consistent between methods even without correction, we should have expected a better correlation because these metrics are considered as not very sensitive to sampling effort [6]. Simpson diversity index, especially, shows an overestimation in HS compared to CE for the most diverse communities. This ends to different community structure evaluation by the two sampling methods. Anecic species as well as *Octolasion cyaneum* are the most frequently missing taxa in HS as compared to CE. Considering those results, chemical extraction (AITC replacing formaldehyde) followed by hand sorting should be favoured in surveys aiming at measuring earthworm biodiversity.

Bibliography

[1] H.R.P Phillips, C.A. Guerra, M.L.C. Bartz et. Al (2019). Global distribution of earthworm diversity. Science. Vol 366, Issue 6464 pp. 480–485, DOI: 10.1126

[2] W.S. Andriuzzi, M.M. Pulleman, D. Cluzeau, G. Pérès, Comparison of two widely used sampling methods in assessing earthworm community responses to agricultural intensification, Applied Soil Ecology. 119 (2017) 145–151

[3] J. Singh, S. Singh, A.P. Vig, Extraction of earthworm from soil by different sampling methods: a review, Environ Dev Sustain. 18 (2016) 1521–1539

[4] ISO 23611-1:2018. Soil quality — Sampling of soil invertebrates — Part 1: Hand-sorting and extraction of earthworms

[5] Jost, L. (2007). Partitioning diversity into independent alpha and beta components. Ecology.88, 2427–2439

[6] Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K. and Ellison, A. M. (2014). Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecological Monographs, 84, 45–67

[7] Hsieh, T.C., Ma, K. H. & Chao, A. (2016) iNEXT: An R package for interpolation and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution, 7, 1451–1456

65 rue de Saint-Brieuc 35000 Rennes
Tél. : + 33 2 23487047
marc.roucaute@inrae.fr
https://www6.rennes.inrae.fr/umrsas/