**Titre : Prédiction de l'abondance, la biomasse et la richesse totale des vers de terre**   
**en France métropolitaine : un outil pour la surveillance biologique du territoire**   
**et la conservation de la biodiversité des sols.**

**Introduction**

La biodiversité des écosystèmes terrestres est majoritairement constituée par la faune vivant dans le sol (FAO, 2020). En effet, la faune du sol fournit une large gamme de services écosystémiques (Bardgett and Van der Putten, 2014; FAO, 2020) dont la plupart résulte du processus conduit par les groupes fonctionnels et taxonomiques de différentes communautés biologiques tels que les lombriciens (Gardi and Jeffery, 2009; Turbé et al., 2010). Les vers de terre sont qualifiés comme étant les « ingénieurs des écosystèmes » car ils participent au développement de la structure du sol (Schon et al., 2017; Sharma et al., 2017), à l'infiltration de l'eau à travers les galeries verticales et horizontales creusés dans le sol (Capowiez et al., 2014; Cunha et al., 2016) et à la minéralisation des nutriments grâce à la dégradation de la matière organique (Van Groenigen et al., 2019). Ces dernières années, plusieurs études ont démontré les rôles importants assurées par les vers de terre, (Blouin et al., 2013; Babu Ojha and Devkota, 2014; Bertrand et al., 2015). Ces travaux ont été justifiés par la prise de conscience croissante du rôle de la diversité lombricienne, mais aussi par le fait qu'on pense que les changements dans la composition des communautés du sol entraîneront une altération des services écosystémiques fournis par la biodiversité du sol (Cardinale et al., 2012; Hooper et al., 2012).

Bien que les vers de terre jouent un rôle important dans les services écosystémiques (Turbé et al., 2010; Blouin et al., 2013; Bertrand et al., 2015), ils sont sensibles à de nombreux facteurs anthropiques et environnementaux (Pelosi et al., 2014; Bai et al., 2018; Gabriac et al., 2022). Les activités humaines notamment la surexploitation des ressources naturelles à cause du développement de l'agriculture intensive et l'industrialisation est la source principale de l’altération des vers de terre (Pelosi et al., 2009). Les effets des facteurs environnementaux sur la communauté lombricienne ont déjà été examinées à l’échelles locale (Marchán et al., 2015; Gabriac et al., 2022), régionale (Marchán et al., 2021, 2016; Marchán and Domínguez, 2022) et territoriale (Palm et al., 2013).

Cependant, peu d’études se sont concentrées sur les effets des facteurs environnementaux et anthropiques qui influencent les communautés lombriciennes (biodiversité et répartition) à des échelles plus larges, c’est-à-dire supra-régional ou national (De Wandeler et al., 2016). Les raisons qui peuvent expliquer ce manque de connaissances sont la faible disponibilité des données à l’échelle des pays, des incohérences taxonomiques et des difficultés pour fusionner les bases de données existantes (Rutgers et al. 2016). Par conséquent, cela pose des défis pour la conservation des vers de terre et donc les fonctions écosystémiques associées. L'un des rares travaux réalisés à des échelles très grandes (continentales) a été effectué par Rutgers et al. (2016) qui ont décrit l’abondance totale et la richesse en espèces sur 3 838 sites échantillonnés dans 8 pays d’Europe. Ils ont observé un gradient latitudinal inverse de l'abondance des vers de terre. De plus, leur étude montre aussi que l'occupation du sol, la texture du sol, la matière organique et le pH du sol sont les facteurs qui affectent plus les vers de terre. Ces résultats sont en accord avec ceux trouvés par (Jänsch et al., 2013; Phillips et al., 2019). Récemment, une autre étude réalisée par Phillips et al. (2019) à l'échelle mondiale a démontré que les facteurs climatiques et le type d’habitat sont les filtres environnementaux les plus importants dans l’altération des paramètres lombriciens que les propriétés du sol. Les études de Rutgers et al. (2016) et de Phillips et al. (2019) sont les pionniers dans le domaine de la prédiction des vers de terre et ils ont permis d’obtenir des cartes de l’abondance et de la richesse des vers de terre à l'échelle continentale (Europe) et mondiale. Toutefois, leurs études sont limitées par l’utilisation d’un seul type d’algorithme de modélisation prédictive : des modèles linéaires généralisés pour Rutgers et al. (2016) et des modèles linéaires généralisés à effets mixtes pour Phillips et al. (2019). Or, il serait pertinent de comparer plusieurs algorithmes afin de trouver le meilleur modèle pour chaque paramètre des vers de terre. Elles sont limitées aussi en termes de nombre d’observations et de prédicteurs, de fiabilité des modèles avec des effets contigus aux frontières des pays, mais aussi d'incertitudes liées à l'extrapolation des cartes de résolutions. Or, lorsque les prédicteurs utilisés lors de la création des modèles ne sont pas précis dans une zone, les prévisions dans cette zone spécifique seront également affectées (Rutgers et al. 2016). De plus, les valeurs de références prédites par les modèles de Rutgers et al. (2016) et de Phillips et al. (2019) sont différents des valeurs observées sur le terrain en France par Cluzeau et al. (2012). Il existe donc un réel besoin de connaissance approfondie de la distribution des vers de terre, et cela peut se faire par le développement des modèles prédictifs performants permettant de fournir des données de référence précise et d’identifier les besoins de conservation spécifiques des vers de terre en France.

Pour déterminer la distribution des communautés, plusieurs outils puissants ont été développés ces dernières années. On distingue les modèles de distribution des espèces « species distribution models (SDMs) » qui sont des modèles basés sur les niches écologique « Niche-based-models » et permet de déterminer directement les relations entre les espèces et leurs environnements pour identifier et quantifier statistiquement les processus qui conditionnent la distribution des communautés (Elith and Leathwick, 2009; Guisan et al., 2017). Néanmoins, ce type de modèle exige d’une part, des données sur les paramètres qui caractérise les communautés et d’autre part des données environnementales spatialisées. Les SDMs sont une approche alternative qui permet d’utiliser efficacement les observations et de prouver ensuite les interprétations ainsi que la compréhension de la distribution des communautés. Divers SDMs existent, chacun présentant des avantages et des inconvénients (Li and Wang, 2013; Valavi et al., 2021). On peut citer les modèles linéaires généralisés (GLM), appropriés lorsque les relations entre les variables explicatives sont linéaires, mais susceptibles d'être influencés par des valeurs aberrantes. Les modèles additifs généralisés (GAM), bien qu'aptes à capturer des non-linéarités dans les variables, présentent une complexité supérieure par rapport aux GLM (Wood, 2017). Les forêts aléatoires (RF), reconnues pour leur performance, sont en revanche moins interprétables (Li and Wang, 2013; Mi et al., 2017). Les modèles de régression boostée généralisée (GBM) démontrent une bonne précision, mais leur paramétrage peut être complexe (Friedman, 2001). Les réseaux de neurones artificiels (ANN), tout en ayant la capacité de modéliser des relations complexes, exigent des volumes importants de données pour éviter le surajustement (Li and Wang, 2013). Les choix méthodologiques exigent donc une évaluation approfondie en fonction des caractéristiques spécifiques des données et des objectifs de modélisation. Selon Maes et al. (2012), les SDMs peuvent être des solutions pour l'aide à la décision pour les programmes de conservation de la biodiversité du sol.

Les approches basées sur les SDMs ont déjà été mises en œuvre avec succès pour prédire la distribution des vers de terre selon divers facteurs environnementaux. En France, la plus récente étude a été réalisée par Fourcade and Vercauteren (2022), qui ont utilisé des arbres de régression boostés pour construire des prévisions spatialement explicites de la diversité fonctionnelle des vers de terre à l’échelle nationale. Cependant, l'étude de Fourcade and Vercauteren (2022) repose uniquement sur des données de présence/absence des espèces de vers de terre en France collectées dans les années 1960 par Bouché (1972). Or, ces données ne reflètent plus l'assemblage actuel des vers de terre en France. De plus, l'étude de la communauté des vers de terre doit inclure les variables décrivant la richesse en espèces, l'abondance et la biomasse totale, qui sont des paramètres vitaux pour les évaluations de la biodiversité. Utiliser seulement la diversité fonctionnelle n'est pas suffisant pour évaluer la biodiversité, car, par exemple, deux communautés peuvent être identiques en termes de richesse, mais différer en densité.

L'objectif généralde notre étude était d'utiliser la base de données LANDWORM (2023-2026 FRB-MTE-OFB) pour développer des modèles de prédictions des vers de terre grâce à des outils statistiques et des systèmes d'information géographique plus avancés, produire des données de référence et identifier les besoins de conservation spécifiques des vers de terre en France. Plus précisément, nous avons cherché à répondre aux objectifs suivants : (i) quantifier et hiérarchiser l'influence des facteurs environnementaux (occupation du sol, propriétés du sol, localisation, données climatiques, …) en fonction de leur contribution aux modèles prédictifs sur l'abondance (individu par m²), la biomasse (g par m²) et la richesse totale des vers de terre (nombre de taxons par m²) sur le territoire national (France métropolitaine) et (ii) prédire et cartographier ces mêmes paramètres lombriciens en fonction des facteurs environnementaux (occupation du sol, propriétés du sol, localisation, données climatiques, …). Nos hypothèses reposent sur le fait que les variables climatiques (précipitation et température) et le type d'occupation du sol seraient les facteurs les plus influençant sur l'abondance et la biomasse des vers de terre, et que le type d'occupation du sol et la texture du sol influenceraient davantage la richesse en espèces. Pour répondre aux objectifs de l'étude, nous avons appliqué et comparé plusieurs algorithmes de modélisation prédictive afin d'identifier le meilleur modèle pour chacun des trois paramètres lombriciens.

**Objectifs détaillés**

1. Quantifier et hiérarchiser l’influence des facteurs environnementaux (occupation et gestion du sol, propriétés de sol, localisation et climatique) sur l’abondance, biomasse et la richesse totales lombricienne sur le territoire métropolitain et sur certaines régions françaises (Bretagne, Dijon métro/Bourgogne ? Aquitaine ? /Poitou-Charentes ?)

*Pour y parvenir nous devrons identifier le meilleur algorithme (RF) pour quantifier et hiérarchiser l’influence des facteurs environnementaux sur les communautés lombriciennes.*

1. Prédire l’abondance, la biomasse et la richesse totales lombriciennes en fonction de différents facteurs environnementaux (occupation et gestion du sol, propriétés de sol, localisation et climatique) sur le territoire métropolitain et sur certaines régions françaises (Bretagne, Dijon métro/Bourgogne ? Aquitaine ? /Poitou-Charentes ?)

*Pour y parvenir nous devrons identifier le meilleur algorithme pour prédire chaque variable des communautés de vers de terre identifiée : abondance totale, biomasse totale et richesse totale*

1. Identifier les meilleures (résolutions des) représentations cartographiques de l’abondance totale, la biomasse totale et la richesse totale des communautés de vers de terre sur le territoire métropolitain et sur certaines régions françaises (Bretagne, Dijon métro/Bourgogne ? Aquitaine ? /Poitou-Charentes ?)

*Pour y parvenir nous utiliserons les différents algorithmes susmentionnés et nous devrons identifier/trouver les données géoréférencées correspondantes aux variables d’entrée des algorithmes (exemple BFC & NA). Nous chercherons à identifier les meilleures représentations cartographique (nbre de classe ?).*

*Cas sup : Model hors Dijon : cas 1*

**Données**

Les données viennent de la base de données LANDWORM (2023-2026 FRB-MTE-OFB).

* Variables de réponses = abondance, biomasse et richesse des vers de terre par parcelle.
* Variables explicatives envisager = occupations du sol, paramètres physico-chimiques du sol (pH, matière organique, texture, carbone, …) et facteurs climatique (précipitation et température, …).
* Autres variables = coordonnées GPS, année, site, …
* Échelle = la France

**Méthodologie / Planning prévisionnel**

**Janvier :** Rédaction du contexte, problématique et objectifs.

**Du 1ER au 15 février :** Analyse exploratoire des données de LANDWORM

-> taux de complétion des données ;

-> exploration des variables quantitative (dispersion, graphe, … ) ;

-> exploration des variables qualitatives (nombre, box plot, … ) ;

-> nettoyage ;

\* Utilisation de base de données externe pour compléter la BDD LANDWORM

-> sélection d’une série de variable explicatives « pertinentes » ;

\* ACP, analyse de variance, … ;

\* à dire d’experts ;

\* selon la bibliographie ;

**Du 16 au 28 février :** Modélisation avec le logiciel R, validation croisée et analyse de sensibilité / qualitative.

**Du 01 au 15 mars :** Validation du modèle avec des données indépendantes. Rédactions de matériels et méthodes.

**Du 16 au 30 mars :** Productions des graphiques et rédaction des résultats.

**Etape XX :** Cartographie avec QGIS et R ?.

**Bibliographie**

Babu Ojha, R., Devkota, D., 2014. Earthworms: “Soil and Ecosystem Engineers” – a Review. World J. Agric. Res. 2, 257–260. https://doi.org/10.12691/wjar-2-6-1

Bai, Z., Caspari, T., Gonzalez, M.R., Batjes, N.H., Mäder, P., Bünemann, E.K., de Goede, R., Brussaard, L., Xu, M., Ferreira, C.S.S., Reintam, E., Fan, H., Mihelič, R., Glavan, M., Tóth, Z., 2018. Effects of agricultural management practices on soil quality: A review of long-term experiments for Europe and China. Agric. Ecosyst. Environ. 265, 1–7. https://doi.org/10.1016/j.agee.2018.05.028

Bardgett, R.D., Van der Putten, W.H., 2014. Belowground biodiversity and ecosystem functioning. Nature 515, 505–511. https://doi.org/10.1038/nature13855

Bertrand, M., Barot, S., Blouin, M., Whalen, J., de Oliveira, T., Roger-Estrade, J., 2015. Earthworm services for cropping systems. A review. Agron. Sustain. Dev. 35, 553–567. https://doi.org/10.1007/s13593-014-0269-7

Blouin, M., Hodson, M.E., Delgado, E.A., Baker, G., Brussaard, L., Butt, K.R., Dai, J., Dendooven, L., Peres, G., Tondoh, J.E., Cluzeau, D., Brun, J.-J., 2013. A review of earthworm impact on soil function and ecosystem services: Earthworm impact on ecosystem services. Eur. J. Soil Sci. 64, 161–182. https://doi.org/10.1111/ejss.12025

Bouché, M.B., 1972. Lombriciens de France. Ecologie et systématique. INRA Editions.

Capowiez, Y., Sammartino, S., Michel, E., 2014. Burrow systems of endogeic earthworms: Effects of earthworm abundance and consequences for soil water infiltration. Pedobiologia 57, 303–309. https://doi.org/10.1016/j.pedobi.2014.04.001

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., Wardle, D.A., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S., Naeem, S., 2012. Biodiversity loss and its impact on humanity. Nature 486, 59–67. https://doi.org/10.1038/nature11148

Cluzeau, D., Guernion, M., Chaussod, R., Martin-Laurent, F., Villenave, C., Cortet, J., Ruiz-Camacho, N., Pernin, C., Mateille, T., Philippot, L., Bellido, A., Rougé, L., Arrouays, D., Bispo, A., Pérès, G., 2012. Integration of biodiversity in soil quality monitoring: Baselines for microbial and soil fauna parameters for different land-use types. Eur. J. Soil Biol., Bioindication in Soil Ecosystems 49, 63–72. https://doi.org/10.1016/j.ejsobi.2011.11.003

Cunha, L., Brown, G.G., Stanton, D.W.G., Da Silva, E., Hansel, F.A., Jorge, G., McKey, D., Vidal-Torrado, P., Macedo, R.S., Velasquez, E., James, S.W., Lavelle, P., Kille, P., Network, the T.P. de I., 2016. Soil Animals and Pedogenesis: The Role of Earthworms in Anthropogenic Soils. Soil Sci. 181, 110–125. https://doi.org/10.1097/SS.0000000000000144

De Wandeler, H., Sousa-Silva, R., Ampoorter, E., Bruelheide, H., Carnol, M., Dawud, S.M., Dănilă, G., Finer, L., Hättenschwiler, S., Hermy, M., Jaroszewicz, B., Joly, F.-X., Müller, S., Pollastrini, M., Ratcliffe, S., Raulund-Rasmussen, K., Selvi, F., Valladares, F., Van Meerbeek, K., Verheyen, K., Vesterdal, L., Muys, B., 2016. Drivers of earthworm incidence and abundance across European forests. Soil Biol. Biochem. 99, 167–178. https://doi.org/10.1016/j.soilbio.2016.05.003

Elith, J., Leathwick, J.R., 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

FAO, I., 2020. State of knowledge of soil biodiversity – Status, challenges and potentialities. Summary for policy makers.

Fourcade, Y., Vercauteren, M., 2022. Predicted changes in the functional structure of earthworm assemblages in France driven by climate change. Divers. Distrib. 28, 1050–1066. https://doi.org/10.1111/ddi.13505

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 29, 1189–1232. https://doi.org/10.1214/aos/1013203451

Gabriac, Q., Ganault, P., Barois, I., Aranda-Delgado, E., Cimetière, E., Cortet, J., Gautier, M., Hedde, M., Marchán, D.F., Reyes, J.C.P., Stokes, A., Decaëns, T., 2022. Environmental drivers of earthworm communities along an altitudinal gradient in the French Alps. https://doi.org/10.1101/2022.10.13.512055

Gardi, C., Jeffery, S., 2009. Soil biodiversity. https://doi.org/10.2788/7831

Guisan, A., Thuiller, W., Zimmermann, N.E., 2017. Habitat Suitability and Distribution Models: With Applications in R, Ecology, Biodiversity and Conservation. Cambridge University Press, Cambridge. https://doi.org/10.1017/9781139028271

Hooper, D.U., Adair, E.C., Cardinale, B.J., Byrnes, J.E.K., Hungate, B.A., Matulich, K.L., Gonzalez, A., Duffy, J.E., Gamfeldt, L., O’Connor, M.I., 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. Nature 486, 105–108. https://doi.org/10.1038/nature11118

Jänsch, S., Steffens, L., Höfer, H., Horak, F., Roß-Nickoll, M., Russell, D., Burkhardt, U., Toschki, A., Römbke, J., 2013. State of knowledge of earthworm communities in German soils as a basis for biological soil quality assessment.

Li, X., Wang, Y., 2013. Applying various algorithms for species distribution modelling. Integr. Zool. 8, 124–135. https://doi.org/10.1111/1749-4877.12000

Maes, J., Egoh, B., Willemen, L., Liquete, C., Vihervaara, P., Schägner, J.P., Grizzetti, B., Drakou, E.G., Notte, A.L., Zulian, G., Bouraoui, F., Luisa Paracchini, M., Braat, L., Bidoglio, G., 2012. Mapping ecosystem services for policy support and decision making in the European Union. Ecosyst. Serv. 1, 31–39. https://doi.org/10.1016/j.ecoser.2012.06.004

Marchán, D.F., Csuzdi, C., Decaëns, T., Szederjesi, T., Pizl, V., Domínguez, J., 2021. The disjunct distribution of relict earthworm genera clarifies the early historical biogeography of the Lumbricidae (Crassiclitellata, Annelida). J. Zool. Syst. Evol. Res. 59, 1703–1717. https://doi.org/10.1111/jzs.12514

Marchán, D.F., Domínguez, J., 2022. Evaluating the Conservation Status of a North-Western Iberian Earthworm (Compostelandrilus cyaneus) with Insight into Its Genetic Diversity and Ecological Preferences. Genes 13, 337. https://doi.org/10.3390/genes13020337

Marchán, D.F., Refoyo, P., Fernández, R., Novo, M., de Sosa, I., Díaz Cosín, D.J., 2016. Macroecological inferences on soil fauna through comparative niche modeling: The case of Hormogastridae (Annelida, Oligochaeta). Eur. J. Soil Biol. 75, 115–122. https://doi.org/10.1016/j.ejsobi.2016.05.003

Marchán, D.F., Refoyo, P., Novo, M., Fernández, R., Trigo, D., Díaz Cosín, D.J., 2015. Predicting soil micro-variables and the distribution of an endogeic earthworm species through a model based on large-scale variables. Soil Biol. Biochem. 81, 124–127. https://doi.org/10.1016/j.soilbio.2014.10.023

Mi, C., Huettmann, F., Guo, Y., Han, X., Wen, L., 2017. Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. PeerJ 5, e2849. https://doi.org/10.7717/peerj.2849

Palm, J., van Schaik, N.L.M.B., Schröder, B., 2013. Modelling distribution patterns of anecic, epigeic and endogeic earthworms at catchment-scale in agro-ecosystems. Pedobiologia 56, 23–31. https://doi.org/10.1016/j.pedobi.2012.08.007

Pelosi, C., Bertrand, M., Roger-Estrade, J., 2009. Earthworm community in conventional, organic and direct seeding with living mulch cropping systems. Agron. Sustain. Dev. 29, 287–295. https://doi.org/10.1051/agro/2008069

Pelosi, C., Pey, B., Hedde, M., Caro, G., Capowiez, Y., Guernion, M., Peigné, J., Piron, D., Bertrand, M., Cluzeau, D., 2014. Reducing tillage in cultivated fields increases earthworm functional diversity. Appl. Soil Ecol., XVI International Colloquium on Soil Zoology & XIII International Colloquium on Apterygota, Coimbra, 2012 – Selected papers 83, 79–87. https://doi.org/10.1016/j.apsoil.2013.10.005

Phillips, H.R.P., Guerra, C.A., Bartz, M.L.C., Briones, M.J.I., Brown, G., Crowther, T.W., Ferlian, O., Gongalsky, K.B., van den Hoogen, J., Krebs, J., Orgiazzi, A., Routh, D., Schwarz, B., Bach, E.M., Bennett, J.M., Brose, U., Decaëns, T., König-Ries, B., Loreau, M., Mathieu, J., …, Eisenhauer, N., 2019. Global distribution of earthworm diversity. Science 366, 480–485. https://doi.org/10.1126/science.aax4851

Rutgers, M., Orgiazzi, A., Gardi, C., Römbke, J., Jänsch, S., Keith, A.M., Neilson, R., Boag, B., Schmidt, O., Murchie, A.K., Blackshaw, R.P., Pérès, G., Cluzeau, D., Guernion, M., Briones, M.J.I., Rodeiro, J., Piñeiro, R., Cosín, D.J.D., Sousa, J.P., Suhadolc, M., Kos, I., Krogh, P.-H., Faber, J.H., Mulder, C., Bogte, J.J., Wijnen, H.J. van, Schouten, A.J., Zwart, D. de, 2016. Mapping earthworm communities in Europe. Appl. Soil Ecol., Soil biodiversity and ecosystem functions across Europe: A transect covering variations in bio-geographical zones, land use and soil properties 97, 98–111. https://doi.org/10.1016/j.apsoil.2015.08.015

Schon, N.L., Mackay, A.D., Gray, R.A., van Koten, C., Dodd, M.B., 2017. Influence of earthworm abundance and diversity on soil structure and the implications for soil services throughout the season. Pedobiologia 62, 41–47. https://doi.org/10.1016/j.pedobi.2017.05.001

Sharma, D.K., Tomar, S., Chakraborty, D., 2017. Role of earthworm in improving soil structure and functioning. Curr. Sci. 113, 1064–1071.

Turbé, A., Toni, A., Benito, P., Lavelle, Patrick, Lavelle, Perrine, Camacho, N., Putten, W., Labouze, E., Mudgal, S., 2010. Soil biodiversity: functions, threats and tools for policy makers. Rep. Eur. Comm.

Van Groenigen, J.W., Van Groenigen, K.J., Koopmans, G.F., Stokkermans, L., Vos, H.M.J., Lubbers, I.M., 2019. How fertile are earthworm casts? A meta-analysis. Geoderma 338, 525–535. https://doi.org/10.1016/j.geoderma.2018.11.001

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R, Second Edition, 2nd ed. Chapman and Hall/CRC, Boca Raton. https://doi.org/10.1201/9781315370279