

Making eDNA more digestible: A pizza analogy for understanding false negatives and occupancy modeling

Jonathan D Bolland, Nathan P Griffiths, Graham S Sellers, Bernd Hänfling, Lori Lawson Handley, Rosalind M Wright

Making eDNA more digestible: A pizza analogy for understanding false negatives and occupancy modeling

Jonathan D. Bolland ^{1,*}, Nathan P. Griffiths ^{1,2,3}, Graham S. Sellers ², Bernd Hänfling ^{2,3}, Lori Lawson Handley ^{2,4} and Rosalind M. Wright ⁵

¹Hull International Fisheries Institute, University of Hull, Hull HU6 7RX, UK

²EvoHull Evolutionary Biology and Environmental Genomics Group, University of Hull, Hull HU6 7RX, UK,

³Institute for Biodiversity and Freshwater Conservation, University of the Highlands and Islands, Inverness IV2 5NA, UK

⁴Lake Ecosystems Group, UK Centre for Ecology and Hydrology, Lancaster Environment Centre, Bailrigg, Lancaster LA1 4AP, UK

⁵Environment Agency, Rivers House, Threshelfords Business Park, Feering CO5 9SE, UK

*Corresponding author: Jonathan D. Bolland, E-mail: j.bolland@hull.ac.uk

Abstract

Environmental DNA (eDNA) sequencing from water samples has emerged as a promising and cost-effective approach to collect comprehensive freshwater biodiversity data. However, critics might highlight potential shortcomings, such as the possibility of false positives (detection of absent species) and false negatives (failing to detect a species that is present). Misconceptions and misunderstandings may also stem from the complexity of the scientific approach and technical language, with implications for decision-makers and potentially hindering conservation. In the present article, we propose an analogy of eating a pizza to simplify messaging and increase understanding of detection probability within typical eDNA metabarcoding workflows. The pizza represents a site, slices represent water samples, bites represent PCR replicates, toppings represent species and olives represent a low-abundance species. Overall, the pizza analogy provides a novel, lighthearted and memorable way to communicate complex eDNA workflows to a broad spectrum of biological scientists and practitioners.

Keywords biodiversity, conservation, endangered species, molecular biology, policy

Environmental DNA (eDNA) sequencing from water samples has emerged as a promising and cost-effective approach to collect comprehensive freshwater biodiversity data (Blancher et al. 2022). Several studies have shown increased sensitivity of eDNA-based monitoring for both entire fish communities and specific priority species when compared with traditional methods (Hänfling et al. 2016, Pont et al. 2018, Harper et al. 2019, Griffiths et al. 2020, Weldon et al. 2020). Substantial work has also been devoted to integrating eDNA-based monitoring into fish-based ecological status assessments in lakes (Willby et al. 2019) and rivers (Hering et al. 2018, Pont et al. 2018). Consequently, it is increasingly important that eDNA workflows and their outputs can be understood by non-molecular ecologists, including traditional fisheries scientists, managers and policy-makers.

In some instances, eDNA-based monitoring is lauded as a game changer (e.g., Carraro et al. 2020) and may lead to an overconfidence in its performance, whereas critics might highlight potential short-

comings, such as the possibility of false positives (detection of absent species; summarized by Cristescu and Hebert 2018), undermining its value and potential. In reality, neither perspective is likely to be completely true, and, like all sampling methods, eDNA has strengths and limitations that must be understood and handled accordingly (Darling 2020, Darling et al. 2020). For example, processes to monitor and mitigate false positives are well established, including the use of blanks or negative controls at each stage of the workflow (Ficetola et al. 2016), library preparation protocols aimed at minimizing cross-contamination (Bohmann et al. 2022), setting low-frequency reads thresholds (Hänfling et al. 2016), and adopting conservative occupancy modeling approaches (Stauffer et al. 2021). Notwithstanding, misconceptions often stem from the complexity of highly specialist scientific approaches, sometimes referred to as *black box science* (Morissette et al. 2021), and the use of technical language that can create a barrier to understanding (Stein et al. 2024). This is particularly problematic when environmental managers and decision-

Received: 20 January 2026. Revised: 17 February 2026. Accepted: 20 February 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of the American Institute of Biological Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. A selected library of relevant multiscale occupancy modeling studies.

Study	Target taxa or context	Key findings
Nichols et al. (2008)	Vertebrates using multiple survey methods	Introduced multiscale occupancy modeling and method-specific detection probabilities
Schmidt et al. (2013)	Chytrid fungus	Demonstrated the need to separate water sampling and PCR detection in some eDNA occupancy models
Ficetola et al. (2015)	Soil earthworms and mammal sediment DNA	Applied occupancy models with replicated PCRs to estimate detection and false positives in eDNA and showed that high PCR replication is often needed for reliable detection
Fukaya et al. (2022)	Freshwater fish communities using eDNA metabarcoding	Developed a multispecies occupancy model that directly models sequence reads and informs survey design
Griffiths et al. (2025)	Freshwater fish communities using eDNA metabarcoding	Applied three-level occupancy modeling to quantify confidence in absence and proposed a framework for decision-making

makers must interpret eDNA data without a clear framework for assessing its reliability and implications (e.g., Darling and Mahon 2011, Lodge 2024).

A key factor in assessing the presence or absence of a species is understanding the probability of false negatives (i.e., failing to detect a species that is actually present; Buxton et al. 2021). This is especially relevant for rare and elusive species that only occupy a subset of sites (i.e., low site occupancy) and may exist in low abundance when present (i.e., low detection probability). False negatives can arise from stochastic sampling processes at both the biological (water sampling) and the technical (PCR—polymerase chain reaction—amplification) stages of the workflow. In the biological stage, false negatives occur when no target species DNA is collected during water sampling, despite presence at the site. In the technical stage, false negatives occur when no target species DNA is present in the PCR replicate, despite being collected during water sampling. Increasing the number of biological samples and technical replicates can help mitigate the rate of false negatives (Ficetola et al. 2015), and multispecies site occupancy models attempt to address it (Doi et al. 2019, McClenaghan et al. 2020), with recent examples integrating read counts into such models (Fukaya et al. 2022; see table 1). However, without a thorough understanding of how false negatives can occur, non-molecular ecologists may erroneously conclude a species was absent from a site, potentially leading to flawed management decisions.

In this article, we propose an analogy of eating a pizza to simplify messaging and increase understanding of detection probability within typical eDNA metabarcoding workflows. For simplicity, we focus on the two most common levels of replication, collecting multiple water samples from a site and carrying out multiple PCRs for each sample. The pizza represents a site, slices represent water samples, bites represent PCR replicates, toppings represent species and olives represent a low-abundance species (figure 1). We provide an example where three water samples were taken from a site, and each had three PCR replicates: Sample 1 is a positive result; the target species is present in two of the three PCR replicates. A single PCR replicate could have yielded a false negative (technical) despite presence in the sample. Sample 2 represents a biological false negative. The target species is absent in the sample despite presence at the site. Additional water samples may have yielded a positive. Sample 3 is a technical false negative. The target species is absent in all PCR replicates despite presence in the sample. Additional PCR replicates may have yielded a positive.

In the example, the combination of three water samples (slices) and three PCR replicates (bites) could have accurately deduced the presence of the low-abundance species (olives) at the site (pizza); that is, the correct result was obtained. However, it also neatly demonstrates how a species can be missed at both water sampling and PCR stages of the workflow. This demonstration furthers the understanding of how the overall probability of occupancy at a site (P_o), the conditional probability of DNA presence in a sample given occupancy at a site (P_s), and the conditional probability of DNA detected in PCR replicates given presence in a sample (P_r) are calculated (figure 1). In combination, these three probabilities allow the overall confidence in species absence at a site to be calculated according to the level of biological and technical replication (Griffiths et al. 2025). Indeed, following project specific applications and conference presentations, the authors have received strong positive feedback from both academic and non-academic stakeholders that the pizza analogy aids interpretation of eDNA workflows and assessments of species presence or absence. For example, it enables managers and funders to appreciate how a lack of detection may not mean a species is absent and why more sampling or lab work may be required.

The pizza analogy can also be extended, with collecting different volumes of water between samples represented by uneven size and shape of pizza slices, and the amount of DNA included in the PCR replicate would be represented in the present article as the size of each bite. It is also possible that low-abundance species are harder to detect with eDNA metabarcoding in complex communities (i.e., a false negative due to masking of target DNA via swamping from non-target DNA). For illustration, it may be harder to detect olives on a pizza with lots of toppings or flavors or where there is a dominant or overwhelming topping (e.g., chillies or pepperoni) than it would be on a simple pizza, such as a Margherita.

Overall, the pizza analogy provides a novel, lighthearted, and memorable way to communicate complex eDNA workflows. Specifically, the analogy simplifies different levels of occupancy modeling to enable a thorough understanding of how false negatives can occur and aid interpretation of how confident we can be in assessments of species absence. This has particular relevance to rare and conservation species and therefore enables non-specialist managers, regulators, and policy-makers to make more informed decisions. More generally, it is believed the pizza analogy will help make eDNA, which has emerged as a fundamental technology to quantify biodiversity, more digestible for a broad spectrum of biological scientists and practitioners.

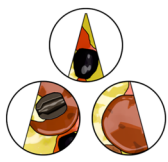
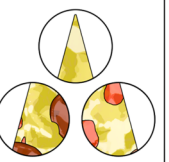
	Pizza analogy			Three-level occupancy model
Site sampling location				P_o probability of occupancy at a site
Sample biological replicate				P_s conditional probability of DNA presence in a sample given occupancy at the site
PCR technical replicate				P_r conditional probability of DNA detection in a replicate given presence in the sample
Result	Positive sample	False negative (sample)	False negative (PCR)	

Figure 1. The pizza analogy. The pizza represents a site, slices represent water samples (biological replicate), bites represent PCR replicates (technical replicate) and olives represent a low-abundance species. Routes to olive detection (i.e., positive sample), and non-detection (i.e., false negatives during water sampling and PCR stages of the workflow) are presented, as is the terminology used during occupancy modeling.

Acknowledgments

This work was funded by the Environment Agency. The authors have no conflicts of interest to disclose

Author contributions

Jonathan Bolland (Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft), Nathan P Griffiths (Conceptualization, Methodology, Writing – original draft), Graham S Sellers (Methodology, Visualization, Writing – review & editing), Lori L Handley (Writing – review & editing), Bernd Hänfling (Methodology, Supervision, Writing – review & editing), Rosalind M Wright (Funding acquisition, Project administration, Resources, Writing – review & editing)

References cited

Blancher P, et al. 2022. A strategy for successful integration of DNA-based methods in aquatic monitoring. *Metabarcoding and Metagenomics* 6: e85652.

Bohmann K, Elbrecht V, Carøe C, Bista I, Leese F, Bunce M, Yu DW, Seymour M, Dumbrell AJ, Creer S. 2022. Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular*

Ecology Resources 22: 1231–1246. <https://doi.org/10.1111/1755-0998.13512>.

Buxton A, Matechou E, Griffin J, Diana A, Griffiths RA. 2021. Optimising sampling and analysis protocols in environmental DNA studies. *Scientific Reports* 11: 11637.

Carraro L, Mächler E, Wüthrich R, Altermatt F. 2020. Environmental DNA allows upscaling spatial patterns of biodiversity in freshwater ecosystems. *Nature Communications* 11: 3585. <https://doi.org/10.1038/s41467-020-17337-8>.

Cristescu ME, Hebert PDN. 2018. Uses and misuses of environmental DNA in biodiversity science and conservation. *Annual Review of Ecology, Evolution, and Systematics* 49: 209–230. <https://doi.org/10.1146/annurev-ecolsys-110617-062306>.

Darling JA, Jerde CL, Sepulveda AJ. 2020. What do you mean by false positive?. *Environmental DNA* 3: 879–883.

Darling JA, Mahon AR. 2011. From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environmental Research* 111: 978–988.

Darling JA. 2020. How to learn to stop worrying and love environmental DNA monitoring. *Aquatic Ecosystem Health and Management* 22: 440–451.

Doi H, Fukaya K, Oka S-I, Sato K, Kondoh M, Miya M. 2019. Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site oc-

- occupancy model. *Scientific Reports* 9: 3581. <https://doi.org/10.1038/s41598-019-40233-1>.
- Ficetola GF, et al. 2015. Replication levels, false presences, and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources* 15: 543–556.
- Ficetola GF, Taberlet P, Coissac E. 2016. How to limit false positives in environmental DNA and metabarcoding? *Molecular Ecology Resources* 16: 604–607.
- Fukaya K, Kondo NI, Matsuzaki S-IS, Kadoya T. 2022. Multispecies site occupancy modelling and study design for spatially replicated environmental DNA metabarcoding. *Methods in Ecology and Evolution* 13: 183–193.
- Griffiths NP, Bolland JD, Wright RM, Murphy LA, Donnelly RK, Watson HV, Hänfling B. 2020. Environmental DNA metabarcoding provides enhanced detection of the European eel *Anguilla anguilla* and fish community structure in pumped river catchments. *Journal of Fish Biology* 97: 1375–1384. <https://doi.org/10.1111/jfb.14680>
- Griffiths NP, Hänfling B, Cattaneo M, Wright RM, Macarthur JA, Peixoto S, Bolland JD. 2025. Proving a negative: Estimating species “confidence in absence for decision-making” using environmental DNA monitoring. *Journal of Applied Ecology* 62: 2409. <https://doi.org/10.1111/1365-2664.70099>
- Hänfling B, Lawson Handley L, Read DS, Hahn C, Li J, Nichols P, Blackman RC, Oliver A, Winfield IJ. 2016. Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology* 25: 3101–3119. <https://doi.org/10.1111/mec.13660>.
- Harper LR, Griffiths NP, Sayer CD, Read DS, Harper KJ, Blackman RC, Li J, Hänfling B. 2019. Development and application of environmental DNA surveillance for the threatened crucian carp (*Carassius carassius*). *Freshwater Biology* 64: 93–107.
- Hering D, et al. 2018. Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research* 138: 192–205.
- Lodge DM. 2024. Lessons learned from eDNA adoption in the management of bigheaded carps in Chicago IL USA Area Waterways. *Environmental DNA* 6: e528.
- McClenaghan B, Compson ZG, Hajibabaei M. 2020. Validating metabarcoding-based biodiversity assessments with multi-species occupancy models: A case study using coastal marine eDNA. *PLOS ONE* 15: e0224119. <https://doi.org/10.1371/journal.pone.0224119>.
- Morisette J, et al. 2021. Strategic considerations for invasive species managers in the utilization of environmental DNA (eDNA): Steps for incorporating this powerful surveillance tool. *Management of Biological Invasions* 12: 747–775.
- Nichols JD et al. 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology* 45: 1321–1329.
- Pont D, Rocle M, Valentini A, Civade R, Jean P, Maire A, Roset N, Schabuss M, Zornig H, Dejean T. 2018. Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific Reports* 8: 10361.
- Schmidt BR, Kéry M, Ursenbacher S, Hyman OJ, Collins JP. 2013. Site occupancy models in the analysis of environmental DNA presence/absence surveys: A case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution* 4: 646–653.
- Stauffer S, et al. 2021. How many replicates to accurately estimate fish biodiversity using environmental DNA on coral reefs? *Ecology and Evolution* 11: 14630–14643.
- Stein ED, et al. 2024. Critical considerations for communicating environmental DNA science. *Environmental DNA* 6: 1–12.
- Weldon L, O’Leary C, Steer M, Newton L, Macdonald H, Sargeant SL. 2020. A comparison of European eel *Anguilla anguilla* eDNA concentrations to fyke net catches in five Irish lakes. *Environmental DNA*: 2: 587–600. <https://doi.org/10.1002/edn3.91>
- Willby N, Law A, Bull C, Hänfling B, Lawson Handley L, Winfield IJ. 2019. A Tool for Classifying the Ecological Status of Lake Fish in Britain Based on eDNA Metabarcoding. Scottish Environment Protection Agency.