Optimal Sampling Density for Catchment-Scale River Biomonitoring with environmental DNA

Master of Science thesis

J. (Jesse) van Leeuwen

Delft University of Technology

TU Delft

Optimal Sampling Density for Catchment-Scale River Biomonitoring with environmental DNA

Master of Science thesis

by

J. (Jesse) van Leeuwen

to obtain the degree of Master of Science
at the Delft University of Technology
to be defended publicly on November 28th, 2024 at 16:00.

Student number: 5659760
Chair: Professor Dr. Thom Bogaard | TU Delft
Daily supervisor: Dr. Luca Carraro | University of Zürich
Third supervisor: Dr. Ir. Astrid Blom | TU Delft

Faculty: Faculty of Civil Engineering and Geosciences
Department: Water Resources Engineering

Date: Thursday 21st November, 2024
Cover: Photo by Tom Fisk on Pexels

TU Delft

Preface

As this work marks the end of an intensive study programme, I would first all like to thank my family for their ongoing support in every aspect of my life whenever I needed it. Secondly, all my friends for their kindness and support, and specifically Casper for proof-reading my work and Jeroen for his help in designing the beautiful front page of my thesis. Third, thanks to my supervisors for sharing their vast knowledge and expertise, being of assistance whenever I reached out to them and their flexibility in accommodating my preferred schedule. Lastly I would like to express my gratitude to Jan Willem, who played a key role in setting up this project. His attitude towards doing science was a big inspiration to me and a significant reason for choosing this topic. Although he is no longer with us, his enthusiasm and encouragement will always be remembered with a smile.

Abstract

River networks are vital ecosystems that face increasing threats from societal pressures and global environmental changes. Effective biomonitoring of these systems is crucial for maintaining water quality and biodiversity by guiding policy and management actions. Over the past two decades, environmental DNA (eDNA) has emerged as a promising tool, offering advantages over traditional biomonitoring methods. However, the downstream transport and decay of genetic material in river networks poses a challenge for the spatial interpretation of eDNA samples. This study follows from work of Carraro et al. (2021) and aimed to find the optimal sampling density to determine taxon distributions at the catchment scale based on eDNA sampling. This was done by means of a virtual experiment, where a large number of sampling campaigns were simulated over multiple synthetic taxon distributions and river networks. This enabled the development of generalized sampling principles irrespective of system-specific confounding factors. Results showed that taking at least one sample per 10 km^2 of catchment area provides a reliable estimate for taxon abundance across the catchment. Taking more samples only marginally improved the accuracy of the predictions when it concerns taxa that are commonly found in different parts of the catchment. In case of finding rare taxa, a minimum of two samples per 10 km^2 is necessary to ensure that the taxon is correctly located. This study also underscores the limited capability of eDNA in capturing fine-scale spatial patterns, suggesting its complementary role alongside traditional surveys. In summary, this study provides a generalized sampling guideline for eDNA biomonitoring as a tool for broad-scale environmental assessments and biodiversity conservation.

Glossary

environmental DNA (eDNA): Genetic material shed by organisms into the environment, used to detect and analyse species presence and distribution in aquatic ecosystems.

Taxon; A classification unit used to group organisms based on shared characteristics, ranging from species (e.g., *Gammarus pulex*, a freshwater amphipod) to higher levels like family (e.g., Baetidae, a family of mayflies), both commonly used in riverine biomonitoring to assess water quality and ecosystem health.

Spatial field; A continuous representation of a spatially varying variable across a defined domain. Here it represents the eDNA production rate across each reach of the river network and is used as a modelling term for the spatial taxon distribution in a given catchment. *see section 2.2.3*

Asymmetric Eigenvector Map (AEM); A spatial modelling technique used to capture directional or flow-based processes by decomposing spatial relationships into orthogonal eigenvectors, often applied in riverine networks to model spatially structured ecological patterns influenced by directional flows, such as water current or species dispersal. *see section 2.2.7*

Contents

1. Introduction

Natural river networks are among the most important and diverse ecosystems (Vörösmarty et al., 2010). Globally, natural river systems face increasing anthropogenic pressures., in order to meet societal needs like irrigation, navigation, and hydropower generation (Grill et al., 2019). The European Water Framework Directive has outlined the importance of maintaining a good ecological status of water bodies, stressing the need for extensive and accurate monitoring at the system level (Voulvoulis, Arpon & Giakoumis, 2017). Biomonitoring is an important tool to assess the health of aquatic ecosystems. Information on the distribution of certain taxa in the environment can guide biodiversity management (e.g. identifying endangered or invasive species), but may also be indicative of water quality. Targeted monitoring schemes have been adequate in addressing known pressures on riverine habitats. However, they are lacking the spatial coverage necessary to identify the effects of emerging stressors induced by global change (Hering et al., 2010). Moreover, differences among EU member states in methodology, frequency, and site density limit a large-scale assessment and response (Hering et al., 2010). A novel way of making such an assessment is by measuring environmental DNA (eDNA), which entails the collection of genetic material shed by these organisms into the environment.

Due to the potential advantages over traditional biomonitoring methods, such as kick-net sampling (macroinvertebrates) and electrofishing (fish), this method has quickly gained traction among researchers over the past decade (Cristescu & Hebert, 2018). Using quantitative PCR (qPCR), the abundance of a single, targeted taxon may be estimated from its eDNA concentration. Metabarcoding, a more recently developed method, enables the simultaneous detection of multiple taxa. Because of the advantages stemming from the simplicity of taking samples and the quality of the collected data, eDNA has the potential to revolutionize the field of biological monitoring. eDNA samples are collected by extracting a small amount of water from the environment and passing it through a filter, after which it is analysed in a laboratory. Because the sample collection is minimally invasive, it leaves a smaller footprint on the environment and reduces the occurrence of field measurement errors. The quick extraction of field data allows more fine differences in space to be captured at a higher frequency with the same level of effort (Altermatt et al., 2020). Moreover, the identification process is fast and easily standardizable. Therefore, eDNA-based biomonitoring can greatly improve surveillance monitoring across large river basins and enable a transboundary approach (Hering et al., 2010).

The relationship between measured eDNA concentrations and their source organisms is shaped by the dynamic interplay of production, transport, and decay processes in streams and rivers. These processes are highly complex, influenced by a range of biotic and abiotic factors (Stewart, 2019). eDNA is produced as organisms shed genetic material into their surroundings (e.g., through skin, saliva, or faeces). This rate of shedding can vary with biotic factors, like life history traits and interactions with other organisms, as well as environmental factors like temperature, pH, salinity, and UV exposure. Once released into stream networks, eDNA undergoes advection, dilution, sedimentation, and resuspension, while simultaneously decaying due to microbial activity and environmental conditions (Stewart, 2019). These dynamic processes can obscure the relationship between the location of an eDNA sample and the actual presence of the organisms it represents (Harrison et al., 2019). Consequently, the spatial interpretation of eDNA samples in flowing water still requires significant development in order to upscale the application of eDNA in research (Blackman et al., 2024b). Understanding and accounting for the transport and decay of eDNA is therefore essential to accurately infer the spatial distribution of taxa in riverine systems.

Consideration of the unique and complex behaviour of eDNA in riverine ecosystems poses both challenges and opportunities in the interpretation of the data. The downstream transport of environmental DNA enables the detection of taxa in the upstream catchment (Deiner et al., 2016). In traditional sampling methods, local hotspots may easily be missed or, through spatial pooling, averaged out. Because the eDNA signal of these hotspots is visible over a much larger area, their detection is more likely. At the same time, the downstream transport and mixing of different eDNA signals blurs out localized taxon abundance patterns (Deiner & Altermatt, 2014). The eDITH (eDNA Integrating Transport and Hydrology) model estimates the origin of eDNA measured at the catchment scale by modelling its fate in the environment (Carraro et al., 2018). Combining first-order decay with hydro-morphological and biogeographical principles, it helps to disentangle the various sources of measured eDNA at the network-level. Studies comparing the performance of eDITH to that of traditional sampling suggest that, while it has been reasonably successful in matching taxon distribution data that was measured with traditional sampling

methods, eDITH could especially complement these practices due to its spatial coverage and different taxonomic coverage (Carraro et al., 2018; Carraro et al., 2020; Blackman et al., 2024a).

Despite the potential of eDNA, supported by vast scientific development of the method (Palowski et al., 2020), its practical application remains limited. One important step toward wider implementation is the development of generalized sampling guidelines (Cristescu & Hebert, 2018; Dickie et al., 2018). To this end, Altermatt et al. (2023) have provided guidelines on sampling volume, considering differences in hydrodynamic dispersal at the reach scale. At the catchment scale, samples should be positioned based on the estimated decay length of eDNA to maximize the recovery of different signals from DNA-shedding taxa (Deiner & Altermatt, 2014). Making use of synthetically generated catchments and taxa, Carraro et al. (2021) have provided concrete guidelines on optimal positioning of samples across catchments. Considering these guidelines, taking more samples may help to better distinguish overlapping signals and accurately tracing their origins. However, the extent to which this is possible may be limited due to the downstream mixing of eDNA and the capability of models to capture these processes well. Insights, into the effect of the number of samples on the quality of their spatial interpretation could complement existing guidelines for eDNA sampling in river networks.

Using a similar setup to Carraro et al. (2021), this study aims to determine the optimal density of eDNA samples to retrieve robust taxon distribution predictions at the catchment scale using eDITH. In addition to previous efforts, this study leverages high-performance computing to simulate a large number of sampling campaigns across a diverse set of computer-generated river networks and taxon distributions. By quantifying the performance of eDITH for each simulation, this study tries to establish generalized sampling guidelines that can be applied across various freshwater ecosystems, independent of the specific environment or taxon. By striking a balance between effort and performance, this research aims to improve current biomonitoring efforts to efficiently determine catchment-scale distribution at large spatial scales. Moreover, by exploring different levels of model complexity in eDITH, this study will assess how these variations affect the accuracy of taxon distribution predictions, thereby aiding practitioners in their efforts to process eDNA data. This may ultimately support more effective strategies for biomonitoring and, in turn, environmental conservation and management.

2. Methods

2.1 eDITH

The eDITH model was introduced by Carraro et al. (2018). Since then, a number of studies (e.g. Carraro et al., 2020; 2021, Blackman et al., 2024a) have been performed to further develop the model, test assumptions, and explore applications. eDITH models the downstream transport and ongoing decay of eDNA in a catchment. It allows the spatial coupling of measured concentrations to the production rate across the catchment. The computation of eDNA concentrations from the production rate is referred to as the 'forward' model. Based on catchment-level approximations of advection and decay, it provides an exact, steady state solution of the observable downstream eDNA concentration. In order to spatially interpret eDNA samples, their upstream sources need to be identified. Due to the dendritic structure of the river network, where different tributaries may contribute differently to any measured eDNA concentration downstream, an inverse modelling approach is required. This 'backward' model estimates the eDNA production across the catchment, such that it best reproduces the observed concentration pattern downstream. Assuming that density of a given taxon is proportional to the shedding rate of eDNA, the estimated pattern may be used to describe taxon distributions. In reality, the relationship between eDNA production rate and taxon actually depends on a wide range of abiotic and biotic factors. However, assuming that this proportionality is homogenous across the catchment, it is sufficient to infer spatial patterns of taxon distribution at a single point in time. In the rest of this study, the terms production rate and taxon abundance will therefore be used interchangeably. Similar approaches to source estimation may be found in the identification of contaminant sources (Abbott et al., 2019).

eDITH discretizes a river network into river reaches (nodes of the network) with homogenous hydraulic characteristics (velocity, width, depth) that follow general scaling laws (Rodriguez-Iturbe & Rinaldo, 1997). A mass balance determines the downstream transport of eDNA through each reach providing a coarse approximation of real hydraulic processes. The scale of this transport is limited by the decay of eDNA, which is summarized by a single first-order rate. Accordingly, the concentration in each reach is given by the sum of all observable eDNA originating from upstream sources, given its gradual decay. The governing equation encapsulating these processes is as follows:

$$C_j = \frac{1}{Q_j} \sum_{i \in \gamma(j)} A_{s,i} \exp\left(-\frac{L_{ij}}{v_{ij}\tau}\right) p_i \tag{1}$$

Where C_j is the concentration of eDNA at node j in units/m^3, Qj is the discharge at node j, p_i is the eDNA production rate of all nodes upstream of j in units per s per m^2 of catchment area, and As,i is the surface area for each upstream node i where eDNA is produced in m^2. Here, 'units' would usually be expressed in moles. The exponential term describes the decay between each node i and j, where Lij is the distance from node i to j in m, v_{ij} the average velocity over L_{ij} in m/s, and τ the decay time in h.

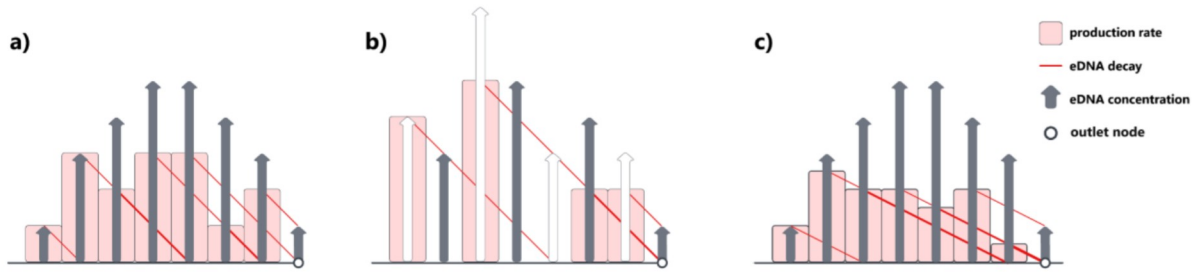


figure 1 The formulation of eDITH allows multiple parameter sets of equation 1 to appear as a correct solution in the sampled reaches. This figure considers a simplified view of different rates of eDNA production along a single river stretch, where eDNA is transported from left to right (x-axis). a) the original production pattern and the corresponding concentration pattern (parabola shape). b) a different production pattern that still results in the same concentration pattern in the second, fourth, sixth, and eighth node (grey arrows, representing sampled sites). c) a different production pattern yields the same concentration pattern in all nodes (parabola shape) when the decay time is also set to a different value (double the original). The relationship between production and concentration is simplified. When considering a dendritic river network, the number of combinations of eDNA production yielding the same 'correct' eDNA concentration in a limited number of sampled reaches further increases

Figure 1 shows how eDITH is used to couple the production of eDNA data with its concentration in each node of the river network. Given that the decay rate is not known beforehand, and the fact that eDNA samples are only taken in a subset of all river reaches, there are multiple solutions to equation 1 that may match the sampled eDNA concentrations (making this an ill-posed problem). Moreover, the accurate estimation of n (= number of reaches) values of production rate p_i is computationally infeasible. Therefore, it is preferred to model the production of eDNA by considering the underlying spatial patterns that describe this distribution. Equation 2 describes a Poisson Generalized Linear Model:

$$p_i = p_0 \exp\left(\beta^T \boldsymbol{x}(i)\right) \tag{2}$$

Where the production rate p_i in each reach is a linear combination of a list of covariates $x(i)$ with weights beta, scaled by a baseline production rate p_0. These covariates can represent any spatial pattern that may explain taxon distribution. This may include environmental covariates, such as land cover and elevation, but also patterns of spatial autocorrelation. Besides aiding in the estimation of distribution patterns, this approach may provide insight into the effect of covariates on taxon distribution. The application of equation 2 in this study will be elaborated on further in the section 2.2.7.

The production rate and decay time were estimated by minimizing the error between the observed eDNA concentration at given locations along the river network (i.e., sampling sites) and the eDNA concentration predicted by the model at the corresponding locations, given the estimated values of the production rate and decay time. In order to differentiate the contribution of various locations across the catchment, an adequate number of sampling locations is necessary. These samples should be positioned in a hierarchical and nested manner, meaning that they should be distributed across various branches while at the same time being close enough along the same path (Carraro & Altermatt, 2024). Using eDITH, this enables the partitioning of different sources and the distinction of along-stream patterns respectively. Within the framework described above, this study investigates the relationship between the density of samples within a given catchment area and the quality of the taxon distribution predictions made by eDITH.

2.2 Virtual experiment

2.2.1 Overview

In order to evaluate the relationship between sampling density and the quality of taxon abundance predictions by eDITH, a virtual experiment is performed. A synthetic catchment with a synthetic taxon distribution is generated, resembling a real system. Over this system, a specific sampling strategy is performed, extracting the eDNA concentration resulting from the synthetic taxon at a specific set of locations. This data serves as input for the backward estimation, which aims to fit a taxon distribution to these points. The outcome is evaluated by comparing the modelled taxon distribution to the original one. By resembling a sampling campaign on a real system, this approach is referred to as the 'virtual ecologist' (Zurell et al., 2010). With the use of computer simulation, a large number of replicate experiments were performed, while controlling for potentially confounding factors. This enables the investigation of specific factors (such as sampling density) in a generalized setting, which may serve as guiding principles for practical applications of eDITH. This approach follows that of Carraro et al. (2021), adopting the conditions that best resemble a real scenario (i.e. as little prior information as possible). Below, each step of this virtual experiment is explained in detail. At the bottom of this section, figure 5 gives a graphical overview of the virtual experiment.

2.2.2 Generation of synthetic river networks

To explicitly model hydrological transport in river networks, it is particularly useful to utilize their universal mathematical relationships and scaling properties (Rodrigues-Iturbe & Rinaldo, 1997). While these may be applied to infer the hydraulic parameters of a real river network with limited measurement locations, these general laws can also be applied to generate idealized synthetic river networks that are geomorphologically analogous to real river networks (Carraro et al., 2020b). These so-called 'Optimal Channel Networks (OCNs) are especially useful in modelling studies, where the generation of a large number of replicate river networks allows to analyse processes that are independent of the specific shape of river networks. OCNs are essentially steady state solutions of landscape evolution (figure 2). They are created by optimizing an initial drainage network, represented by a spanning tree, to minimize the total energy loss throughout the network (Carraro et al., 2020). Similar to the real

landscape-forming process, an algorithm randomly searches for different flow paths, which it accepts when they result in a lower energy expenditure along the stream network. The result is a network that stretches into every pixel of the lattice, which may be aggregated into a river network at any desired scale (by adjusting the threshold catchment area for a river to form. The topology of the landscape is such that its drainage network following the steepest descent directions would be the same as the OCN. Using this method randomly generates a synthetic catchment with realistic geomorphological features.

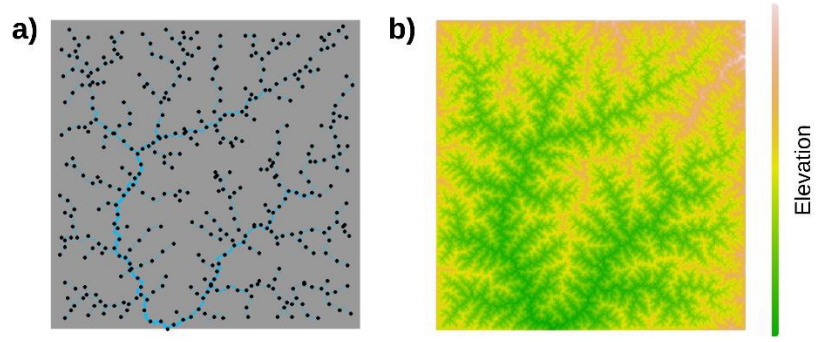


figure 2 a) example optimal channel network, following the discretization in nodes (black dots) of this study. b) the corresponding elevational landscape. This image only serves to illustrate the morphological implications of OCNs.

In order to generate a large number of replicate systems, this study used a set of fifty OCNs (which were generated for Carraro & Altermatt (2022), from which it randomly selects one for each simulation. This number is deemed sufficiently large to mitigate the effect that a specific catchment shape may have on the results, while limiting the computational demand required. These OCNs were generated on a 200 by 200 pixel lattice, of which the size is not determined a priori. Essentially, every cell of the lattice represents a flow direction corresponding to the optimal outcome of the random searching algorithm. For the purpose of this study, the pixel size is set to 100 by 100 m, implying a total catchment size of 400 km^2 per OCN. The minimum drainage area for a pixel to be classified as part of the river network is set to 1 km^2. The network is discretized into reaches with a maximum length of 1 km. This length may be cut short due to confluences or at the ends of the network (headwaters or outlet) but is equalized as much as possible as to obtain reaches of uniform lentgh. The hydrological conditions are specified at the outlet, and scaled across the network following the following scaling laws (Rodriguez-Iturbe & Rinaldo, 1997):

$$Q_i = Q_o \left(\frac{A_i}{A_o}\right); \quad w_i = w_o \left(\frac{A_i}{A_o}\right)^{0.5}; \quad v_i = v_o \left(\frac{A_i}{A_o}\right)^{0.1} \tag{3}$$

Where $Q, A, w,$ and v represent the discharge, contributing area, channel width, and flow velocity at any location i as opposed to the outlet o. An outlet discharge Q_o of 10 m^3/s is used, which is deemed typical for pre-alpine catchments of this size (Carraro et al., 2021). This flow is facilitated by a channel width of 10 m, a depth of 1 m, and a velocity of 1 m/s.

2.2.3 Generation of synthetic taxa (spatial fields)

A synthetic taxon distribution, represented by random spatially autocorrelated values is generated over the river network. These 'spatial fields' were generated with a multivariate lognormal distribution, where the value in each node reflects the eDNA production rate. The covariance between each pair of nodes depends on their Euclidean distance. This covariance is scaled to generate spatial fields with spatially clustered areas of high values, representing abundance hotspots, and a high frequency of low values. Equation 4 is used to generate the spatial fields:

$$p^{\mathrm{sol}} = \exp(z), \quad z \sim \mathcal{N}(\mu, \Sigma), \quad \Sigma = k_{\mathrm{cov}} \cdot \exp\left(-\frac{\boldsymbol{D}}{\lambda}\right), \quad \boldsymbol{D} = [d_{ij}] \tag{4}$$

Where p^{sol} is an array of values for the production rate in each node ('sol' implying this is the 'solution' of eDITH). p^{sol} is drawn from an exponentiated multivariate normal distribution, with mean μ and covariance matrix Σ. The covariance matrix is based on exponential, which is obtained by dividing the Euclidean distance matrix $\boldsymbol{D}$ by a decay length λ and is scaled by a covariance strength k_{cov}. $\boldsymbol{D}$ is a matrix with n rows and n columns, whose entries are pairwise distances between nodes. The values used for μ, λ and k_{cov} are arbitrarily defined to create realistic-

looking taxon distributions. In a univariate lognormal distribution defined by μ and Σ (= mean and standard deviation of the corresponding normal distribution), the mean depends on both μ and Σ, while the median does only depend on μ. The multivariate lognormal distribution thus generates spatial fields with varying magnitudes and frequencies of high values (i.e. hotspots), representing a wide range of varying taxon distributions. Similarly to using multiple OCNs, this enables the analysis of processes independent of a taxon's distribution characteristics and its specific realization. As this study aims to evaluate the influence of sampling density in a general case, the taxon distribution will be randomized for each simulation run. Figure 3 gives an example taxon distribution.

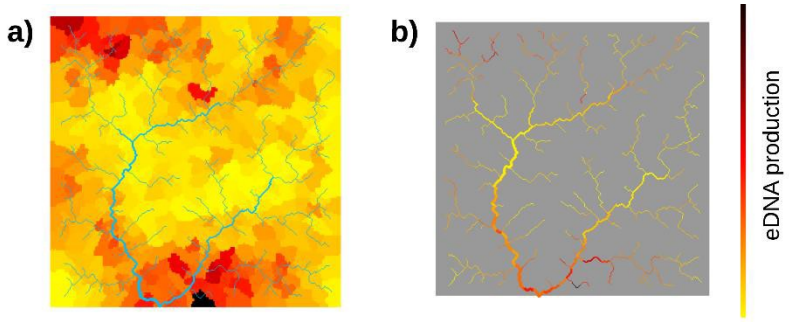


Figure 3 Example taxon distribution. a) the production rate in each sub catchment. b) the production rate in the corresponding river reach.

2.2.4 Computing observable eDNA concentration over the network

A forward model of eDITH was run to determine the steady state distribution of eDNA throughout the network corresponding to the production rate in each node. It is thus the application of equation 1, where all quantities on the right-hand side are known. The decay time is set to 5 hours, which is deemed a realistic estimate (e.g. Deiner et al., 2016, Stewart, 2018). In order to better mimic a sampling campaign, a non-detection probability and measurement error are added to the observable eDNA concentrations (Carraro et al., 2021). The former is given by:

$$p_{\mathrm{nondet},C} = \exp\left(-\frac{C_{\mathrm{sol}}}{100}\right) \tag{5}$$

Where $p_{\mathrm{nondet},C}$ is the probability that the eDNA concentration C_{sol} (being the solution of equation 1). This probability increases exponentially for low C_{sol} and is scaled by an arbitrary value of 100. Afterwards, the measurement error is incorporated as follows:

$$\mathrm{C_{obs}} = \exp(\mathrm{W}), \quad \mathrm{W} \sim \mathcal{N}(\ln(\mathrm{C_{sol}}), 0.25) \tag{6}$$

Where C_{obs} is a perturbed version of C_{sol}, where a lognormally distributed error has been added. As these values are used to calibrate the model, this introduces some error when the modelled production rate is compared to the ground truth.

2.2.5 Virtual sample collection (densities, positioning)

The optimal sampling density is determined by taking samples from the nodes of the network, where concentration of eDNA has been determined with the forward run. The sampling process imitates a real situation where only parts of the network are sampled. Sampling will be performed at different, discrete degrees of intensity, which are given by the percentage of nodes. Based on explorative model runs, these densities are 5%, 10%, 15%, 20%, 25%, and 30%. Given that each node represents a river reach with a specified maximum length, the percentage of nodes is proportional to the number of samples per km^2 of catchment area. Given that the virtual catchments that are used in this study represent have area of 400 km^2 and have a contributing area of roughly 1 km^2 per river network node, the percentage of nodes directly reflects the number of samples that are taken per 100 km of catchment area (10% ~10 samples per 100 km^2, so ~40 samples across a synthetic catchment spanning 400 km^2). Samples are positioned randomly across the network, but constrained such that they are equally divided over the upstream and downstream half of the river network. This was found to be the most robust ratio of upstream to downstream samples in Carraro et al. (2021). The upstream and downstream portion of the river network are defined by the

portion of reaches where the total upstream catchment area either does or does not exceed the median of all catchment areas. In addition to this, a sampling site is added at the outlet node of each river network. This is logical since the downstream transport of eDNA in river systems only provides information of the area upstream of a sampling site.

2.2.6 Estimation of production rate

The sampled values are used as input for the backward run. Here, parameters such as the decay time (eq. 1), baseline eDNA production (eq. 1) and covariate weights in the species distribution model (eq. 2) are fitted, such that the resulting concentration pattern matches the eDNA concentration in the sampled river reaches. The free parameters of eDITH may be estimated using a Markov-Chain Monte Carlo (MCMC) algorithm, which converges towards a posterior distribution for each of the free parameters. This posterior distribution is an empirical collection of all preceding posterior distributions from the Markov chain. One may either use the maximum a posteriori (MAP, mode of the posterior) or the median of the posterior as the final parameter set. While the MAP provides the single, most probable parameter set, the median is determined for each parameter separately, and thus more robust when there is a possibility of multimodality (multiple probable parameter sets). Vrugt et al. (2009) developed DREAM$_{ZS}$, an adaptation of the metropolis algorithm that runs multiple chains simultaneously, leading to higher efficiency (implying a higher acceptance ratio). This method is incorporated into *run_eDITH_BT* (hereafter referred to as BT). BT keeps updating the posterior distribution (using Bayes' rule) until a convergence criterion is met. The specified maximum number of iterations is set to 10^6. The likelihood function used to calculate the posterior distribution in each iteration is set to lognormal, which agrees with the assumed distribution of errors across the spatial field. This allows the random walk to accept parameter sets with a small number of large errors into the posterior distribution, which is necessary to reproduce the lognormally distributed spatial fields.

2.2.7 Modelling taxon distribution (AEMs)

In order to model the spatial distribution of taxa to fit the observed eDNA samples, it is necessary to consider the causes of spatial patterns in these distributions. Spatial variation (structure) in the distribution of species may be explained by environmental factors or neutral processes (i.e. distance-based dispersal, birth, mortality) (Fortin & Dale, 2005). In the absence of environmental data influencing taxon distribution, as is the case with the synthetically generated taxa in this study, eq. 2 may be applied by only using patterns of spatial autocorrelation. While the use of environmental covariates may explain the distribution of taxa reasonably well, spatial autocorrelation at different scales has been shown to explain a larger portion of the spatial variation. This implies that a large part of spatial structure in data, which only partly results from environmental covariates, is reproducible by a weighted combination of spatial patterns. In river networks, Asymmetric Eigenvector Maps (AEMs) may be used to represent spatial structure (Blanchet et al., 2008). These are a special kind of distance-based Moran's Eigenvector Maps (MEMs, Fortin & Dale, 2005), modified for river networks (**see appendix 1**). Rather than Euclidean distance, AEMs consider the along-stream distance to model structure in the data. The main advantage of using AEMs is that they consider the specific behaviour of river ecosystems, being that the flow of matter mainly occurs towards downstream connecting nodes (Peterson et al., 2013). Because of this, AEMs effectively model the relative effect of branches of dendritic networks on taxon distribution. When more AEMs are used, finer differences between branches can be modelled. This comes at the expense of computational time and equifinality, as well as the risk of adding unnecessary noise to the model.

While spatial autocorrelation may explain a larger portion of spatial patterns in data than environmental covariates, a portion of the data observed in real systems will not follow these general patterns and thus not be explained by it (Fortin & Dale, 2005). AEMs are especially helpful for this study, being one with a large number of simulations, because they are quickly determined based on network structure only. In real catchments, however, environmental covariates structure communities in such a way that cannot be explained by distance-based autocorrelation alone. Within the framework of this study, it is therefore important to make sure that the spatial fields do not follow the exact same principles as the AEMs used to estimate them, making predictions too 'easy' to resemble a real-world application of eDITH. While it is theoretically possible to incorporate spatially structured environmental covariates in the generation of taxon abundance maps and use these same environmental covariates (or a subset thereof) to try and reproduce these patterns, this would essentially result in the same artificial match between the original and fitted covariates. However, by generating the spatial fields that are spatially autocorrelated based on the Euclidean distance, rather than the along-stream distance employed by eDITH, one essentially produces patterns that only partly resembles taxon distributions that are autocorrelated along-stream. This leads to patterns that are more resemblant to those of terrestrial taxa, or of taxa whose distribution is not solely explained by network connectivity

(e.g. taxa that are only found in streams with a certain surrounding vegetation whose distribution is not limited by the structure of the river network, but rather the whole terrestrial landscape). Accordingly, eDITH is tested in a more complex, realistic context, where an unavoidable mismatch between the modelled and true spatial fields is known a priori.

2.2.8 Quantifying performance of predictions

Any simulation outcome may be evaluated qualitatively through visual inspection or quantitatively by using an error metric. It is necessary to quantify the performance of a single simulation in order to summarize the outcome over a larger number of simulations. Regression-based error metrics summarize the total error of model predictions. Here it is important to choose the error metric that best fits the type of data and the analysis. Since this study evaluates the use of eDITH to transform eDNA concentration data to a production rate, a quantification of the difference between the original and modelled p values across the river network may yield the most. This section goes over a number of metrics that could be relevant for this analysis. By evaluating multiple error metrics, each highlighting different types of errors, differences among these metrics reveal differences among model settings. The Mean Absolute Error (MAE) gives the overall deviation of the simulated production rate from the ground truth:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|p_i^{\mathrm{mod}} - p_i^{\mathrm{sol}}\right| \tag{7}$$

Where p^{mod} is the modelled producstion rate. Across simulations, however, the metric overlooks differences in the distribution (mean and variance) of the production rate. As a result, it gives more weight to model runs where the spatial fields have a higher mean and variance. To address this, the Mean Absolute Scaled Error (MASE) scales the MAE by the variance in the modelled data:

$$\mathrm{MASE} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left|p^{\mathrm{mod}}i - p_i^{\mathrm{sol}}\right|}{\frac{1}{n}\sum_{i=1}^{n}\left|p_i^{\mathrm{sol}} - \bar{p}^{\mathrm{sol}}\right|} \tag{8}$$

Where $\bar{p}^{sol}$ is the mean of p^{sol}. While the MASE appears to be more suited to draw conclusions independent of the specific taxon, scaling the error of each simulation comes at the expense of potentially overestimating the weight of minor errors when the true spatial field is relatively homogenous, and underestimating the error when the model wrongfully predicts a very homogeneous spatial field (i.e. missing hotspots). In order to facilitate a more local error assessment, independent of characteristics across the whole network, the symmetric Mean Absolute Percentage Error (sMAPE) may be used:

$$\mathrm{sMAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{\left|p_i^{\mathrm{mod}} - p_i^{\mathrm{sol}}\right|}{\left(\left|p_i^{\mathrm{mod}}\right| + \left|p_i^{\mathrm{sol}}\right|\right)/2} \tag{9}$$

While the sMAPE is able to weigh the relative error in each reach of the river network separately, the risk is that large errors are assigned to reaches where the true production rate is close to zero, while eDITH also predicts a relatively low value. A remaining issue with regression-based error metrics is that they treat the error in each node of the network equally. However, in reality, predictions in one node may be more important than the other. This issue is particularly apparent when the goal is to identify hotspots of taxon distribution. One may decide to evaluate the squared error to increase the weights of large errors, believing that these are more likely to occur near hotspots where the actual values are also largest. However, in the case of the correct prediction of a hotspot, this effect will over-emphasize large errors outside of these hotspots, thus reversing the desired effect.

For these types of scenarios it may therefore be more informative to use a classification-based error metric. To this end, the True Skill Statistic (TSS) gives the overall model performance as a combination of the ratio of correctly predicted presences and absences (Allouche, Tsoar & Kadmon, 2006):

$$TSS = \mathrm{Sensitivity} + \mathrm{Specificity} - 1,$$

$$\text{Sensitivity} = \textit{True Positive Ratio (TPR)} = \frac{TP}{P}, \tag{9}$$

$$\text{Specificity} = \textit{True Negative Ratio (TNR)} = \frac{TN}{A}$$

Where TP and TN are the true positives and true negatives (i.e. the quantity of correctly predicted presences and absences respectively), and P and A the total number presences and absences in the original spatial field. By weighing the TPR and TNR separately, this metric is better able to quantify eDITH capacity to correctly predict presences and hotspots that only occur in a small portion of the catchment. Since this metric treats the data as binary (being either present/positive or absent/negative), it lends itself particularly for datasets that include absences. Since the spatial fields for this study are generated according to a lognormal distribution, both the true and modelled spatial fields contain a large portion of low values. This makes it difficult to impose a threshold that separates presence or absence beforehand, without the true and predicted value each falling just on the other side of this threshold. As will be explained in section 2.3.2, this study therefore transforms the dataset to enable the use of the TSS

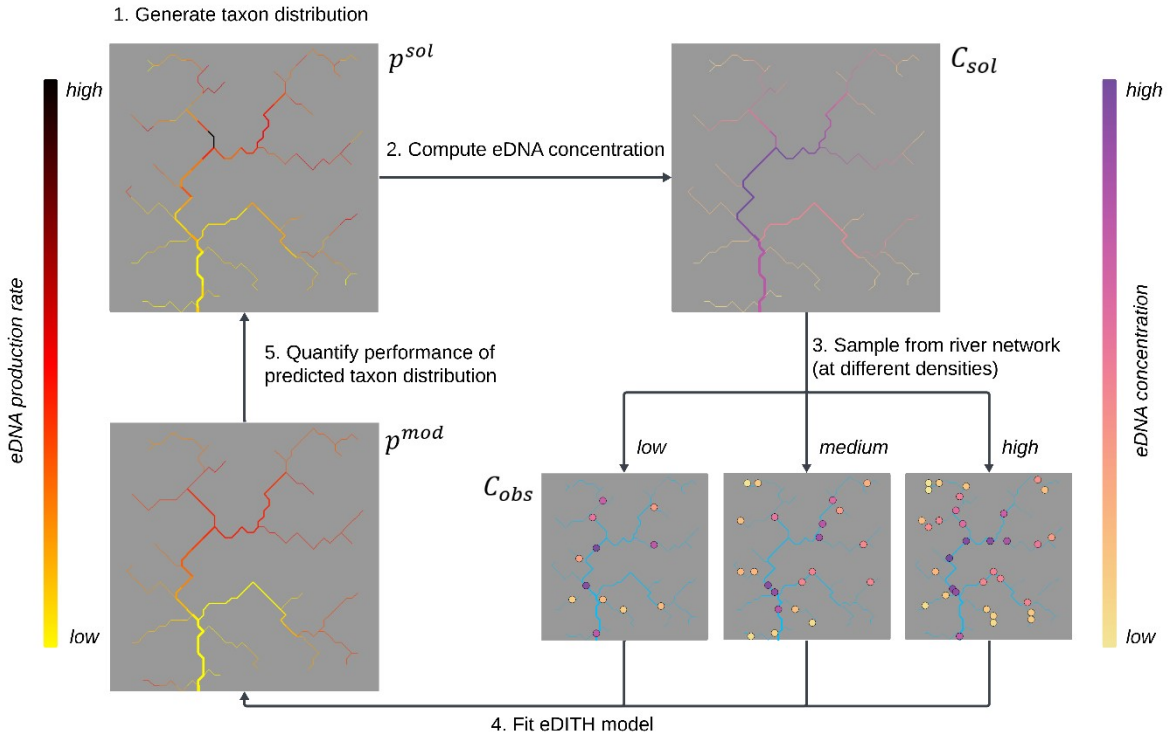


figure 4 Overview of the virtual experiment described in section 2.2 used to assess the performance of different degrees of sampling densities (adaptation of figure in Carraro et al. (2021)). For illustrative purposes, this image shows a simplified version of the virtual river networks and taxon distributions used in this study.

2.3 Study design

2.3.1 Overview

In order to address the research question, finding the optimal sampling density for taxon distribution predictions with eDITH, it is necessary to define the exact setup of the analysis. This setup was informed by initial simulations, as briefly described in appendix A2. By repeating the procedure of the virtual experiment a large number of times, the performance across simulations can be determined while changing certain factors. To investigate how changes in sampling density affect the performance, the error metrics of the individual simulations were grouped according to the degree of sampling density. Accordingly, it may be evaluated to which extent the lack of sufficient sampling

limits model performance, and at which point adding more samples does not meaningfully improve it. To achieve this, the distribution of these scores is visualized for each group by means of a boxplot. Additionally, a Wilcoxon signed rank test is used to compare each degree of sampling intensity to the next one. This test has the advantage of being non-parametric, making it suitable for lognormally distributed data. Since the large number of replicates leads to an artificially low p-value, the effect sizes were explored to make the comparison (White et al., 2014). Besides the sampling density, this study also investigates how the specified level of detail of the predictions affects the performance for each degree of sampling density. Additionally, the analysis is repeated for three different taxon distribution types. By combining these three components in a 'full factorial' study design (table 1), it is possible to study how they are interrelated. The following subsections explain each part in more detail.

2.3.2 Spatial resolution of the predictions

Besides investigating the relationship between sampling density and the performance of eDITH, this study also considers the extent to which spatially complex taxon distribution patterns may be modelled by eDITH. Inquiries into the effect of model complexity are common within the virtual ecologist framework (Zurell et al., 2010). As explained in section 2.2.7, the amount of AEMs governs the level of detail at which the optimization algorithm seeks to recover the original signal. While more fine-scale predictions may better approximate the level of detail that is found in the synthetic taxon distributions, this may only be achieved when sufficient sampling points are present. Thus, investigating the relationship between the spatial resolution of the taxon distribution predictions and the performance may shed light on the optimal resolution at different sampling densities. Moreover, this analysis may reveal how other factors, such as the behaviour of eDNA in river networks, the optimization method, or the setup of the virtual experiment itself, limit the resolution at which complex taxon distribution patterns may be determined. These insights can help eDITH users to determine the appropriate scale at which they may interpret eDNA data, balancing accuracy and precision. As shown in table 1, the virtual experiment is repeated for five levels of pattern resolution, ranging from a relatively coarse to fine scale.

2.3.2 Variations in taxon distribution

One may also use eDNA to recover the distribution of a rare or invasive species that only resides in a small number of nodes across the river network. While the spatial fields generated with the multivariate lognormal distribution represent a wide range of virtual taxon distributions, it does not contain cases where a taxon is only present in a part of the catchment, and absent elsewhere. To address this, the lognormal distribution was 'filtered' such that all values below a specified percentile were converted into zeros, representing the absence of the taxon. This was done to two degrees, one distribution retaining the 25% highest values, and one retaining 5% of the highest values. Essentially, these 'filtered' spatial fields represent taxa that are more sparsely distributed across the catchment, with the 25th percentile resembling a more scattered version of the original spatial field, and the 5th percentile a situation where only a small number of hotspots remain (figure 5). While, in reality, a rare taxon may also be less abundant within these remaining reaches, the spatial fields represent arbitrary production rates that may be scaled to any magnitude. However, it has to be noted that this method assumes that the production rates remaining in the filtered spatial fields lie well above the threshold of non-detection, while this may not always be the case in a real environment.

Creating different taxon distribution types also allows a different assessment of the prediction skill of eDITH. While the regression-based error metrics provide a more complete indication of the overall 'closeness' of the predictions, the TSS enable the focus on specific areas The scattered taxon enables the use of the TSS on spatial fields with similar patterns to the general taxon, and may thus be seen as facilitating the use of a different error metric. The rare taxon represents an entirely different situation, where the actual goal of eDITH becomes to recover each presence, rather than precisely quantifying the abundance across the catchment. Here, the TSS, being a combination of the TPR and TNR, enables the performance analysis to focus on the correct prediction of presences, without interference from the large number of absences across the catchment. In order to compute the TSS, p^{sol} and p^{mod} were converted to presences and absences using a threshold of ¼. This value was chosen low enough such that all nonzero values of the scattered and rare taxon distributions were treated as presences, but high enough to prevent the false indication of presences in the predictions.

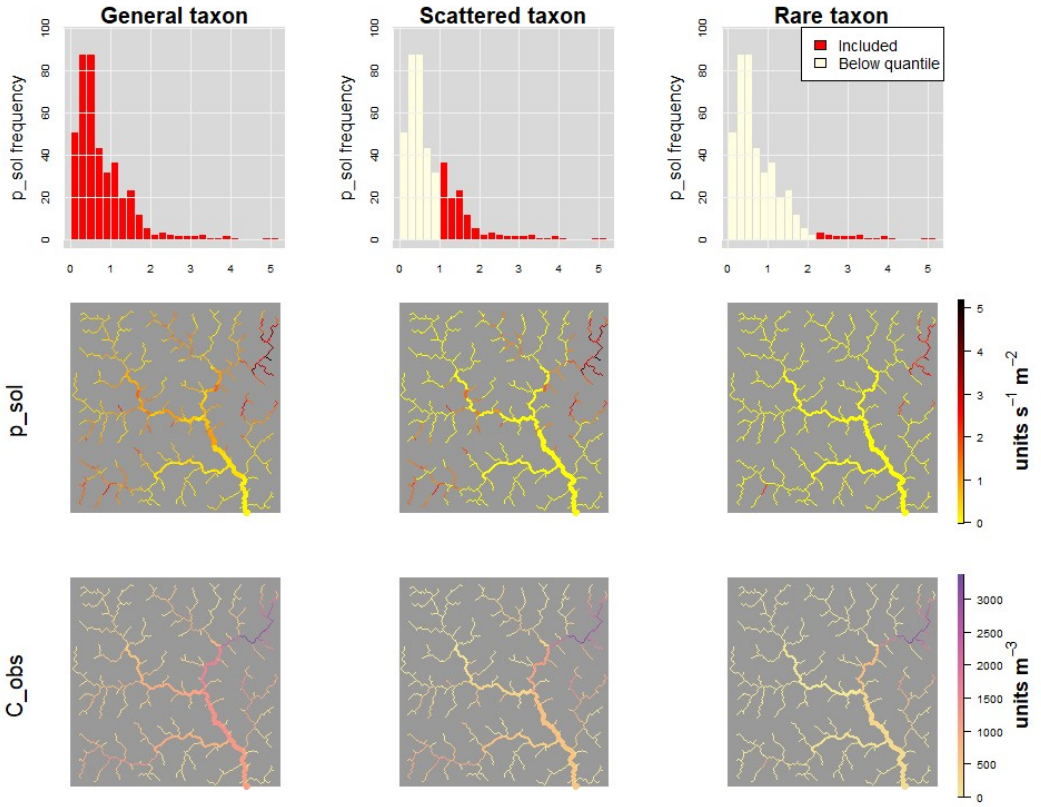


Figure 5 – The different taxon distribution types that were used in the analysis. A subset of the original spatial field is created by converting all values below a certain percentile to zero. This results in spatial fields that represent general, scattered, and rare taxon distributions. The bottom row shows how these filtered spatial fields (scattered and rare) produce an eDNA distribution characterized by smaller values in the main stem downstream of the hotspot.

Table 1 Overview of the full factorial experiment. The number of realizations represents the amount of simulations that is performed per level. Note that for the three taxon distribution types and five levels of AEMs used, all 6000 simulations across sampling intensities were repeated. This table thus represents a total of 90000 simulations.

Factor	Number of levels	Levels	Realizations	Wording
Sampling density	6	5%, 10%, 15%, 20%, 25%, 30%	1000	low-high density
Taxon distribution type	3	Full field, 25th percentile, 5th percentile	1000 per sampling intensity	general, scattered, rare
Model complexity (number of AEMs)	5	9, 12, 15, 18, 30	1000 per taxon distribution type	coarse-fine scale patterns; low-high resolution

3. Results

3.1 Optimization

The results were obtained following the full factorial approach from table 1. Details on the computation process are provided in appendix A3. For each degree of sampling density, the original taxon distribution was estimated at five degrees of spatial resolution. This process was repeated a thousand times for the general, scattered, and rare taxon distributions. Using the *BayesianTools* optimization method, each simulation yielded a series of parameter sets for the unknown covariate effects $x(i)$ that make up spatial taxon distribution predictions, expressed as the p^{sol} (eq. 2). The median of the posterior distribution of each single parameter led to better average performance than the Maximum A Posteriori (MAP). This suggests that there is indeed multimodality present in the solution space, which deems the use of the median of the posterior parameter distributions more suitable. The error between the observed and fitted eDNA concentrations is visualized in figure 6. The performance of the optimization itself was relatively equal across the various degrees of sampling intensity and model complexity. Most notably, the error is somewhat larger for more complex model variants when the sampling density is low.

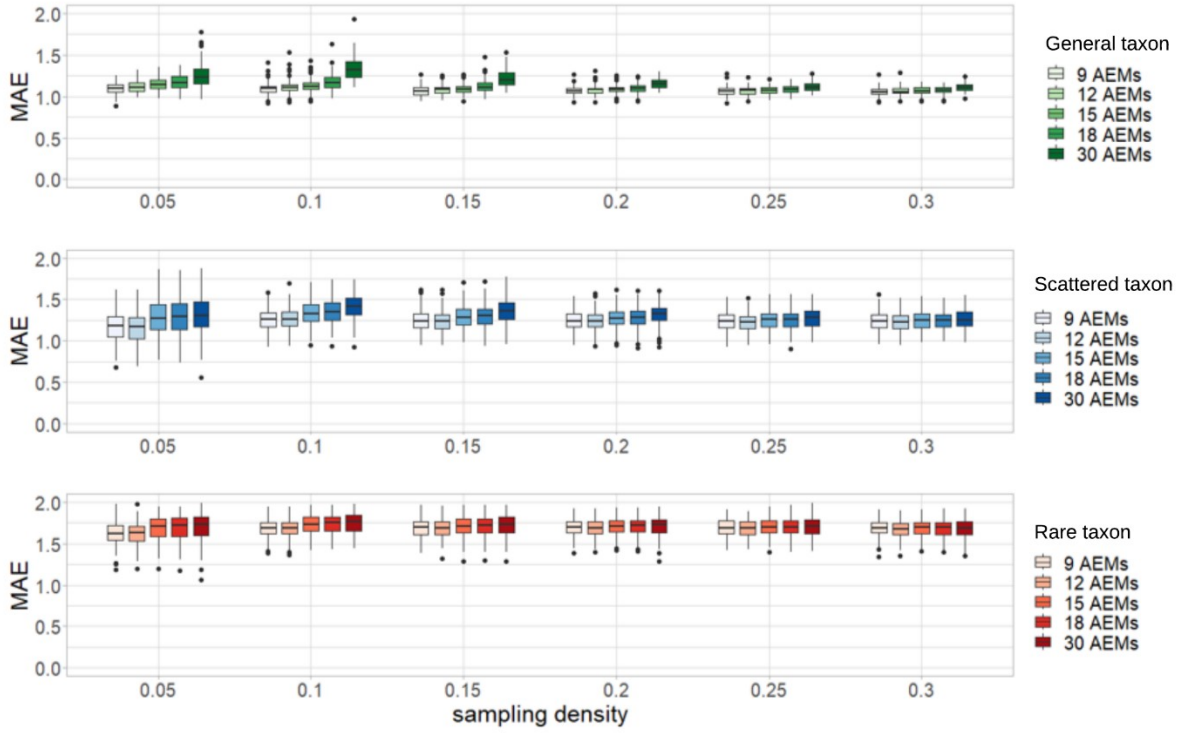


Figure 6 Mean absolute error of log-scaled values of C_{obs} and C_{mod}.

3.2 General taxa

Figure 7 provides insight into the different steps that were involved following the virtual ecologist approach described in section 2.2 and figure 5 for a subset of these simulations. This approach was successful in creating a large number of virtual taxon distributions, and enabled testing different sampling strategies. The performance of each of these simulation runs is evaluated using the error metrics described in 2.2.8. The performance of eDITH in reproducing general taxa was evaluated by means of the mean absolute error (MAE), mean absolute scaled error (MASE), and scaled mean absolute percentage error (sMAPE).

The results of the virtual experiment of the general taxa are summarized in figure 8. As expected, increasing the sampling intensity leads to a notable increase in performance. While each of the three metrics have a different magnitude, they all show a similar dynamic between sampling intensity and the spatial resolution of the predictions in relation to performance. At low sampling densities, the performance is highest when a relatively small number

of AEMs is used, implying that estimations of p^{sol} with a lower spatial resolution more closely match the original taxon distribution. This tendency weakens at higher sampling densities, up to a point where more detailed predictions perform slightly better. However, using 30 AEMs, representing the highest degree of spatial resolution that is asked of eDITH, never yields the best average performance. The optimal amount of AEMs used shifts upwards, as increasing the density of samples across the catchment allows defining more fine-scale differences in taxon distribution. Within the explored range of 5 to 30 percent of the nodes, this shift appears to be from 9 AEMs up to 18 (or possibly more, but less than 30) AEMs. As sampling intensity increases, the performance gains, both between one degree of intensity and the next and between the amount of AEMs within each degree of intensity, appear to increase asymptotically. The evaluation of different error metrics was intended to highlight different features of the dissimilarity between the original and modelled taxon distribution. The similarity between these scores, however, suggests that these differences are small across a large set of simulations.

Multiple examples of the simulations are provided in figure 9, providing a more intuitive understanding of the effect of sampling intensity on performance. Here, each notable decrease in error score represents a visible improvement in the predictions. Moreover, among the higher sampling intensities, there are only small differences in scores. This is also visible when looking at the maps; they look roughly the same. This suggests that the error metrics give a good indication of the relative performance between model runs. Interestingly, the relative scores of MAE and MASE are reversed between these three sets of simulations: the middle one has the lowest MAE and the highest MASE, the bottom one has the highest MAE and the lowest MASE, and for the top one both the MAE and MASE lie in the middle. None of the metrics particularly penalize missing the hotspot in the bottom-center of the second row at a sampling intensity of 5%. On the other hand, all consecutive degrees of sampling intensity are missing the hotspot next to it. In any case, each simulation appears to perform comparably well in reproducing the original taxon distribution, each missing a hotspot while correctly identifying broad spatial patterns. The different error metrics do not appear to perform differently in terms of accounting for errors in identifying hotspots.

Figure 10 explores the varying results for a different degree of sampling intensity and number of AEMs for a specific taxon distribution. Figure 8 is essentially summarizes this figure over a thousand replications. Appendix A3 gives the effect sizes between each degree of sampling intensity and the next, as well as each degree of spatial complexity (given by the number of AEMs) and the next. Comparison between these two factors suggests that, while increasing the sampling density will very likely lead to some increase in prediction performance, the effect of changing the spatial complexity of the predictions is accompanied by considerable uncertainty.

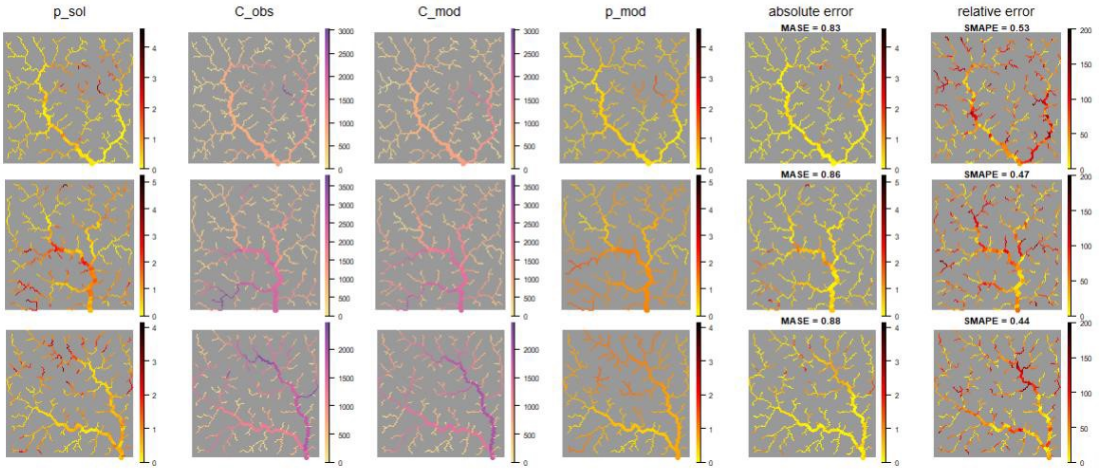


Figure 7 - From left to right - from generated taxon (p_sol) to prediction and performance. Three example model runs at the same settings (15 AEMs and 10 percent of the reaches sampled) are displayed. The performance evaluation is visualized with the absolute and relative error, which are quantified using the MASE and sMAPE respectively. The values of p^{sol} and p^{mod} are in units/s/m^2, the values of C_{obs} and C_{mod} are in units/m^3.

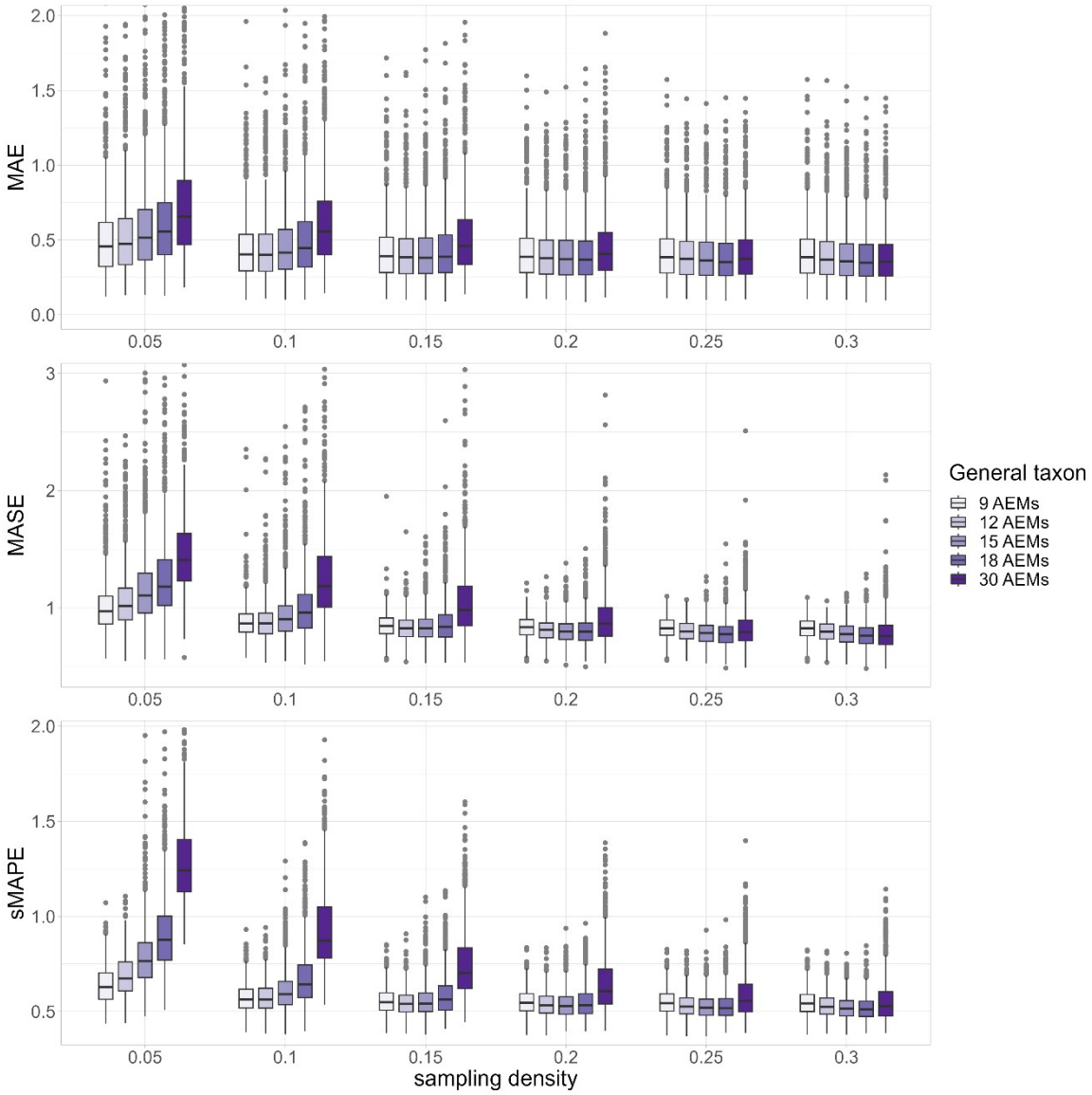


figure 8 MAE, MASE, and sMAPE of the simulations of the general taxon, for varying degrees of sampling intensity and model complexity. The y-axis has been truncated to make sure that fine differences between the boxplots can be visually assessed. Within each degree of sampling density, the actual range of the outliers above the whisker is roughly proportional to the length of the whisker.

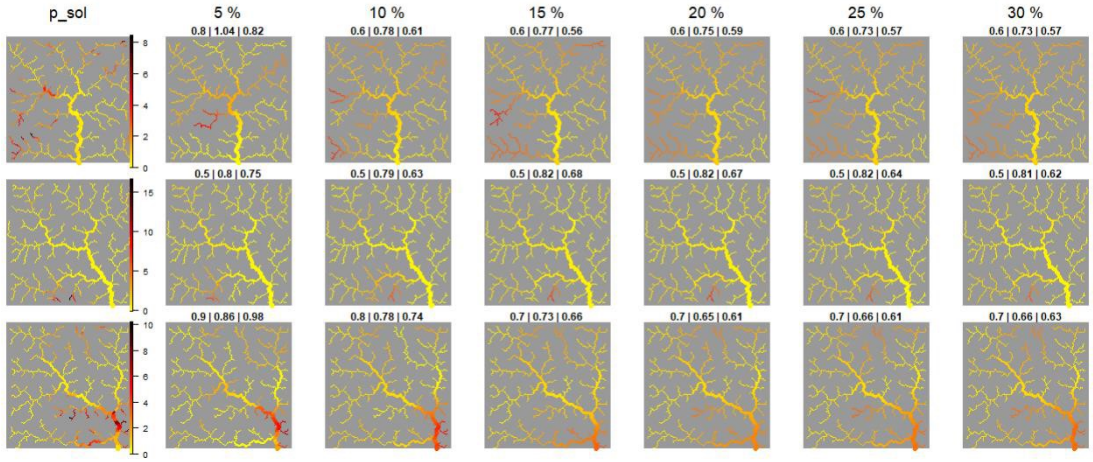


Figure 9 Original vs modelled distribution of the general taxon. Illustrative examples of simulation results at different sampling intensities and intermediate spatial resolution (15 AEMs). Color bars show the production rate in units/s/m^3 and has the same definition for map in the same row. The numbers above each plot give the MAE, MASE, and sMAPE respectively.

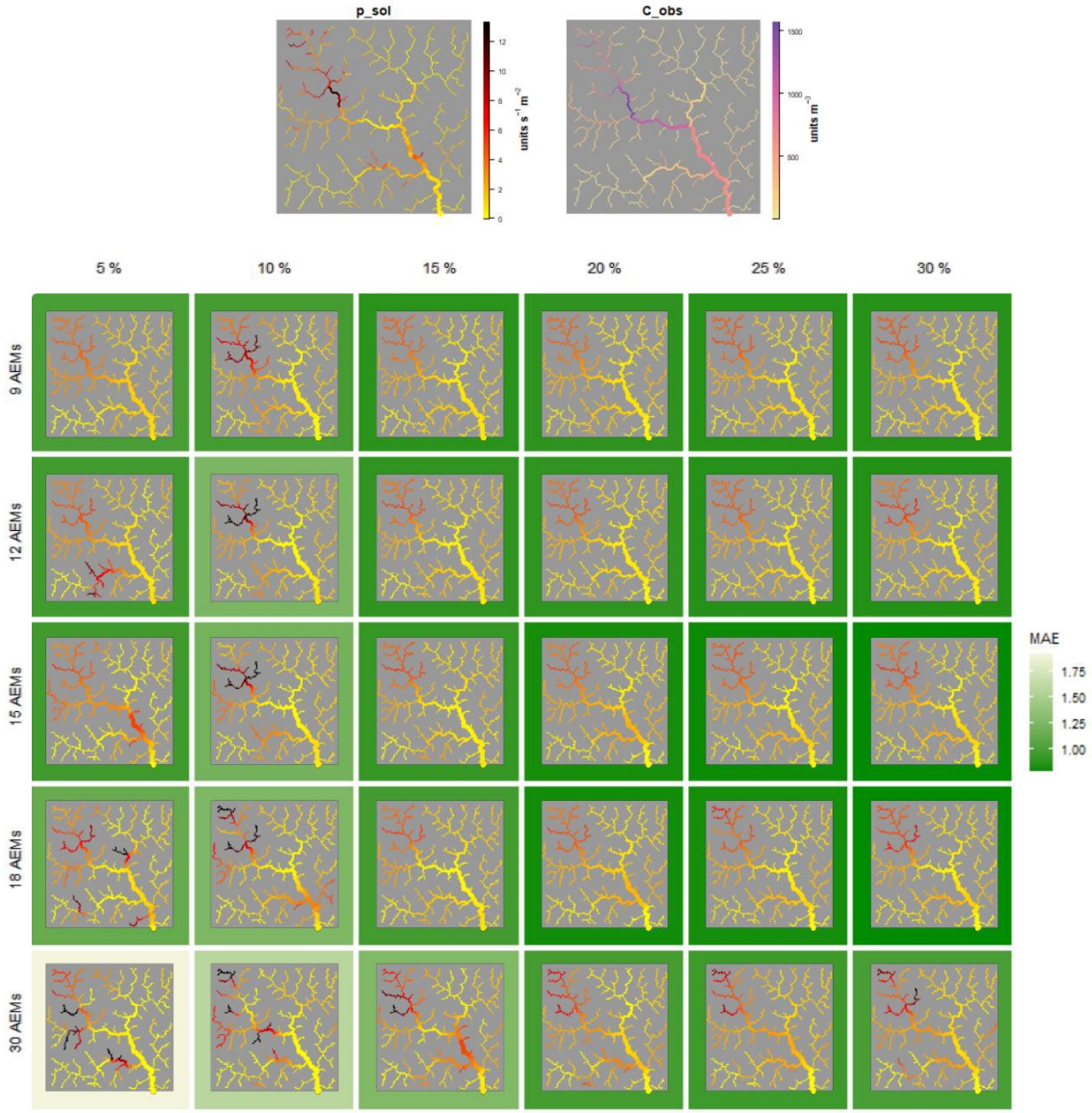


Figure 10 Original vs. modelled production rate across the explored degrees of sampling intensity and model complexity (number of AEMs) for a specific realization of the general taxon. For this taxon distribution, it can be seen that sampling more generally improves the performance. Furthermore, fine-scale predictions (represented by a larger amount of AEMs) are more likely to wrongly predict hotspots in small tributaries. This is especially evident at low sampling densities, suggesting that there are no samples taken these specific tributaries. Such anomalies do not appear to occur at low levels of spatial resolution for this specific taxon.

3.3 Scattered and rare taxa

Figure 11 and 12 display the true skill statistic (TSS), as well as its constituents true positive rate (TPR) and true negative rate (TNR). These results show both similarities and differences between the scattered and rare taxa. For both the taxon distribution types, the positive effect of sampling density on performance appears to be stronger for more detailed estimations of the taxon distribution (using a larger number of AEMs). While the quantification of the error was different here, overlooking the accuracy within single reaches, this interplay between sampling density and spatial resolution coincides with the results from the general taxa. The TPR is most strongly affected by the sampling density, implying that an increase in sampling effort strongly reduces the occurrence of false negatives in the predictions. The TNR, on the other hand, is less affected by sampling density. Logically, an increase in the correct prediction of presences comes at the expense of the correct prediction of absences. For both taxon distribution types, this balance appears to be controlled by the number of AEMs, with more detailed predictions having a high specificity but low sensitivity. This likely happens when the production rate is correctly estimated to be high in a certain section of the river, but the spatial extent of this production (and thereby the spatial extent of the taxon) is underestimated. Thus, in order to correctly locate as many reaches as possible that contain the scattered and rare taxa, it is necessary to define larger sections of the catchment that may contain these taxa. This requires choosing predictions with a low spatial resolution, and accepting a considerable amount of false positives.

For the scattered taxon, the TSS of predictions with a low spatial resolution (using a small number of AEMs) are generally not affected much by the sampling density. The opposite is true for more detailed predictions (using a larger number of AEMs), where more samples strongly increase the TSS. Following this metric, the relationship between spatial resolution and performance shifts from negative to positive at a sampling density of 15% and higher. Thus, in contrast to the evaluation of the general taxa, this figure indicates that at high sampling intensities, a large number of AEMs (here 30) yield better performance. While this results in an increase in performance up to a sampling density of 30 percent, the TPR is already close to its peak at 15 percent. At this point, using a low spatial resolution, eDITH correctly locates approximately 90 percent of the reaches containing the rare taxon, while incorrectly suggesting its presence in 50 percent of the reaches where it is not located. Example simulation results are provided in figure 13. For the scattered taxon, the TNR gets especially low when the hotspots are scattered out over the map. In order to capture all the hotspots that are noted at higher sampling intensities, it is necessary to approximate larger regions to contain the taxon. However, this is only achieved when estimates are made at a low spatial resolution. This is supported by the results from figure 11, which shows that simulations with a low number AEMs contain considerably more false positives.

For the rare taxa, the increased performance for more detailed predictions is notably weaker, as a larger number of AEMs does not attain a higher TSS at high sampling intensities. As a result, a similar trend to that of the general taxon can be observed. However, the TPR increases notably until a sampling density of 20 percent. At this point, using a low spatial resolution, eDITH correctly locates approximately 85 percent of the reaches containing the rare taxon, while incorrectly suggesting its presence in 15 percent of the reaches where it is not located. Example simulation results are provided in figure 14. A low sampling density likely favours absence predictions because the limited eDNA production across the catchment makes it hard to be observed in randomly sampled sites. Sampling a larger portion of the reaches leads to an increase in presence predictions, as each signal becomes more likely to be noted when more reaches are sampled. The association between a high TPR and a reduced TNR appears to be either the consequence of a limited number of AEMs, an overestimation of the decay time, or a combination of the two. At the explored sampling intensities, using more AEMs (towards 30) leads to less estimations of presences, at the cost of a lower TPR. Thus, with insufficient sampling, false negatives are more likely.

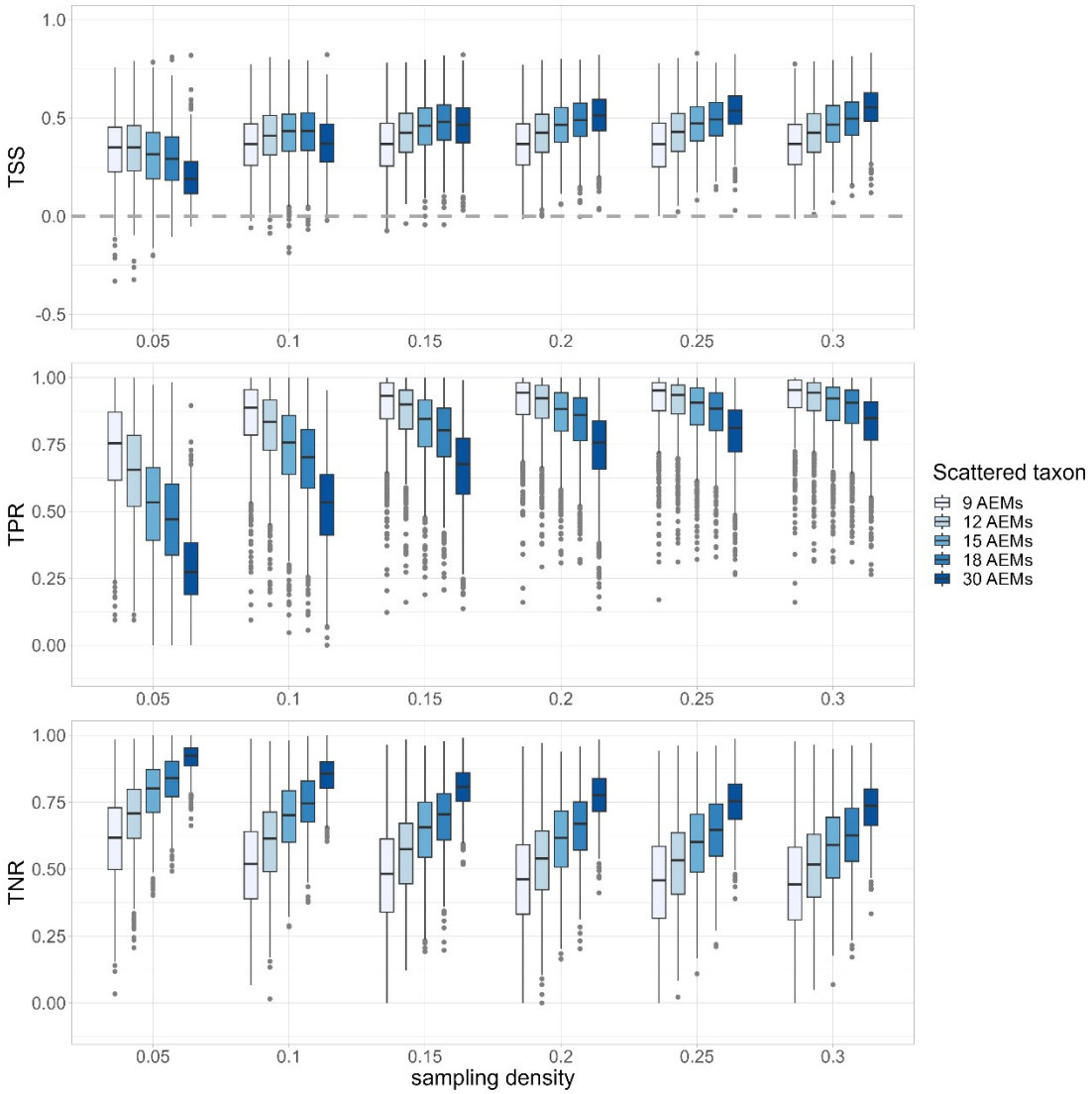


Figure 11 TSS, TPR, and TNR for the scattered taxa.

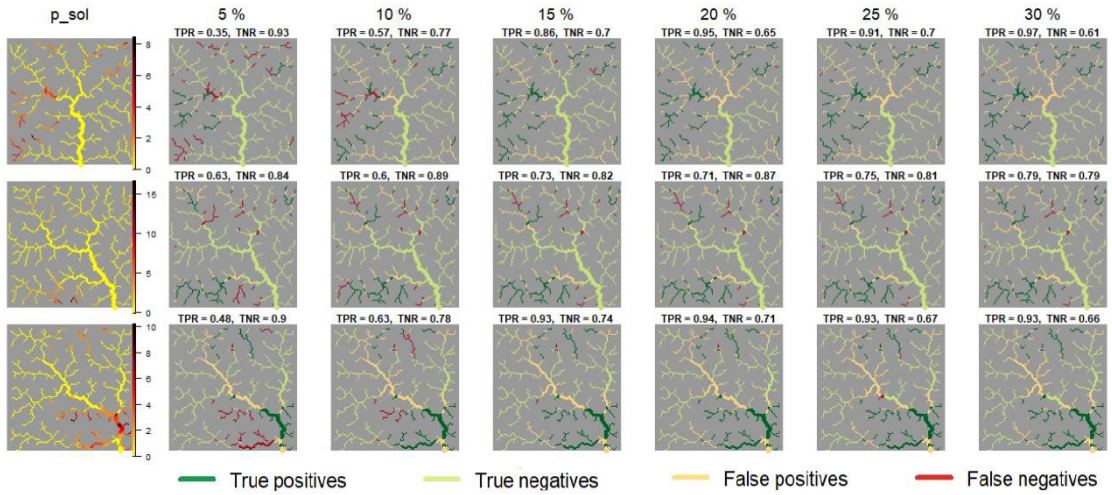


Figure 12 Original production rate compared to the performance of modelled presence/absence distributions for the scattered taxon. Illustrative examples of simulation results at different sampling intensities and intermediate spatial resolution (15 AEMs). The colorbars show the production rate in units/s/m^3.

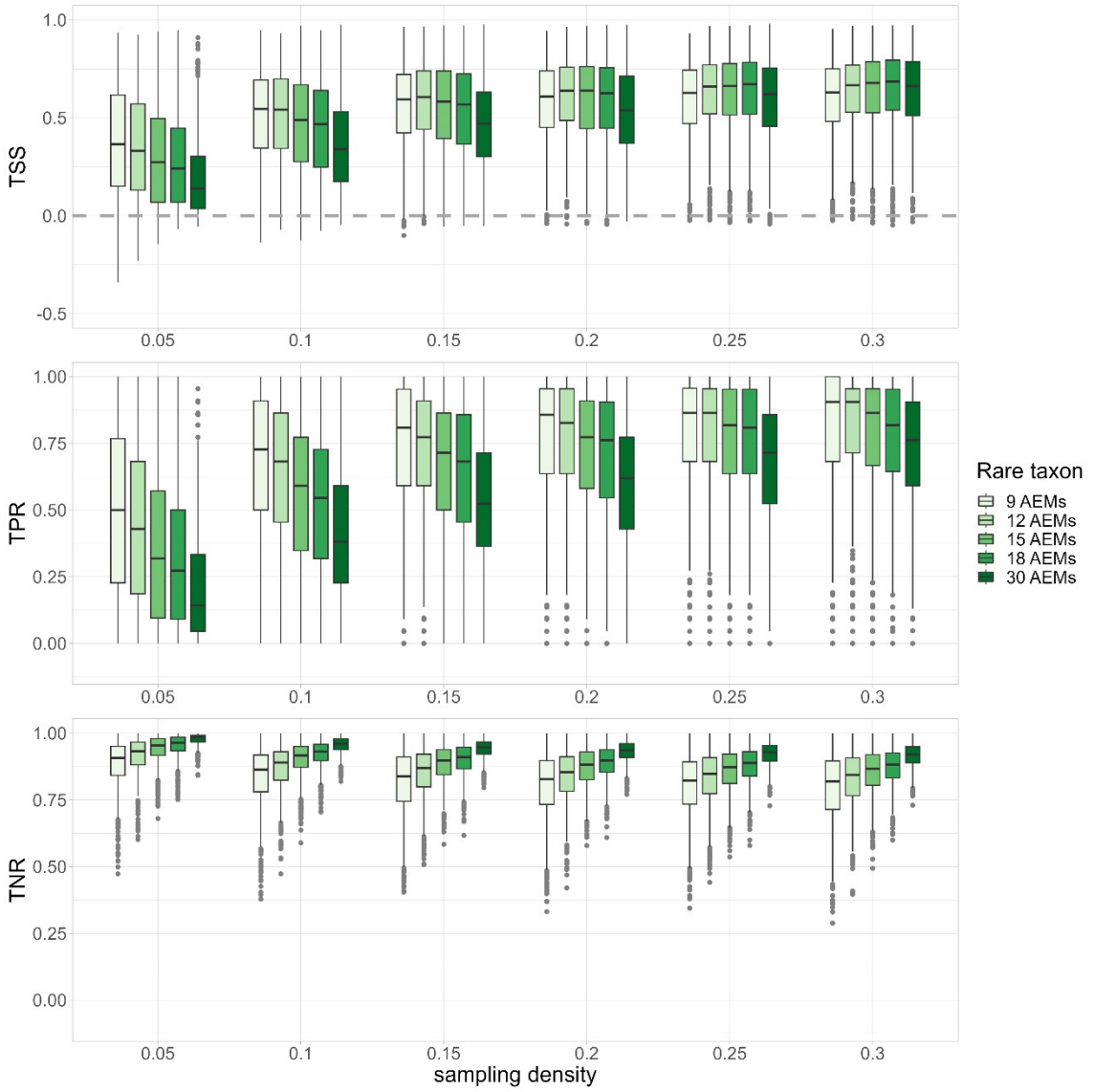


Figure 13 TSS, TPR, and TNR for the rare taxa

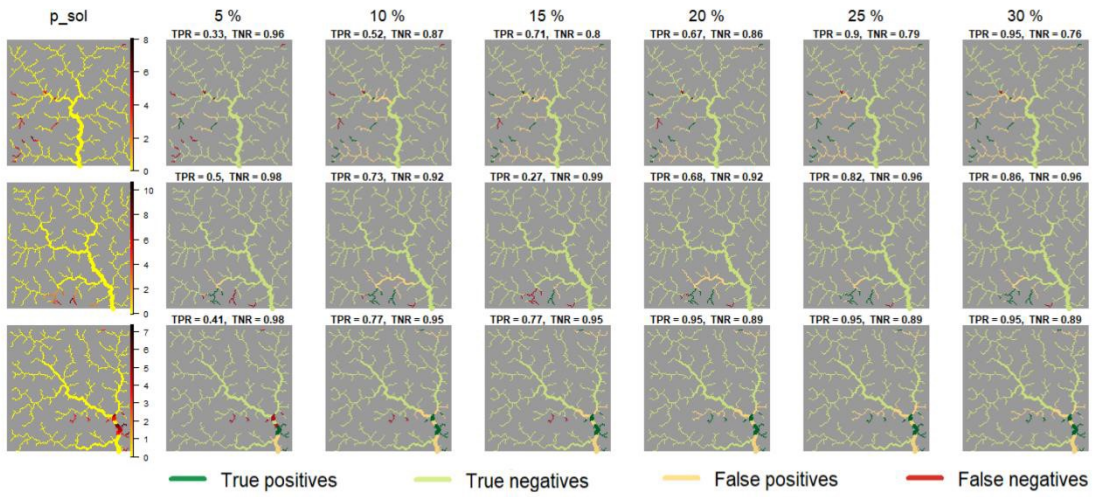


Figure 14 Original production rate compared to the performance of modelled presence/absence distributions for the rare taxon. Illustrative examples of simulation results at different sampling intensities and intermediate spatial resolution (15 AEMs). The colorbars show the production rate in units/s/m^3.

4. Discussion

By modelling the fate of eDNA in river networks, it is possible to transform pointwise measurements into maps of spatial taxon abundance resolved at the reach scale. Using eDITH, this study aimed to build on this methodology by investigating the relationship between the eDNA sampling intensity and the quality of the taxon distribution predictions. Moreover, the full-factorial design of this study has shed light on the role of the spatial resolution of these predictions and differences between taxon distribution types. This chapter starts with an analysis of the results, then goes on to discuss limitations of the study design, after which it places these findings in a practical perspective and provides concluding remarks.

4.1 Analysis of results

This study found that both the sampling intensity and the specified spatial resolution affect the quality of taxon distribution predictions with eDNA. The analysis primarily focussed accurate estimation of a wide array of general taxon distributions. Here, a sampling density of at least 10% of the reaches strikes an efficient balance between sampling effort and accuracy. This equals taking one sample per 10 km^2 of catchment area. When it specifically concerns the detection of taxa that only occur in a small portion of the catchment, doubling the amount of sampling sites yields notable gains in performance. This aligns with previous findings from Carraro et al. (2021), who suggest a stronger relationship between sampling density and performance for such taxa. Among all taxon distribution types, the performance may be maximized marginally by further increasing the sampling density, as finer patterns in taxon distribution may be recovered. By summarizing the performance across a wide range of catchments and taxa, this study directly addresses the lack of standardized sampling guidelines for eDNA-based biomonitoring. While previous work has made significant advances in the understanding of eDNA transport and decay, and its link to sampling positioning (e.g., Deiner & Altermatt, 2014; Carraro et al., 2020), these findings provide a clear perspective on the optimal sampling density necessary to ensure accurate taxon distribution predictions across catchments and taxa. Based on the objective, target taxon, and budgetary constraints, it may be decided how many samples are needed to obtain a taxon's distribution from its eDNA.

This study employed different error metrics to assess how well eDITH predicts taxon distributions under different conditions. Regression metrics, such as the Mean Absolute Error were used to quantify the 'closeness' of the predictions to the actual values. The relative performance across simulations was fairly consistent between the different metrics, and visual comparisons suggest that the choice of metric did not significantly distort the overall performance quantification. Likely, the large number of simulations employed also smoothened out differences between these metrics that may be more apparent when only a small number of taxon distributions were analysed. This suggests that the findings on the general taxa provide a robust estimate of the optimal sampling density when the goal is to maximize overall accuracy across the catchment. The True Skill Statistic (TSS), by highlighting different types of errors, enabled a more focused assessment of the model's capability to correctly locate hotspots across the catchment. The findings concerning the scattered taxon suggest that taking more samples improves the overall precision of the model, while coarse predictions with a lower sampling density suffice when the priority is to first identify larger areas that contain the taxon. However, in determining the optimal sampling strategy, it should be considered that doubling the sampling density may on average only reduce the number of false positives by 20 percent. In contrast, the strong relationship between sampling intensity and True Positive Ratio (TPR) for the rare taxa suggest that at least 2 samples per 10 km^2 are necessary to determine their full spatial extent across the catchment. These findings demonstrate that, while the TSS may offer a less complete summary of the performance of a single simulation, it enables a more practical interpretation of the quality of the predictions. The TPR enables a specific focus on locating a limited number of hotspots across the catchment.

While the main findings show that a sampling intensity of 1 to 2 samples per 10 km^2 of catchment area balances sampling effort and prediction quality, the residual error observed in the model predictions after optimizing the sampling intensity highlights inherent limitations. Insights into the effect of model complexity may inform whether the setup allows for convergence towards the right solution. While more AEMs could allow the model to better reproduce the taxon distributions responsible for the observed concentration patterns, they may also introduce equifinality, meaning that multiple finer patterns of eDNA production may fit the concentration data comparably well. The latter seems particularly evident at low sampling intensities where, similarly to the production rate, a larger amount of AEMs leads to larger errors between the observed and modelled C's at the sampling sites. Most likely, there are simply not enough samples for the larger number of AEMs to be properly fitted to, causing them to distort the optimization. The observation that neither of the proposed effects of model complexity is apparent at higher

sampling intensities suggests that increasing the model complexity further does not improve the quality of the predictions in a meaningful way. While future analyses could explore this relationship at higher sampling intensities, the observed trends suggest that improvements in the performance are expected to be marginal at best. The remaining error that is observed across simulations is then likely a combination of the inability of the AEMs to precisely reproduce the generated spatial patterns, the optimization method itself and a distortion of the signal inherent to river solutes. While the latter may be addressed by taking additional samples in combination with more AEMs, moving in this direction would defeat the purpose of this study, which is not to achieve perfect predictions but rather finding an optimal trade-off between sampling effort and model accuracy to facilitate efficient biomonitoring.

Comparisons with similar studies may help to further evaluate the efficacy of this study and the predictions obtained with eDITH. Using a similar setup, Carraro et al. (2021) found that increasing towards higher sampling densities kept improving performance. This is likely the case because they predicted the abundance in each reach separately. The resulting relationship between sampling density and performance is similar to what would be expected in this study when predictions are made at a very high spatial resolution. Similarly to this study, their analysis only facilitates a relative comparison of different strategies. Alternatively, relating this study to those comparing taxon distribution data obtained with eDITH to other forms of in-situ data may provide insight into the quality of the predictions made with eDITH. However, as these studies each take an approach to sampling eDNA, estimating the taxon distribution, and quantifying performance that is different from this study, it is difficult to connect optimal sampling strategies to absolute performance. Carraro (2018) found that observed and modelled eDNA concentrations correlated well ($r = 0.8$), but did not quantitatively summarize the performance of predictions across the catchment. Blackman et al. (2024a) found a decent relationship between modelled and observed taxon richness using metabarcoding data ($R^2 = 0.33$). As opposed to this study, which used the measured concentration using qPCR, metabarcoding requires the conversion of the read-count of each taxon to a detection probability, which is then converted to a binary presence/absence prediction. Using a similar approach, Carraro et al. (2020) found an average accuracy of 82.4% across macroinvertebrate genera and Carraro & Altermatt (2022) suggest a complex interplay between species richness and different environmental covariates. While several studies have highlighted the potential of eDITH, evidence validating the approach of Carraro (2021) and this study to determine taxon abundance from eDNA is still missing.

4.2 Assumptions and limitations in representing a real system

This study is based solely on virtually generated data and computer simulations to answer a research question about the real natural environment. Therefore, relies on significant assumptions that need to be discussed. Looking at eDITH specifically, the model approximates the advection and decay of eDNA through a mass balance resolved at the reach scale. As indicated by Carraro et al. (2018), the assumption of homogeneous hydrological conditions along individual reaches is especially wavering in headwaters, where the stream's morphology may stray far away from these assumptions. More importantly, however, the assumption of a spatially homogeneous decay rate overlooks differences that are found in real catchments (Barnes et al., 2014). While differences between taxa are accounted for by the model (being that any unknown value of tau may be estimated), spatial heterogeneities across different parts of the catchment are not accounted for. Analogously, eDITH assumes proportionality between species biomass and eDNA production, while this is known to differ between taxa (Palowski et al., 2020). These simplifications do not confound the inferences that this study makes based on synthetic realities. Nevertheless, it should be considered whether these results also apply to real catchments. Studies have shown that eDITH is capable of predicting the distribution of various taxa (e.g. Carraro et al., 2020a). However, they have not analysed the extent to which eDNA production or its fate in the environment may confound these results. Here, studies comparing the production rate determined with eDITH to actual abundance data may again prove valuable. They could both evaluate the quality of taxon abundance predictions from the modelled production rate, as well as quantify this mismatch and identify its causes. Additionally, further research could evaluate the sensitivity of the results to incorporating a spatially heterogeneous decay time within the framework of this modelling study.

It should also be considered whether the virtual catchments provide an adequate representation of reality. While the generic nature of OCNs is especially useful for large-scale simulation studies, their idealized morphology should be taken into account. First of all, this generalization overlooks inconsistencies that may be found in many real catchments. As long as these disturbances only occur locally, this should not significantly influence the approximate effect of sampling intensity at the catchment scale that was found in this study. Ponds and dams may locally slow down the flow and channels may speed up or redirect the downstream transport of eDNA to such an extent that the flow assumptions of OCNs do not hold. This could be the case in polders or catchments that are heavily

exploited for hydropower or industrial water use. Given that such catchments are likely to benefit from frequent eDNA monitoring to meet environmental objectives (Hering et al., 2010), considering these effects is especially important. To a certain extent, eDITH may be modified to account for such inconsistencies, e.g. by manually adjusting the mass balance through each reach. Additionally, when natural processes like sediment deposition or glaciation, rather than flow-driven erosion, are dominant in shaping the catchment, their network topology may not satisfy the dendritic structure of OCNs (Rodriguez-Iturbe & Rinaldo, 1997). More specifically, the findings of this study concern river networks with a constant branching ratio. Given the role of confluences in the mixing of eDNA signals, river networks that do not branch out like this may require a different sampling strategy. This may entail prioritizing the main stem, taking more samples to resolve eDNA dynamics locally or excluding certain parts of the catchment from the analysis. Delta systems form another important exception, as they mainly consist of stream bifurcations, rather than confluences in the downstream direction. In such systems, eDNA signals may originate from a single upstream stretch, simplifying source estimation but requiring tailored sampling strategies.

Another consideration is the assumption on the transport distance of eDNA at the catchment level. This study was performed with a decay time of 5h, which is on the high end of what is expected of riverine eDNA (Deiner et al., 2016). Combined with the hydrological conditions used in this study, which reflect relatively high flow velocities of a typical pre-alpine catchment, this leads to a long transport distance. However, in systems with slower flow velocities, such catchments with small elevational gradients and those influenced by the backwater effect of lakes or oceans, eDNA transport distances may be significantly lower. This makes each signal is observable over a relatively long distance downstream of its production. While this increases the probability of taxa being observed in a water sample, this comes at the expense of increased uncertainty about the exact origin of this eDNA. This may explain why the results suggest combining a relatively low sampling density and spatial resolution to predict taxon distributions. A sensitivity analysis may help to determine how the sampling guidelines from this study would translate to taxa or environments with faster-decaying eDNA. Likely, more samples would be necessary to capture each signal, while less mixing allows more fine patterns in taxon distribution to be distinguished at this sampling density. This implies that a combination of increased sampling and spatial resolution would become optimal. Additionally, a study setup similar to this one could make the optimal sampling strategy more inclusive across taxa by assigning a random decay for each simulation (Palowski et al., 2020).

The generation of virtual taxa to study to obtain insights across various types of taxa relies on the assumption that the spatial fields actually represent realistic taxon distributions. As explained in 2.7.7, the virtual taxa were spatially autocorrelated based on the Euclidean distance between nodes, in order to challenge the model to reproduce more complex and varying taxa, which may disperse over land or be influenced by terrestrial landscape attributes. The residual error and the observation that more detailed predictions may only marginally improve the performance suggest that this approach was successful in creating a challenge for eDITH. In fact, the inability of additional AEMs to accurately model more complex patterns suggest that the taxon distributions may have even been ‘too’ complicated for eDITH to reproduce. Moreover, this shifts the focus of this study away from riverine species whose distribution is governed by river network structure (Altermatt, 2013). Thus, this study may underestimate the positive effect of increasing model complexity, and thereby the effect of increasing the number of samples, on performance. On the other hand, this mismatch may be seen as a safe margin of error when making sure that the findings of this study apply to a wide variety of taxa. Alternatively, taxon distributions could have been generated according to along-stream autocorrelation, while using a large number of AEMs to predict them. The setup could still include randomness by limiting the number of AEMs used or adding noise to the data according to overland correlation. Within the current framework, one could include a limited number of Moran’s Eigenvector Maps in the predictions to better model overland autocorrelation. In this way, there would still be inherent randomness coming from the lognormal distribution, and unaccounted spatial autocorrelation coming from the limited number of predictor variables. Exploring these changes could further refine the virtual setup of this study and thereby improve the robustness of its results.

4.3 Towards better sampling guidelines

Besides considering the potential sources of the residual error and address uncertainties in representing a real system, moving towards better eDNA sampling guidelines also requires viewing the results in a practical context. This study used a relative comparison to determine the optimal sampling density. To enable a more absolute interpretation that values how good the predictions really are, it may be helpful to consider the spatial resolution of the predictions. The spatial resolution of the predictions is primarily limited by the discretization of the river network (being the 1 km reach-scale). However, the actual scale at which predictions may be made is determined by the optimal spatial resolution of the predicted taxon distribution. Increasing the desired level of detail requires eDITH

to fit more fine distribution patterns to the observed eDNA samples. When the taxon distribution predictions at a coarser scale on average attains the same performance as the more fine-scale predictions, it is more likely that the latter attain this error at a finer scale as well. In other words, the fine-scale predictions may succeed in predicting detailed patterns in some portion of the network, but fail to do so in other locations. Conversely, the error from the more coarse predictions more likely stems from a mismatch in resolution of the predictions, where the average abundance over a larger area is correct but fine-scale patterns are overlooked. In this case, the conservative choice would be to accept the coarser predictions, and apply a more thorough investigation in specific subsections of the catchment. This resolution roughly corresponds to making predictions in second-order streams, but excluding differences within these streams and first-order tributaries. Using a similar setup to predict taxon distributions, eDITH users should use a similar resolution in when processing eDNA data. Moreover, the strong effect of the specified spatial resolution and the uncertainty associated with it should be taken into account.

Such a broad-scale assessment is in line with current ideas of eDNA use in river networks (Deiner et al., 2016). In water quality monitoring and biodiversity conservation, where fish and macroinvertebrates often serve as key bioindicators, species distribution patterns may vary greatly in spatial scale. While fish are mobile and may occupy large habitats, their presence in specific spawning sites relies on microhabitats, creating localized distribution patterns (Bretzel et al., 2023). Macroinvertebrates also show high habitat specificity and depend on suitable substrates, potentially leading to more fragmented distributions (Verdonschot et al., 2016). However, the downstream transport of eDNA enables the detection of species presence over longer upstream stretches, reducing the likelihood of overlooking taxa due to fragmented distributions (Deiner et al., 2016). Moreover, knowledge on large-scale ecological dynamics is necessary to understand the effect of emerging environmental stressors and to apply appropriate management interventions (Fausch et al., 2002; Vörösmarty et al., 2010). Because of the simplicity of the sampling procedure, frequent measurements allow the monitoring of taxa with a quick turnover, such that changes in environmental conditions are understood better (Altermatt et al., 2020). In this context, by favoring accuracy over precision, eDITH can offer robust coarse distribution information that may be complemented by local, targeted monitoring. In data-scarce regions, eDNA may serve as a good starting point. Moreover, setting up citizen science campaigns to sample eDNA are a promising way of increasing data collection and raising public awareness for the biodiversity crisis (Deiner et al, 2017).

The taxon distribution that is obtained with eDITH may be used in various applications. Findings on the general taxon suggested an optimal sampling density that maximizes the accuracy across the catchment. The abundance of common taxa across the catchment can directly inform biodiversity conservation, but may also be useful to compare the ecological status of different areas and monitor subtle changes over time. For example, the abundance of taxa that are tolerant to organic pollution can provide an assessment of the level of eutrophication across a catchment (Johnson et al., 1993). However, in practice, eDNA may be more promising for broad presence/absence assessments. Despite the added bias of quantifying abundance from read count data (Blackman et al., 2024ba), presence/absence data of multiple species may be combined into a quantitative richness estimate, which can in turn be used to quantify the ecological status across the catchment (e.g. Blackman et al., 2024a). Moreover, metabarcoding can be used to locate rare or elusive taxa that were not initially considered as the target, when measured eDNA matches sequences from a reference library (i.e. passive surveillance, Deiner et al., 2017). Such an approach has been shown to facilitate the detection of invasive species in river networks (Blackman et al., 2022). While this study focused on quantitative eDNA measurements (qPCR-based), findings on the scattered and rare taxon may also reflect data that could be obtained using metabarcoding, where a rough estimate of the presence of eDNA may be used to determine the presence of taxa. While further studies could explore the optimal sampling density with datasets that reflect read-count data, frequent mismatches between the richness in the environment and what is observed in a water sample remain an issue that requires further development (Blackman et al., 2024a). Regarding the overall difficulties of translating eDNA production into abundance, an interesting workaround would be to use the production rate of eDNA itself as an indicator of the state of the environment or the taxa residing therein (Blackman, 2024b). Harrison et al. (2019) provide the example where a change in eDNA concentration over time could rather describe the change in conditions affecting the production of genetic material (such as stress or temperature) than the actual abundance.

Besides the density of samples, there are other considerations that are important when setting up a sampling campaign. While samples were equally divided over the upstream and downstream half of the catchment in this study, real sampling campaigns should maximize their efficiency by sampling around confluences and spacing them equally apart (Altermatt et al., 2020). Such an approach may allow predictions to be more precise, although this effect should not exceed what the results of this study show at high sampling intensities. Thus, optimal positioning may rather be seen as increasing the robustness of the same predictions that are expected with random

positioning. At the same time, however, it should be taken into account that some upstream sections of the catchment may not be easily accessible, making it hard to sample equally in the upstream and downstream half of the catchment. Sampling more in the downstream portion of the catchment may negatively affect the overall accuracy of the predictions (Carraro et al., 2021). When it only concerns a limited number of tributaries, the effect of environmental covariates on taxon distribution may be fitted to other sections of the catchment in order to predict the taxon distribution in the unsampled regions (e.g. Carraro & Altermatt, 2022, Blackman et al., 2024a). However, to predict the taxon abundance in any reach with more certainty, it is necessary to sample at such a distance that the eDNA that is produced in this reach is recovered (Deiner et al., 2016). Lastly, more so than traditional sampling methods, due to limited (lateral) hydrodynamic dispersal in large river reaches, eDNA signals may be missed in a single water sample (Altermatt et al., 2023). It is thus advised to sample in multiple locations along a transect crossing larger river reaches and extract an appropriate volume of water. This ensures that each eDNA sample properly represents its environment, integrating all potential upstream sources.

4.4 Conclusions

This study has successfully simulated a large number of sampling campaigns on virtual catchments and taxa to further refine existing guidelines on the sampling density of eDNA in catchments. Doing so, it was found that taking at least one sample per 10 km^2 of catchment area gives a pragmatic optimum for species abundance estimations across the catchment. Increasing the sampling density may yield some benefits, especially when the combined presence or absence of multiple taxa is combined into a biodiversity estimate. When it concerns finding rare taxa, taking two samples per 10 km^2 is necessary to ensure that the species is located. These recommendations complement existing guidelines that strive to optimize the interpretation of eDNA data. The findings of this study suggest that eDITH itself provides an appropriate simplification of catchment-level eDNA dynamics. However, further refinement of the virtual experiment to represent real taxon distributions could increase the robustness of the results. A sensitivity analysis into the effect of changing transport distance would be particularly insightful, as this is expected to strongly influence the optimal sampling density. Lastly, further empirical validation of eDITH with actual taxon distribution data, particularly their abundance, is necessary enable a more widescale adoption of eDNA in various biomonitoring applications. The results provide an adaptable framework for setting up biomonitoring campaigns at the catchment scale. Here, eDNA serves as a valuable complementary method to traditional biomonitoring practices by providing broad-scale taxon distribution estimates. These may be complemented by more localized surveying, using traditional sampling methods. By further exploring its applicability across disciplines, eDNA may be able to fulfil its potential for environmental assessment and biodiversity conservation.

References

Abbott, B. W., Baranov, V., Mendoza-Lera, C., Nikolakopoulou, M., Harjung, A., Kolbe, T., ... & Pinay, G. (2016). Using multi-tracer inference to move beyond single-catchment ecohydrology. *Earth-Science Reviews*, *160*, 19-42.

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology, 43(6), 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Altermatt, F. (2013). Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology*, *47*, 365-377.

Altermatt, F., Carraro, L., Antonetti, M., Albouy, C., Zhang, Y., Lyet, A., ... & Pellissier, L. (2023). Quantifying biodiversity using eDNA from water bodies: General principles and recommendations for sampling designs. *Environmental DNA*, *5*(4), 671-682.

Blackman, R. C., Brantschen, J., Walser, J. C., Wüthrich, R., & Altermatt, F. (2022). Monitoring invasive alien macroinvertebrate species with environmental DNA. *River Research and Applications*, *38*(8), 1400-1412.

Blackman, R. C., Carraro, L., Keck, F., & Altermatt, F. (2024a). Measuring the state of aquatic environments using eDNA - Upscaling spatial resolution of biotic indices. Philosophical Transactions of the Royal Society B: Biological Sciences, 379(1904). https://doi.org/10.1098/rstb.2023.0121

Blackman, R., Couton, M., Keck, F., Kirschner, D., Carraro, L., Cereghetti, E., Perrelet, K., Bossart, R., Brantschen, J., Zhang, Y., & Altermatt, F. (2024b). Environmental DNA: The next chapter. Molecular Ecology. https://doi.org/10.1111/mec.17355

Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Modelling directional spatial processes in ecological data. Ecological Modelling, 215(4), 325–336. https://doi.org/10.1016/j.ecolmodel.2008.04.001

Bretzel, J. B., Pulg, U., & Geist, J. (2024). Juvenile salmonid abundance in a diamictic semi-fluvial stream in Norway—does stream bed shelter beat large woody debris?. *River Research and Applications*, *40*(5), 780-790.

Bruce, K., Blackman, R. C., Bourlat, S. J., Hellström, M., Bakker, J., Bista, I., Bohmann, K., Bouchez, A., Brys, R., Clark, K., Elbrecht, V., Fazi, S., Fonseca, V. G., Hänfling, B., Leese, F., Mächler, E., Mahon, A. R., Meissner, K., Panksep, K., … Deiner, K. (2021). A practical guide to DNA-based methods for biodiversity assessment.

Bylemans, J., Gleeson, D. M., Lintermans, M., Hardy, C. M., Beitzel, M., Gilligan, D. M., & Furlan, E. M. (2018). Monitoring riverine fish communities through eDNA metabarcoding: Determining optimal sampling strategies along an altitudinal and biodiversity gradient. Metabarcoding and Metagenomics, 2, 1–12. https://doi.org/10.3897/mbmg.2.30457

Carraro, L., & Altermatt, F. (2022). Optimal Channel Networks accurately model ecologically-relevant geomorphological features of branching river networks. Communications Earth and Environment, 3(1). https://doi.org/10.1038/s43247-022-00454-1

Carraro, L., & Altermatt, F. (2024). eDITH: an R-package to spatially project eDNA-based biodiversity across river networks with minimal prior 2 information 3. https://doi.org/10.1101/2024.01.16.575835

Carraro, L., Bertuzzo, E., Fronhofer, E. A., Furrer, R., Gounand, I., Rinaldo, A., & Altermatt, F. (2020b). Generation and application of river network analogues for use in ecology and evolution. Ecology and Evolution, 10(14), 7537–7550. https://doi.org/10.1002/ece3.6479

Carraro, L., Blackman, R. C., & Altermatt, F. (2023). Modelling environmental DNA transport in rivers reveals highly resolved spatio-temporal biodiversity patterns. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-35614-6

Carraro, L., Hartikainen, H., Jokela, J., Bertuzzo, E., & Rinaldo, A. (2018). Estimating species distribution and abundance in river networks using environmental DNA. Proceedings of the National Academy of Sciences of the United States of America, 115(46), 11724–11729. https://doi.org/10.1073/pnas.1813843115

Carraro, L., Mächler, E., Wüthrich, R., & Altermatt, F. (2020a). Environmental DNA allows upscaling spatial patterns of biodiversity in freshwater ecosystems. Nature Communications, 11(1). https://doi.org/10.1038/s41467-020-17337-8

Carraro, L., Stauffer, J. B., & Altermatt, F. (2021). How to design optimal eDNA sampling strategies for biomonitoring in river networks. Environmental DNA, 3(1), 157–172. https://doi.org/10.1002/edn3.137

Cristescu, M. E., & Hebert, P. D. N. (2018). Uses and Misuses of Environmental DNA in Biodiversity Science and Conservation. Annu. Rev. Ecol. Evol. Syst, 49, 209–230. https://doi.org/10.1146/annurev-ecolsys-110617

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. PLoS ONE, 9(2). https://doi.org/10.1371/journal.pone.0088786

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., de Vere, N., Pfrender, M. E., & Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. In Molecular Ecology (Vol. 26, Issue 21, pp. 5872–5895). Blackwell Publishing Ltd. https://doi.org/10.1111/mec.14350

Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016). Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nature Communications, 7. https://doi.org/10.1038/ncomms12544

Fausch, K. D., Torgersen, C. E., Baxter, C. V., & Li, H. W. (2002). Landscapes to riverscapes: Bridging the gap between research and conservation of stream fishes. BioScience, 52(6), 483-498.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., Cheng, L., Crochetiere, H., Ehalt Macedo, H., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B., McClain, M. E., Meng, J., Mulligan, M., … Zarfl, C. (2019). Mapping the world's free-flowing rivers. Nature, 569(7755), 215–221. https://doi.org/10.1038/s41586-019-1111-9

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C. K., ... & van de Bund, W. (2010). The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. *Science of the total Environment*, *408*(19), 4007-4019.

Peterson, E. E., ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. Ecology Letters, 16(5), 707–719. https://doi.org/10.1111/ele.12084

Rodriguez-Iturbe, I., & Rinaldo, A. (1997). *Fractal river basins: chance and self-organization*. Cambridge University Press.

Stewart, K. A. (2019). Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. In Biodiversity and Conservation (Vol. 28, Issue 5, pp. 983–1001). Springer Netherlands. https://doi.org/10.1007/s10531-019-01709-8

Verdonschot, R. C. M., Kail, J., McKie, B. G., & Verdonschot, P. F. M. (2016). The role of benthic microhabitats in determining the effects of hydromorphological river restoration on macroinvertebrates. Hydrobiologia, 769(1), 55–66. https://doi.org/10.1007/s10750-015-2575-8

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Liermann, C. R., & Davies, P. M. (2010). Global threats to human water security and river biodiversity. Nature, 467(7315), 555–561. https://doi.org/10.1038/nature09440

Voulvoulis, N., Arpon, K. D., & Giakoumis, T. (2017). The EU Water Framework Directive: From great expectations to problems with implementation. Science of the Total Environment, 575, 358–366. https://doi.org/10.1016/j.scitotenv.2016.09.228

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., & Higdon, D. (2009). Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. International Journal of Nonlinear Sciences and Numerical Simulation, 10(3), 273–290. https://doi.org/10.1515/IJNSNS.2009.10.3.273

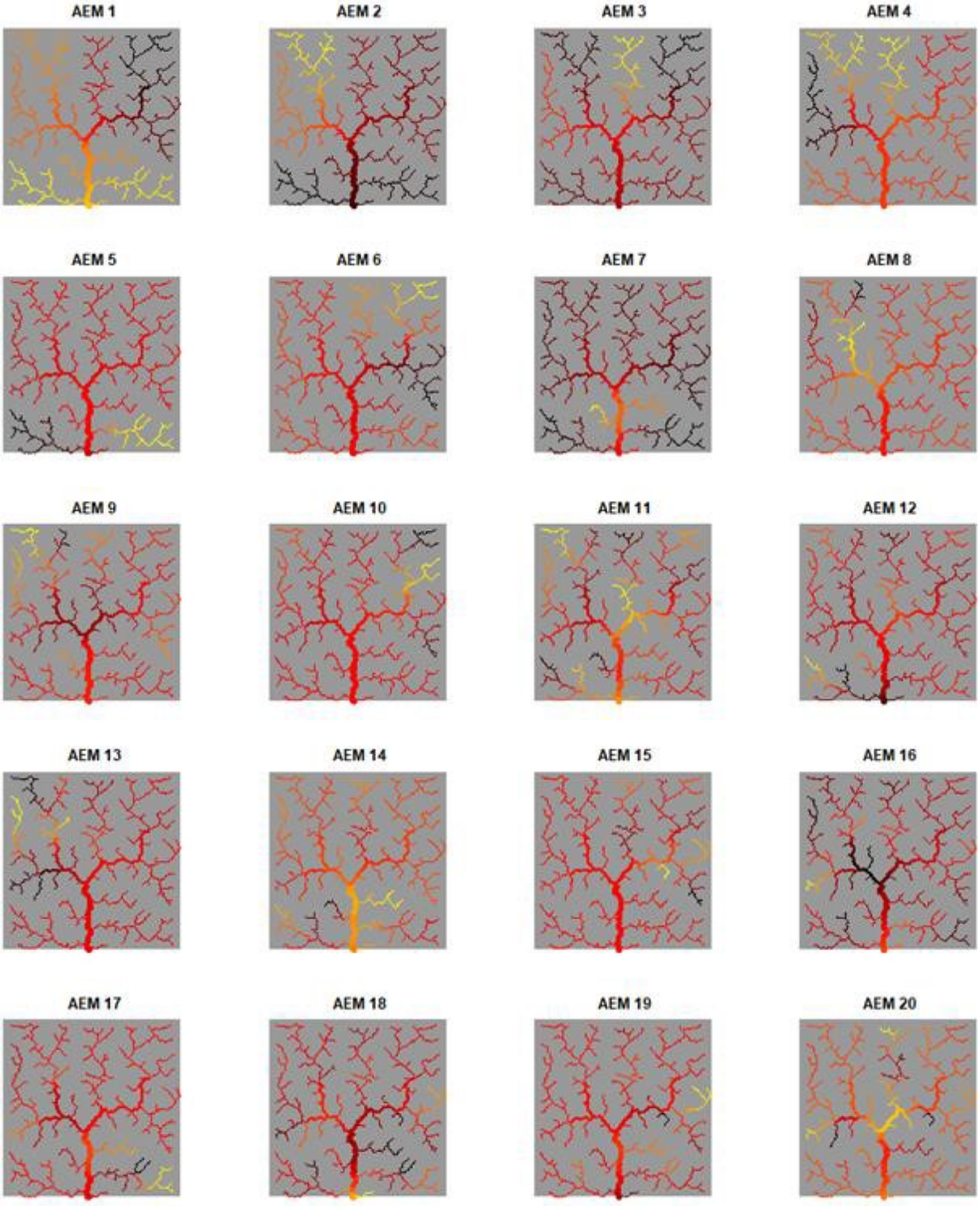
White, J. W., Rassweiler, A., Samhouri, J. F., Stier, A. C., and White, C. (2014). Ecologists should not use statistical significance tests to interpret simulation model results. Oikos 123, 385–388. doi: 10.1111/j.1600-0706.2013.01073.x

Zurell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B., & Grimm, V. (2010). The virtual ecologist approach: Simulating data and observers. *Oikos*, *119*(4), 622–635. https://doi.org/10.1111/j.1600-0706.2009.18284.x

Appendix

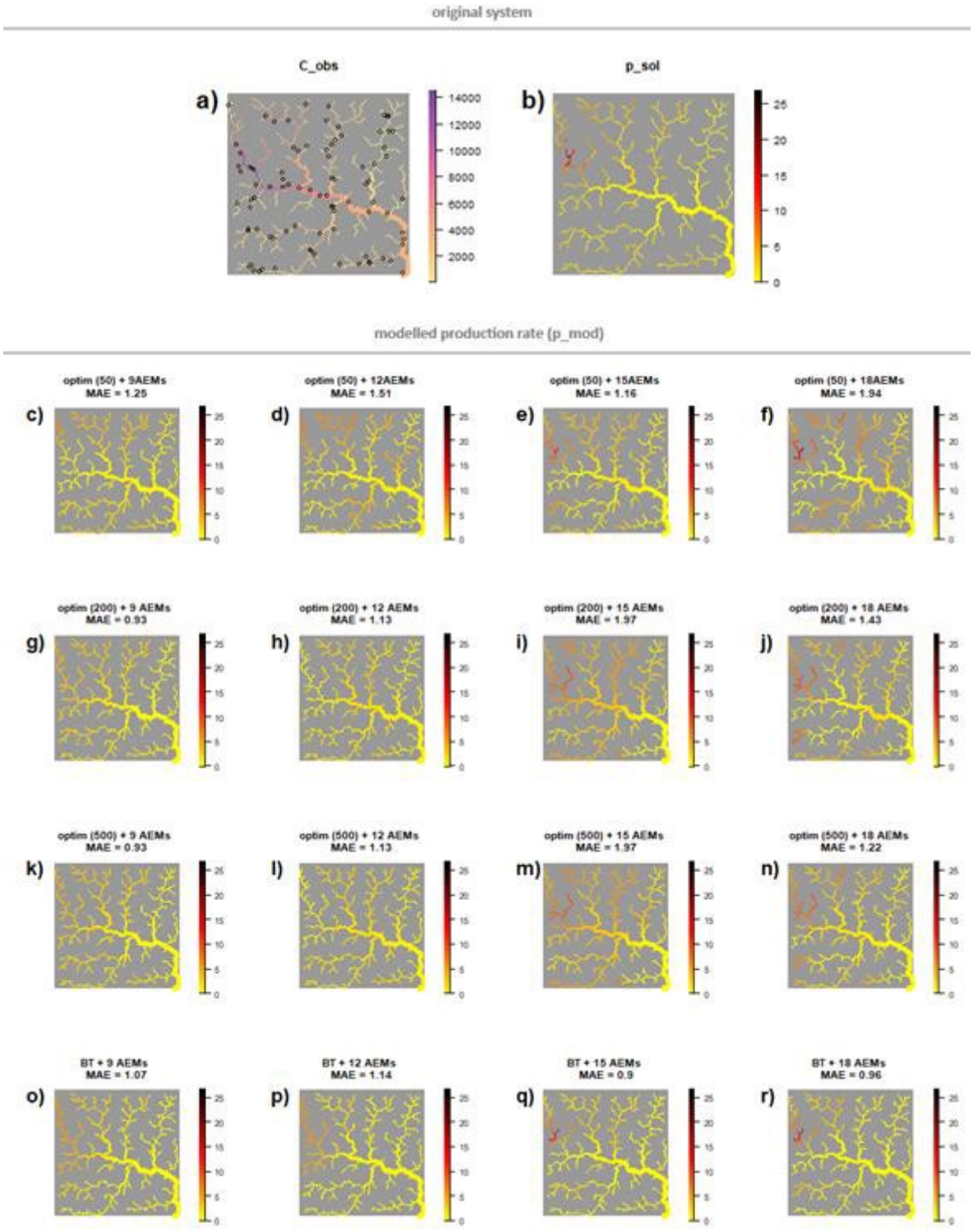
A1 AEMs

This figure shows the first 20 Asymmetric Eigenvector Maps (AEMs) for a generic river network (Optimal Channel Network). These figures show how each consecutive AEM describes a distribution pattern that is different from (and, in fact, orthogonal to) the previous one. A combination of AEMs is used to fit a number of eDNA samples. In doing so, smooth patterns of species distribution are generated.



A2 Initial simulations (pilot)

The setup of this study (as given in section 2.2 and 2.3) was determined based on initial simulation runs. First, this 'pilot' was used to determine the parameters used to generate synthetic taxa. Second, it pointed out that the BayesianTools optimization was preferable to a less computationally expensive simplex optimization. Third, it suggested a strong interaction between the density of samples and the spatial resolution of the predictions (given by the number of AEMs), which led to its further investigation in this study. The figure below shows a visual comparison between the 'true' eDNA production rate (b) and corresponding eDNA concentration that is measured with samples (a, black circles represent sampled nodes) and the estimated production rate from model runs with varying calibration methods and spatial complexity (c-r). Each plot represents a single simulation using a sampling density of 20 percent.



A3 Supplementary results

Rank-biserial correlation

The rank-biserial correlation coefficient between each degree of sampling density and spatial complexity was computed for the general taxon. Most interestingly, this revealed that, at a low sampling density, while the boxplots show the largest differences in performance. Moreover, the effect sizes between the number of AEMs are generally weaker than between each degree of sampling density. Logically, adding more samples will lead to slightly better performance in most cases. This supports the claim that each increase in sampling intensity leads to an increase in performance, but that this increase is marginal. However, changing the degree of spatial resolution, be it upward or downward, carries more uncertainty regarding the change in performance.

Table A4.1 These values represent paired the effect sizes between each degree of sampling density for a given degree of spatial complexity (i.e. boxplots of the same shade of purple in figure 8)

	5-10 %	**10-15 %**	**15-20 %**	**20-25 %**	**25-30 %**
9 AEMs	0.873	0.671	0.665	0.584	0.564
12 AEMs	0.891	0.736	0.68	0.631	0.593
15 AEMs	0.89	0.831	0.771	0.684	0.623
18 AEMs	0.852	0.882	0.81	0.755	0.68
30 AEMs	0.739	0.825	0.866	0.836	0.776

Table A4.2 These values represent paired the effect sizes between each degree of spatial complexity for a given degree of spatial complexity

	9 – 12 AEMs	**12 – 15 AEMs**	**15 – 18 AEMs**	**18 vs 30 AEMs**
5%	0.341	0.274	0.307	0.153
10%	0.534	0.324	0.267	0.108
15%	0.692	0.538	0.4	0.114
20%	0.812	0.655	0.536	0.182
25%	0.869	0.75	0.638	0.322
30%	0.905	0.794	0.712	0.489

A4 Computation

The TU Delft offers the use of a high performance computer for programming tasks (Delft High Performance Computing Centre (DHPC), 2022). Using a 'student share', it is possible to perform up to 8 tasks simultaneously, each using up to 48 cores and 192 GB of memory. Output data is stored on a temporary drive, such that it can be accessed for post processing on the cluster itself. While the cluster usage drastically increases computational power, computational demand is minimized by optimizing the code and determining a priori how specific inferences can be made most efficiently. Using the full capacity of a DHPC student share, the computer cluster allowed computations to run at roughly 100 times the speed of a normal PC. However, each job may only take up to 24 hours and enters a queue after being submitted. By dividing simulations into chunks of 16 hours, it was possible to ensure a near-daily simulation round between starting a task and submitting the next one. Because of fluctuations in demand for cluster usage, the queue time for large tasks varied from hours to days.