PhD subject Modelling environmental DNA to derive abundance indices

Supervisors: Verena Trenkel (verena.trenkel@ifremer.fr, + 33 240374157) and Robin Faillettaz (robin.faillettaz@ifremer.fr)

Location: UMR DECOD (https://www.umr-decod.fr/en), Ecology and Models for Fisheries Science laboratory, Ifremer Nantes, France

Summary

Methods for assessing fish stocks require significant resources to provide reliable abundance estimates, but do not allow the assessment of rarer species; they are also destructive. To fill these gaps, a new methodology seems promising, the analysis of abundance by environmental DNA (eDNA). However, statistical models for using eDNA observations for estimating abundance indices accounting for the drift and degradation of eDNA et making use of other available information (habitats, catches) still need to be developed and tested for the marine environment. The objectives of the thesis are to i) develop a diffusion-advection model of DNA molecules accounting for the degradation of DNA as a function of temperature and the hydrology of the study area using an individual-based model coupled with a hydrodynamic model (e.g. CROCO); ii) develop a spatial state-space model with a latent layer for abundance and several observation models which includes the results from the diffusion-advection model and the molecular analysis processes for eDNA data; iii) apply the models to available data from the Bays of Biscay.

Key-words: eDNA, Lagrangian transport, exploited species, state-space model, fish.

Background and scientific context

Human pressures on ecosystems and climate change will alter marine life and ecosystem services (Diaz et al., 2019). Monitoring the impacts of these increasingly rapid changes on exploited fish populations is challenging given the extent of the areas to monitor. For several decades, IFREMER has been carrying out bottom trawl surveys to obtain abundance time series and other biological information (Rochet et al., 2005). Indeed, abundance time series underpin the assessment and management of exploited marine fish stocks worldwide, derived from a range of observation methods, including bottom trawl surveys (Trenkel et al 2019). Despite the great efforts and costs deployed, current observation methods do not provide reliable abundance time series for several commercially exploited species, mainly due to their low abundance, lack of detectability by the survey method or lack of spatial coverage by routine surveys. These surveys also increasingly raise ethical questions due to their invasive nature (Trenkel et al. 2019).

Environmental DNA (eDNA) has attracted attention in recent years as a novel observation method for fisheries management (e.g., Lacoursiere-Roussel et al., 2016, Hansen et al., 2018) and wider biodiversity monitoring (e.g. Dejean et al., 2012; Thomsen and Willerslev, 2015; Deiner et al., 2017; Makiola et al., 2020). This emerging method called "barcoding/metabarcoding" is the result of the coupling of high-throughput sequencing methods to single/multiple taxonomic assignments of DNA sequences from eDNA samples. Environmental DNA has allowed the scientific community to rapidly generate inventories of species communities at a given time and place (e.g., Polanco et al., 2020) and could be used for fisheries management (Hansen et al., 2018). While positive correlations between eDNA concentrations and species abundance have been demonstrated in vitro and for freshwater environments (Yates et al., 2019), certain studies have demonstrated such a positive correlation in the marine environment (e.g. Thomsen et al., 2016, Salter et al. 2019, Shelton et al., 2022), while others have found none (Knudsen et al., 2019). Confident use of eDNA for deriving abundance proxies for marine species is currently hindered by prevailing lack of knowledge regarding key biotic and abiotic factors impacting DNA production and shedding, degradation and transport of DNA in the ocean (Hansen et al. 2018, Harrison et al. 2019, Stewart 2019). These factors act together to determine the spatio-temporal footprint of marine eDNA samples.

Transport dynamics of eDNA have been primarily studied in streams and seem to resemble that of fine particulate matter (Stewart, 2019). Given the size of most eDNA particles (> 1 µm), they are expected to sink in the water column (Hansen et al., 2018). Marine systems, on the contrary, are characterised by complex multidirectional flow conditions comprising directional regional currents, bidirectional tidal currents and local eddies. One of the rare marine modelling studies using a Lagrangian particle tracking model including settling and decay for Monterey Bay concluded that if eDNA was detectable for 4 days, sampled eDNA would originate from within 40 km (Andruszkiewicz et al., 2019). Few marine studies have integrated transport and degradation processes into comprehensive spatial abundance estimation models (but see Fukaya et al., 2021). Recently, factors relating to the eDNA analysis process have also been accounted for when modelling eDNA concentrations (Shelton et al., 2022).

Scientific objectives

The three objectives of the thesis are to i) develop a diffusion-advection model of DNA molecules accounting for the degradation of DNA as a function of temperature and the hydrology of the study area using a hydrodynamic model coupled with an individual-based model; ii) develop a spatial state-space model with a latent layer for abundance and several observation models which includes the results from the diffusion-advection model and the molecular analysis processes for eDNA data for eDNA data; iii) apply the models to available data from the Bays of Biscay.

Methodology

The first step of the PhD will be to conduct a Systematic Literature Review (Siddaway et al., 2019) of diffusion-advection models (e.g. Lam et al., 2020) and their application to DNA (e.g. Andruszkiewicz et al., 2019) as well as of adult and larval dispersion models (e.g. Lett et al. 2008). The review will include a meta-analysis of model parameters.

Based on the review, the PhD student will propose a modelling solution for the movement of DNA molecules in three dimensions taking account of currents (diffusion-advection model). The modelling platform CROCO (www.croco-ocean.org) dedicated to coastal oceanographic modelling and integrating tidal currents will be used for environmental forcing. It includes a Lagrangian transport module (ICHTHYOP; www.ichthyop.org) which could be used to study diffusion and advection phenomena or alternatively the Connectivity Modeling System (University of Miami; Paris, 2013). The forced diffusion/advection model will allow the student to address DNA dispersion issues at different spatial scales and for contrasted hydrological conditions. The model will be applied to the areas and times of available eDNA data to estimate their spatial footprint (table 1). To account for fish movements, fish dispersion patterns, obtained from fish tracking studies (published, carried out by Ifremer), will be coupled by the student to the DNA advection-diffusion model through a probabilistic approach. The effects of fish movements and DNA drift on eDNA footprints will be compared to appraise the relative importance of each phenomenon for different species. To parameterise the behaviour of DNA molecules in the marine environment in the model (degradation rates, etc.) the PhD student will review the literature (e.g. Andruszkiewicz et al., 2021).

A spatial state-space abundance index model will be developed using a Gaussian random field for the underlying latent abundance distributions and spatial autocorrelations (e.g. Cressie, 1993). The use of habitat covariates (bathymetry, SST, etc.) for describing abundance distributions will be tested. Several nested observation models will be formulated to account for the eDNA sampling process, including DNA shedding, diffusion and fish dispersal as well as molecular analysis processes (Shelton et al., 2022). For parameterising the observation model for the Bay of Biscay eDNA data (table 1), the results from the diffusion-advection and dispersal models will be used. The form of the relationship between abundance (biomass) and DNA concentrations will be based on a literature review of relevant experimental studies. For comparison and improving model fit, bottom trawl survey data will be included as a second data source using a distinct observation model. Model fitting will be by Maximum Likelihood using the TMB software (Template Model Builder – Kristensen et al. (2016)); spatial random effects will be estimated through the SPDE approach (Lindgren et al., 2011). A similar approach was recently used by Alglave et al. (2022) for fitting spatial models to survey data and commercial landings. Integrating abundance estimates over the modelled area will provide abundance indices. Model evaluation will carried out by simulation before fitting the models to species with sufficient data, pooling years to ensure sufficient data. Results for different eDNA approaches will be compared: number of reads for metabarcoding, number of copies for metabarcording qPCR and number of droplets from ddPCR.

Expected results

The PhD work will lead to at least three publications and will be presented in various national and international scientific and general public fora.

Paper 1: Systematic review on diffusion-advection, individual-based models and their application in ecology for the vertical diffusion of DNA as well as adult and larval dispersion and experimental procedures for estimating model rates.

Paper 2: Modelling the movements of DNA molecules in 3D, integrating hydrodynamics and active dispersion potential (i.e. movement) of fish at different life stages. Footprint results for the Bay of Biscay.

Paper 3: Modelling eDNA data using spatial state-space models for estimating abundance indices for selected species from the Bay of Biscay.

Provisional schedule

	Year 1				Year 2				Year 3			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
Part I: Review of diffusion-advection models												
I-1 Bibliography and model parameter selection	█	█										
I-2 MS 1 writing			█	█								
Part II : Oceanographic modelling												
II-1 Data acquisition (current data)			█	█								
II-2 Getting started with CROCO and Ichtyops/CMS models					█							
II-3 Application to Bay of Biscay						█	█					
II-4 MS 2 writing							█	█				
Part III: Abundance modelling												
III-1 Bibliography	█	█					█	█				
III-2 Model development and simulation testing							█	█	█			
III-3 Application to Bay of Biscay									█	█		
III-4 MS 3 writing										█	█	
PhD writing											█	█

International collaborations

Claire B. Paris (University of Miami) probabilistic IBM for dispersal processes; Camille Albouy (ETH Zürich) describing eDNA sampling and analyses observation processes.

References

Alglave et al. (2022) Combining scientific survey and commercial catch data to map fish distribution. ICES Journal of Marine Science, 79: 1133–1149.

Andruszkiewicz et al. (2019) Modeling Environmental DNA Transport in the Coastal Ocean Using Lagrangian Particle Tracking. Frontiers in Marine Science 6.

Andruszkiewicz et al. (2021) Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. Environmental DNA, 3: 492-514.

Cressie (1993) Statistics for spatial data. John Willy and Sons. Inc., New York.

Díaz et al. (2019). Pervasive human-driven decline of life on Earth points to the need for transformative change. Science, 366(6471).

Deiner et al. (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology 26:5872–5895.

Dejean et al. (2012) Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. Journal of Applied Ecology 49:953–959.

Hansen et al. (2018) The sceptical optimist: challenges and perspectives for the application of environmental DNA in marine fisheries. Fish and Fisheries 19:751-768.

Harrison et al. (2019) Predicting the fate of eDNA in the environment and implications for studying biodiversity. Proceedings of the Royal Society B-Biological Sciences 286.

Fukaya et al. (2021) Estimating fish population abundance by integrating quantitative data on environmental DNA and hydrodynamic modelling. Molecular Ecology 30:3057–3067.

Kristensen et al. 2016. TMB: Automatic differentiation and Laplace approximation. Journal of Statistical Software, 70: 1–21.

Lacoursiere-Roussel et al. (2016) Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. Journal of Applied Ecology 53:1148-1157.

Lam et al. (2020) Selected topics on reaction-diffusion-advection models from spatial ecology. Math. Appl. Sci. Eng. 1 No. 2, pp. 150-180.

Lett C, Verley P, Mullon C, Parada C, Brochier T, Penven P, Blanke B (2008) A Lagrangian tool for modelling ichthyoplankton dynamics. Environmental Modelling & Software 23:1210-1214.

Lindgren et al. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73: 423–498.

Makiola et al. (2020) Key questions for next-generation biomonitoring. Front. Environ. Sci. 7.

Paris et al. (2013) Connectivity Modeling System: A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean, Environmental Modelling & Software, 42: 47-54.

Polanco et al. (2020) Comparing environmental DNA metabarcoding and underwater visual census to monitor tropical reef fishes. Environmental DNA 3: 142.

Rochet et al. (2005) Combining indicator trends to assess ongoing changes in exploited fish communities: Diagnostic of communities off the coasts of France. ICES Journal of Marine Science, 62, 1647–1664.

Salter et al. (2019) Environmental DNA concentrations are correlated with regional biomass of Atlantic cod in oceanic waters. Communications Biology 2.

Shelton et al. 2022. Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. Proceedings of the Royal Society B-Biological Sciences, 289.

Siddaway et al. (2019) How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. Annual Review of Psychology 70: 747–770.

Stewart (2019) Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. Biodiversity and Conservation 28:983-1001.

Thomsen et al. (2016) Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater fishes. PLoS ONE 11.

Thomsen et al. (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183: 4–18.

Trenkel et al. (2019) We can reduce the impact of monitoring on marine living resources. Marine Ecology Progress Series 609:277–282.

Yates et al. (2019) Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. Environmental DNA 1:5–13.