# UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Ingegneria Civile Edile e Ambientale

Master Degree in Environmental Engineering

Final Dissertation

# Genetic Material Dispersion: Integrating CFD Modeling and eDNA Techniques in the Case Study of Aquarium Diergaarde Blijdorp, Rotterdam.

Thesis supervisor:
**Professor Paolo Peruzzo**
**Università degli Studi di Padova**

Interniship supervisor:
**Professor Laura Stancanelli**
**Technische Universiteit Delft**

Candidate:
**Greta Morganti**

Academic Year 2023/2024

*A mia cugina Federica*

# Contents

# List of Figures

# Acknowledgements

# Introduction

Global biodiversity has recently undergone a concerning decline, impacting both terrestrial and aquatic ecosystems, with freshwater biodiversity facing particularly severe threats [1]. This trend poses significant challenges to ecosystem resilience and human well-being, including the provision of essential ecosystem services such as clean water, food security, and climate regulation. Addressing this crisis requires accurate biomonitoring strategies capable of capturing the intricate dynamics of biodiversity across diverse habitats and landscapes.

Traditional biomonitoring approaches, relying on localized point estimates, offer limited spatial and temporal coverage, hindering our ability to detect and effectively address biodiversity loss.

According to the Intergovernmental Science-Policy Platform a Biodiversity and Ecosystem Services (IPBES), unless action is taken promptly, about one million animal and plant species could be at risk of extinction in the next decades. Habitat destruction and degradation, primarily driven by human activities such as deforestation, urbanization, and conversion of natural habitats for agriculture and infrastructure development, are among the leading causes of biodiversity loss. Additionally, pollution, including water pollution from agricultural runoff, industrial discharge, and plastic waste, has significantly impacted aquatic ecosystems, threatening the survival of numerous species.

Overexploitation of natural resources, such as overfishing and illegal wildlife trade, further exacerbates biodiversity decline, contributing to the depletion of populations and alteration of ecological equilibrium. Climate change, characterized by rising temperatures, changing precipitation patterns, and more frequent extreme weather events, poses additional challenges to biodiversity conservation, altering habitat suitability and disrupting species distributions.

Finally, the introduction of invasive species, often facilitated by human activities such as international trade and travel, can exert profound impacts on native biodiversity. This includes the displacement of allogenous species through competition and predation on local fauna, with complex repercussions on the fragile ecosystem dynamics.

In response to these challenges, there is a pressing need to adopt innovative approaches that can enhance our capacity to monitor biodiversity accurately and efficiently. In this frame environmental DNA (eDNA) technology has emerged as a promising tool, offering a non-invasive and high-performing method for detecting

and quantifying biodiversity from various environmental samples. By analyzing traces of genetic material shed by organisms into their surroundings, eDNA enables researchers to obtain comprehensive insights into species' presence and abundance across different ecosystems.

The use of eDNA technology in biomonitoring practices holds immense potential to revolutionize our understanding of biodiversity dynamics. By providing high spatiotemporal resolution data, eDNA allows for a more accurate assessment of biodiversity trends and ecosystem health. Moreover, its non-invasive nature minimizes disturbance to natural habitats, making it particularly suitable for monitoring sensitive or protected areas.

This study focuses on the biodiversity dynamics within a controlled tank placed in the Aquarium Diergaarde Blijdorp in Rotterdam. The unique setting of the tank provides the opportunity to investigate the complex interactions between fish spatial distribution and environmental DNA concentration in water. The overarching objective is to elucidate correlations between these factors and understand how fluid dynamics influences the transport of genetic material. The integration of computational fluid dynamics (CFD) modeling and eDNA techniques offers a comprehensive approach, which provides a predictive guide for designing, sampling protocol, and monitoring activities to trace biodiversity in aquatic environments. The results obtained from CFD simulations can then be compared with experimental data collected during direct observations and in situ measurements. This comparison allows us to assess the consistency between the predictions of the numerical model and the observed data, providing valuable insights into water flow dynamics and spatial distributions of genetic material in the aquatic environment.

The conducted experimental work involves several phases, including the setup of the experimental apparatus at the Aquarium Diergaarde Blijdorp in Rotterdam, the acquisition of fish tracking data, and the collection of water samples for eDNA concentration analysis.

The thesis is structured as follows: Chapter 1 provides a theoretical and scientific context regarding aquatic biodiversity and monitoring methodologies. Chapter 2 describes in detail the materials and methods employed in the study, including experimental details, data collection, and analysis techniques used. Chapter 3 presents the results of the analyses conducted, including data collected from the tracking software, analysis of eDNA concentrations, and CFD simulations as well as the implications of the findings and their potential applications in the context of aquatic biodiversity conservation.

By exploring the relationship between fish behavior, fluid dynamics, and eDNA dispersion, this research aims to contribute to a deeper understanding of ecosystem dynamics and to enhance our ability to monitor and preserve aquatic biodiversity.

# Chapter 1

# Biomonitoring technique: environmental DNA

The term *environmental DNA* was originally coined in microbiology to describe a method for extracting DNA from a soil sample without isolating the target microorganism [2]. Since the 1980s, researchers are employing eDNA to explore microbial diversity in natural environments. The use of eDNA to detect organisms was first proposed in 1986 as a method of obtaining microorganism nucleic acids from environmental samples [3]. However, only since the first decade of this century, researchers started using eDNA to evaluate macroorganisms abundance. The first adoption of eDNA to macroorganisms was to determine the diversity of mammals, birds, and plants in ancient sediments [4].

In this application, environmental DNA refers to the genetic material found in environmental samples such as sediment, water, and air. eDNA consists of various components as whole cells and extracellular DNA [5]. An intuitive sketch of the origine of eDNA is shown in Figure 1.1. eDNA can be extracted from environmental samples, preserved, amplified, sequenced, and categorized based on its genetic sequence [6]. This process allows one to detect and classify potentially all the species that live in the survey environment. The sources of eDNA are several and can include skin, mucous, saliva, secretions, eggs, feces, urine, blood, roots, leaves, pollen, and decaying bodies of larger organisms [7]. At the same time the production of eDNA is influenced by factors such as biomass, organism age, feeding activity, physiology, life history, and spatial distribution [8].

Despite being a relatively recent surveying method, eDNA has demonstrated potential advantages in biological monitoring. Traditional surveying methods for assessing species richness and abundance are often constrained by taxonomic identification challenges, habitat disturbance or destruction, and limitations in detecting small or elusive species. As a complementary approach, eDNA targeting permits the sampling of a broader diversity of species, enhancing taxonomic resolution, and providing a more comprehensive understanding of entire communities [8]. Moreover, environmental DNA is a valuable resource for researchers and professionals to monitor biodiversity and identify invasive species in water and organic sediments.
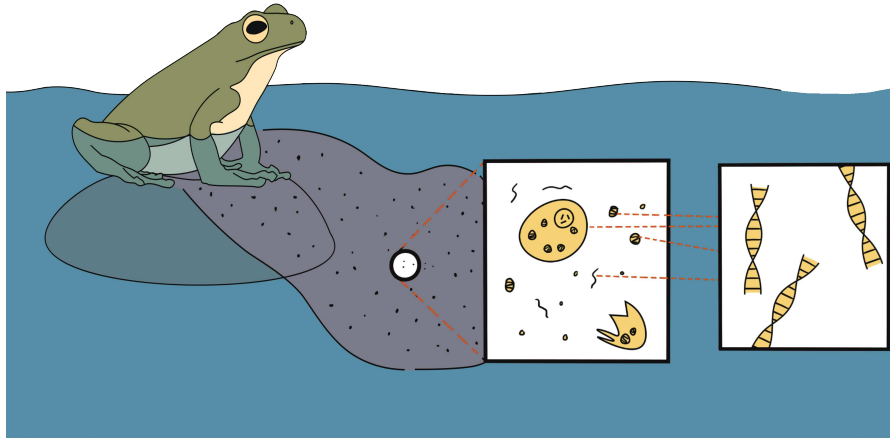
**Figure 1.1:** *eDNA origine representation*

The procedures of direct monitoring fishes may be costly and time-consuming, with potential damage to the organisms. The eDNA methodology is a genuinely non-invasive method that causes no harm to the species or environments under study, even while sampling. It never hurts ecosystems or threatens species. Environmental DNA research has grown in popularity over the last years, thanks to a rise in the number of studies that gather DNA from the environment, particularly in freshwater and marine ecosystems [9] [10] [11].

However, various factors, both abiotic and biotic, can influence fish eDNA and their detection surveys. In addition, eDNA degradation occurs during transport due to processes like sedimentation and biofilm formation, but few studies paid attention to the role of the hydrodynamics of the eDNA fate. Accordingly, the eDNA-based analysis deserves specific studies to clarify the interaction between the genetic material, hydrodynamics, and surrounding environment.

## 1.1    Principle of operation of eDNA in aquatic ecosystem

Environmental DNA can be mitochondrial or nuclear DNA that has been released into the environment by animals. DNA traces are short, degraded fragments of less than 500 bp (base pair) that can be extracted from fresh or seawater, bottom sediments, and ice cores [5]. The process of evaluating the environmental DNA of aquatic species from water samples involves multiple fundamental processes, including sample collection, eDNA extraction and capture, polymerase chain reaction (PCR) and sequencing, bioinformatics analysis, and statistical analysis afterward. Fish eDNA abundance in the water may be correlated with sample volume. All kinds of equipment need to be used with a certain attention and they must be thoroughly disinfected before sampling to prevent contamination. Temperature, pH, and several other environmental factors should be assessed while collecting water since they might affect how long eDNA stays in the water [12].

The three basic procedures for obtaining eDNA from aquatic environments are (i)

filtration, (ii) precipitation, and (iii) centrifugation. Filtration is the most prevalent method of collecting eDNA; it involves passing a known quantity of water through a porous filter, which traps the eDNA particles in the water. Once the water has passed through the filter, the eDNA remains trapped on its surface or within the pores. Subsequently, the eDNA is extracted from the filter using chemical or thermal methods. Membrane filters are mostly used for this purpose. Glass fiber, cellulose nitrate, and polyether sulfone filters are often employed for efficient eDNA collection due to their ability to absorb DNA and proteins [13]. In all cases, filter paper can be refrigerated before extraction or stored in ethanol for up to four days without impacting DNA copy numbers.

Precipitation entails extracting eDNA by adding a precipitating agent to the aqueous solution containing the eDNA. The precipitating agent coagulates and traps the eDNA, which can then be collected through centrifugation or filtration. The precipitation approach reduces the possibility of detection, especially for rare and invasive species, because it involves the addition of ethanol or isopropanol together with sodium acetate to water samples, but it is only practical for small amounts of water (30 ml). Using an ethanol and sodium acetate solution, the eDNA can be concentrated after precipitation and left at room temperature for up to seven days.

Centrifugation is a method for separating particles of different densities suspended in a solution. In eDNA extraction, centrifugation is used to separate the eDNA from other particles present in the solution. After centrifugation, the precipitated eDNA can be collected and further purified for analysis. However, it is challenging to employ centrifugation with large amounts of water [13].

Furthermore, combinations of filtration and precipitation methods or precipitation and ultrafiltration can also be employed for eDNA collection.

After capturing eDNA, commercial kits such as the DNeasy Blood & Tissue Kit or the DNeasy PowerWater Kit (Qiagen) are commonly utilized for extraction and purification. The choice of extraction kit can impact the analysis and relevance of data.

Influencing factors affecting eDNA stability include characteristics of the discharged eDNA (such as molecular conformation and length), environmental abiotic factors (like temperature, UV radiation, dissolved oxygen, pH, salinity, and water flow), and biotic factors related to the organisms themselves and their biological environment [14].

Two are the approaches for eDNA analysis: a targeted, single-species approach and a comprehensive, multi-species approach (Figure 1.2). The species-specific approach, also known as target-species analysis or eDNA barcoding, utilizes Polymerase Chain Reaction (PCR), a technique used to amplify specific DNA fragments in vitro, utilizing repeated thermal cycling and the heat-resistant enzyme DNA polymerase, with species-specific primers to amplify and identify short fragments of eDNA released by a particular species [13]. This method primarily employs the following techniques to detect the eDNA of the target fish species: (1) gel electrophoresis of conventional PCR (cPCR) amplicons, (2) quantitative PCR (qPCR),
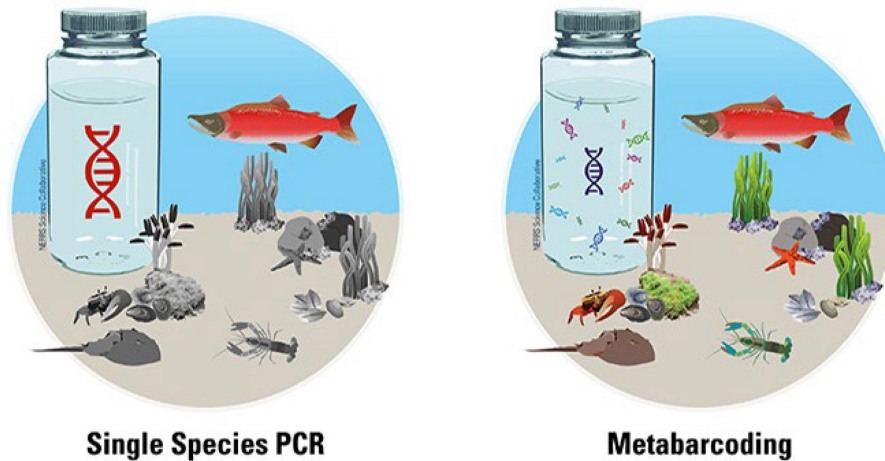
**Figure 1.2:** *Single-species vs multi-species (metabarcoding) approach (Source NOAA).*

(3) nested PCR (nPCR), and (4) digital PCR (dPCR). Conventional PCR methods can be employed to investigate the presence or absence of fishes in the ecosystem, while real-time qPCR serves as the primary method for species-specific detection in eDNA analysis.

The multi-species approach, also known as environmental DNA metabarcoding, is a powerful technique for identifying multiple species present in water samples. By sequencing a standardized gene fragment, this method allows for the simultaneous detection of multiple species in a single analysis. Similar to traditional barcoding methods, metabarcoding utilizes primers designed to target DNA regions shared across species within a group, such as different species of fishes sharing their habitat, while still revealing unique genetic signatures for each species. This approach is particularly useful for the studying of community dynamics and can provide valuable insights into the taxonomic composition of fish populations across multiple samples [15].

Prediction models concerning the spatio-temporal distribution of fish populations offer valuable insights into their regional and temporal dynamics, even using powerful tools such as machine learning techniques [16]. Diverse models can be confronted [17] for this purpose as shown in Table 1.1.

### 1.1.1 Advantages and disadvantages of using eDNA technique

The benefits of using eDNA technique are numerous [25], resulting from its autonomous nature that enables sample collection in inhospitable environments, such as the Arctic, deep sea, or even other planets, without the need for expert scientists. This capability results in substantial reductions in costs and processing time. While eDNA cannot entirely replace traditional sampling methods and taxonomic expertise, it offers numerous advantages for rapidly and efficiently generating biodiversity information, particularly in natural lotic and lentic aquatic systems. These advantages include:

**Table 1.1:** *Different types of Spatio-temporal distribution prediction models used in eDNA studies.*

| Spatio-temporal distribution prediction models | References |
|---|---|
| Evolving Izhikevich spiking neuron model | Demertzis and Iliadis (2017) [16] |
| Hierarchical model | Shelton et al. (2019) [18] |
| Bayesian model, Logistic regression models | Griffin et al. (2020) [19] |
| Intrinsic potential model (IP) | Matter (2016) [20] |
| Geographic information systems (GISs) model | Sakai et al. (2019) [21] |
| Time series model | Thalinger et al. (2019) [22] |
| Species distribution models (SDMs) | Muha et al. (2017) [23] |
| Phenomenological models | Gallien et al. (2010) [24] |

  (i) enhanced sensitivity, as eDNA is less influenced by environmental factors such as time of day or wind speed;

 (ii) accessibility to non-experts, allowing personnel without extensive experience in species-specific survey techniques to perform the analysis;

(iii) cost-effectiveness compared to traditional methods;

(iv) expedited sample collection and generation of results;

 (v) high accuracy in species identification, resulting in a high detection probability;

(vi) non-invasive sampling techniques;

(vii) ability to detect the presence of pathogens responsible for infectious diseases;

(viii) capacity for detecting and quantifying biodiversity;

(ix) reduction of pathogen transfer risks between sites;

 (x) improvement of animal welfare due to non-destructive sampling;

(xi) mitigation of environmental impact.

Furthermore, as stated in section 1.1, eDNA technology is demonstrated to be suitable for detecting multiple species in a single sample [7]).

Another advantage is the ability to preserve samples in laboratory storage for extended periods, allowing for subsequent evaluation of additional species not initially targeted at the time of collection.

While the adoption of eDNA technologies has significant advantages, it is not a one-size-fits-all solution for species inventory. One limitation is that the eDNA approach only provides information on the presence or absence of the target species.

**Figure 1.3:** *Conceptual diagram depicting the three governing processes of eDNA movement: 1) Transport, 2) Retention, and 3) Resuspension, and associated hypotheses for eDNA movement*

It does not offer insights into factors such as the species' life stage, reproductive status, or overall health. When employing eDNA technologies, several critical considerations must be taken into account, including contamination, inhibition, and errors [7].

In particular, contamination can occur during sample collection, filtration, preservation, and extraction, whereas cross-contamination at sampling sites is possible if gloves are not changed between sites or if filtration equipment is not properly sterilized. Laboratory environments also pose contamination risks during extraction and qPCR preparation. There are two types of errors in eDNA analysis: false positive (Type 1) and false negative (Type 2) error rates. Environmental DNA approaches also face these potential errors, but they are less likely to result in Type 2 errors due to the high sensitivity of PCR analysis methods. Inhibition of the PCR reaction can occur during the analysis phase, potentially leading to false negative results. Non-target components present during sample collection may also contribute to inhibition. While many substances are eliminated during the filtering and extraction process, some may co-extract with the DNA and inhibit the PCR reaction. Failure to detect inhibition can result in inaccurate conclusions [25].

## 1.2   Hydrodynamics of eDNA

The hydrodynamics of environmental DNA refers to the transport and behavior of DNA molecules in aquatic environments. eDNA can be transported through water bodies mainly by advection, however, wakes and interaction with vegetation can strongly influence the flow velocity and persistence over time. Furthermore once released, eDNA undergoes complex transport processes influenced by both abiotic and biotic factors (Figure 1.3) [8].

One key factor influencing eDNA hydrocory is water flow dynamics. Flow velocity, turbulence, and directionality can primarily determine the fate of eDNA particles within a water body [26]. For example, in fast-flowing streams or rivers, eDNA particles may be rapidly transported, leading to downstream accumulation and dilution effects. Conversely, in stagnant or slow-moving water bodies, eDNA dispersal may be limited, resulting in localized concentrations near the source.

Sedimentation is another important process affecting eDNA dynamics. In still or slow-moving water, eDNA particles may settle onto the sediment surface over time [27]. This can lead to the accumulation of eDNA in sediment layers, where it may persist for extended periods, providing a record of past biological activity.

Biofilm formation on submerged surfaces can also influence eDNA transport and concentration. Biofilms, composed of microbial communities and organic matter, can act as both sinks and sources of eDNA [28]. eDNA may become incorporated into biofilms, where it can be protected from degradation and contribute to the genetic diversity of the biofilm community. Conversely, biofilms can release eDNA back into the water column through processes such as sloughing or microbial activity (Figure 1.4) [29].

The behavior of eDNA in water also depends on its molecular characteristics, such as size and conformation. Larger DNA fragments may settle more quickly due to their higher mass, while smaller fragments may remain suspended in the water column for longer periods [6]. The conformation of eDNA molecules, including their association with cell or organelle membranes, can influence their stability and transport properties [14]. Environmental factors such as temperature, pH, and salinity can further influence eDNA hydrodynamics. Changes in these parameters can affect the stability and degradation rate of eDNA molecules, altering their transport capability. For example, high temperatures and low pH levels may accelerate eDNA degradation, reducing its persistence in the environment [8].

Numerical models can be used to describe the hydrodynamics and thus eDNA transport [30]. The common basis of these models is the advection-diffusion, which reads:

$$\frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{u}C) = \nabla \cdot (D\nabla C) + S \qquad (1.1)$$

where $C$ is the concentration of the substance, $t$ the time, $\mathbf{u}$ the velocity vector of the fluid flow, $D$ the diffusion coefficient and $S$ additional source or sink terms.
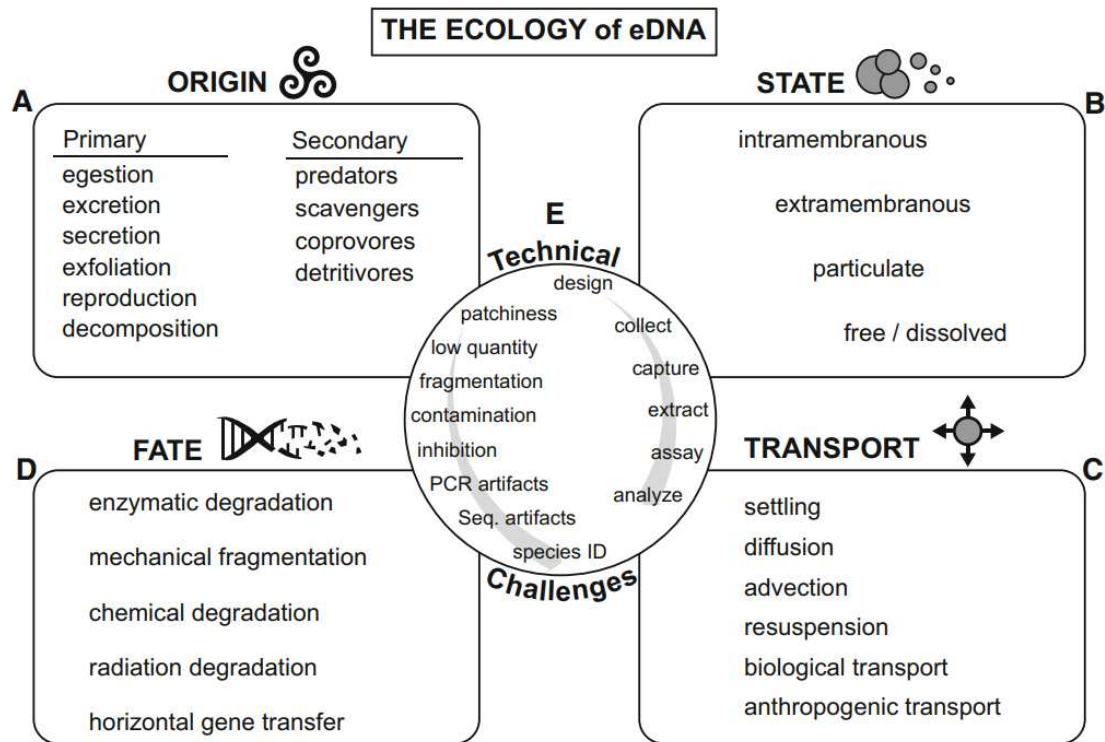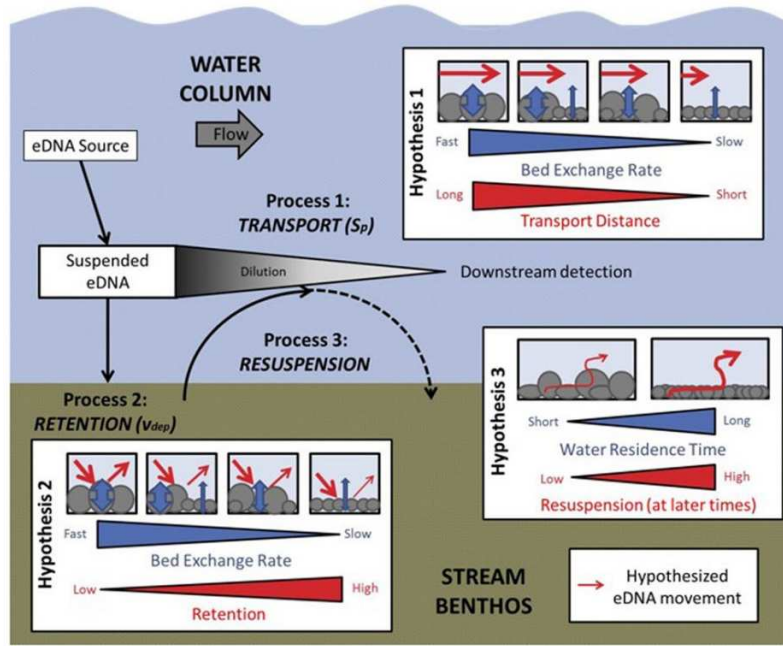
**Figure 1.4:** *Conceptual diagram depicting the three governing processes of eDNA movement: 1) Transport, 2) Retention, and 3) Resuspension, and associated hypotheses for eDNA movement.*

Equation (1.1) accounts for both advection (transport due to bulk water movement) and diffusion (random movement of particles) processes. The A-D equation can be used to simulate the spread of eDNA in aquatic systems under different flow conditions and environmental scenarios.

Understanding the hydrodynamics of eDNA for every specific case of study is central to designing efficient sampling strategies and interpreting survey results accurately. By considering the interactions between eDNA and its surrounding environment, researchers can optimize detection methods and enhance the reliability of eDNA-based surveys.

# Chapter 2

# Materials and methods

This Chapter outlines the methodologies employed to investigate the correlation between the spatial distribution of fish and the concentration of eDNA. The methods detailed here are part of an coherent workflow, in which each process contributes to the ultimate goal.

Following the introduction of the case study and the presentation of the experiments site, the chapter delves into the water sampling procedure, elucidating the methodology used for first collecting the water and then filtering it.

Subsequently, the section on fish tracking elucidates the technique employed to record and track fish movements within the experimental tank, through the utilization of a tracking software.

Moving forward, the numerical modeling section describes the computational fluid dynamics (CFD) approach utilized to simulate the hydrodynamics of the experimental tank and presenting the software used for the CFD.

Afterwards, the domain was defined by means of photogrammetry techniques, and the techniques used to collect the images and process them in order to obtain the three-dimensional reconstruction of the domain were explained. The section on hydrodynamic modeling provides insights into the modeling process of our specific case study, detailing the methodology employed to simulate the fluid flow within the tank.

Lastly, the chapter explores particle tracing techniques, explaining how the movement of particles within the simulated fluid was tracked to understand the dispersion patterns of eDNA.

## 2.1 Case Study Overview

The research delineated in this thesis studies a specified tank located within the confines of the Oceanium of Diergaarde Blijdorp. This tank served as the primary site for our experimental observations. Nestled within a sector designated for housing marine species accustomed to colder climates, this setting presents a suitable environment for our investigative pursuits 2.1. Adjacent to our selected tank lies
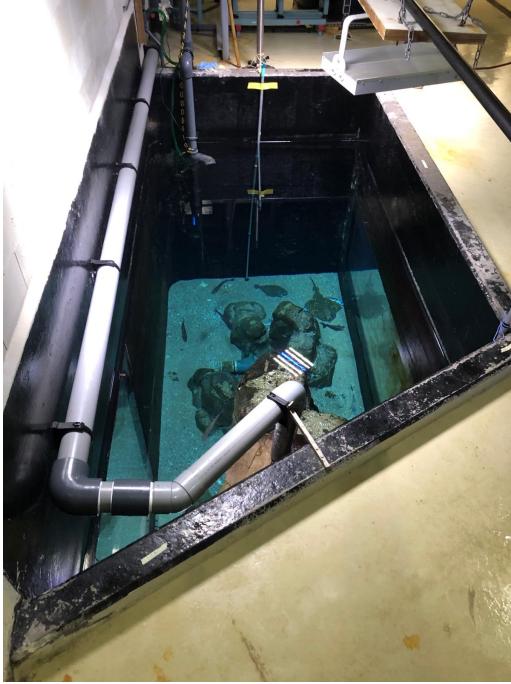
(A) *Photo from distal side.*                    (B) *Photo from proximal side.*

**Figure 2.1:** *Photos of the tank from the two short sides.*

the 'North Sea tank', a larger aquatic enclosure that hosts a diverse assembly of marine life, including but not limited to sharks, turtles, and rays.

The selection of this particular tank for our experimental endeavors was informed by several critical factors. Foremost among these was the tank's geometric simplicity, a characteristic that significantly facilitated the creation of a Computational Fluid Dynamics (CFD) model. Additionally, the habitual vacancy of the tank—typically reserved for ailing fish, those exhibiting behavioral issues, or simply as a supplementary space—rendered it an ideal candidate for our study. This absence of marine inhabitants ensured that our experimental procedures, which included photogrammetric filming and dye introduction for hydrodynamic analysis, could be conducted without any undue impact on the aquatic fauna.

This setup provided a unique opportunity to examine the hydrodynamic behavior within the tank's confines, enabling a thorough investigation free from the complexities and ethical considerations associated with disturbing live specimens.

In the tank the renew of the water is possible by regulating the inflow by an inlet positioned on one side near the free surface, while an outlet centrally located at the bottom of the tank, allows for the removal of water and filtration of settled material. This arrangement maintains a constant water flow. Generally, the discharges varies between 7.5 m$^3$/s and 9.0 m$^3$/s, while the water temperature is maintained at a temperature of 16°C and the conductivity 48.9 mS/cm, resulting in an estimated water density of 1029 kg/m$^3$.

This tank is characterized by a rectangular trapezoidal geometry, whose dimensions and the shape are shown in Figure 2.2. The tank contains several rocks to reproduce a natural aquatic habitats as much as possible. Within the confines of the tank, there are four main rock structures: three of these are approximately at 40 cm height and 40 ÷ 80 cm in width, and are located close the outlet of the tank. The fourth rock is located near the tank's oblique side, distinctively larger, is approximately 1.80 m in height.

In the forthcoming experiments, five fish of pelagic species (mullets),living in the water column, and two fish of benthic species (flounders), living close to the bottom, will be introduced. This choice is based on the expectation that the flounders will be primarily sources of eDNA on the bottom of the tank, while the mullets will release genetic maerial at lower depths. Subsequently, eDNA sampling will be conducted according to a predetermined timetable.

## 2.2 Diergaarde Blijdorp zoo: the Oceanium

The Oceanium is an acclaimed public aquarium within Diergaarde Blijdorp, a zoo located in Rotterdam (The Netherlands). This area of the zoo was inaugurated in 2001 and represents the most ambitious initiative undertaken by the zoo so far. Strategically situated in the zoo's expansion area, the Oceanium is dedicated to showcasing a wide range of different spiecies.

Over the years the Oceanium has been place of scientific researches, focused particularly on the field of coral conservation. Additionally, the Oceanium stands out for its impressive solar power installation, called 'the Sunport'. This installation, notable for being the largest solar power plant in the Netherlands within a built-up area, encompasses approximately 3400 solar panels spread over 5000 square meters, generating about 325,000 kWh per year.

All aquariums within the Oceanium are filled with clean seawater, amounting to more than eight million liters in total. The water refill of such an amount of water is possible thanks to the direct connection of Rotterdam to the North Sea by the river Maas. However an advanced filtration system allows for minimal water replacement, ensuring the wellbeing of the diverse marine species hosted, which vary from puffins and sea turtles to a variety of sharks.

This comprehensive approach to showcasing marine biodiversity, coupled with a strong emphasis on sustainability and scientific research, makes the Oceanium a unique and invaluable resource for both public education and conservation efforts.

## 2.3 Water sampling

The water sampling was conducted to determine the concentrations of eDNA at various points within the tank. Sampling was performed as depicted in Figure 2.3. At each position shown in the plan view (Figure 2.3 A), sampling was conducted through the water column at three different depths: (i) just below the water surface,

(A) *Top view of the tank.*



(B) *Side view of the tank.*



(C) *Perspective view of the tank.*
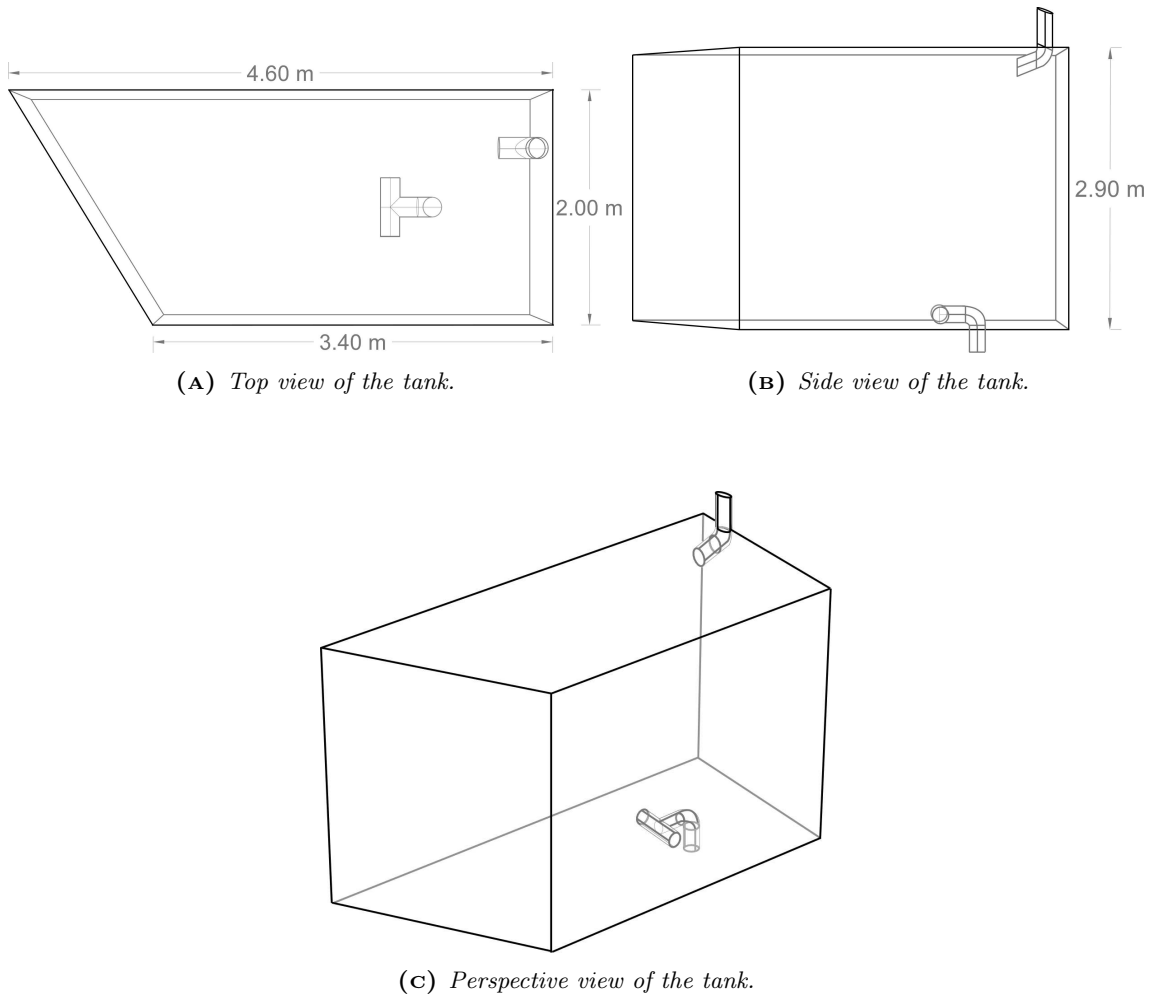
**Figure 2.2:** *Tank geometry*

(ii) at mid-depth, and (iii) at the bottom (Figure 2.3 B), following this schedule: the first sampling was conducted one hour after fish introduction, followed by intervals of 24 hours over a period of five days. Sampling was carried out using 3 Niskin bottles (Figure 2.4). The Niskin bottles are equipped with two openings at the top and bottom. The caps of the bottles feature two lanyards so that once opened, they can be held open by tying the lanyards to a small metal cylinder positioned halfway along the bottle. This metal cylinder is sliding and can be opened by depressing a piston located next to the upper cap, which is necessary for sealing the bottles. To immerse the bottles in water, an approximately 5 meters long rope is used. When the bottles are submerged in water, the piston is depressed with a weight connected to a "messenger" rope parallel to that used to dive the Niskin bottle into the water. When the weight touches the piston, the bottles close, and the water contained inside the Niskin bottle is trapped by the two hermetic caps. The rope is then pulled up slowly, ensuring that the process is carried out gradually to minimize disturbance to the tank's hydrodynamics.
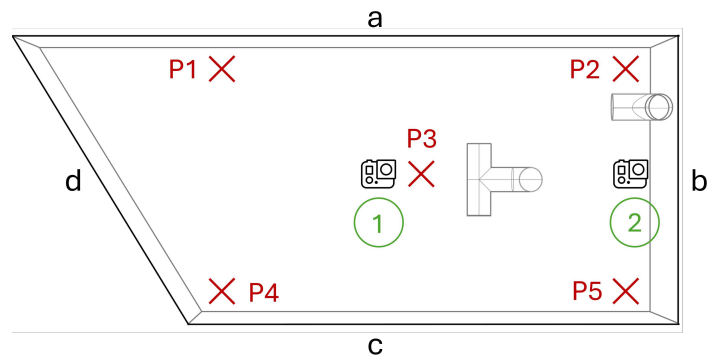
(A) *Top view of sampling sections P1-P5.*
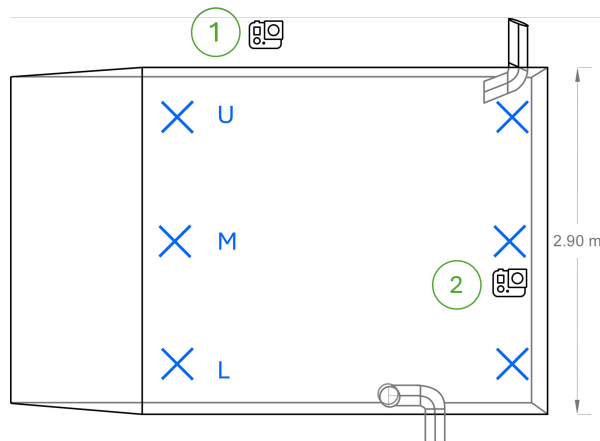


(B) *Side view of sampling and cameras positions.*

**Figure 2.3:** *Top (A) and side (B) view of sampling points of each sections (blue cross) and cameras positions (green numbers).*

One water column per time was sampled, collecting at three different depths. The water was subsequently filtered, and finally the three empties niskin bottles were sterilised with bleach to remove residual eDNA and rinsed thoroughly for subsequent sampling.

### 2.3.1 Water filtering

After sampling a water column, filtration was performed. Filtration is a delicate process that needs to keep the working environment as sterile as possible to prevent contamination of the water. For this purpose, all work surfaces were sterilised with bleach, rinsed and sterilised again with ethanol. Each process was carried out with sterilised gloves.

For each sample, we filtered a volume of 1000 mL through a PES (polyether-sulfone) membrane filter of 47 mm diamater provided by Millipore Express. This

(A) *Closed bottle.*    (B) *Opened bottle.*
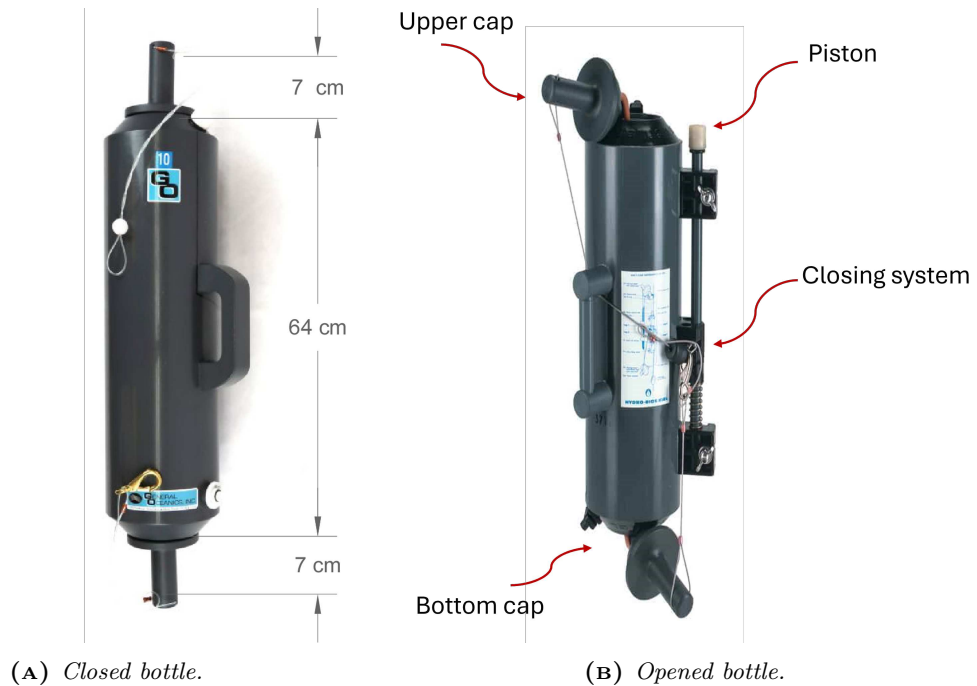
**Figure 2.4:** *Niskin Bottles*

membrane ensures chemical and mechanical resistance properties, essential for maintaining the integrity of the filtration process. The membrane has a nominal pore size of 0.45 µm, defining the maximum size of particles that can pass through it. This ensures effective retention of particles larger in size, thereby maintaining the purity of the filtered solution. To execute the filtration, a sterile Nalgene filter column was employed, designed to maintain the filtration system aseptic and securely contain the filter membrane during the process. The necessary pressure to drive the liquid through the membrane was applied using a Laboport N816 vacuum pump. Creating a negative pressure inside the column allowed the liquid to flow through the membrane, separating unwanted particles from the clean liquid.

Through this process, a high-quality filtered solution was obtained, free from impurities and ready for analysis or further processing in the experiment. (Figure 2.5).

## 2.4  Fish tracking

Animal tracking algorithm are used to study wildlife behavior, and are adopted by researchers to monitor the movements and interactions of animals in their natural habitats and thus gaining insights into animal migration patterns, social structures, habitat use, and responses to environmental changes. Different tools can be employed to collect data on animal locations over time, including GPS collars, radio tracking, and automated video or photographic systems. By analyzing these

**Figure 2.5:** *Filtering station.*

data, researchers can better understand the dynamics of wildlife populations in their ecosystems, contributing to conservation efforts and the management of natural resources. Animal tracking allows for the quantitative analysis of an organism's detection and decision-making processes across various ecological settings.

In the present work, we employed animal tracking algorithm to investigate behavioral patterns; specifically, we used the software *TRex* which allow us to track of multiple individuals simultaneously for this purpose [31]. *TRex* stands out due to its real-time tracking capabilities, handling up to 256 individuals and providing detailed analytics such as 2D visual fields and identification through deep learning. Its efficiency, requiring significantly less memory and being faster compared to similar tools, alongside an intuitive graphical user interface, makes *TRex* an ideal choice for our study, offering a combination of precision and ease-of-use, useful for analyzing complex behavioral data over the time.

### 2.4.1   Fish recording

To track the movements of the two introduced fish species, we used two GoPro Hero 10 cameras, capable of underwater operation up to a depth of 10 meters without the need for protective case. The cameras were mounted on telescopic poles equipped with sliding supports designed to position the GoPros at the desired depth. One camera was positioned outside the tank to capture the top view (Figure 2.3 A). The second camera was placed inside the tank, approximately 1.5 meters above the bottom, to record the movement of the fish along the vertical axis (Figure 2.3 B). Both cameras were set to linear mode, with a resolution of 1080p and a frame rate of 24 fps. The cameras started to record separately for practical reasons, and synchronized in post processing thanks to a light source that was turned off at the beginning of each recording. A video was recorded in the first hour after

the introduction of the fish into the tank, before to initiating the water sampling
described in Section 2.3. In the next days recording were always made 1 hour before
the sampling. Finally, the submerged camera was carefully extracted to minimize
hydrodynamics disturbance enhancing an unwanted artificial mixing.

### 2.4.2   TRex: a multi-animal tracking system

Tracking animal movements is pivotal for disciplines such as behavioral science and
neuroscience. Numerous algorithms have emerged, yet many are limited by species
specificity or struggle with high-resolution footage. Fast-tracking solutions often
fail to simultaneously track a large number of individuals, restricting their appli-
cation in complex studies. Additionally, the market is fragmented, with each tool
offering a niche set of features. This fragmentation and the limitations of current
tools underscore the need for a more versatile and efficient tracking solution able
of accommodating diverse research requirements across multiple species and exper-
imental conditions.
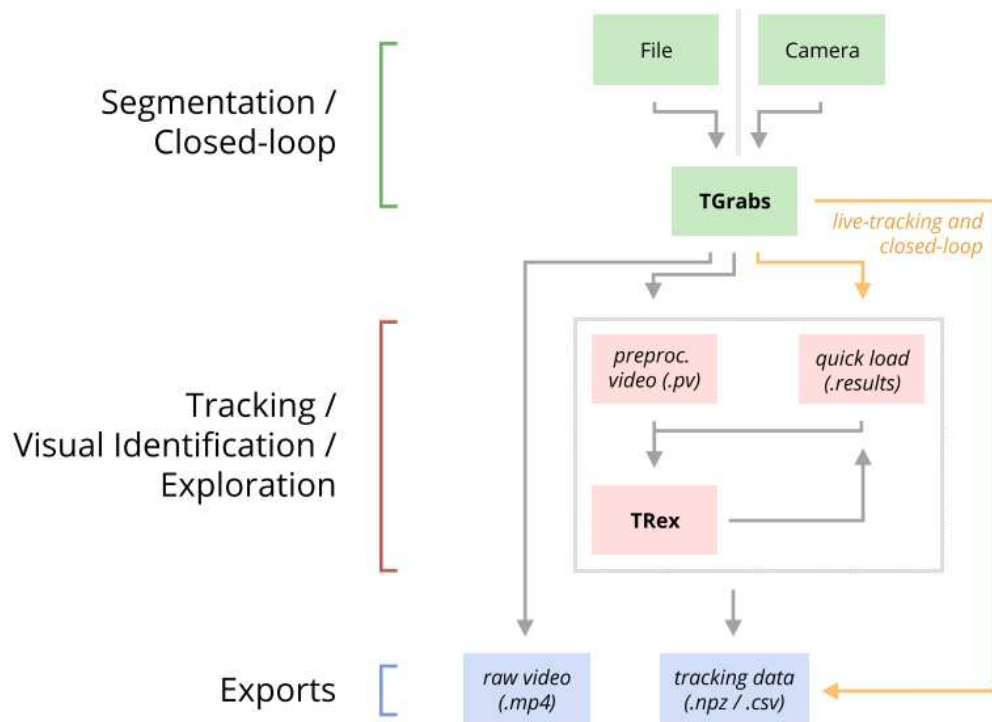


**Figure 2.6:** *TRex workflow.*

TRex is an open-source software designed for advanced behavioral studies, en-
abling efficient tracking of moving objects or animals with background differentia-
tion capabilities. It provides detailed metrics per individual, such as body shape,
and head/tail positions, through posture analysis adaptable for various organisms.
Moreover, TRex offers computational reconstruction of visual fields for all tracked

individuals, enhancing understanding of visual interactions and distance estimations between subjects.

The TRex software package includes TGrabs, a tool for pre-processing videos and live-streaming from cameras, supporting a wide range of tracking features like position, posture, and visual fields. It allows for immediate access to tracking results, enabling efficient closed-loop experiments with up to 128 individuals. TRex and TGrabs, developed in C++, offer a Python-based scripting interface for customization and compatibility with external programs, facilitating consistent integration and expansion for various research needs.

TRex utilizes kinematic data for fast tracking and identity maintenance of organisms, essential for analyzing group dynamics. However, maintaining accurate identity in videos, especially with numerous individuals or when they exit the frame, poses challenges. TRex addresses these issues by implementing a machine-learning approach to recognize up to 100 unmarked individuals based on visual appearance, favoring the automatic correction of tracking errors without manual intervention. This visual identification feature significantly enhances the reliability of behavioral analysis by ensuring consistent identity tracking throughout the study.

TRex is an open-source, platform-independent tracking software compatible with Linux, Windows, and macOS. It streamlines batch processing for large video datasets and offers multiple ways to adjust settings via GUI, command-line, or settings files. The software supports remote access through a web-server within a VPN for added security. Results are exportable to NPZ or CSV formats for further analysis. The workflow is depicted in Figure 2.6 Detailed documentation and a comprehensive guide to its features are available at `trex.run/docs`.

### 2.4.3   Application of TRex in the Case Study

The procedure for implementing TRex in our study starts by launching TRex from the Python prompt. In the pre-processing phase, the video recorded with the GoPro camera is converted to PV format. With the PV-converted video, we set preliminary parameters such as `track_max_individuals` to specify the maximum number of individuals to be tracked, which is equal to the number of individuals we have within the tank. Another parameter is `blob_size_range` that regulates the filter potential noise; in our case, this parameter was set equal to the minimum to avoid considering floating particles, whereas we did not set any maximum threshold limits since no other individuals are present in the tank besides the two species introduced. Furthermore, we employed the option `track_threshold` is employed for aiding background subtraction. Although `manual_matches` allows for manual identity assignments, this feature was not necessary for our goal because we are not really interested in maintaining identities within the same species but just distinguishing between the two species. Finally, `Output_prefix` is utilized to label all output files consistently.

When tracking starts, the video is automatically segmented into frames, and each

| Variable | Description |
|---|---|
| X, Y | Position on the X/Y-axis. Y starts on the top-left. Both are in cm (if cm_per_pixel() is set correctly, otherwise the conversion is based on an assumed width of 30cm for the arena). |
| SPEED | Magnitude of the first derivative of position (in cm/s). |
| VX, VY | X/Y-component of the first derivative of position (in cm/s). |
| AX, AY | X/Y-component of the acceleration in cm/s (if cm_per_pixel() is set correctly, otherwise, the conversion is based on an assumed width of 30cm for the arena). |
| time | Time in seconds (floating point) from the start of the video, consistent across individuals. |
| timestamp | The time as a UNIX timestamp. |
| frame | Monotonically increasing integer number, consistent across individuals. |
| num_pixels | The number of pixels in the blob (after thresholding). |
| ANGLE | The absolute angle in radians with respect to the X-axis of the image. |
| BORDER_DISTANCE | Usually this is the distance from the centroid of the individual to the edge of the image - in case a use_circle_mask() has been used during conversion, the circle is used as an approximation of the tank size. |
| NEIGHBOR_DISTANCE | Average Euclidean distance to neighbors (in cm). |

**Table 2.1:** *Description of the main variables exportable*

frame is cross-correlated with the subsequent one (Figure 2.7). The software provides multiple display options, such as `#individuals` which indicates the number of tracked individuals over time, and `texts` that assigns numbers to each tracked fish. While numerous additional features are available, they are extraneous for our objectives. Utilizing the `exported options`, we elect to export specific parameters like horizontal and vertical positions (`X` and `Y`), velocities (`VX` and `VY`), and accelerations (`AX` and `AY`).

Subsequently, from the menu, several options are available, including `save tracking data`, here used for saving selected parameters within `exported options`. Furthermore, the software facilitates saving the operational state and configurations, permitting their reuse in the next tracking endeavors. This, along with a multitude of additional features, affords a high degree of customization in the analytical process.

Once the tracking data has been exported, a number of files NPZ equal to the number of fish tracked are created, every file corresponds to a single fish. These files contain the selected information exported. These files are read in Python and contain all the variables explained in table 2.1.

The variables of interest for the case study considered and for the aim of this thesis are `X` and `Y`, which are respectively the position on the horizontal and vertical axis. The axes' origin is on the top-left of the picture. Both the coordinates are in cm

(A) *An example frame captured by camera 1.*



(B) *Same frame analysed with TRex.*

**Figure 2.7:** *Example of a converted frame in TRex*

(if `cm_per_pixel()` is set correctly, otherwise the conversion is based on an assumed width of 30 cm for the window of interest). If `output_interpolate_positions()` is set, X and Y values can be interpolated.

## 2.5   Numerical model

Computational fluid dynamics (CFD) is emerging as a fundamental discipline in both engineering and scientific research, which increasingly requires quantitative predictions and analysis of fluid dynamic phenomena through the use of computational resources. Fundamentally, CFD solves numerically the equations governing fluid motion, favoring the comprehension of complex fluid dynamic models in various practical contexts.

CFD has been more and more adopted due to rapid technological advancements that have made increasingly powerful computers available at affordable costs. This has opened the possibility for engineers to quickly conduct complex simulations, significantly reducing the cost and time associated with physical testing. The accuracy of the mathematical models used in CFD has also improved over time, enabling more reliable predictions of fluid behavior. Due to its wide applicability in sectors such as meteorology, aerospace, and automotive, CFD has become a fundamental resource for engineers and researchers. The development of specialized software has made it more accessible to perform complex simulations, expanding the number of professionals capable of using this technology. CFD therefore continues to gain popularity due to its ability to optimise the design and analysis of fluid dynamic systems, helping to improve efficiency and safety in a wide range of industrial applications.

In a CFD analysis, fluid flow and its physical attributes (e.g. velocity, pressure, viscosity, density, and temperature) are calculated based on defined operating conditions. To obtain an accurate physical solution, these variables must be calculated simultaneously. Either using commercial or open-source CFD tools, each employs mathematical models and numerical methods to simulate the desired flow problems. Predominantly, the governing equations are based on the Navier-Stokes equations, the cornerstone of fluid dynamic modeling. Although the fundamental structure of these equations remains intact, additional terms are introduced or omitted based on specific physical considerations. For example, low Reynolds number problems can be modeled neglecting the convective acceleration, therefore reducing the Navier-Stokes equations to the simpler Stokes equation.

The accuracy of the operating conditions, numerical methodologies, and physical parameters is the basis for successful CFD analyses. When performed correctly, CFD analyses provide rapid insights, leading to even highly accurate analyses and predictions.

To investigate a fluid problem using CFD, the following stages must be followed. First, the mathematical equations that describe the fluid flow are defined. Usually, they are a set of partial differential equations, that are discretized to produce a numerical counterpart. The domain is split into small grids or elements. Ultimately, the initial and boundary conditions for the specific problem are prescribed to solve these equations. Solution methods might be direct or iterative. Furthermore, specific control parameters are used to assess the method's convergence, stability, and accuracy.

Every CFD code includes three essential elements: (i) A pre-processor, that gives in input the problem geometry, generates the grid, and defines the flow parameter and the boundary conditions. (ii) A flow solver, used to solve the governing equations of the flow subject to the conditions provided. Different methods are developed for this purpose. Among others, we mention the finite difference method, finite element method, finite volume method, and spectral method. (iii) A post-processor, which graphically displays the model outputs.
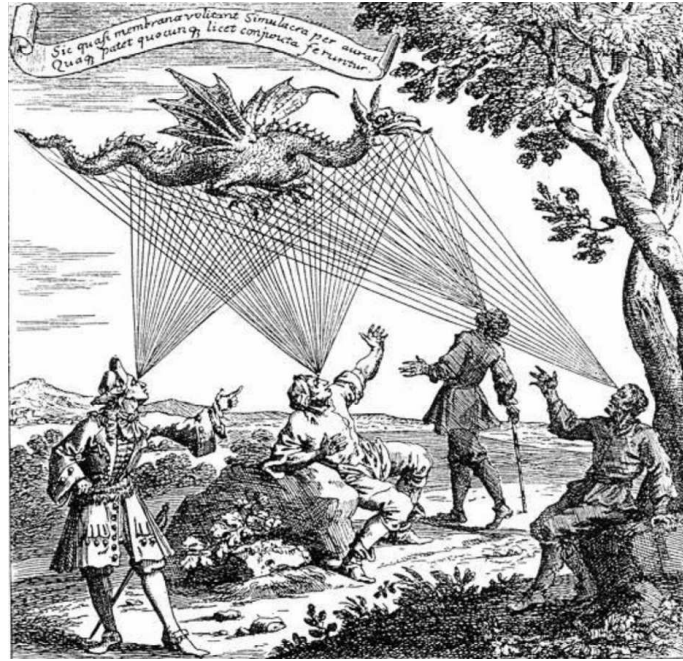
**Figure 2.8:** *Illustration from Oculus artificialis teledioptricus sive Telescopium, 1685*

### 2.5.1   COMSOL Multiphysics software

COMSOL Multiphysics is a software designed for real-world applications, aiming to closely mimic observed effects in both engineering and scientific contexts. This software, used for modeling and simulating physical systems across various disciplines, provides a flexible and user-friendly environment for solving complex multiphysics problems. With COMSOL Multiphysics, users can create detailed models involving coupled physics phenomena, such as fluid flow, heat transfer, structural mechanics, electromagnetics, and chemical reactions. The software offers a wide range of predefined physics interfaces and material libraries, allowing for efficient model setup. It supports both finite element analysis and finite volume methods, to ensure accurate simulations of diverse systems. COMSOL Multiphysics includes robust solvers capable of handling nonlinearities, transient phenomena, and multiphysics interactions. It offers comprehensive post-processing tools for visualizing and analyzing simulation results, aiding in the interpretation of complex data. The software supports customization and automation through its scripting and programming interfaces, facilitating advanced modeling techniques and workflow automation.

### 2.5.2   Domain definition: photogrammetry

As highlighted in the previous section, the proper geometric representation and the associated numerical grid of the domain of analysis are necessary to reproduce the flow field in the tank.

In the present work, the presence of rocks strongly complicates the manual execution of this task. To tackle this challenge, photogrammetry has been utilized to

reconstruct the geometry of the rock elements.

Photogrammetry is a technique for extracting reliable information on the natural environment or physical objects by recording, measuring, and interpreting photography or other images obtained by electromagnetic radiation or other phenomena.

The technique dates back as early as 1480 with Leonardo da Vinci's research on perspective. In 1685, Johann Zahn published his Oculus Artificialis Teledioptricus Sive Telescopium (The Artificial Long Distance Eye, or Telescope). In the work, which provides an early and comprehensive account of the function and use of a range of optical instruments, early images depicted the idea behind photogrammetry (Figure 2.8). In 1851, the French inventor Aimé Laussedat saw the possibility of using the newly invented camera in mapping. While in 1867, Prussian architect Albrecht Medenbauer coined the name photogrammetry in his article "Die Photometrophie."

To understand the purpose of photogrammetry, it is sufficient to decompose the term; "photos" refers to light, "gramma" means to draw, and "metron" to measure.

Through photogrammetry, it is possible to make measurements and infer the geometry of a scene from a set of unordered photos or videos. Conversely, to photography, which projects a 3D scene onto a 2D plane, losing depth information, photogrammetry is based on the reverse process.

Exploiting the transformation from 2D to 3D, photogrammetry empowers diverse fields, from land surveying and cartography to architecture and archaeology, facilitating advanced analyses and informed decision-making processes.

The following pages outline the workflow related to photogrammetry, aiming to utilize the output of this process for Computational Fluid Dynamics simulations.

**Images aquisition**

The photogrammetry was adopted to obtain three-dimensional models of the rock structures within the tank, after capturing images from videos recorded with a GoPro camera. This methodology was implemented to capture precise and complete details of the rocks, thereby accurately reconstructing the domain.

To record the videos, a GoPro camera was mounted on a 4.5-meter telescopic pole, providing a high degree of flexibility for the acquisition. The telescopic pole, appropriately equipped with a GoPro holder, allowed for controlled extension within the tank, ensuring footage at constant distances from the rock structures. In order to capture the object of interest in its entirety, different camera angles and perspectives were adopted. The GoPro's 'wide' mode was selected to widen the viewing angle, accepting slight distortions to ensure a wider view of the rocks within the confined environment of the tank.

The recorded videos were then processed through the *ffmpeg* video processing software, extracted at a frame rate of 3 fps to obtain a series of images. Subsequently, the most suitable frames that ensured the clarity and integrity of the object in each shot were selected manually. The resulting images were used to perform four separate photogrammetries, each focused on a specific rock structure within the tank. For
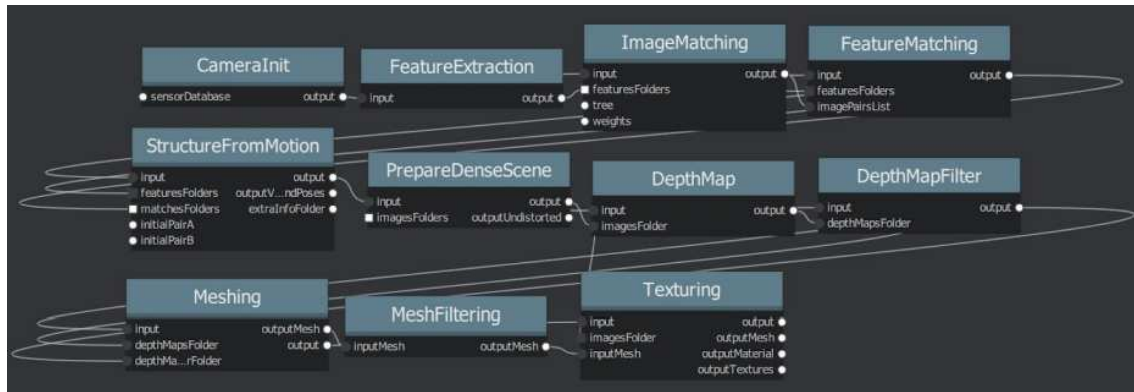
**Figure 2.9:** *Workflow of meshroom*

each rock structure, we used from 200 to 350 images to extract 3D geometry, thus ensuring the coverage of the rocks and surrounding environmental details.

**Usage of Meshroom software and Rhino 3D**

We extracted the 3D geometries using the software Meshroom. Meshroom is a free, open-source 3D Reconstruction Software based on the AliceVision framework. The advantage of this software lies in its node-based structure, which permits a customizable workflow at each step. This structure retains information for each node, enabling multiple pipelines simultaneously. With this structure, new images of the object under analysis can be added without losing previously acquired information. Accordingly, one can easily correct parts of the mesh that are not accurately reconstructed by adding new photos, or interrupting the pipeline if modification of a node becomes necessary. Subsequently, restarting from the interrupted node without repeating the entire pipeline is feasible. The default pipeline (Figure 2.9) includes the following steps:

1. Feature extraction:
   extract distinctive groups of pixels that are, to some extent, invariant to changing camera viewpoints during image acquisition using SIFT (Scale-invariant feature transform) algorithm

2. Image matching:
   identifies images capturing the same scene area, associating a feature descriptor with each image to simplify distance calculation between cameras.

3. Feature matching:
   finds correspondences between images using feature descriptors.

4. Structure from motion:
   computes geometric relationships among viewpoints from input images, yielding a point cloud of the object and determining the orientation and position of all cameras.

5. Prepare dense scene:
   corrects image distortion to expedite the subsequent depth calculation.

6. Depth map:
   computes the depth level of each pixel for images with known camera positions.

7. Depth map filter:
   filters occlusions among different depth maps.

8. Meshing:
   generates an initial version of the mesh.

9. Mesh filtering:
   refines the obtained mesh, involving operations like mesh smoothing, removal of large triangles, and elimination of secondary meshes (elements near the object also reconstructed).

10. Texturing:
    computes UV maps, that define how the texture will be positioned on each surface of the model, and texture for the mesh.

The complete guide is available on the website [32].

The algorithm output was a point-cloud of the rocks inside the tank, extracted as file *.obj* (Figure 2.10b).

Subsequently, `Rhinoceros` —a commercial 3D graphics and computer—aided design (CAD) application and Grasshopper —a visual programming language that is deeply integrated with Rhino's 3D modeling capabilities— were employed. First Grasshopper was used to convert the point-cloud into a mesh (Figure 2.11) and then into a closed surface, suitable to be inserted in the numerical model. However, such a 3D reconstructions generated by Meshroom are very detailed, resulting in localized regions of exceedingly dense meshes. To circumvent this limitation, a smoothing process was undertaken utilizing Rhino 3D. As a final result, the complexity of the meshes to be inserted in the numerical model was reduced up to 80%. This combined approach facilitated the manipulation and optimization of the rock models, ensuring a mesh suitable to the computational constraints of COMSOL without sacrificing the essential geometric fidelity of the structures. The 3D shaped models were exported in .igs format from Rhino and subsequently, in COMSOL, these models were scaled and positioned.

### 2.5.3 Hydrodynamic Modeling of the Diergaarde Blijdorp Zoo Tank

The CFD model created through COMSOL Multiphysics aims to reproduce the hydrodynamic conditions of the tank used for the present experimental work conducted at the zoo Dierdgaarde Blijdorp in Rotterdam. Analysing the water flow behaviour inside the tank providing an informative basis for modeling the eDNA transport dynamics.

(A) *Photo of the real rock.*

(B) *Reconstruction of the rock tro-hugh meshroom.*



(C) *Shaping of the rock through Rhino 3D.*
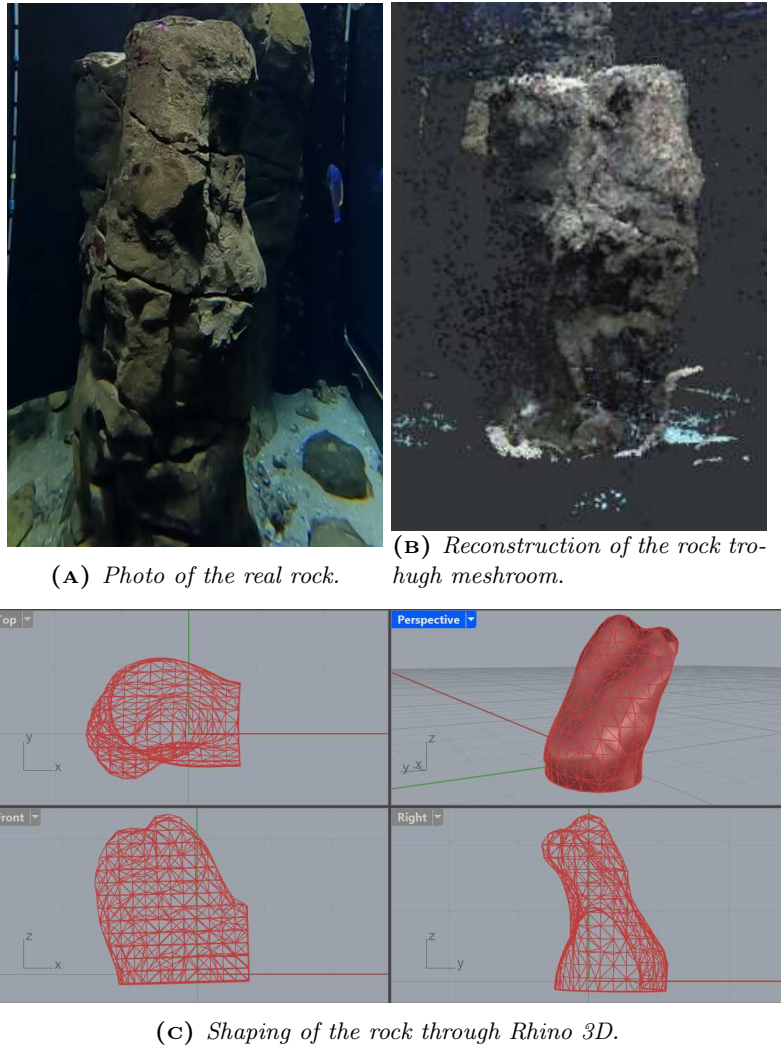
**Figure 2.10:** *Reconstruction of one of the rocks inside the tank..*

First, the geometry was drawn in COMSOL by reproducing the shape of the tank, inlet, and outlet, as shown in Figure 2.12. The construction of the computational mesh presented considerable difficulties, mainly attributed to the complexity of the rock within it. Once these challenges were overcome, as illustrated in the previous subsection, the rocks were placed within the created model and were integrated with the domain.

Then, the boundary conditions were set: the inlet flow rate was fixed, and the pressure boundary condition was specified in the outlet as shown in Figure 2.12 (A). A no-slip condition was applied to all walls except the free surface, which was treated as a slip surface. We initially considered the flow at rest and applied a flow rate that linearly increased from 0 to the target value of 9.0 $m^3/s$ in 100 s, after that, we kept this value constant for 1800 s to ensure the steady flow condition in the tank. The mesh generated with Comsol has the following characteristics: number of vertex
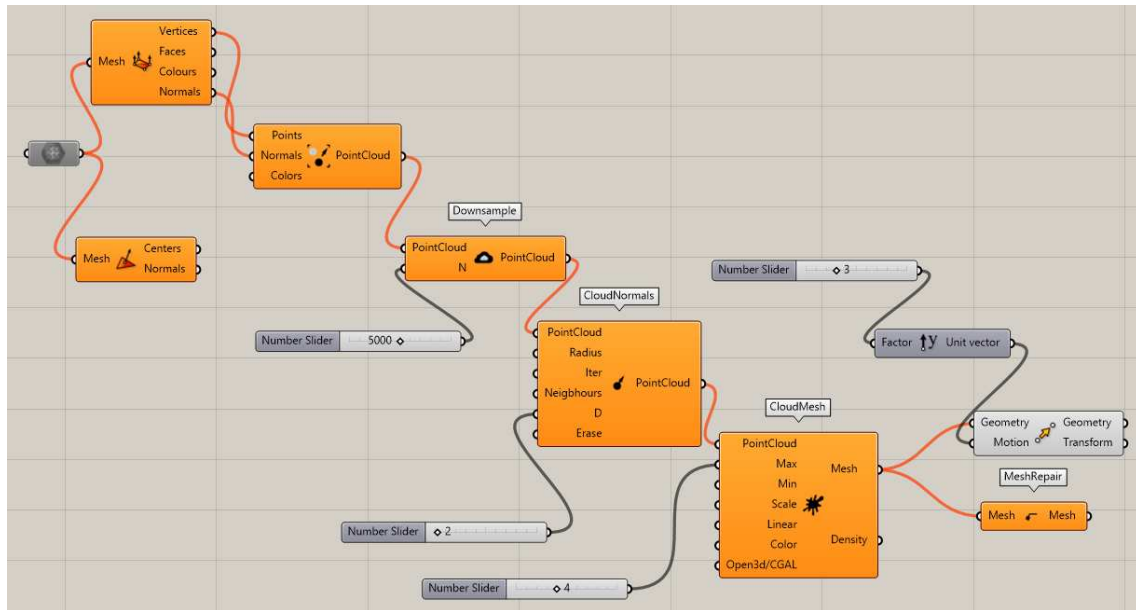
**Figure 2.11:** *Workflow of Grasshopper*

elements: 751, number of edge elements: 9958, number of boundary elements: 5322, number of elements: 596153. (Figure 2.12 (B)).
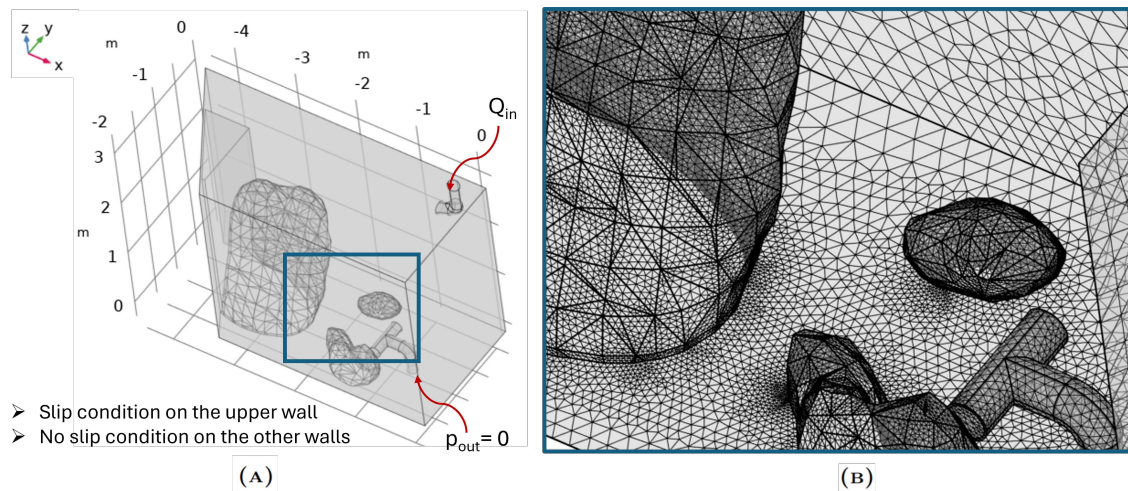


**Figure 2.12:** *Reconstruction of the domain in Comsol Multiphysics (A). Mesh detail (B).*

### 2.5.4 Particle tracing

Particle tracing was used in our study to track the movement of particles in the simulated velocity field, allowing us to better understand the dispersion of eDNA traces in the aquatic environment. This approach allowed us to assess the spatial distribution of particles in the system and to analyse eDNA transport patterns.

To implement particle tracing, the particle tracing function of COMSOL was em-

ployed. A Lagrangian approach was adopted, which allows the individual tracking of each particle in the system. This method proves particularly suitable for simulating the transport of eDNA within the simulated tank, allowing the identification of potential dispersion paths depending also on the residence time of each particle.

In the process of tracking the particles, the various parameters influencing their movement in the fluid were considered, such as the intrinsic parameters of the particles and fluid, the external forces applied to the system, and the boundary conditions of the model. Using COMSOL Multiphysics software, a simulation of the particles' motion in the water system was performed, taking into account the aforementioned parameters. During the simulation, particles were released at specific points (or areas) in the tank and their movement was tracked over time to understand their behaviour in response to environmental conditions and model parameters. For further details and how to set the parameters, please refer to the COMSOL Particle Tracing User's Guide [33].

It is important to note that despite the accuracy of the particle tracing model, there may be some limitations and additional considerations to take into account. For example, the simulation may not take into account mixing due to fish swimming in the environment.

Concerning particle properties, we assigned the same characteristics for the benthic species and for the pelagic species, with density $\rho_p = 1010 \ kg/m^3$ and particle diameter $d_p = 0.01 \ mm$. The release of the particles was performed considering the tracking of the fish: as shown in the results, the flounders remained practically motionless during the recording time, so the points occupied by the flounders were chosen as release points. For the pelagic species, a higher probability density of presence was identified in certain areas, so the tank was divided into 3 volumes, and different numbers of particles released over time were assigned according to the proportions of the probability densities found.

the forces considered in the eDNA particles dynamics are the force of gravity and the drag force, which are described by the following equations:

$$\mathbf{F}_g = m_p \mathbf{g} \frac{\rho_p - \rho}{\rho_p} \tag{2.1}$$

$$\mathbf{F}_D = \frac{1}{\tau_p} m_p (\mathbf{u} - \mathbf{v}) \tag{2.2a}$$

$$\tau_p = \frac{4 \rho_p d_p^2}{3 \mu C_D Re_r} \tag{2.2b}$$

$$C_D = \frac{24}{Re_r} \left( a + \frac{3}{16} Re_r \right) \tag{2.2c}$$

$$Re_r = \frac{\rho |\mathbf{u} - \mathbf{v}| d_p}{\mu} \tag{2.2d}$$

As a boundary condition on the outlet, the condition 'disappear' was set, meaning that particles that are sucked out of the outlet will disappear from the statistics.

A time-dependent study was launched at the initial instant $t_0 = 200$ s, and final instant $t_f = 3800$ s with a time step $\Delta t = 10$ s. The particle tracing study was coupled with the solution from the time-dependent simulation done to solve the flow.

# Chapter 3

# Results and discussion

This chapter provides a comprehensive overview of the findings derived from the activities described in Chapter 2.

We begin by examining the distribution of fish in the tank, which allows us to identify preferential zones where likely more genetic material can be released.

Subsequently, we delve into the hydrodynamics of the tank, considering factors such as velocity flow and vorticity distribution. We explore how these dynamics influence the transport of eDNA particles within the tank, examining their trajectories over time or predicting how they can be spread in the simulated environment to study potential regions where eDNA can accumulate.

Finally, we conclude by comparing sampling data with numerical results. This comparison between preliminary experimental findings and computational simulations provides us with a first insight into the accuracy and reliability of the numerical model in predicting eDNA distribution patterns in the simulated marine environment.

## 3.1   Probability distribution of fish

The primary goal of the tank monitoring is to identify if any zones of the tank may be sources of eDNA. A thorough examination of the spatial coordinates reveals the patterns of probability density for the pelagic species, with high densities clustered near the tank outlet and intermediate densities in a second portion of the domain; while for the benthic species, two specific points have been identified as preferential positions. Identified mistakes, such as reflections and artifacts, are acknowledged and taken into account in further investigations. The quantification of probability density facilitates the determination of targeted particle release strategies for numerical simulations, ensuring a proportional representation of high- and low-density regions.
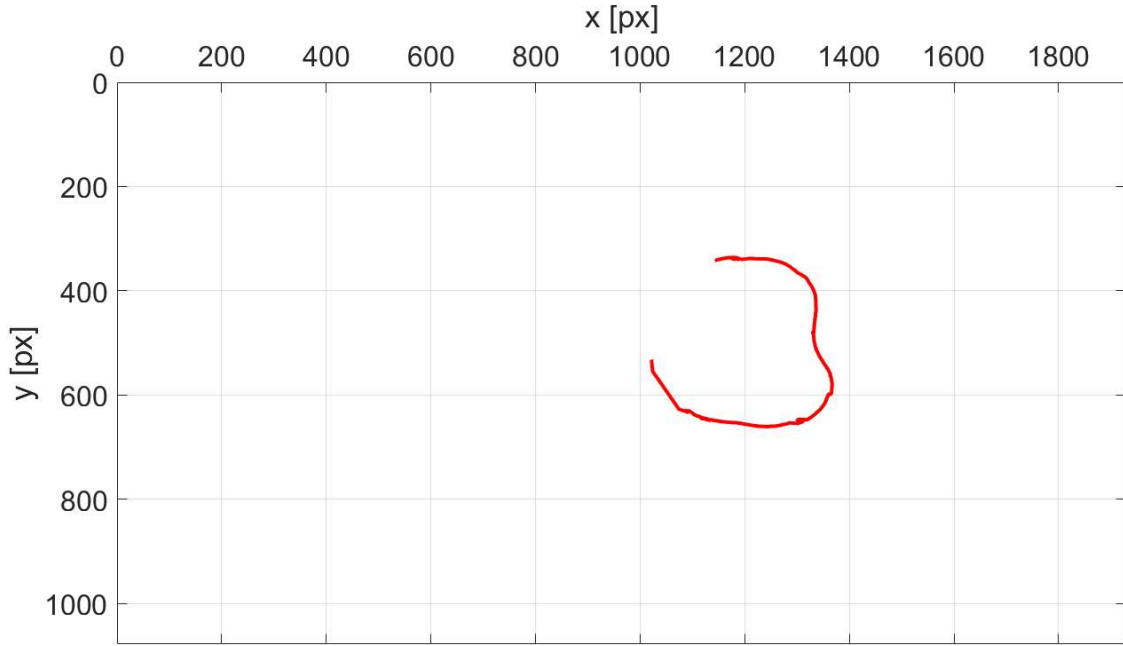
**Figure 3.1:** *Trajectory of fish 5 for t: t1<t<t3.*

### 3.1.1   Probability distribution of the pelagic species

The objective of pelagic species tracking is to identify preferential locations of the 5 individuals present in the tank, which we expect to release the most of the eDNA. For each individual fish that have been recognised, TRex exports .NPZ file, which contains the spatial information preset on 'export options' as explained in the previous Section 2.4.2. Using the Python code in Appendix, it was possible to extract position vectors in the plane (top view) for each recognised fish. There were a total of 9 fish in the tank, 7 of which were included in our case study. TRex exported a total of 30 .NPZ files. This discrepancy between the number of fish present in the tank and the number of fish recognised by TRex can be attributed to the following errors in tracking:

- When the software 'loses' a fish and newly detect it, it may not recognise the original fish it and assign it another number;

- When two or more fish come into contact or are very close they are sometimes identified as one fish and TRex momentarily assigns it a new number;

- The software has incorrectly identified reflections as real fish;

A solution to correct these errors consists to use fish `track_max_individuals` option and assign the maximum number of individuals to be tracked. This wasn't done because reflections were often mistaken as fish and many manual corrections would have to be made. In addition, considering we are interested in extracting preferential zones of spatial positions of fish, punctual and random errors can be tolerated.
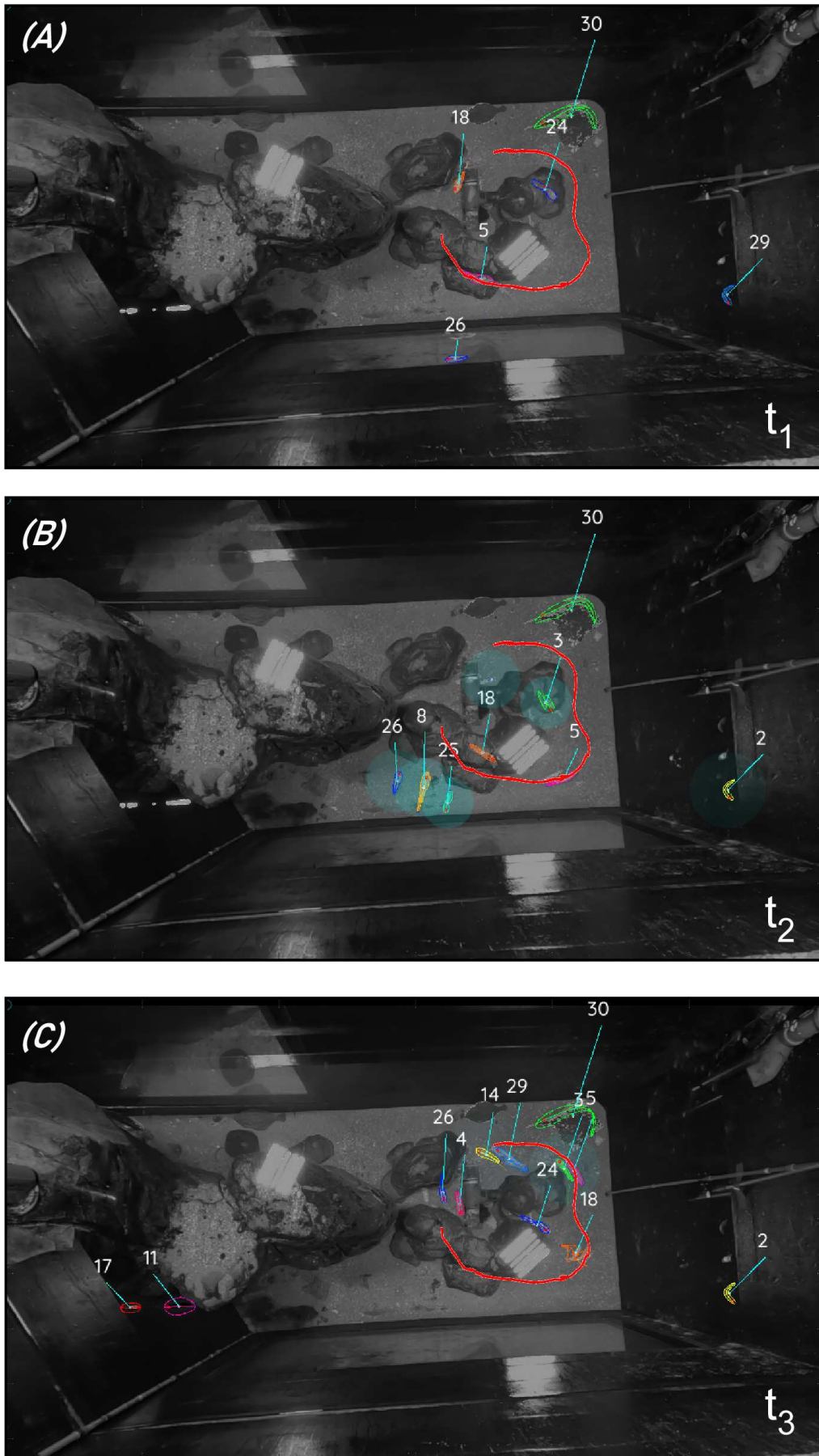
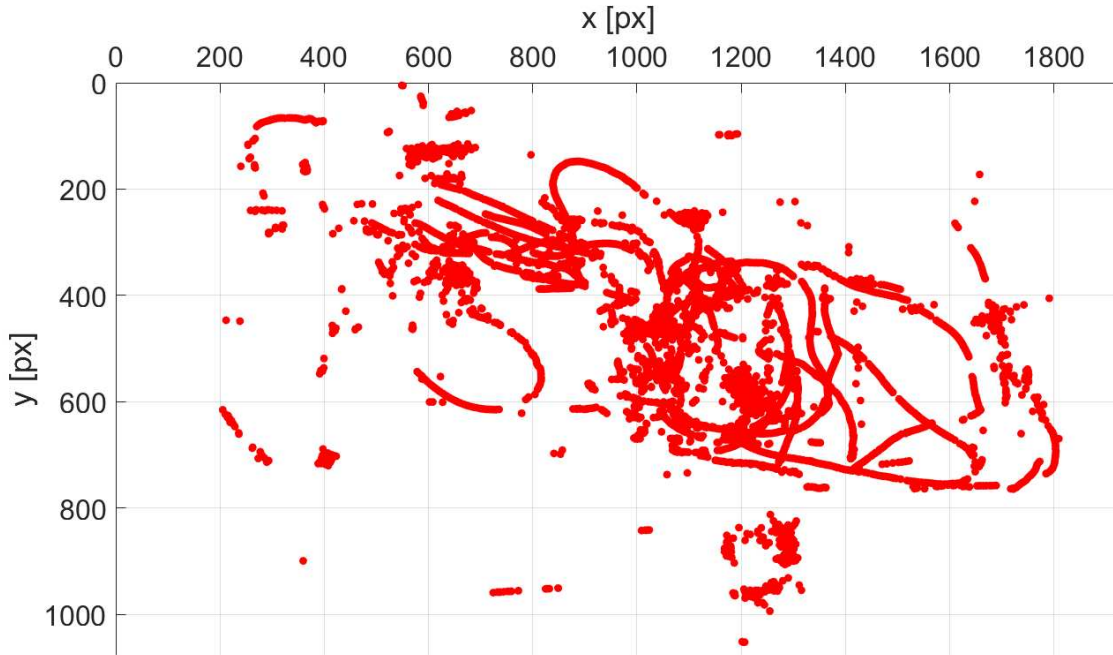**Figure 3.2:** *Overlay between trajectory and centroid of Fish 5 at $t_1$, $t_2$ and $t_3$.*

**Figure 3.3:** *Dot plot of the positions of fish 5 for t: 0<t<1 h.*

In Figure 2.7 one can see how we tried to correct the identity of the fish as much as possible by manual matching. The pink lines at the top of the frame bar represent the frames in which a manual match was applied. The positions and corresponding numbers associated by TRex to the 2 individuals of benthic species were identified in order to differentiate them in tracking from the pelagic ones. The breed in the tank was excluded by setting an appropriate blob size range.

Despite these limitations in the tracking analysis, which do not significantly impact the statistics of the fish positions in the tank (especially the pelagic species), we are able in some case to construct partial trajectories for some fish.

Figure 3.1 shows the trajectory of Fish 5 (pelagic fish) between time $t_1 = 217.3$ s and time $t_3 = 230.6$ s of the recorded video after the fishes release in the tank. The domain of this graph has been scaled to match the query window considered with TRex, allowing us to superimpose the graphs on the tank photo without further scaling. To show that the tracking works properly, this trajectory was assessed at three different times, namely $t_1$, $t_2$, $t_3$, with the corresponding frames on TRex ($t_2$=226.7 s). As can be seen in Figure 3.2 the red trajectory aligns perfectly with the centre of Fish 5 centroid.

We then proceeded to collect all the positional data in the xy plane (top view) relating to the numbers identified as belonging to the pelagic species fish. The files identified are those corresponding to the following numbers: [0, 3, 5, 8, 10, 13, 14, 16, 18, 24, 28]. An example of the spatial positions (X centroid and Y centroid) occupied throughout the hour by fish number 5 is depicted in the dotplot in Figure 3.3.
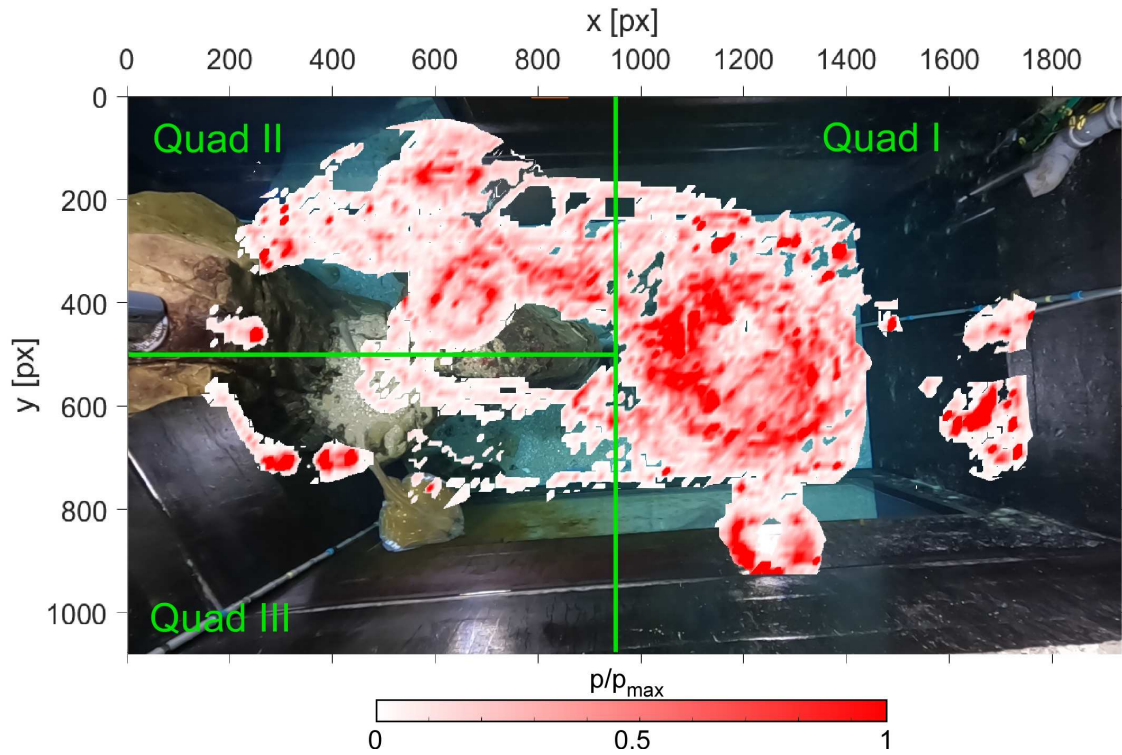
**Figure 3.4:** *Bidimensional density plot in the top view plane of the 5 fishes of pelagic species in the hour of recording. Partition of the domain into quadrants according to the integral of the probability density (green lines).*

The probability density plot was superimposed on the top view image of the tank (Figure 3.4), where one can recognise a zone near the outlet and a zone in the quadrant limited by sides a-d, where the probability to detect a fish is high and moderate, respectively. In contrast, the zone in the quadrant identified by sides c-d shows a low probability to have fishes, being few the detection in this area. The high-density spots near side b, c and d were identified as errors. The reflection near side b is due to the submerged GoPro, while those near sides c and d are given by reflections of lights mounted on the ceiling.

To simplify the numerical simulation of the eDNA transport we decided to divide the tank plan into three zones as shown in Figure 3.4 and we computed the integral of probability density within these zones. The proportion of the total detections between quadrants I and II was approximately 4:1, while quadrant III ha a relatively small number of fishes compared to the other zones, therefore we decided to exclude the presences of the fishes, and then the release of eDNA in this area. Thereby we will proceed to release the particles relating to the pelagic species in the zones, in the proportions 4:1 as to respect the values of integration of the density function (see matlab codes in Appendix). This operation should ensure a release of eDNA consistent with the video analysis, catching the fundamental information deriving from the tracking procedure.
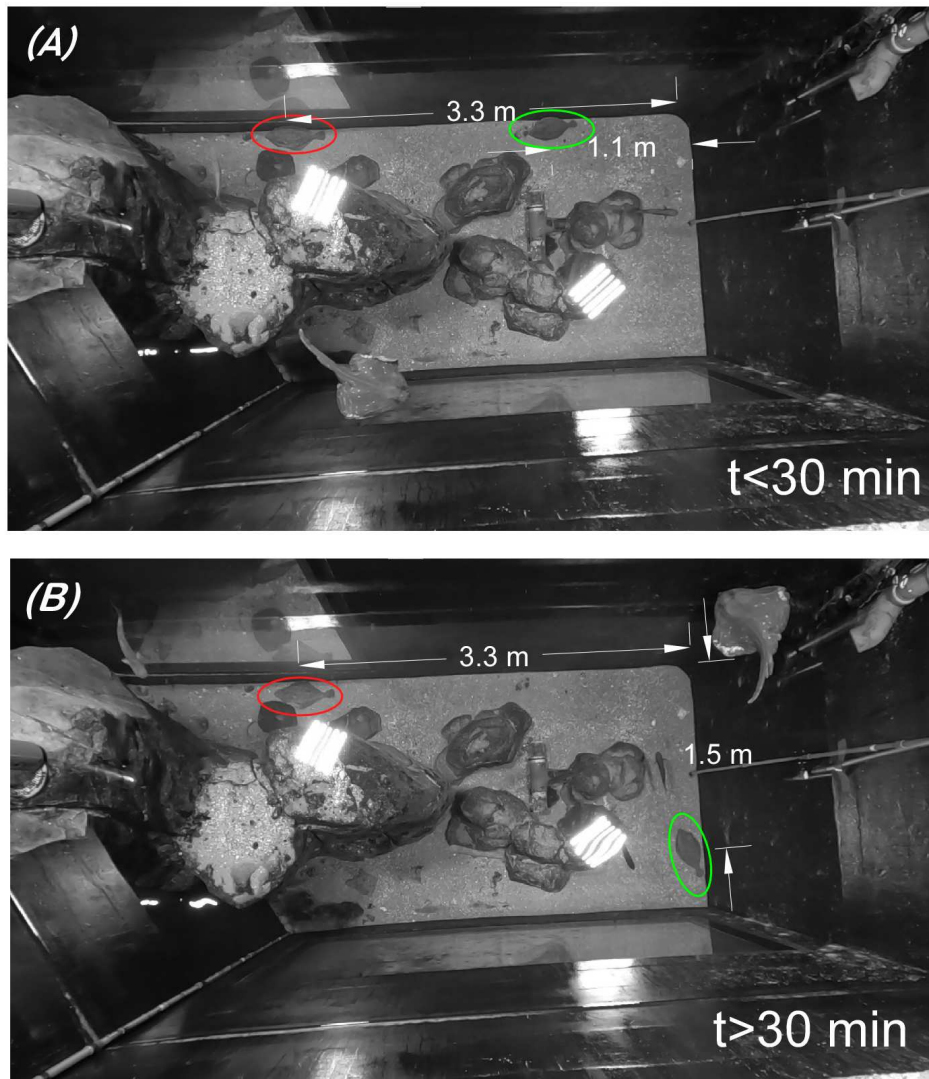
**Figure 3.5:** *Identification of preferential positions of the two flounders.*

### 3.1.2 Positions of the benthic species

The tracking of the two individuals of the benthic species in the tank during the hour of recording under consideration was very simple: one of the two flounders never moved from its initial position, the other changed position only once at minute 30 of recording (Figure 3.5). We thus assume a stationary source of eDNA in the corresponding positions of the centroidof the two benthic fishes (flounders).

## 3.2 Hydrodynamics of eDNA trace

Aquatic flow dynamics determine the dispersion of substances and materials in the aquatic environment. Understanding these dynamics therefore plays a major role when eDNA is used as a monitoring technique. For this purpose, velocity fields and

particle trajectories were analysed to elucidate the primary drivers of eDNA trace transport in the simulated environment.

### 3.2.1 Velocity field in the tank

Assessing the flow dynamics within the tank is fundamental for understanding the transport dynamics of eDNA. While the overall flow appears to be steady, the complex geometry of the tank generates periodic circulations with timescales of approximately 60 seconds, resulting in temporal fluctuations in velocity. These fluctuations may play a significant role in driving transport dynamics within the tank. Velocity magnitudes typically range in the order of millimiters to centimeters per second (Figure 3.6 and 3.7). Near the inlet, velocities are higher due to the jet inflow, whereas velocities near the tank bottom are lower, partly attributed to the presence of small rocks that slow down the flow and prevent the direct outflow through the outlet. Additionally, stagnant zones with velocities close to zero are observed in the tank's four corners at the bottom, suggesting minimal transport of the particles that are released in these areas. In contrast, in the regions closer to the free surface, velocities are higher, indicating that particles released in these zone can be easily moved away by the current.

Plotting the velocity vectors within the tank volume provides a clearer visualization of the vortical structures (Figure 3.8). A primary vortex is observed, affecting a significant portion of the tank, which is expected to ensure the mixing of water contained in the lower layers with that present in the upper layers.

Comparison of simulated velocity fields with experimental measurements, for example through ADV, could have been usefull to validate the numerical model and assess its accuracy in capturing flow behavior.

### 3.2.2 eDNA particle trajectories

The objective of this section is to gain a deeper understanding of the particle transport dynamics within the tank, aiming to identify accumulation zones and potential dead zones where particle transport is negligible.

From the particle release simulation conducted in COMSOL, positional and temporal data of all particles were exported. These data were processed using matlab (code provided in Appendix), and the spatial positions of the particles were plotted on the xy-plane (top view of the tank) and the xz-plane (side view of the tank) at two different time instances (t=1000 s and t=3800 s) (Figure 3.9).

Although the plotted graphs do not distinguish between particles assigned to the pelagic species and those assigned to the benthic species, the latter are easily identifiable. They exhibit by the trajectories close to the bottom that from the source point extends towards the outlet. This behavior can be attributed primarily to their release from the tank bottom, where they experience lower fluid velocities, making their trajectories more traceable. Additionally, being released in a position
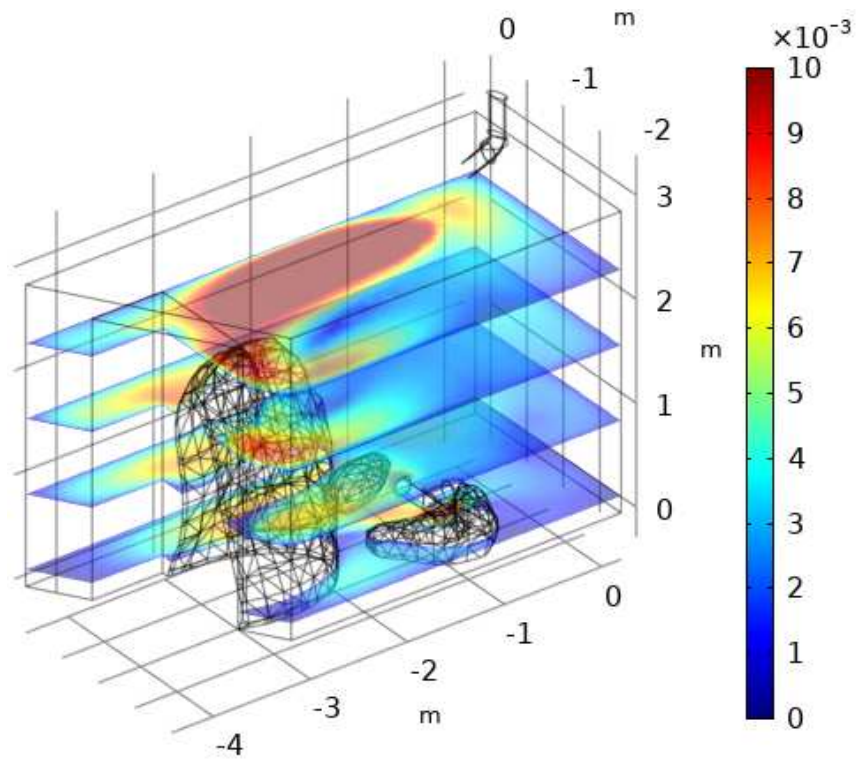
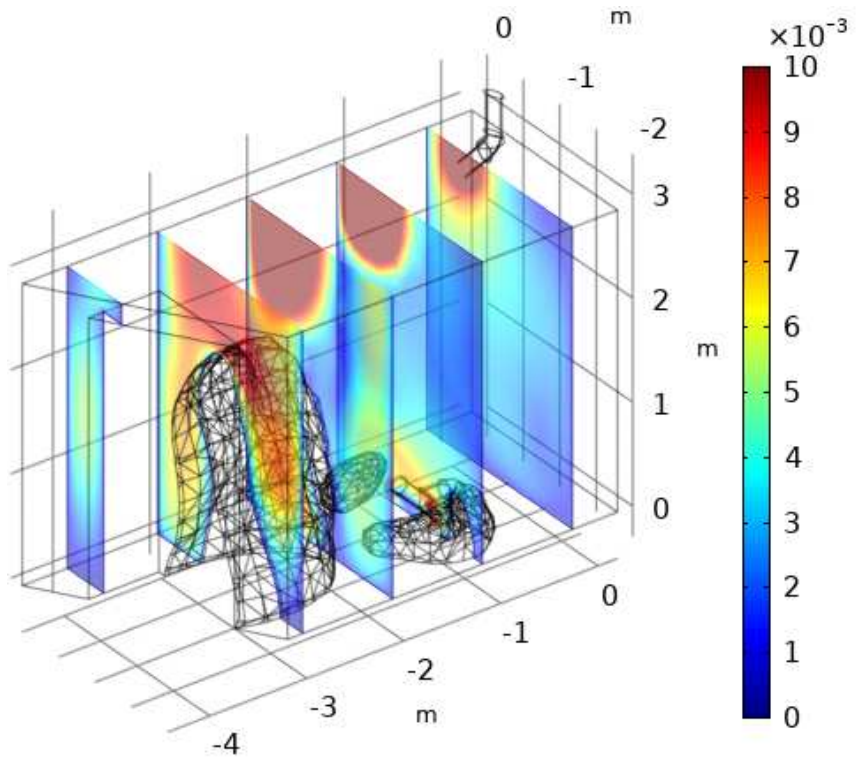**Figure 3.6:** *Velocity field in m/s along horizontal planes at t = 1800 s.*



**Figure 3.7:** *Velocity field in m/s along vertical planes at t = 1800 s.*
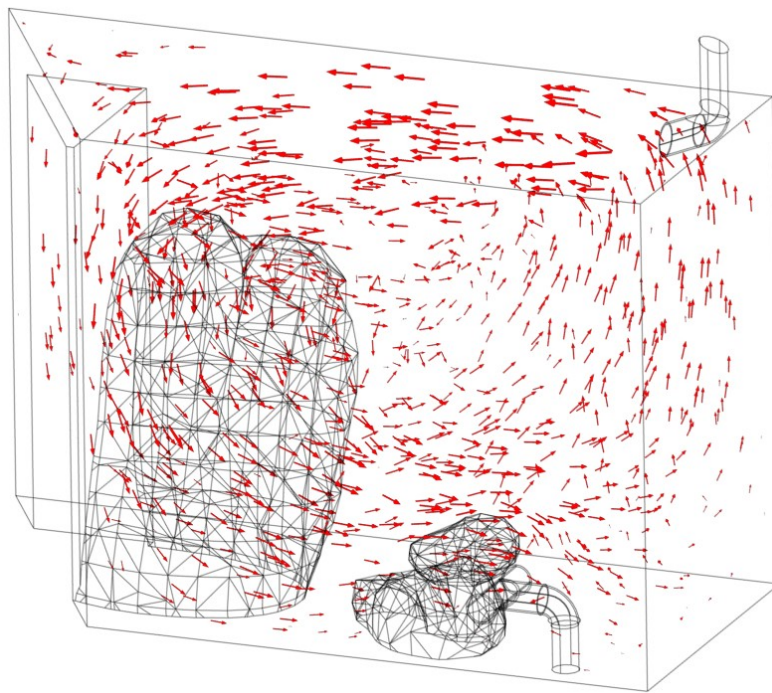
**Figure 3.8:** *Velocity vector field at t=1800 s. Arrow lengths are proportional to the velocity module on a logarithmic scale.*
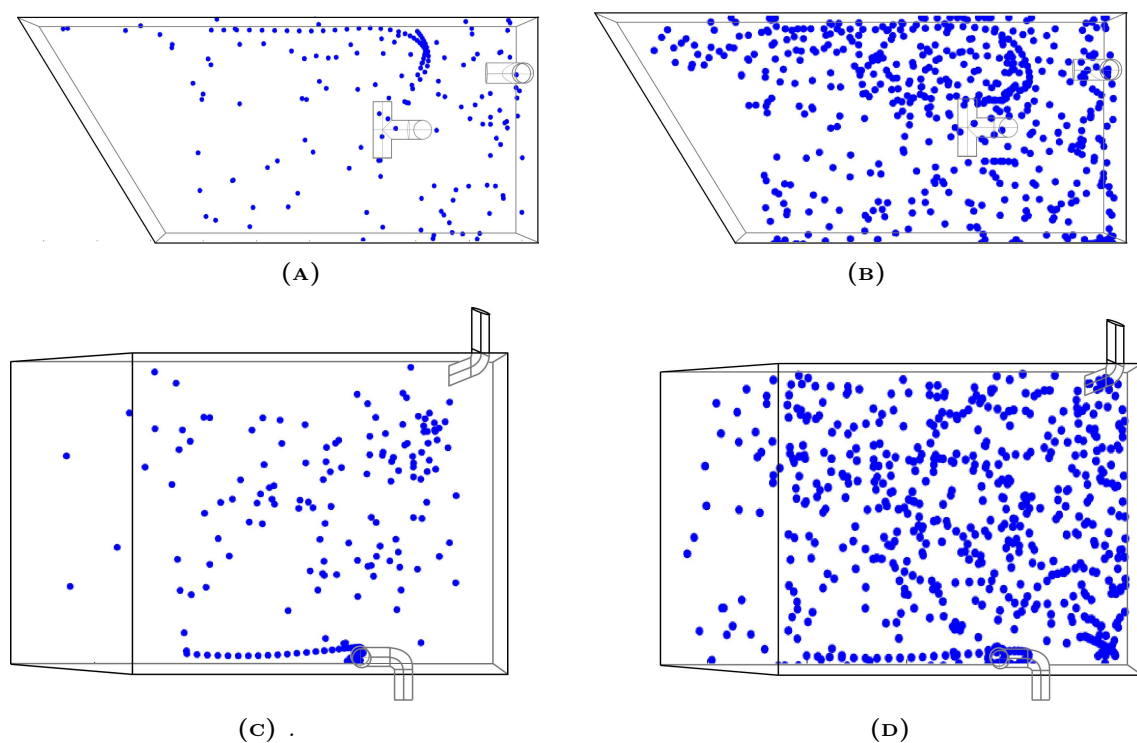


**Figure 3.9:** *Top view (A) and side view (C) at t=1000 s positions of the particles released. Top view (B) and side view (D) at t=3800 s.*

of the tank with negative vertical velocities, they are more susceptible to being drawn towards the outlet rather than mixing vertically.

A qualitative assessment of Figure 3.9B indicates a slightly higher particle density near side "a", while the area near the side "d" exhibits lower particle densities, suggesting reduces particle transport toward these areas (Figure 3.9D). Examining Figure 3.9D, the trajectories of benthic species simulated eDNA leading towards the outlet, an accumulation zone is evident on the bottom, behind the outlet and beneath the inlet.

From a quantitative standpoint, at the end of the simulation, a total of 111 particles exited through the outlet, out of the initial 847 particles released (one per fish every 30 s for the 3600 s simulation period). Consequently, 13% of the particles exited before the simulation's completion. Among these, most belong to benthic species being eDNA easily transported toward the outlet.

Conducting an extended simulation would provide valuable insights into the temporal evolution of particle distribution within the tank. It would elucidate whether a plateau is eventually reached as both time and the number of particles within the domain increase. The Lagrangian model developed can be supportive in interpreting the results of the conducted sampling. Consistency between the numerical model and experimental data would validate the assumptions made in this study and give the opportunities to implement an numerical to multi-scenario analysis.

In Figure 3.10, the trajectories of all released particles are depicted. The colors of the trajectories correspond to the release times: shades closer to blue indicate older particles released, while closer to red indicate particles released more recently.

In addition to analyzing the distribution density of the particles, it is also useful to assess their age in relation to their spatial distribution. Correlating residence times with spatial distribution allows for extracting additional and meaningful information, as:

(i) The areas where we find a majority of new eDNA particles are regions of low persistence. This indicates that by sampling in these areas, we are likely to find fresh DNA, which can be useful for extracting information about the current status of one or more species. For example, stress condition of the animals is directly related to eDNA release, and monitoring this zone offers the opportunity to have information on the status of the fishes.

(ii) The areas where we find a majority of old particles are stagnant zones where either fish have not passed or have passed for an insufficient time to release enough eDNA. The eDNA is transported by advection in these areas and tends to accumulate there. This means that by sampling in these areas, we will find older DNA, which can provide insights into the species dating back in time.

(iii) Examining the so-called "dead zones," where no eDNA transport is detected, is still significant. The presence of sampled DNA in these areas suggests that eDNA extracted here cannot be attributed to fishes that occupy other zone of the water body.
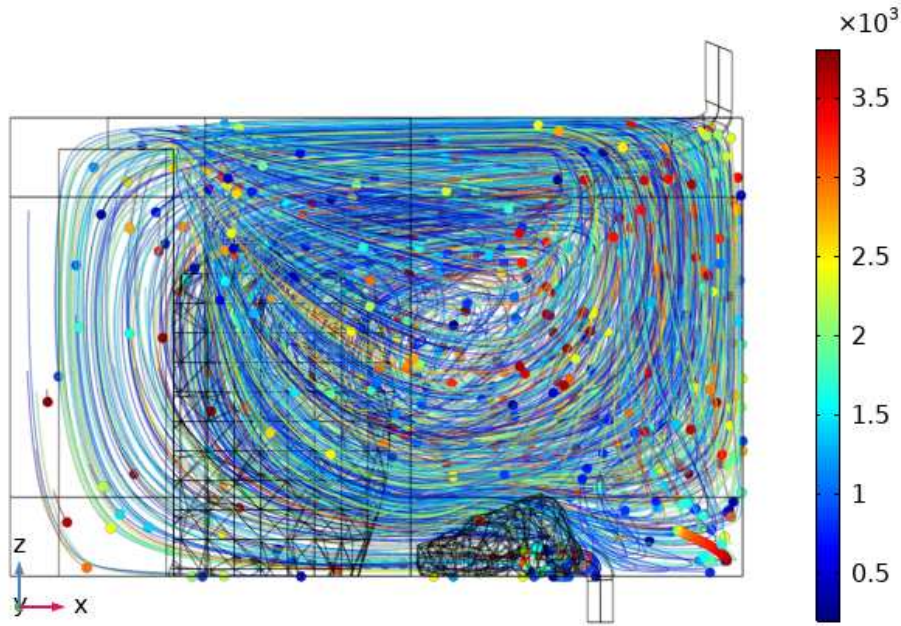
**Figure 3.10:** *Trajectories of all the particle released.*

By examining the probability density plot in Figure 3.11 and comparing it with the probability density plot derived from the tracking data (Figure 3.4), a discernible shift is observed in the denser cluster of points towards the 'd' side of the tank., i.e. where the velocities are highest at the surface (Figure 3.7). This suggests that the observed accumulation is likely situated at shallow depths and is induced by downdrafts.

## 3.3 eDNA maps: comparison between sampling data and numerical results

Due to time limitations, the analysis is based on a partial sample of data collected during the 96-hour sampling period, focusing exclusively on the pelagic species. Consequently, it is not possible to establish a direct correlation between fish tracking results and these concentration data. However, by relying on the CFD simulation results, it is possible to make some meaningful considerations and draw conclusions.

In Table we report 3.1 the eDNA concentration data for the different sampling positions. The position to which they refer is defined by the letter 'P' followed by a number indicating the position on the xy plane (top view) and a letter (U, M, L, standing for upper, middle, lower) indicating the position on the vertical, as explained in Chapter 2.

By this preliminary analysis some concentration anomalies emerge, particularly at the P1M position and the P4L position.

| Sample | Concentration [$\mu$g/L] |
|:------:|:------------------------:|
| P1U    | 5.5                      |
| P1M    | 26.2                     |
| P1L    | 3.8                      |
| P2U    | 3.8                      |
| P2M    | 4.7                      |
| P2L    | 4.3                      |
| P3U    | 3.6                      |
| P3M    | 6.6                      |
| P4U    | 4                        |
| P4M    | 4.4                      |
| P4L    | 15.7                     |
| P5U    | 3.9                      |
| P5M    | 3.2                      |
| P5L    | 4.9                      |

**Table 3.1:** *Values of concentrations measured at t=96 h for fish of the pelagic species at the sampling locations (Figure 2.3)*
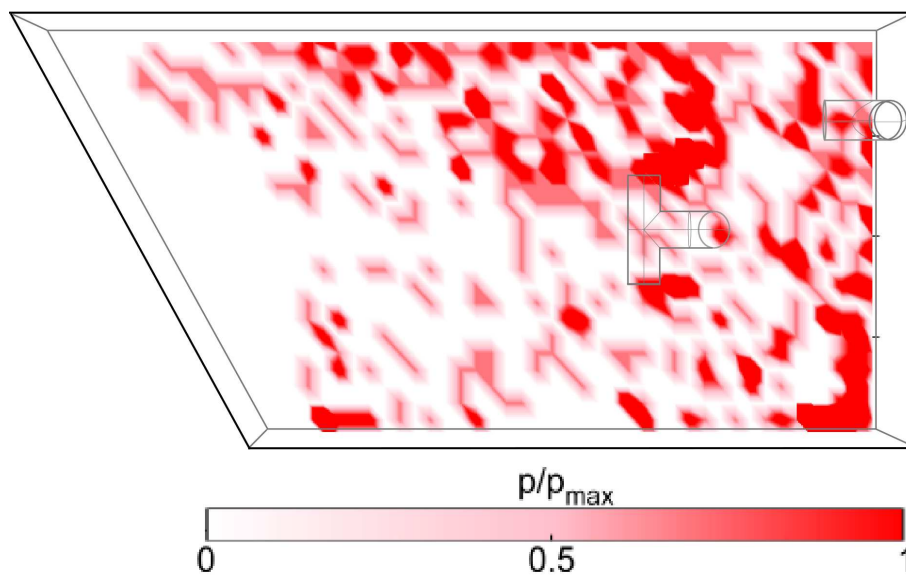


**Figure 3.11:** *Probability density function of the particles in the hydrodynamic model after 1 hour.*

The P1M zone was identified by the hydrodynamic model as a 'dead zone', characterised by the absence of water exchange preventing eDNA transport. While the condition of zero-particle estimated by the numerical model and the high concentration of eDNA seems apparently contradictory, these results actually confirm indirectly the validity of the numerical analysis, and suggest that eDNA collected belongs only to the fishes that swam in this wake zone during the 96 hours. Accordingly, in this zone, any genetic material released by a fish tends to accumulate without dispersal. Similarly, the P4L zone, also identified as a 'dead zone' by the

model, shows an anomalous datum concentrated at the bottom. This phenomenon could be attributed to downdrafts or particle sedimentation. In particular, sedimentation appears to be a plausible explanation, considering the period of time elapsed, during which particle deposition may have occurred on the bottom.

These few data partially confirm what we found in the hydrodynamic model and are therefore a good encouragement to continue and improve the present study.

# Conclusions and outlooks

In the present study, video fish monitoring techniques, Computational Fluid Dynamics techniques and water sampling techniques were integrated to attempt to define eDNA transport dynamics of two species of fishes contained within a tank. Initial results showed that this technique is promising as it allows one to identify both the persistence of fish within the tank and the dynamics of small particles modelled according a Lagrangian approach observations emerge from the results.

First, we identified distinct patterns in the distribution of pelagic and benthic fish species. Pelagic species showed higher densities near the outlet of the tank, while specific preferred locations were identified for benthic species.

Secondly, the analysis of hydrodynamic dynamics revealed, despite the steady inflow, the presence of periodic circulations that influence eDNA dispersal. Areas of accumulation near the outlet of the reservoir and areas of minimal transport near the corners of the reservoir were indicated by the numerical simulation. These observations allowed us to make predictions about the transport of particles such as eDNA that we were able to partially confirm with the eDNA extracted from the water samples taken. The latter correlation is very limited, as the data extracted from the tracking and hydrodynamic simulation refer to different period of time (1 hour vs 4 days). Dead zones', characterised by a lack of water exchange and particle sedimentation, emerged as critical points of eDNA accumulation.

These initial perspectives provide an initial overview, when monitoring animal species in natural habitats, of the possible results to be expected and their interpretation. Areas in which a very high concentration of eDNA is found could mistakenly be interpreted as areas rich in biodiversity; they can merely be areas of accumulation. At the same time, areas where there is a low concentration could be characterized by strong mixing and thus lead us to underestimate the biodiversity.

Effective monitoring techniques in the aquatic environment based on the eDNA approach cannot do without a comprehensive understanding of these dynamics.

In this experimental work, a methodology was also developed, reproducible in a controlled environment, which integrates spatial tracking of fish, CFD with water sampling and all necessary operations for its storage in order to subsequently extract eDNA concentrations in the laboratory.

Further research is encouraged to refine and validate the model, explore the temporal evolution of eDNA transport in the aquatic environment and thus improve understanding of the complex dynamics of the aquatic ecosystem.

# Appendix

## 3.4 Python codes

**Analysis of .NPZ files**

```python
import numpy as np
import os
import matplotlib.pyplot as plt

# Define the directory containing the .npz files
directory = '/Users/GMorganti/Videos/data_TOP_VIDEO2/'

# List of .npz files in the directory
files = [file for file in os.listdir(directory)
        if file.endswith('.npz')]

# Iterate over all .npz files in the directory
for file in files:
    # Load the .npz file
    data = np.load(os.path.join(directory, file))

    # Get the names of arrays inside the file
    array_names = data.files

    # Print the names of arrays
    for name in array_names:
        print(name)

    # Save the y coordinates in an array
    y = data["Y"]

    # Plot variables
    time = data["time"]
    x = data["X"]
    frame = data["frame"]
```

```python
# Create the dot plot
plt.plot(x, y, 'o')

# Add labels to the axes and a title
plt.xlabel('x')
plt.ylabel('y')
plt.title('')

# Show the plot
plt.show()
```

## 3.5   Matlab codes

**Bidimensional probability density plot**

```matlab
clc
clear all
close all

x = load('x_top.txt');
y = load('y_top.txt');

% bins for the histogram
edges_x = linspace(min(x), max(x), 200); % 100 bins
edges_y = linspace(min(y), max(y), 100); % 100 bins

% Compute the histogram
density = histcounts2(x, y, edges_x, edges_y);

% Add a row and a column of zeros to match the dimensions
% of density with edges_x and edges_y
density_padded = padarray(density, [1, 1], 0, 'pre');

% Custom colormap
num_colors = 256; % Number of colors in the colormap
white_color = [1 1 1]; % White color
red_color = [1 0 0]; % Red color
custom_colormap = [...
linspace(white_color(1), red_color(1), num_colors)', ...
linspace(white_color(2), red_color(2), num_colors)', ...
linspace(white_color(3), red_color(3), num_colors)'...
];
```

```matlab
figure;
pcolor(edges_x, edges_y, density_padded');
shading interp; % To interpolate colors
colormap(custom_colormap); % Use the custom colormap
%colormap('hot');
colorbar;
xlabel('x');
ylabel('y');
title('Probability_Density_Function_of_Pelagic_species');
caxis([min(density_padded(:)),...
max(density_padded(:))]*0.01);
% Mirror the x-axis
set(gca, 'YDir', 'reverse');
% Move the x-axis to the top of the plot
set(gca, 'XAxisLocation', 'top');
% Set the tick marks on the x-axis
set(gca, 'TickDir', 'out');
box on

% Probability density
[f_x, xi_x] = ksdensity(x);
[f_y, yi_y] = ksdensity(y);

% Bidimensional grids for xi_x and yi_y
[X, Y] = meshgrid(xi_x, yi_y);

pdf_2d = f_x'*f_y;

figure;
contour(X, Y, pdf_2d);
xlabel('x');
ylabel('y');
title('Probability_Density_Function_of_Pelagic_species');

% Mirror the x-axis
set(gca, 'YDir', 'reverse');
% Move the x-axis to the top of the plot
set(gca, 'XAxisLocation', 'top');

% Find the indices of x and y values
idx_x = find(edges_x >= 0 & edges_x <= 15);
idx_y = find(edges_y >= 0 & edges_y <= 17);
```

```matlab
integral_value = sum(sum(density_padded(idx_x, idx_y)));
area_rectangles = ...
(edges_x(2) - edges_x(1)) * (edges_y(2) - edges_y(1));

disp(['Value_int_top_left:', num2str(integral_value)]);


idx_x = find(edges_x >= 0 & edges_x <= 15);
idx_y = find(edges_y >= 8 & edges_y <= 17);

integral_value = sum(sum(density_padded(idx_x, idx_y)));
area_rectangles = ...
(edges_x(2) - edges_x(1)) * (edges_y(2) - edges_y(1));

disp(['Value_int_bottom_left:', num2str(integral_value)]);


% Find the indices of x and y values
idx_x = find(edges_x >= 15 & edges_x <= 25);
idx_y = find(edges_y >= 0 & edges_y <= 11);

integral_value = sum(sum(density_padded(idx_x, idx_y)));
area_rectangles = ...
(edges_x(2) - edges_x(1)) * (edges_y(2) - edges_y(1));

disp(['Value_integral_tot:_', num2str(integral_value)]);
```

**Trajectories of fish 5 plot**

```matlab
clc
clear all
close all

% Load the data
x = load('x5.txt');
y = load('y5.txt');

% Remove inf and NaN
x_cleaned = x(~isinf(x) & ~isnan(x));
y_cleaned = y(~isinf(y) & ~isnan(y));

% Conversion from default measures for
```

```matlab
% the interrogation area of TRex to pixel
x_cleaned = 64 * x_cleaned;
y_cleaned = 64 * y_cleaned;

% Plot the trajectory
plot(x_cleaned, y_cleaned, 'r', 'LineWidth', 2);
xlabel('x [px]');
ylabel('y [px]');
title('Trajectory of fish 5 for t_{1} < t < t_{3}');
xlim([0 1929]);
ylim([0 1078]);
grid on;

% Mirror the y-axis
set(gca, 'YDir', 'reverse');
% Move the x-axis to the top of the plot
set(gca, 'XAxisLocation', 'top');
% Set the font size of the axes
fontSize = 16; % Desired font size
set(gca, 'FontSize', fontSize);

% Set the tick marks on the axis to be outward
set(gca, 'TickDir', 'out');
set(gca, 'YTickDir', 'out');

% Load the data
x = load('x5_tot.txt');
y = load('y5_tot.txt');

x = 64 * x;
y = 64 * y;

% Scatter plot
figure;
scatter(x, y, 20, 'filled', 'MarkerFaceColor', 'r');
xlabel('x [px]');
ylabel('y [px]');
title('Positions of fish 5 for 0 < t < 1 h');
xlim([0 1929]);
ylim([0 1078]);
grid on;

% Mirror the y-axis
```

```matlab
set(gca, 'YDir', 'reverse');
% Move the x-axis to the top of the plot
set(gca, 'XAxisLocation', 'top');
% Set the font size of the axes
set(gca, 'FontSize', fontSize);
```

**Particle tracing analysis**

```matlab
    clc
clear all
close all

% File path
file_path = 'C:\Users\GMorganti\particletracing_greta.xlsx';

% Import data from Excel file
data = xlsread(file_path);

% Find unique indices of particles
unique_particles = unique(data(:, 4));

% Initialize a cell to store particle data
particle_data = cell(length(unique_particles), 1);

% Loop to extract particle data
for i = 1:length(unique_particles)
    % Find rows corresponding to the current particle
    idx = data(:, 4) == unique_particles(i);

    % Extract corresponding particle data
    particle_data{i} = data(idx, :);
end

% Initialize vectors for final positions
x_final = [];
y_final = [];
z_final = [];

% Loop to extract final positions of particles
for i = 1:length(particle_data)
    % Extract last x value
    last_x = particle_data{i}(end, 1);
    x_final = [x_final; last_x];
```

```matlab
    % Extract last y value
    last_y = particle_data{i}(end, 2);
    y_final = [y_final; last_y];

    % Extract last z value
    last_z = particle_data{i}(end, 3);
    z_final = [z_final; last_z];
end

% Scatter plot on the xy plane
figure;
scatter(x_final, y_final, 15, 'b', 'filled');
xlabel('X');
ylabel('Y');
title('Final_positions_on_xy_of_the_particles_released');
xlim([-4.8, 0]);
ylim([-2, 0]);
set(gca, 'Color', 'none');

% Scatter plot on the xz plane
figure;
scatter(x_final, z_final, 15, 'b', 'filled');
xlabel('X');
ylabel('Z');
title('Final_positions_on_xz_of_the_particles_released');
set(gca, 'Color', 'none');

% Scatter plot on the yz plane
figure;
scatter(y_final, z_final, 15, 'b', 'filled');
xlabel('y');
ylabel('Z');
title('Final_positions_on_yz_of_the_particles_released');
set(gca, 'Color', 'none');

% Initialize vectors for positions at time t=1000 seconds
x_final_t1000 = [];
y_final_t1000 = [];
z_final_t1000 = [];

% Loop to extract positions at time t=1000 seconds
for i = 1:length(particle_data)
```

```matlab
    % Find index of time t=1000 seconds in the time column
    idx_t1000 = find(particle_data{i}(:, 5) == 1000);

% Check if time t=1000 seconds is present in the time column
if ~isempty(idx_t1000)
    % Extract corresponding positions
    x_final_t1000 = [x_final_t1000; particle_data{i}
                        (idx_t1000, 1)];
    y_final_t1000 = [y_final_t1000; particle_data{i}
                        (idx_t1000, 2)];
    z_final_t1000 = [z_final_t1000; particle_data{i}
                        (idx_t1000, 3)];
end
end

% Scatter plot on the xy plane at time t=1000 seconds
figure;
scatter(x_final_t1000, y_final_t1000, 15, 'b', 'filled');
xlabel('X');
ylabel('Y');
xlim([-4.8, 0]);
ylim([-2, 0]);
title('Scatter Plot on the XY Plane at Time t=1000 seconds');

% Scatter plot on the xz plane at time t=1000 seconds
figure;
scatter(x_final_t1000, z_final_t1000, 15, 'b', 'filled');
xlabel('X');
ylabel('Y');
title('Scatter Plot on the XZ Plane at Time t=1000 seconds');

x = x_final;
y = y_final;
z = z_final;

% Find indices of NaN elements in vectors x, y, and z
nan_indices = isnan(x) | isnan(y) | isnan(z);

% Remove NaN elements from vectors x, y, and z
x_clean = x(~nan_indices);
y_clean = y(~nan_indices);
z_clean = z(~nan_indices);
```

```matlab
% Histogram bins
edges_x = linspace(min(x_clean), max(x_clean), 50); %
edges_y = linspace(min(y_clean), max(y_clean), 25); %

% Compute histogram
density = histcounts2(x_clean, y_clean, edges_x, edges_y);

% Add a row and column of zeros to match the
% dimensions of density with edges_x and edges_y
density_padded = padarray(density, [1, 1], 0, 'pre');

% Custom colormap
num_colors = 256; % Number of colors in the colormap
white_color = [1 1 1]; % White color
red_color = [1 0 0]; % Red color
custom_colormap = [...
linspace(white_color(1), red_color(1), num_colors)', ...
linspace(white_color(2), red_color(2), num_colors)', ...
linspace(white_color(3), red_color(3), num_colors)'...
];

% Normalize data for correct color visualization
normalized_density = density_padded/max(density_padded(:));

figure;
pcolor(edges_x, edges_y, normalized_density');
shading interp; % Interpolate colors
colormap(custom_colormap); % Use custom colormap
colorbar;
xlabel('x');
ylabel('y');
title('PDF_of_pelagic_and_benthic_species_on_xy');
caxis([min(normalized_density(:)),...
max(normalized_density(:))*0.018]);
% Move x-axis to the top of the plot
set(gca, 'XAxisLocation', 'top');
% Set tick marks on the x-axis
set(gca, 'TickDir', 'out');
box on

% Histogram bins
edges_x = linspace(min(x_clean), max(x_clean), 50);
```

```matlab
edges_z = linspace(min(z_clean), max(z_clean), 25);

% Compute histogram
density = histcounts2(x_clean, z_clean, edges_x, edges_z);

% Add a row and column of zeros to match the...
% dimensions of density with edges_x and edges_z
density_padded = padarray(density, [1, 1], 0, 'pre');

% Custom colormap
num_colors = 256; % Number of colors in the colormap
white_color = [1 1 1]; % White color
pink_color = [1 0.75 0.8]; % Light pink color
red_color = [1 0 0]; % Red color
custom_colormap = [linspace(white_color(1),
```

# Bibliography

[1] J. Smith, "Global biodiversity decline," *Nature*, vol. 500, no. 7464, pp. 20–25, 2020.

[2] A. Ogram, G. S. Sayler, and T. Barkay, "The extraction and purification of microbial dna from sediments," *Journal of Microbiological Methods*, vol. 7, no. 2, pp. 57–66, 1987. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016770128790025X

[3] G. J. Olsen, D. J. Lane, S. J. Giovannoni, N. R. Pace, and D. A. Stahl, "Microbial ecology and evolution: A ribosomal rna approach," *Annual Review of Microbiology*, vol. 40, no. 1, pp. 337–365, 1986, pMID: 2430518. [Online]. Available: https://doi.org/10.1146/annurev.mi.40.100186.002005

[4] M. Ariza, B. Fouks, Q. Mauvisseau, R. Halvorsen, I. G. Alsos, and H. J. de Boer, "Plant biodiversity assessment through soil edna reflects temporal and local diversity," *Methods in Ecology and Evolution*, vol. 14, no. 2, pp. 415–430, 2023. [Online]. Available: https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13865

[5] G. F. Ficetola, J. Pansu, A. Bonin, E. Coissac, C. Giguet-Covex, M. De Barba, L. Gielly, C. M. Lopes, F. Boyer, F. Pompanon, G. Rayé, and P. Taberlet, "Replication levels, false presences and the estimation of the presence/absence from edna metabarcoding data," *Molecular Ecology Resources*, vol. 15, no. 3, pp. 543–556, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/1755-0998.12338

[6] K. Deiner, H. M. Bik, E. Mächler, M. Seymour, A. Lacoursière-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, and L. Bernatchez, "Environmental dna metabarcoding: Transforming how we survey animal and plant communities," *Molecular Ecology*, vol. 26, no. 21, pp. 5872–5895, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.14350

[7] P. F. Thomsen, J. Kielgast, L. L. Iversen, P. R. Møller, M. Rasmussen, and E. Willerslev, "Detection of a diverse marine fish fauna using environmental dna from seawater samples," *PLOS ONE*, vol. 7, no. 8, pp. 1–9, 08 2012. [Online]. Available: https://doi.org/10.1371/journal.pone.0041732

[8] M. Barnes and C. Turner, "The ecology of environmental dna and implications for conservation genetics," *Conservation Genetics*, vol. 17, 09 2015.

[9] C. Goldberg, K. Strickler, and D. Pilliod, "Moving environmental dna methods from concept to practice for monitoring aquatic macroorganisms," *Biological Conservation*, vol. 183, pp. 1–3, 2015.

[10] P. Taberlet, A. Bonin, L. Zinger, and E. Coissac, *Environmental DNA: For biodiversity research and monitoring*. Oxford University Press, 2018.

[11] P. F. Thomsen and E. Willerslev, "Environmental dna – an emerging tool in conservation for monitoring past and present biodiversity," *Biological Conservation*, vol. 183, pp. 4–18, 2015.

[12] H. C. Rees, K. Bishop, D. J. Middleditch, J. R. M. Patmore, B. C. Maddison, and K. C. Gough, "The application of edna for monitoring of the great crested newt in the uk," *Ecology and Evolution*, vol. 4, no. 21, pp. 4023–4032, 2014. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.1272

[13] S. Tsuji, T. Takahara, H. Doi, N. Shibata, and H. Yamanaka, "The detection of aquatic macroorganisms using environmental dna analysis—a review of methods for collection, extraction, and detection," *Environmental DNA*, vol. 1, no. 2, pp. 99–108, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/edn3.21

[14] A. J. Shogren and J. L. Tank, "Understanding and managing environmental dna dynamics in aqueous ecosystems," *Environmental DNA*, vol. 1, no. 3, pp. 141–156, 2018.

[15] S. Matsuoka, Y. Sugiyama, M. Nagano, and H. Doi, "Influence of dna extraction kits on freshwater fungal dna metabarcoding," *bioRxiv*, 2022. [Online]. Available: https://www.biorxiv.org/content/early/2022/04/18/2021.10.19.464972

[16] K. Demertzis and L. Iliadis, "Detecting invasive species with a bio-inspired semi-supervised neurocomputing approach: the case of *Lagocephalus sceleratus*," *Neural Computing and Applications*, vol. 28, no. 6, pp. 1225–1234, 2017. [Online]. Available: https://doi.org/10.1007/s00521-016-2591-2

[17] D. Wang, J. Liu, W. Shao, C. Mei, X. Su, and H. Wang, "Comparison of cmip5 and cmip6 multi-model ensemble for precipitation downscaling results and observational data: The case of hanjiang river basin," *Atmosphere*, vol. 12, no. 7, 2021. [Online]. Available: https://www.mdpi.com/2073-4433/12/7/867

[18] A. O. Shelton, J. P. Vetter, J. Novembre, and M. R. Stephens, "Environmental dna provides quantitative estimates of a threatened salmon species," *Conservation Letters*, vol. 12, no. 5, p. e12635, 2019.

[19] R. H. Griffin, K. D. Lafferty, K. D. Feller, and R. R. Gentry, "Modelling environmental dna data; bayesian variable selection accounting for false positive and false negative errors," *Methods in Ecology and Evolution*, vol. 11, no. 2, pp. 267–279, 2020.

[20] S. P. Matter and D. W. Schirokauer, "A rapid assessment method to estimate the distribution of juvenile chinook salmon (*Oncorhynchus tshawytscha*) in an interior alaska river basin," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 73, no. 11, pp. 1697–1707, 2016.

[21] H. Sakai, M. Inoue, A. Harada, Y. Kaneko, and A. Shoji, "Discovery of an unrecorded population of yamato salamander (*Hynobius vandenburghi*) by gis and edna analysis," *Conservation Genetics*, vol. 20, no. 5, pp. 1019–1031, 2019.

[22] B. Thalinger, M. Hofer, and M. Traugott, "Monitoring spawning migrations of potamodromous fish species via edna," *Freshwater Biology*, vol. 64, no. 10, pp. 1762–1772, 2019.

[23] T. P. Muha, C. R. Turner, R. B. Chandler, B. L. McGlynn, J. P. Wares, and S. B. Munch, "Using environmental dna to improve species distribution models for freshwater invaders," *Ecological Applications*, vol. 27, no. 1, pp. 313–325, 2017.

[24] L. Gallien, R. Douzet, S. Pratte, N. E. Zimmermann, and W. Thuiller, "Predicting potential distributions of invasive species: where to go from here?" *Diversity and Distributions*, vol. 16, no. 3, pp. 331–342, 2010.

[25] A. Sahu, N. Kumar, C. Pal Singh, and M. Singh, "Environmental dna (edna): Powerful technique for biodiversity conservation," *Journal for Nature Conservation*, vol. 71, p. 126325, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1617138122001984

[26] X. Liu, Y. Pei, X. Zhou, F. Liu, S. Lai, J. Xiong, Z. Zhong, M. Li, Y. Chen, X. Guo *et al.*, "Combining electrostatic enrichment with pcr-free nanopore-based metagenomic sequencing for rapid environmental bacterial monitoring," *Environmental Science & Technology*, vol. 53, no. 17, pp. 10 337–10 340, 2019.

[27] C. R. Turner, K. L. Uy, and R. C. Everhart, "Fish environmental dna is more concentrated in aquatic sediments than surface water," *Biological Conservation*, vol. 183, pp. 93–102, 2015, special Issue: Environmental DNA: A powerful new tool for biological conservation.

[28] J. J. Piggott and C. R. Primmer, "Evolutionary consequences of dna methylation in a parasite," *Biology Letters*, vol. 11, no. 1, p. 20141085, 2015.

[29] A. J. Shogren, J. L. Tank, E. Andruszkiewicz, B. Olds, A. R. Mahon, C. L. Jerde, and D. Bolster, "Controls on edna movement in streams: Transport, retention, and resuspension," *Journal of Environmental Management*, vol. 264, p. 110451, 2020.

[30] T. M. Wilcox, K. S. McKelvey, M. K. Young, W. H. Lowe, and M. K. Schwartz, "A novel approach to environmental dna metabarcoding: edna observation and capture of aquatic macroinvertebrates using light," *Journal of Fisheries and Aquatic Science*, vol. 11, no. 4, pp. 701–710, 2016.

[31] T. Walter and I. D. Couzin, "Trex, a fast multi-animal tracking system with markerless identification, and 2d estimation of posture and visual fields," *eLife*, vol. 10, p. e64000, feb 2021. [Online]. Available: https://doi.org/10.7554/eLife.64000

[32] AliceVision. (2024) AliceVision Photogrammetry Pipeline (Meshroom). [Online]. Available: https://alicevision.org/#meshroom

[33] COMSOL, "Particle tracing module user's guide," https://doc.comsol.com/5.4/doc/com.comsol.help.particle/ParticleTracingModuleUsersGuide.pdf, 2017.