

Eco-AlpsWater

Innovative Ecological Assessment and Water Management Strategy for the
Protection of Ecosystem Services in Alpine Lakes and Rivers

Priority 3: Liveable Alpine Space. SO3.2 - Enhance the protection, the
conservation and the ecological connectivity of Alpine Space

Project Eco-AlpsWater

Work Package WPT1

Activity A.T1.2

Deliverable D.T1.2.1

Version 2.0

Date June 2021

Coordination of the WP1: I Domaizon

Deliverable D.T1.2.1

Data storage in cloud storage services

Interreg Alpine Space - Eco-AlpsWater project – WP1

Giulia Riccioni¹, Adriano Boscaini¹, Nico Salmaso¹

¹ *Edmund Mach Foundation(FEM), San Michele all'Adige, Trento, Italy*

ABSTRACT

The aim of this document is to provide a concise description of the general policy recommendations adopted within the Eco-AlpsWater project to store all the data produced and made them accessible to all the partners involved in the project.

Brief outline description of recommendations and policy guidelines for data storage, transnational sharing and retrieval of information.

Within the Eco-AlpsWater project up-to-date approaches have been used to assess and describe biodiversity of aquatic biota in the Alpine Space region. These new approaches produce a huge amount of data which need to be managed carefully to allow an easy access to all the project partners. For this intent a web-based platform has been conceived to store, analyse and transfer huge datasets. Input and structure of data have to be consistent with successive steps, including data pre-processing, bioinformatic pipelines, downstream analyses (merging reads with metadata such as field data and laboratory data, Figure 1).

This data storage network was built following EU ELIXIR standards. ELIXIR is an intergovernmental organisation that brings together life science resources from across Europe. These resources include databases, software tools, training materials, cloud storage and supercomputers. ELIXIR has compiled a list of resources that it recommends for the deposition of experimental data. The scientific community has a shared responsibility to ensure long-term data preservation and accessibility. The purpose of Deposition Databases list is to provide guidance to those who formulate policy and working practices about the appropriate repositories for publishing open data in the life sciences. Following ELIXIR guidelines, data produced within Eco-AlpsWater consortium are structured according to the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) and will be deposited into institutional repositories (e.g. ENA, Zenodo) to allow long-term endurance of the WPT1 outputs.

The Eco-AlpsWater consortium has developed a mandatory and a recommended list of variables that should be measured/collected during the sampling campaigns (Appendix A - Table 1). These variables have been included following formatted rules and measurement units.

LAKE PELAGIC STATION 16S and 18S rDNA

- PP1: Partner in charge of samplings and field measurements
- PP2: Partner in charge of extracting DNA
- PP3(=PP2): Partner in charge of sequencing

- (a): Decision to be adopted by each partner
- (b): Decision to be adopted by each partner

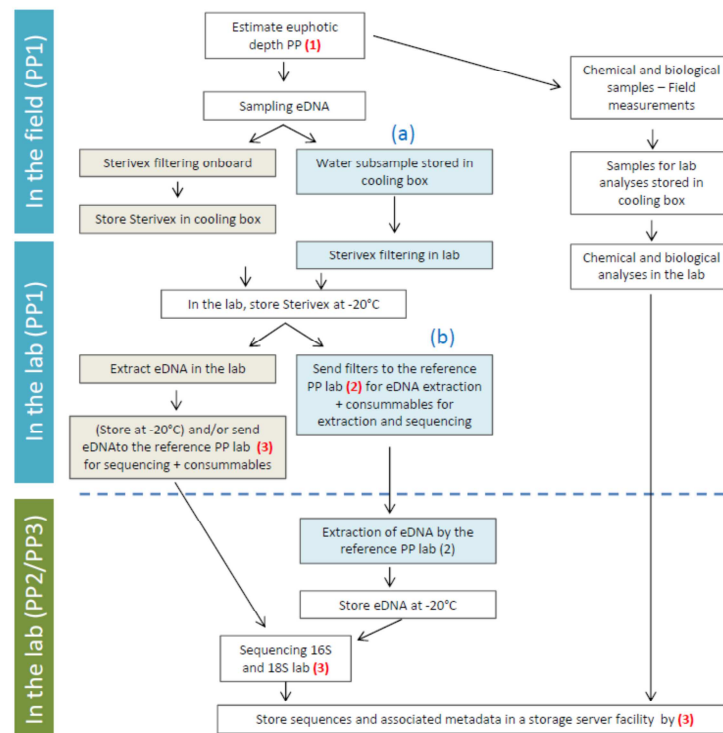


Figure 1. Sampling and processing procedures, from the field to the laboratory, followed by the Eco-AlpsWater consortium during collection activities.

The web-based platform (described in detail in the deliverable DT1.2.2) has been designed to be controlled by a database manager, following rigid rules to upload data information and avoid potential errors. Each sample is uniquely identified by a sample code and a sample ID automatically generated by the system. The HTS files obtained during the project have been named accordingly. Several metadata for each sampling station, describing the HTS data (FASTQ files), need to be recorded, notably the user name, the waterbody, the station name, the date and the characteristics of the sampling site (sampling matrix, sampling depth, depth type...). The laboratory data, which include also temperature, pH, conductivity, oxygen concentration, and chemical analyses of water (bicarbonate, nitrate nitrogen, sulphates, chloride...), represent important information useful to couple with biodiversity estimates obtained from HTS data (Table 1, Appendix A). This web-based platform ensure long-term data preservation and accessibility.

The Eco-AlpsWater (EAW) website is hosted at the Fondazione Edmund Mach at the URL <https://eco-alpswater.fmach.it>. A detailed description of the informatic platform and database used for the data input have been described in the Deliverable DT1.2.2. The data saved in this storage database have been extracted and exported to a Microsoft Access database (Deliverables D.T1.2.3, D.T4.1.1). In this format, the data will be carefully examined for consistency of numerical inputs, checking of outliers, and general checking.

Appendix A

Environmental variable	Measurement units
sampling_depth_min	m
sampling_depth_max	m
sampling_volume	ml
depth_type	
mean_river_outflow	m ³ /s
water_renewal_time	year
mixing_type	
catchment_area	km ²
temperature	°C
field_ph	
field_conductivity	μS/cm
light_attenuation_coefficient	1/m
secchi_disk_depth	m
euphotic_layer	m
oxygen_concentration	mg/l
oxygen_percentage	%
laboratory_ph	
laboratory_conductivity	μS/cm at 25°C
total_alkalinity	mg CaCO ₃ /L
bicarbonates	mg/L
nitrate_nitrogen	μgN/L
sulphates	mg/L
chloride	mg/L
calcium	mg/L
magnesium	mg/L
sodium	mg/L
potassium	mg/L
ammonium	μgN/L
total_nitrogen	μgN/L
soluble_reactive_phosphorus	μgP/L
total_phosphorus	μgP/L
reactive_silica	mgSi/L
dry_weight	mg/L
chlorophyll_a	μg/L
dna_extraction_kit	
dna_extraction_date	
dna_quantity	ng/μL
dna_quantification_method	
dna_quality_a260_230	
dna_quality_a260_280	

vertical_temperature_profiles
phytoplankton_countings
cyanotoxin_samples
sequence_set

Table 1. Environmental variables collected during the sampling campaign and corresponding measurement units.