Bloowater project - Deliverables

Water Joint Programming Initiative
2018 Joint Call
Closing the water cycle gap - Sustainable management of water resources

BlooWater

D 4.1 Data collection guidelines

Workflow scheme to build a dataset for remote sensing HABs monitoring

1. Remote Sensing Data:

    Drones and satellites capture remote sensing data to monitor cyanobacteria. This data includes[1]:

    a. Multispectral Imagery: Images captured in different bands of the electromagnetic spectrum, such as red, green, blue, and near-infrared. These images allow for the identification of chlorophyll-a concentrations, which can indicate the presence of cyanobacteria.
    b. Hyperspectral Imagery: Images captured in numerous narrow bands across the electromagnetic spectrum, enabling more detailed analysis of water quality parameters, including the detection of specific pigments associated with cyanobacteria.
    c. Thermal Imagery: Infrared images that measure water temperature, which can help identify areas where cyanobacterial blooms might occur due to increased temperatures.

2. Water Quality Data: In addition to remote sensing data, collecting on-site water quality data is crucial for calibration and validation of remote sensing measurements. This data includes:

    a. Chlorophyll-a Concentrations: Measuring the concentration of chlorophyll-a, a pigment found in cyanobacteria, can provide an estimate of the biomass and abundance of cyanobacterial blooms.
    b. Nutrient Levels: Assessing the levels of nutrients such as nitrogen and phosphorus is important since they can contribute to the growth of cyanobacteria.
    c. Dissolved Oxygen: Monitoring dissolved oxygen levels helps identify areas where cyanobacterial blooms might lead to oxygen depletion and negatively impact aquatic life.
    d. pH and Conductivity: Measuring pH and conductivity provides information about the water's chemical composition and can help understand the environmental conditions favorable for cyanobacterial growth.
    e. Secchi Depth: Assessing the transparency of the water using a Secchi disk can indicate changes in water clarity caused by cyanobacterial blooms.

3. Spatial and Temporal Data: It is essential to collect data on the location and time of observations to track the dynamics and movement of cyanobacterial blooms. This includes:

    a. Geographic Information System (GIS) Data: Spatial data, including lake boundaries, shoreline characteristics, and land use patterns, which can help identify potential sources of nutrient pollution.
    b. Time Series Data: Regular monitoring over time allows for the detection of trends and seasonal variations in cyanobacterial bloom occurrences.

---

[1] In this project we considered only to use images in the multispectral range which are free of charge specifically from the satellite Sentinel-2

4. Meteorological Data: Collecting meteorological data for understanding the environmental factors influencing cyanobacterial growth. This includes:

   a. Temperature: Monitoring air temperature can help identify favorable conditions for cyanobacterial blooms.

   b. Precipitation: Assessing rainfall patterns can provide insights into the inflow of nutrients into lakes, which can influence cyanobacterial growth.

   c. Wind Speed and Direction: Understanding wind patterns helps predict the movement and dispersal of cyanobacterial blooms.

Data Request to Support Lake and Reservoir Modeling of Phytoplankton and Cyanobacteria

In order to simulate changes in the concentration of cyanobacteria, using mechanistic lake models requires data from several sources that fulfill several different functions in the modelling process.

- Input or forcing data
  These data are what drive the model or make its output change over time. They can be further broken down into
    o Meteorological data that affect the simulation of lake thermal structure
    o Watershed hydrologic and nutrient inputs that affect the water balance of the lake and the amount of nutrients that are available to support phytoplankton growth.

- Data used for parametrizing the model.
  Most important would be information on lake bathymetry. It is also possible to include other information on biogeochemical processes that could be used to set or constrain model parameters. This could include things like ratios of C:Chl or rates of sediment accumulation, but often such information is not available and it is not absolutely required.

- Data used to calibrate the model and evaluate its performance.
  A part of the modeling process is to compare simulated and measured values of the things (state variables) that the model simulates. Measured data can also be used to calibrate the model, by adjusting parameters so that the model state variables match as closely as possible to the measured data. Data needed will include water temperature lake chemistry, and measurements of phytoplankton.

In case specific data is not available, it may be possible to get some data from other sources, for example, meteorological data may be available from gridded global data sets if data from a local meteorological station is not available. It may also be possible (although not ideal) to run some simulations with minimal calibration with the idea of using the model output as a relative indicator rather than an actual prediction. In the case of phytoplankton, it may be especially challenging to obtain sufficient data to calibrate models that predict seasonal and short-term variations in cyanobacteria. The most common and reliable way to measured cyanobacteria is by microscopic counts, which are expensive and time consuming obtain. Thus, data that is frequent enough to calibrate models (especially machine learning models) of cyanobacteria may be difficult to obtain.

Below a list of the requirements for these three data categories in more detail.

Model data requirements

Input forcing data the following data will be needed to make the model run

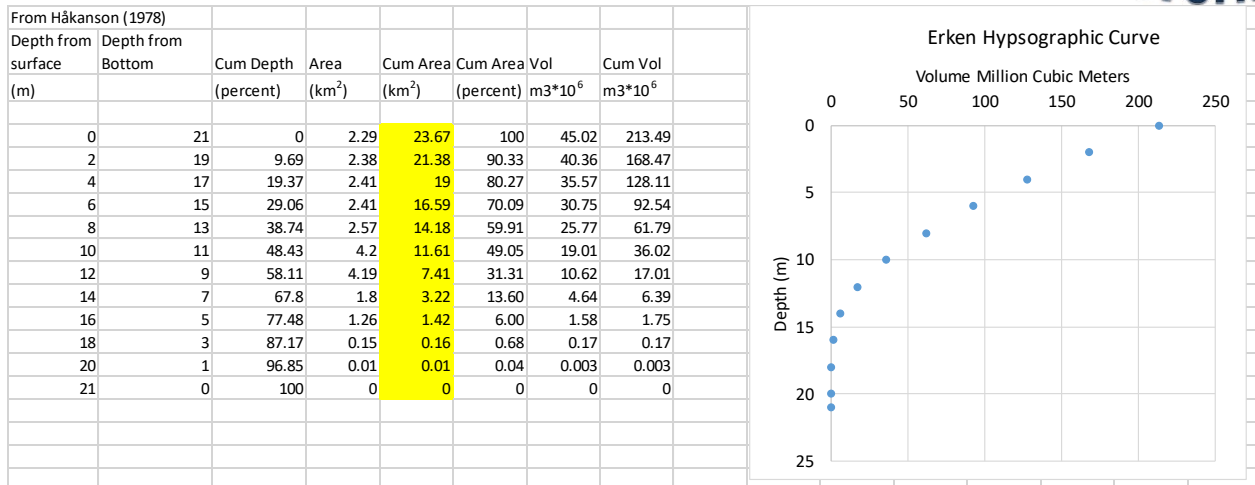Meteorology - usually at an hourly or daily time step

- Wind speed (m/s)
- Wind direction (optional)
- Air Pressure_hPa Air pressure (hPa)
- Air Temperature Air temperature (C)
- Incoming solar radiation (watts/m2)
- Relative Humidity (percent)
- Precipitation (mm)
- Cloud Cover Fraction of sky (0-1 can also be estimated from solar radiation)

Hydrology

- The discharge of the major streams/rivers entering the lake (m3/s usually at a daily time step)
- The concentrations of the major nutrients in the stream water these are generally measured much less frequently than stream discharge.
  - Nitrogen (NOx and NH4) also measurements of total N
  - Phosphorous Soluble reactive phosphorus and Total phosphorus
  - Silica (optional)
  - DOC (optional)
- For reservoirs it would also be good to know information on operations such as withdrawal (volume and depth) spill over the dam, and controlled releases from the dam (daily or longer time step)
- Measurements of lake water level or surface water elevation. This can be used as either model input (forcing the lake to have the specified elevation) or for calibration to confirm that the simulated water level is correct. (Daily or longer time step).

Data used for parametrizing the model

Lake bathymetry . What is needed is the surface area of a horizontal plane passing through the lake at different depths. The area at 0 meters depth is the lake surface area. This data is needed to calculate the volume of the lake and of each horizontal layer of the lake that is simulated by the 1D model. Below is an example from lake Erken.

| From Håkanson (1978) | | | | | | | |
Depth from surface (m)	Depth from Bottom	Cum Depth (percent)	Area (km²)	Cum Area (km²)	Cum Area (percent)	Vol m3*10⁶	Cum Vol m3*10⁶
0	21	0	2.29	23.67	100	45.02	213.49
2	19	9.69	2.38	21.38	90.33	40.36	168.47
4	17	19.37	2.41	19	80.27	35.57	128.11
6	15	29.06	2.41	16.59	70.09	30.75	92.54
8	13	38.74	2.57	14.18	59.91	25.77	61.79
10	11	48.43	4.2	11.61	49.05	19.01	36.02
12	9	58.11	4.19	7.41	31.31	10.62	17.01
14	7	67.8	1.8	3.22	13.60	4.64	6.39
16	5	77.48	1.26	1.42	6.00	1.58	1.75
18	3	87.17	0.15	0.16	0.68	0.17	0.17
20	1	96.85	0.01	0.01	0.04	0.003	0.003
21	0	100	0	0	0	0	0

Erken Hypsographic Curve

The data in yellow is what is actually used in the model. (see WP 2)

Data used to calibrate the model

These are data that are used to compare with state variables from model simulations. For example, comparing simulated lake water temperature to that actually measured in the lake. These data are used both to verify that the model is preforming correctly, and to calibrate the model to make it more closely match the measured data. Consequently, these data are quite important.

- **Water temperature** is needed to calibrate the hydro thermal calculations made by the model. In the GOTM model there are about 5 parameters we can adjust that affect the inputs of wind energy, the surface heat fluxes, and mixing processes. We will need as many water temperature measurements as possible. Ideally they will be collected as full profiles or at least from several different depth. We should have measurements spread over several years, but they do not need to that frequent ie daily –weekly would be great, but monthly would also work
- **Water chemistry**. We will need measurements of the same major nutrients (N, P, and Optionally Si and DOC) as required for the inflowing stream chemistry. Ideal also be collected at multiple depths allowing separate estimations of the concentrations in the epilimnion and hypolimnion during stratified conditions. Estimations of dissolved oxygen are also very useful, and often more plentiful since they can be measured with in situ sensors as well as more time consuming laboratory measurements.
- **Phytoplankton concentration**. Ideally this would be based on microscopic counts. But this is not always available since these are time consuming to collect. Chlorophyll concentrations are also very valuable as a general indicator of phytoplankton biomass and are usually much more available. Chlorophyll is typically the parameter used to calibrate the phytoplankton components of most lake water quality models. To get even more frequent measurements there are in situ probes that measure chlorophyll fluorescence and phycocyanin fluorescence, with phycocyanin being a more specific indicator of cyanobacteria.

Data request to design a Decision Support System

Data collection and acquisition should comprise a list of:

• the stakeholder, at local (regional) level, has been identify (public health, consumers and environmental associations, administrations, plant management and water supply companies, waste management companies, engineering companies, etc);
• the data concerning the characteristics of the pilot action area (water surface, surface of the phenomenon, evolution time, bacteria species, cyanotoxins concentration evolution, nutrient contents, pH, TSS, etc);
the process parameters (water inflow, pH, temperature, abatement kinetic, polymer concentration, etc);
• the water supply needs;
• the legislation (local, national and international) about the drinkable water management and characteristic.
Taking the above into consideration, the data to be collected is categorized as the following scheme:

1. **Statutory Legislation;**
2. **Bibliography/projects concerning, in particular, the cyanotoxin problem and the applicable technologies;**
3. **Phenology and bacteria species;**
4. **Stakeholders;**
5. **Water related ancillary data;**
6. **Ancillary data miscellanea;**
7. **Water quality and characteristics;**
8. **Drinkable water treatment plants;**
9. **Cartography of the pilot action site;**
10. **GIS**

1.Legislation

In this category, partners must address all the statutory legislation on water protection / water management that currently exist in their countries.

If currently exists, the parameters that need to be examined under this category are the following:

1) Water Management Legislation for:

• Urban Wastewater (alternative resource for irrigation);

• Use of drinkable water in the industries;

• Drinking Water;

• Drain/Sewer.

2) Water Protection Regulations

- provide necessary strategic and technical tools for the decision makers to make informed decisions on the use of non-conventional water resources based on economic, environmental and social justification;

- provide the necessary rules and regulations for the protection of surface water resources and reduce health exposure and hazards;

- establishing institutional and consultation mechanisms for the management of these resources in a transparent, and accountable manner.

2. Bibliography/projects

This category records all the previous projects performed and the bibliography concerning the drinkable water and the relative technologies applied for in order to emphasize the previous experiences.

The information that needs to be collected for each project is as follows:

- Name of the project and data such as partners involved, main objectives, level of interest (local, national, internatiomnal), identification of the stakeholders of each project, impact/relevance, sponsor, amount of founding, the site, the year, etc;

- Bibliography: list the references of the past projects

- Website / Link Hyperlink: list the website (if any) of project/ equipments/ technical records.

3. Phenology and bacteria species

A great effort on these data will be devoted. Firstly data obtained from the scientific literature has been collected but, naturally, the data of the pilot sites will be of primary importance.

4. Stakeholders See D 4.2

5. Water Related Ancillary Data

The Water Related Ancillary Data section is about collecting data regarding the water usage, the water discharges and the water distribution network, in order to estimate the quantity of water that is used (and to treat) per day. This collection data can be improved to increase the DSS application to the whole water cycle management.

The parameters that are needed to be taken into consideration for the collection of data, are the following:

- Water use - identification of the amount and sources of water usage for:

1. Residential area (water consumption per capita);
2. Industrial / Commercial/ Institutional;
3. Agricultural;
4. Municipal;
5. Tourism.
- Untreated water:

• If a drinkable water treatment plant exist then data will be filled in in the specific excel worksheet

• If a water treatment plant does not exist, please give a characterization for each inflow water in terms of the most important parameters, such as:

  ✓ Flow rate: It depends upon population density, water consumption, and the extent of the commercial or industrial activity in the community.
  ✓ Chemical Oxygen Demand (COD) and Biological Oxygen Demand (BOD);
  ✓ Principal pollutants concentration: suspended solids, nutrients (Nitrogen and phosphorus), heavy metals;
  ✓ Cyanobacteria.

Water distribution network, this is to measure the amount of water that distributed by the following water networks:

  • Drinking water distribution network;
  • Sewage network;
  • Reservoirs;
  • Wells;
  • Rivers.

6. Ancillary Data Miscellanea

In this category the partner has to estimate the quantity of water used from each interested area (local and regional). Population estimates are one of the greatest data challenges for water supply planners. It is critical to know how many people a utility serves and to project how many may be served in the future to ensure adequate water supply. Accurate and consistent estimates of population are a necessary component of calculating metrics such as gallons per capita per day.

The Parameters that the partner needs to take into consideration are the following:

Population: Estimate the seasonal equivalent inhabitants (or population) living in each study area in order to be able to calculate the amount of water used per person.

Industrial plant: water use for the process purposes (i.e. textile, agroindustry, etc). The characteristics of the to wastewater treatment plant (if any) should be recorded.

7. Water quality

As water quality depends on all the substances that compose the hydric solute system therefore for water management purposes it is important to examine the water quality of each study area, in order to be able to suggest ways of improvements related with the water cycle management. Partners have to take into consideration the parameters according to the legislation limits and, of course, relating to the cyanobacteria presence. Moreover the monitoring systems (if any) should be recorded.

8. Drinkable Water Treatment Plant

It is important to collect information about any existing Drinkable Water Treatment Plant in order to optimize the process allowing the removal of cyanotoxins.

The information that needs to be collected is the following: adopted technology, the year the water treatment plant was constructed, maintenance records of the plant, water flow rate, water characteristics, seasonality, produced waste, water use after treatment, etc.

9. Cartography

The data that will be collected in this category will be combined also with the GIS section. The data needed is as follows:

• Landuse, a landuse map illustrates types and intensities of different land uses in a particular area;

• Hydrographical Network, hydrography is the mapping of water features;

• Water distribution network, a map of water distribution illustrates the routes that the water networks follow ending up into different water bodies;

• Wastewater network, a map of wastewater network illustrates the routes that the wastewater follows ending up into different water bodies;

• Temperature, a temperature map could present the past, current and future temperature of the area of interest;

• Precipitation/rainfall, precipitation/ rainfall map presents the rate of precipitation in areas of interest;

• Administrative boundary, subdivisions of areas/territories/jurisdictions recognized by governments or other organizations for administrative purposes;

• Terrain morphology, for example, to generate elevation, area, and volume curves for a set of selected drainage areas;

• Geology, a geology map is a special-purpose map made to show geological features;

• Protected areas, a map could be prepared illustrating if there are any protected areas in specific locations of interest (i.e. NATURA 2000 areas, National Parks etc.);

• Water Treatment plant, all the existing industrial plants for water treatment, locating in specific areas of interest, could be presented in a map;

• Reservoirs, locations of water reservoirs.

10. GIS

A Geographic Information System (GIS) enables us to visually represent, inquire, analyze, and interpret data, leading to a deeper understanding of relationships, patterns, and trends. It is specifically designed to capture, manage, analyze, and display various types of geographically referenced information. Through GIS, we gain insights into our world by revealing connections, patterns, and trends in the form of maps, globes, reports, and charts. By presenting data on a map,

GIS software facilitates easy comprehension and sharing, helping us answer questions and address problems effectively. To ensure the project's future sustainability, the initial data collection system should be upgraded at the national level.