INTERNSHIP 8: DERIVING FUNCTIONAL PLANKTON METRICS FROM IMAGING FLOW CYTOMETRY: TOWARDS A FAIR APPROACH

Supervisors: Isabelle Fournier (UKCEH), Mike Bowes (UKCEH), Emilie Poisson Caillault (Université du Littoral Côte d'opale)

Internship Location: UKCEH Wallingford, Maclean Building, Crowmarsh Gifford, Wallingford, OX10 8BB

Project Description:

Flow cytometry is a technique that groups particles based on their interaction with light (scattering and fluorescence), often through fixed wavelength lasers. The molecules that allow phytoplankton to conduct photosynthesis, the photosynthetic pigments, naturally interact with light, producing fluorescence. As different groups of phytoplankton contain different pigments in varying ratios, the signal acquires by flow cytometry can be analysed to quantify the abundance of phytoplankton groups.

Flow cytometry data analysis is currently undertaken by experts through flow cytometry software developed by the instrument companies. This leads to variability in settings, the grouping, the name given to them and the abundance of cells in each group, depending on the instrument\software and the expert that created the analysis protocol. The variability makes it difficult to compare data from different laboratories and between different instruments, and limit the use of this technique in large scale studies. An increasing number of agencies and scientists are developing frameworks to normalise the analysis of phytoplankton through flow cytometry, moving towards a FAIR approach (Findable, Accessible, Interoperable, Re-usable). In the phytoplankton field of research, it can be achieved by gaining consensus on and applying open-source classification on raw data.

Flow cytometry data is a rich source of information on phytoplankton, particularly when the technique is coupled with imagery. Imaging flow cytometry allows insight into natural phytoplanktonic communities at unprecedented levels, because it produces a coupled image\fluorescence dataset quicker and cheaper than any traditional approaches. One of the key gaps that imaging flow cytometry can fill is the ability to quantify non-photosynthetic plankton.

UKCEH has developed an expertise on imaging flow cytometry and has used it to increase the understanding of river ecology and the causes of phytoplankton blooms. This project aims to develop or adapt an open-source image clustering algorithm and an open-source classification workflow of fluorescence data to automate the quantification of the functional plankton metrics derived from imaging flow cytometry. This work will involve metrics such as the classic photosynthetic plankton abundance and the novel non-photosynthetic plankton abundance.

Tasks:

- Manual counting of non-photosynthetic plankton through flow cytometry images.
- Develop/adapt and apply an open-source image clustering algorithm to automate the non-photosynthetic plankton counting.
- Develop/adapt and apply an open-source classification workflow to automate the analysis of flow cytometry data for photosynthetic plankton.

Expected Outcomes:

- A workflow specifying the required steps towards an automated analysis of flow cytometry data for photosynthetic plankton from the Attune Cytpix.
- Extracted data from the Attune Cytpix, formatted and ready to be tested. Tested if possible.
- An image clustering algorithm to automate the non-photosynthetic plankton counting from the Attune Cytpix images, on GitHub.

- Given the novelty of the approach and the data that it is going to generate, this project can lead the intern to be part of multiple scientific papers.

Required Skills and Background:

Essential:

- Currently in postgraduate (including PhD) education at university or have graduated within the six months prior to the start of the internship (30th June 2025) - either informatics/data science with an interest in aquatic ecosystems OR ecology/biology with a working knowledge of machine learning
- Experience with machine learning (class imbalance, parameters selection and tuning, etc.)
- Coding with Python or R

Desirable:

- Familiarity with flow cytometry, community ecology statistics, or freshwater ecology.