Title: Deciphering hidden patterns in carbon cycling from molecules to landscapes

Supervisor: Prof Andrew Tanentzap, Dr Lucas Braga and Dr Jeremy Fonvielle

Department/Institution: Plant Sciences / Conservation Research Institute

Brief description of background to project:

Viruses are the most abundant and diverse biological entities on Earth, but their contribution to global carbon cycling remains poorly understood. By killing their hosts, viruses release highly labile, intracellular compounds into the environment. These compounds can then be used by other microbes that respire greenhouse gases. An estimated 145 Gt C per year – more than 15-times that released by burning fossil fuels – comes from virus-mediated cell lysis in the oceans alone. However, our understanding of this massive carbon flux is hampered by limited knowledge about viral diversity and the composition of these intracellular compounds, which are not all equally labile. This project will discover hidden patterns in carbon cycling using artificial intelligence. The goal is to identify how viruses influence the greenhouse gas emissions of freshwaters by changing the chemical diversity of the environment.

Background introductory reading:


Brief description of what the student will do, the skills they will gain and the outcome expected from the project:

First, the student will analyse an existing dataset of viral and chemical diversity from 119 lakes across Europe. They will spend 4 weeks designing and validating machine learning (ML) models (e.g. random forest) that predict viral and chemical diversity from satellite remote sensing data. Viral and chemical diversity cannot be estimated directly from a specific absorbance signature. Instead, the student will develop proxies of viral and chemical diversity, including from surrounding land use and Sentinel satellite bands related to turbidity, photosynthetic pigments, and sediment loading. The student will then spend 2 weeks statistically testing the links among viral diversity and chemical diversity, estimated from both observations and the ML models, and in-situ measurements of greenhouse gas emissions. The last 2 weeks will be dedicated to writing up the results. The student will gain skills in machine learning, geospatial analysis and knowledge of the global carbon cycle.

Required academic background/skills of student:
- A good knowledge of R (preferred) or Python is essential for designing the model.
- Basic experience with geographical analysis (not necessarily GIS or qGIS) is desirable.
- Basic knowledge of Sentinel images and the SNAP software would be a plus.

| Dates within which the project could be carried out and supervision could be offered: |
| April to October |

| Logistic arrangements: |
| The placement can be undertaken in person, though it is not necessary for this project. |