EarthServer2 : The Marine Data Service - Web based and Programmatic Access to Ocean Colour Open Data

Oliver Clements and Peter Walker

Plymouth Marine Laboratory, Plymouth, United Kingdom (olcl@pml.ac.uk)

The ESA Ocean Colour - Climate Change Initiative (ESA OC-CCI) has produced a long-term high quality global dataset with associated per-pixel uncertainty data. This dataset has now grown to several hundred terabytes (uncompressed) and is freely available to download. However, the sheer size of the dataset can act as a barrier to many users; large network bandwidth, local storage and processing requirements can prevent researchers without the backing of a large organisation from taking advantage of this raw data.

The EC H2020 project, EarthServer2, aims to create a federated data service providing access to more than 1 petabyte of earth science data. Within this federation the Marine Data Service already provides an innovative on-line tool-kit for filtering, analysing and visualising OC-CCI data. Data are made available, filtered and processed at source through a standards-based interface, the Open Geospatial Consortium Web Coverage Service and Web Coverage Processing Service. This work was initiated in the EC FP7 EarthServer project where it was found that the unfamiliarity and complexity of these interfaces itself created a barrier to wider uptake. The continuation project, EarthServer2, addresses these issues by providing higher level tools for working with these data. We will present some examples of these tools.

Many researchers wish to extract time series data from discrete points of interest. We will present a web based interface, based on NASA/ESA WebWorldWind, for selecting points of interest and plotting time series from a chosen dataset. In addition, a CSV file of locations and times, such as a ship's track, can be uploaded and these points extracted and returned in a CSV file allowing researchers to work with the extract locally, such as a spreadsheet.

We will also present a set of Python and JavaScript APIs that have been created to complement and extend the web based GUI. These APIs allow the selection of single points and areas for extraction. The extracted data is returned as structured data (for instance a Python array) which can then be passed directly to local processing code. We will highlight how the libraries can be used by the community and integrated into existing systems, for instance by the use of Jupyter notebooks to share Python code examples which can then be used by other researchers as a basis for their own work.