Felyx : A Free Open Software Solution for the Analysis of Large Earth Observation Datasets

Jean-Francois Piolle (1), Jamie Shutler (2), David Poulter (3), Veronica Guidetti (4), and Craig Donlon (5)

(1) IFREMER, Brest, France (Jean.Francois.Piolle@ifremer.fr), (2) Plymouth Marine Laboratory, United Kingdom (jams@pml.ac.uk), (3) Pelamis Ltd, United Kingdom (david.poulter@pelamis.co.uk), (4) European Space Agency, ESRIN, Italy (veronica.guidetti@esa.int), (5) European Space Agency, ESTEC, The Netherlands (Craig.Donlon@esa.int)

GHRSST project, by assembling large collections of earth observation data from various sources and agencies, has also raised the need for providing the user community with tools to inter-compare them, assess and monitor their quality. The ESA /Medspiration project, which implemented the first operating node of GHRSST system for Europe, also paved the way successfully towards such generic analytics tools by developing the High Resolution Diagnostic Dataset System (HR-DDS) and Satellite to In situ Multi-sensor Match-up Databases. Building on this heritage, ESA is now funding the development by IFREMER, PML and Pelamis of felyx, a web tool merging the two capabilities into a single software solution. It will consist in a free open software solution, written in python and javascript, whose aim is to provide Earth Observation data producers and users with an open-source, flexible and reusable tool to allow the quality and performance of data streams (satellite, in situ and model) to be easily monitored and studied.

The primary concept of Felyx is to work as an extraction tool, subsetting source data over predefined target areas (which can be static or moving) : these data subsets, and associated metrics, can then be accessed by users or client applications either as raw files, automatic alerts and reports generated periodically, or through a flexible web interface enabling statistical analysis and visualization.
Felyx presents itself as an open-source suite of tools, written in python and javascript, enabling :

* subsetting large local or remote collections of Earth Observation data over predefined sites (geographical boxes) or moving targets (ship, buoy, hurricane), storing locally the extracted data (refered as miniProds). These miniProds constitute a much smaller representative subset of the original collection on which one can perform any kind of processing or assessment without having to cope with heavy volumes of data.

* computing statistical metrics over these miniProds using for instance a set of usual statistical operators (mean, median, rms, ...), fully extensible and applicable to any variable of a dataset. These metrics are stored in a fast search engine, queryable by humans and automated applications.

* reporting or alerting, based on user-defined inference rules, through various media (emails, twitter feeds,..) and devices (phones, tablets).

* analysing miniProds and metrics through a web interface allowing to dig into this base of information and extracting useful knowledge through multidimensional interactive display functions (time series, scatterplots, histograms, maps).

The services provided by felyx will be generic, deployable at users own premises and adaptable enough to integrate any kind of parameters. Users will be able to operate their own felyx instance at any location, on datasets and parameters of their own interest, and the various instances will be able to interact with each other, creating a web of felyx systems enabling aggregation and cross comparison of miniProds and metrics from multiple sources.

Initially two instances will be operated simultaneously during a 6 months demonstration phase, at IFREMER - on sea surface temperature (for GHRSST community) and ocean waves datasets - and PML - on ocean colour. We will present results from the Felyx project, demonstrate how the GHRSST community can exploit Felyx and demonstrate how the wider community can make use of the GHRSST data within Felyx.