ARTMO's retrieval toolboxes for optimizing parametric, non-parametric and physically-based biophysical variable mapping

J. Verrelst[*], *J.P. Rivera, G. Camps-Valls & J. Moreno*

Laboratory of Earth Observation (LEO), Image Processing Laboratory (IPL), University of Valencia, Spain.
* *jochem.verrelst@uv.es*, T: +34 96 354 40 67; F: + 34 96 354 32 61

In all generality, retrieval methods dedicated to the quantification of terrestrial biophysical variables can be categorized into three main domains: 1) *parametric regression*, 2) *non-parametric regression*, and 3) *physically-based methods*. For the last few years, we have made significant advances in all these domains, including the development of software to automate these methods. It eventually led to a scientific software package ARTMO (Automated Radiative Transfer Models Operator) that embodies multiple toolboxes and a suite of leaf and canopy radiative transfer models (RTMs); http://ipl.uv.es/artmo/. The following toolboxes enable to fully exploit the three retrieval domains:

Parametric methods refer to the use of regression models through spectral indices. The 'Spectral Indices' toolbox allows systematic calculation of all possible band combinations of a sensor according to the formulation of an established index. The prediction efficiency of each index can be automatically evaluated against in-situ data or input data coming from a RTM by using a fitting curve (e.g. linear, exponential, power). Options to add noise, to control calibration/validation partitioning and various goodness-of-fit measures to assess the performances (e.g., r^2, RMSE) are provided. The best performing regression model can subsequently be applied to an imagery, which leads to instantaneous mapping of the targeted biophysical variable.

Non-parametric methods refer to the use of machine learning regression algorithms (MLRA). The 'MLRA' toolbox encompasses a collection of linear and non-linear nonparametric regression techniques such as partial least squares (PLS), neural networks (NN), support vector regression (SVR), kernel ridge regression (KRR) and Gaussian processes regression (GPR) and others. Depending on the chosen MLRA, multi-output is possible (PLS, NN, KRR) or associated uncertainty estimates are delivered (GPR). This toolbox is designed in a similar way as the Spectral Indices toolbox; with the same type of calibration and validation options and goodness-of-fit measures provided. The best performing MLRA model can subsequently be applied to an imagery which leads to instantaneous mapping of the targeted biophysical variable(s).

Physically-based methods refer to the inversion of Lookup-table (LUT)-based RTMs through cost functions. This method is considered a physically-sound and robust to retrieve biophysical variables, but regularization strategies are required to mitigate the drawback of ill-posedness. The 'Inversion' toolbox encompasses a collection of more than 60 cost functions, originating from three different mathematical families, being: *information measures*, *M-estimates* and *minimum contrast methods*. Various regularization options can be introduced in the inversion, being: adding noise, multiple solutions, and data normalizing. Simultaneous retrieval of multiple variables is possible. Additional uncertainty estimates can be provided in the form of standard deviation and residuals. The best assessed inversion strategy can subsequently be applied to an imagery, which leads to mapping of the targeted biophysical variable(s).

In this work, all these methods were evaluated by using Simulated Sentinel-2 data against ground-based validation data from the ESA campaign SPARC (Barrax, Spain). Results will be presented for leaf area index (LAI) retrieval. Apart from retrieval accuracy also processing speed was analyzed. This work will close with consolidated guidelines towards powerful retrieval methods that are implementable in operational processing chains.

Keywords: biophysical parameter retrieval, ARTMO, Sentinel-2, LAI

Presentation preference: Oral