EGU2020-21952
https://doi.org/10.5194/egusphere-egu2020-21952
EGU General Assembly 2020

A machine learning software framework for extraction of phenology indicators from multi-temporal sentinel-2 images

Dounia arezki[1,2], Hadria Fizazi[1], Santiago Belda[2], Charlotte De Grave[2], Luca Pipia[2], and Jochem Verrelst[2]
[1]University of science and technology of Oran-MB, Mathematics and computer science, computer science, Algeria (dounia.arezki@univ-usto.dz)
[2]Image Processing Laboratory (IPL), University of Valencia, C/Catedratico Jose Beltran 2, 46980, Paterna, Valencia, Spain

Optical Earth observation satellites provide spatially-explicit data that are necessary to study trends in vegetation dynamics. However, more of often than not optical data are discontinuous in time, due to persistent cloud cover and instrumental noises. Hence, the operating constraints of these data require several essential pre-processing steps, especially when aiming to reach towards monitoring of vegetation seasonal trends. To facilitate this task, here we present an end-to-end processing software framework applied to Sentinel-2 images.

To do so, first biophysical retrieval models were generated by means of a trained machine learning regression algorithm (MLRA) using simulated data coming from radiative transfer models. Among various tested MLRAs, the variational heteroscedastic Gaussian process regression (VHGPR) was evaluated as best performing. to train the retrieval model. The training and retrieval were conducted in the Automated Radiative Transfer Models Operator (ARTMO) software framework.

Subsequently, in view of retrieving the phenological parameters from the obtained vegetation products, a novel times series toolbox as part of the ARTMO framework was used, called: Decomposition and Analysis of Time Series software (DATimeS). DATimeS provides temporal interpolation among other functionalities with several advanced MLRAs for gap filling, smoothing functions and subsequent calculation of phenology indicators. Various MLRAs were tested for gap filling to reconstruct cloud-free maps of biophysical variables at a step of 10 days.

A demonstration case is presented involving the retrieval of Leaf area index (LAI), fraction of Absorbed Photosynthetically Active Radiation (FAPAR) from sentinel-2 time series. A large agricultural Algerian site of 143, 75 km² including Oued Rhiou, Ouarizane, Djidioua (1,345,075 pixels) was chosen for this study. A reference image was excluded from the time series in order to evaluate the reconstruction accuracy over a 40-day artificial gap.

The reference vs. Reconstructed maps produced by the gap-filling methods were compared with statistical goodness-of-fit metrics. Considering both accuracy and processing speed, the fitting algorithms Gaussian process regression (GPR) and Next neighbour interpolation (R^2= 0.90 / 0.081

sec per pixel and R^2=0.88 / 0.001 sec per pixel respectively) interpolations proved to reconstruct the vegetation products the most efficient, with GPR as more accurate but Next faster by a factor of 70.

Finally, we evaluated of the phenology indicators such as start-of-season and end-of-season based on LAI and FAPAR. The obtained maps provide valid information of the vegetation dynamics. Altogether, the ARTMO-DATimeS software framework enabled seamless processing of all essential steps: (1) from L2A sentinel-2 images converted to vegetation products, (2) to cloud-free composite products, and finally (3) converted into vegetation phenology indicators.