Improvement of Existing and Development of Future Copernicus Land Monitoring Products - The ECoLaSS Project

Sevillano Marco, E. [1], Schwab, K.[1], Sommer, C.[1], Probeck, M.[1]
1 GAF AG, Munich, Germany

KEY WORDS: agriculture, forest, grassland, Copernicus, CLMS, ECoLaSS, Earth Observation, DIAS, forest, grassland, HRL, LCLU, Sentinel, time series analysis

ABSTRACT:
The Copernicus Programme, managed and coordinated by the European Commission (EC) and implemented in partnership with the European Space Agency (ESA), member states and various EU agencies, offers services mainly based on Earth Observation (EO) data provided by ESA through the Copernicus Space Component. Complementing the operational-phase implementation, the Horizon 2020 project ECoLaSS (Evolution of Copernicus Land Services based on Sentinel data) contributes to improving existing and developing next-generation Copernicus Land Monitoring Service (CLMS) products of the pan-European and global components, with a suggested implementation schedule from 2020 onwards.

Among the key products of the pan-European component of the CLMS are the so-called High Resolution Layers (HRLs), which are thematic products currently targeting land cover characteristics of 5 main classes: Imperviousness, Forest, Grassland, Water/Wetness and Small Woody Features. These layers are produced in regular 3-year intervals from multi-temporal EO data at 10-20 meter spatial resolution for 39 European countries (EEA 39). Further similar and other products (such as a HR Arable Land Layer) are in the debate, as well as possibly improved product specifications, timeliness, etc.

Therefore, the rapidly evolving scientific developments as well as emerging additional user requirements are continuously analysed by the project in a close stakeholder interaction process, targeting a future pan-European roll-out of new/improved CLMS products, and assessing the potential transferability to global applications. Within ECoLaSS, selected innovative improved or novel products and methods are being developed, tested and prototypically demonstrated, based on advanced dense time series processing of optical (Sentinel-2) and SAR (Sentinel-1) satellite data. In addition, several ancillary data are employed for the classification (e.g. LUCAS, IACS, VHR), be it for algorithm training or validation purposes. The project outcomes will be made fully publicly available, including extensive methodological documentation.

This work introduces the ECoLaSS project setup, together with a comprehensive overview of the collected user and stakeholder requirements, developed processing methodologies and the candidate prototypes for operational service implementation. The focus is laid on the established prototypes for potential next-generation HRL products on Forest, Grassland and Agriculture, based on dense Sentinel-1 and Sentinel-2 time series analytics. Throughout Europe, representative demo sites of approx. 60,000 - 90,000 km² size each were selected, covering boreal, Mediterranean, steppic, Atlantic, alpine and continental parts of 13 European countries. The improvement endeavours of existing products comprise assessing prototypes of *enhanced* forest types and forest cover change layers as yearly updates, as well as better-quality permanent grassland classification and use categorisation together with a yearly change identification approach. The project also targets entirely novel products, such as a possible future HRL on arable land, for which a prototype is presented, targeting a pan-European crop mask of high precision and a representative, robust classification of crop types. In ECoLaSS, these and all other candidate prototypes are being produced with a view to a potential European-wide implementation and the integration into the operational Copernicus service infrastructure, in case of being selected for service roll-out by the procuring entity(ies).

From a methodological point of view, the temporal analysis techniques applied make use of a multitude of temporal feature descriptors (multiple temporal statistical metrics), as well as derived indices (e.g. VV/VH ratio, NDVVVH from SAR data and NDWI, NDVI, Brightness and IRECI from optical data) extracted from dense Sentinel-1 and -2 time series. Temporal feature descriptors are able to depict and quantify a surface's status and phenological behaviour over time as well as to capture the intensity and significance of change information and time series related statistical properties. Thus, they constitute

powerful input features for various classification or regression analysis tasks. Not being directly related to image acquisition dates, neither customised scene selection efforts nor prior knowledge of change event dates are required. Hence, feature descriptors can be flexibly computed from reflectance or derived index data. Iterative calculation over a time stack allows for implementing suitable dynamic change detection systems, which require to be frequently updated, such as it would be the case for the existing HRL forest and grassland services as well as for a potential future arable land service. Seven "simple" temporal feature descriptors represent standard statistical temporal metrics of a time series: maximum value (max), minimum value (min), average value (mean), median, standard deviation (std) and coefficient of variation (cv). Using a random forest-based classification approach on several hundred temporal feature descriptors, a feature selection is performed to identify the most meaningful input features for classification, ensuring a high resultant product quality as well as a cost and time efficient processing with high accuracy. The prototypes are fully built upon the capabilities of state-of-the-art cloud processing solutions, exploring implementation on, amongst others, the Copernicus Data and Information Access Service (DIAS) platforms.

The overall thematic accuracies for all Grassland, Forest, and Arable Land layer prototype implementations produced so far are promising, ranging from 77-98%. The lessons-learned from the first temporal feature descriptors extraction and selection tests, together with the added value of the combined use of both optical and radar data, lead now to refined workflows and results. Besides quantitative statistical accuracy assessments and other quality metrics, a rigorous benchmarking procedure is applied to assess the prototypes' overall operational readiness and technical maturity for integration into the CLMS overall setup and architecture. This comprises assessing, amongst others, automation levels, portfolio complementarity, or the state of the art/innovation degree. Moreover, an integration plan for the most suitable prototypes will be set up.