The Earth in a Box: A light-weight data cube approach to empower the study of land-surface processes and interactions

Fabian Gans (1), Miguel D. Mahecha (1), Markus Reichstein (1), Gunnar Brandt (2), Norman Fomferra (2), Hans Permana (2), and Carsten Brockmann (2)

(1) Max Planck Institute for Biogeochemistry, Jena, Germany, (2) Brockmann Consult GmbH, 21502 Geesthacht, Germany

Scientists today are confronted with a plethora of Earth observations (EO) that are available through a multitude of data platforms, monitoring initiatives, and model-data integration approaches. To tap into the full potential of joint exploitations of these datasets, the scientist has to deal with different data formats, retrieval methods, data usage policies, inconsistent resolutions and formatting as well as getting access to suitable computation and data storage facilities. This can be time consuming and impede scientific progress. The "Earth in a Box" or "Earth System Data Cube" is a novel infrastructure for scientists. It not only provides a set of harmonized data cubes, but also developed a strong data analytics toolkit in tandem with access to computational facilities to jointly analyze the datasets. The data cube consists on a selection of diverse Earth observation datasets including multiple atmospheric variables, Land-Atmosphere fluxes, biophysical parameters and is currently extended to ocean variables. These datasets are analysis-ready, harmonized on different temporal and spatial resolutions to allow rapid exploitation. Visualization can be performed through the CCI Toolbox. The core data analytics toolkit allows users, proficient in a language of scientific computing like Python or Julia, to apply their own analysis methods on the multivariate data cube. Examples implemented include nonlinear time series analytic tools, dimensionality reduction approaches, or anomaly detectors, among others. To showcase the potential of the approach we will present some scientific case studies demonstrating the full potential of the infrastructure and seek for participants in our early-adopter user engagement call.