Geophysical Research Abstracts
Vol. 20, EGU2018-12702, 2018
EGU General Assembly 2018

Estimating information in Earth data cubes

Juan Emmanuel Johnson (1), Emiliano Diaz (1), Valero Laparra (1), Miguel Mahecha (2), Diego Miralles (3), and Gustau Camps-Valls (1)

(1) Universitat de València - ESQ4618001D, Image Processing Laboratory (IPL), Image Processing Laboratory (IPL), Paterna, València. Spain., Spain (gustau.camps@uv.es), (2) Department Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany, (3) Laboratory of Hydrology and Water Management, Ghent University

The Earth is a highly complex networked system, and from a pure statistical point of view, it can be viewed as a multidimensional distribution in a representation space. In recent years, a number of initiatives have collected richful Earth data cubes representing our planet status explicitly in space and time, and across many variables describing the atmosphere, biosphere and anthroposphere. Projects like CABLAB, http://earthsystemdatacube.net/, Open DATACUBE, https://www.opendatacube.org/, and SAT-Ex, http://www.sat-ex.ugent.be, are just some examples. Such unprecedented data volume poses challenges and research opportunities.

In this work we will explore the information contained in the data cubes. We will introduce a thorough analysis of the information content, intrinsic dimensionality, and nonlinear feature relations across time and space. We aim at answering relevant questions about what variables and time/space scales carry most of the information, and how much information is conveyed by considering higher order feature relations beyond covariance/correlation analysis. We will resort to standard tools in information theory (such as total correlation, KL divergences and entropy) as well as novel techniques of multivariate Gaussianization.