



Xarray: multi-dimensional data analysis in Python

Stephan Hoyer (1), Joe Hamman (2), and Fabien Maussion (3)

(1) Google Research, Mountain View, California, USA, (2) National Center for Atmospheric Research (NCAR), Boulder, Colorado, USA, (3) Institute of Atmospheric and Cryospheric Sciences, University of Innsbruck, Austria
(fabien.maussion@uibk.ac.at)

xarray (<http://xarray.pydata.org>) is an open source project and Python package that provides a toolkit and data structures for N-dimensional labeled arrays, which are the bread and butter of modern geoscientific data analysis. Key features of the package include label-based indexing and arithmetic, interoperability with the core scientific Python packages (e.g., pandas, NumPy, Matplotlib, Cartopy), out-of-core computation on datasets that don't fit into memory, a wide range of input/output options, and advanced multi-dimensional data manipulation tools such as group-by and resampling. In this contribution we will present the key features of the library and demonstrate its great potential for a wide range of applications, from (big-)data processing on super computers to data exploration in front of a classroom.