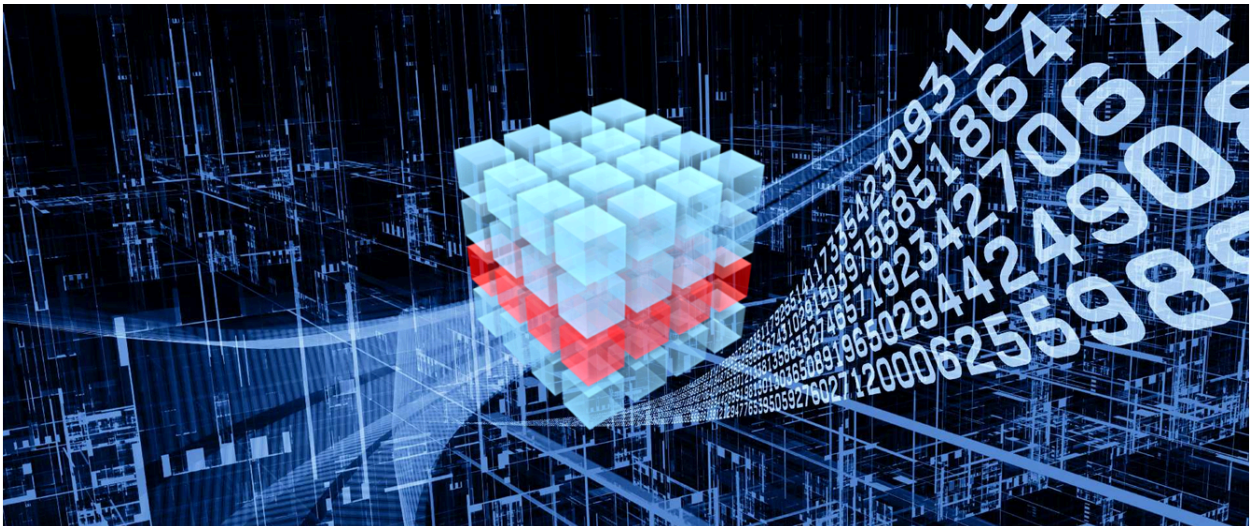

Earth Datacubes: Concepts, Standards, Services

Peter Baumann, Jacobs University, Germany

Workshop Description

Datacubes are emerging in Earth science as an enabling paradigm for offering massive spatio-temporal Earth data in an analysis- (and visualization-) ready way by combining individual files into single, homogenized objects, thereby easing access, extraction, analysis, and fusion. Essentially, datacubes unify spatio-temporal sensor, image (timeseries, simulation, and statistics data under a common modelling and servicing paradigm, independent from the variety of raster encodings utilized. Also, the datacube paradigm serves to homogenize across dimensions, allowing unified wrangling of 1-D sensor data, 2-D imagery, 3-D x/y/t image timeseries and x/y/z geophysics voxel data, and 4-D x/y/z/t climate and weather data. On the side, there is a close relationship to statistical datacubes as common in OLAP, for example.

In OGC and ISO standardization, coverages provide the unifying concept for spatio-temporal datacubes, with the streamlined service model of Web Coverage Service (WCS) including Web Coverage Processing Service (WCPS), OGC's geo datacube analytics language. A large, continuously growing number of open-source and proprietary tools support the coverage standards. In parallel, in 2019 the SQL standard has been enhanced with datacube functionality.

In this comprehensive tutorial, which is suitable for newcomers and remote sensing experts alike, we present the concept of datacubes and their contribution to analysis-ready data mentioning, e.g., the

CARD4L initiative, relevant standards established in various bodies, as well as interoperability successes and issues existing. We inspect implementations and discuss their individual shortfalls and benefits. Based on the OGC reference implementation, rasdaman, live demos accessing existing services and real-life examples which participants can recap and modify on their Internet-connected laptop will play a key role. Slides will be made available to participants, including extensive link lists.

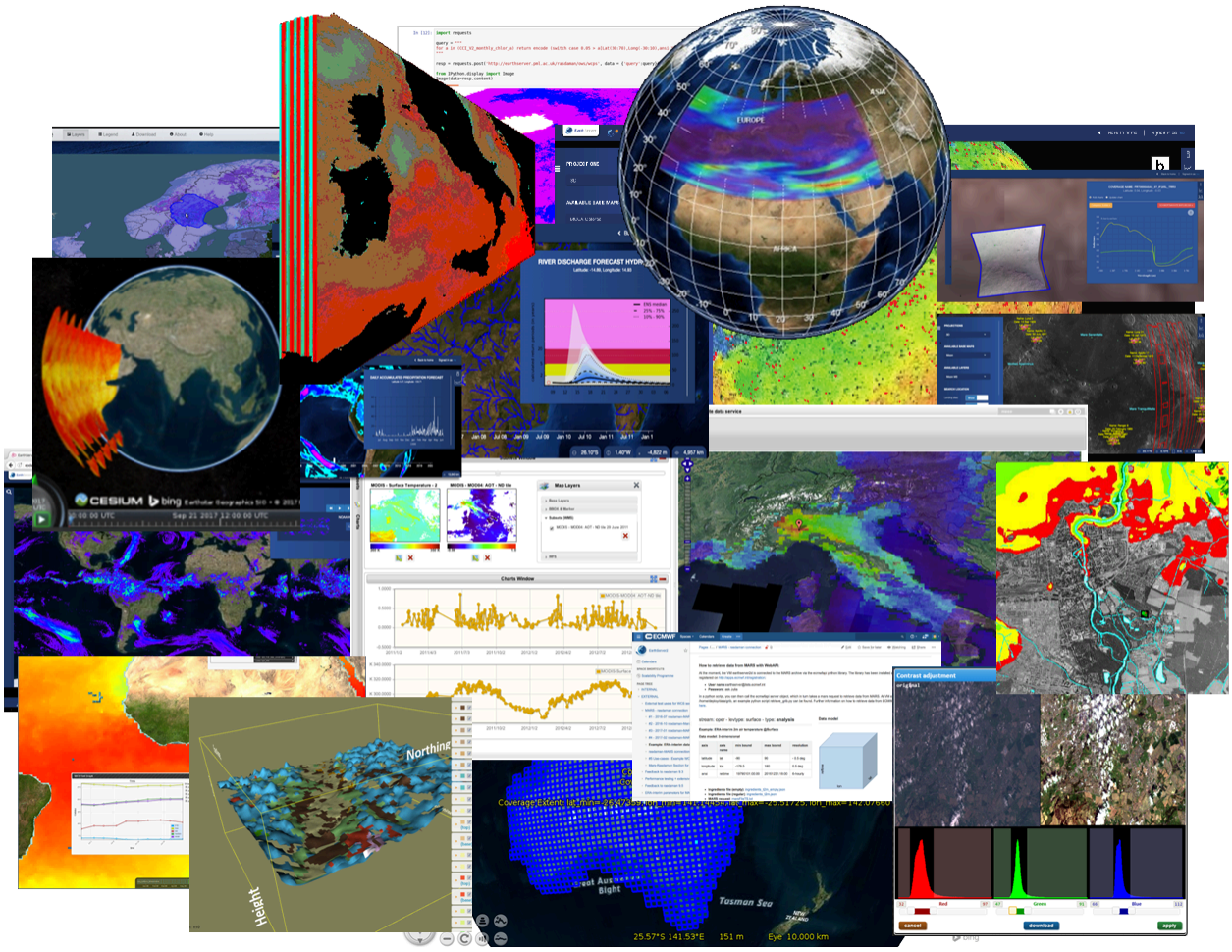

Figure: Screenshots of sample datacube portals

Proposed Agenda

- Introduction to datacubes: concepts, benefits
- Datacube standards in OGC, ISO, INSPIRE
- OGC / ISO Coverage data & service model
- Getting hands dirty: how to model common raster types as coverages
- Datacube technology survey
- Datacube architecture deep dive: the OGC datacube reference implementation, rasdaman
- Outlook: Spatio-temporal Machine Learning boosted by datacubes
- Wrap-Up & discussion

Proposed Duration

Half day or full day

Expected Target Audience and Number:

The tutorial can host up to 30 participants. We invite participation from diverse areas:

- Students who want to get a glance at latest developments in Earth science service technology
- Experts in remote sensing who want to keep abreast of recent developments
- Data providers seeking to enhance the quality of their offerings to better serve existing and attract new users, based on open standards
- Service developers seeking guidance on the implementation of standards
- Scientists seeking more efficient and less time consuming methods of understanding Big Data

Learning Objectives

After this workshop, participants will be able to

- Describe the concept of spatio-temporal data cube services and their added value for data providers and users.
- Correlate the terms homogenized data, analysis-ready data, and datacubes.
- Summarize the state of standardization in datacubes, centered around the notion of spatio-temporal coverages
- Model common raster data types as coverages
- Formulate common datacube tasks as OGC WCS / WCPS requests
- Understand core implementation considerations for achieving flexibility and scalability, and differences between various datacube engines
- Interact with sample standards-based datacube services in 1-D through 4-D scenarios.

About the presenter

Peter Baumann is Professor of Computer Science at Jacobs University, Germany, and founder and CEO of research spin-off rasdaman GmbH. He has pioneered actionable datacubes and the rasdaman (“raster data manager”) Array DBMS which is in operational use internationally. Baumann is editor of the OGC and ISO datacube standards which are implemented by the major open-source and proprietary raster server tools.

In OGC and ISO TC211 he is editor of the "Big Earth Datacube" data and service model standards. The modular Web Coverage Service (WCS) datacube standards suite, including the WCPS which also have been adopted by the European common Spatial Data Infrastructure, INSPIRE. Upon his initiative, ISO SC32 has established SQL/MDA (Multi-Dimensional Arrays) extending the SQL standard with domain-independent datacubes, based on the rasdaman array query language. Throughout his 15+ years

of shaping standards he is actively pursuing outreach and education about the standards and their practical use.

OGC has honored his contribution to Big Data standardization with the prestigious Kenneth Gardels Award for "significant and enduring advances" in technical standards. See his homepage for more information, including the 160+ publication list.