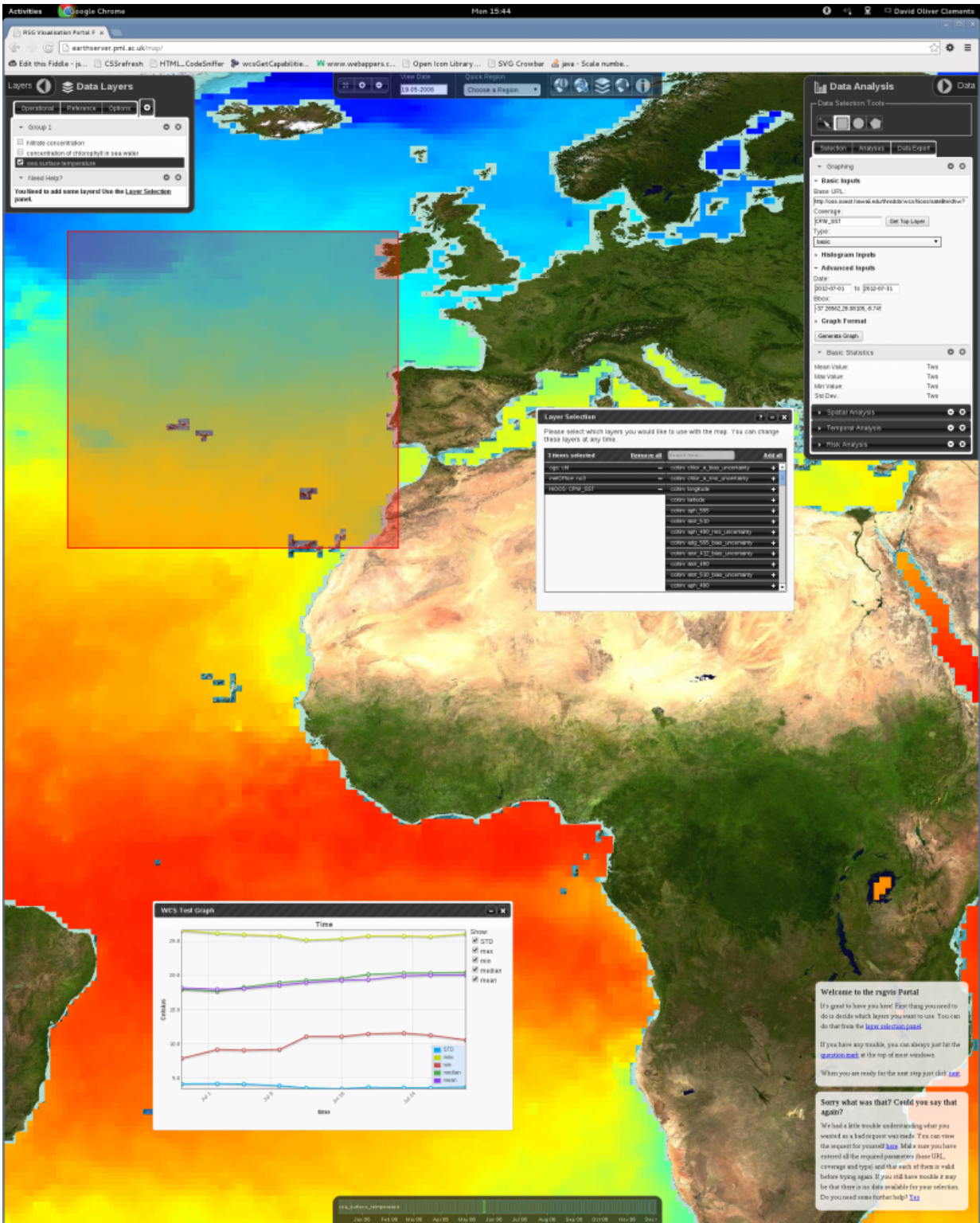


Big Earth data at your fingertips becomes a reality

June 24 2015



Credit: EARTHSERVER

Pushing the boundaries of Big Earth Data processing, the EARTHSERVER project allows researchers access and analyse multi-dimensional data from a wide range of sources.

The [earth sciences](#), like geology, oceanography and astronomy, generate vast quantities of Big Data. Yet without the right tools scientists either drown in this sea of Big Earth Data or it sits in an archive, barely used.

The vision of the EARTHSERVER [project](#) is to offer researchers 'Big Earth Data at your fingertips' so that they can access and manipulate enormous data sets with just a few mouseclicks.

'The project was the result of a 'push' and a 'pull',' says project coordinator Peter Baumann, Professor of Computer Science at Jacobs University in Bremen, Germany. 'On the demand side there was a need for new concepts to handle the wave of data crashing down on us. On the supply side we had a data cube technology that is well-suited to this domain.' A data cube is a three- (or higher) dimensional array of values, commonly used to describe a time series of image data.

Data cubes help researchers access and visualise data

EARTHSERVER built advanced data cubes and custom web portals to make it possible for researchers to extract and visualise earth sciences data as 3-D cubes, 2-D maps or 1-D diagrams. The British Geological Survey, for example, used EARTHSERVER technology to drill down through different layers of the earth in 3-D.

'For the user, data cubes hide the unnecessary complexity of the data,' says Professor Baumann. 'As a user, I don't want to see a million files: I want to see a few data cubes.'

The massive data in the earth sciences is represented by sensor, image,

simulation, and statistics data, often with a time dimension. The data typically form regular or irregular grid values with space/time coordinates. EARTHSERVER made these arrays available as data cubes.

Aside from ease-of-use, the data cubes also made it possible to integrate data from different disciplines, and scientists could combine measurement data with data generated from simulations.

Building on existing technologies

To handle Big Earth Data efficiently, EARTHSERVER needed to extend existing technologies and standards. The SQL database query language, for example, is more oriented towards the manipulation of alphanumeric data.

To enable data cubes, the project was built upon rasdaman, a new type of database management system specialised in multi-dimensional gridded data, called rasters or arrays. Rasdaman enables the flexible, fast extraction of data from Big Earth Data arrays of any size.

'Essentially, we have married the SQL database language with image processing,' says Professor Baumann. 'This is now becoming part of the ISO SQL standard.'

In addition, the project has strongly influenced the Big Earth Data standards of the Open Geospatial Consortium and INSPIRE, the European Spatial Data Infrastructure.

EARTHSERVER's researchers also developed a 'semantic parallelisation' technology that sub-divides a single database query into multiple sub-queries. These are sent to other database servers for processing.

This method allows EARTHSERVER to distribute a single incoming query over more than 1 000 cloud nodes and rapidly answer queries on hundreds of Terabytes in less than a second.

Provided by CORDIS

Citation: Big Earth data at your fingertips becomes a reality (2015, June 24) retrieved 26 April 2024 from <https://phys.org/news/2015-06-big-earth-fingertips-reality.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.