

Fostering Cross-Disciplinary Earth Science Through Datacube Analytics

Peter Baumann, Angelo Pio Rossi, Brennan Bell, Oliver Clements, Ben Evans, Heike Hoenig, Patrick Hogan, George Kakalettris, Panagiota Koltsida, Simone Mantovani, Ramiro Marco Figuera, Vlad Merticariu, Dimitar Misev, Huu Bang Pham, Stephan Siemen, and Julia Wagemann

With the unprecedented increase of orbital sensor, in situ measurement, and simulation data there is a rich, yet not leveraged potential for obtaining insights from dissecting datasets and rejoining them with other datasets. Obviously, goal is to allow users to “ask any question, any time, on any size”, thereby enabling them to “build their own product on the go”.

One of the most influential initiatives in EO is EarthServer which has demonstrated new directions for flexible, scalable EO services based on innovative NoSQL

P. Baumann • V. Merticariu • D. Misev
rasdaman GmbH, Bremen, Germany

Jacobs University, Bremen, Germany
e-mail: p.baumann@jacobs-university.de; v.merticariu@jacobs-university.de;
d.misev@jacobs-university.de

A.P. Rossi • B. Bell • R. Marco Figuera • H.B. Pham
Jacobs University, Bremen, Germany
e-mail: an.rossi@jacobs-university.de; b.bell@jacobs-university.de;
r.marcofiguera@jacobs-university.de; b.phamhuu@jacobs-university.de

O. Clements
Plymouth Marine Laboratory, Plymouth, UK
e-mail: olcl@pml.ac.uk

B. Evans
National Computational Infrastructure (NCI), Australian National University, Canberra, ACT,
Australia
e-mail: Ben.Evans@anu.edu.au

H. Hoenig
rasdaman GmbH, Bremen, Germany
e-mail: hoenig@rasdaman.com

P. Hogan
NASA Ames, Moffett Field, CA, USA
e-mail: Patrick.Hogan@nasa.gov

technology. Researchers from Europe, the USA and Australia have teamed up to rigorously materialize the concept of the datacube. Such a datacube may have spatial and temporal dimensions (such as an $x/y/t$ satellite image time series) and may unite an unlimited number of scenes. Independently from whatever efficient data structuring a server network may perform internally, users will always see just a few datacubes they can slice and dice.

EarthServer has established client and server technology for such spatio-temporal datacubes. The underlying scalable array engine, *rasdaman*, enables direct interaction, including 3D visualization, what-if scenarios, common Earth Observation data processing, and general analytics. Services exclusively rely on the open OGC “Big Geo Data” standards suite, the Web Coverage Service (WCS). Phase 1 of EarthServer has advanced scalable array database technology into 100+ TB services; in Phase 2, Petabyte datacubes are being built for ad-hoc extraction, processing, and fusion.

But EarthServer has not only used, but also shaped several Big Data standards. This includes OGC coverage data and service standards, INSPIRE WCS, and the ISO Array SQL candidate standard.

We present the current state of EarthServer in terms of services and technology and outline its impact on the international standards landscape.

Introduction

The term “Big Data” is a contemporary shorthand characterizing data which are too large, fast-lived, heterogeneous, or complex to be understood and exploited. Technologically, this is a cross-cutting challenge affecting storage and processing, data and metadata, servers and clients and mash-ups. Further, making new, substantially more powerful tools available for simple use by non-experts while not constraining complex tasks for experts just adds to the complexity. All this holds for many application domains, but specifically so for the field of Earth Observation (EO). With the unprecedented increase of orbital sensor, in situ measurement, and simulation data there is a rich, yet not leveraged potential for acquiring insights from dissecting datasets and rejoining them with other datasets. The stated goal is

G. Kakaletis • P. Koltsida
CITE s.a, Attiki, Greece
e-mail: g.kakaletis@cite.gr; p.koltsida@di.uoa.gr

S. Mantovani
MEEO s.r.l., Ferrara, Italy
e-mail: mantovani@meeo.it

S. Siemen • J. Wagemann
ECMWF, Reading, UK
e-mail: Stephan.Siemen@ecmwf.int; julia.wagemann@ecmwf.int

to enable users to “ask any question, any time, on any volume” thereby enabling them to “build their own product on the go”.

In the field of EO, one of the most influential initiatives towards this goal is EarthServer (Baumann et al. 2015a; EarthServer 2015) which has demonstrated new directions for flexible, scalable EO services based on innovative NoSQL technology. Researchers from Europe, the USA and Australia have teamed up to rigorously materialize the concept of the datacube. Such a datacube can have spatial and temporal dimensions (such as a satellite image timeseries) and is able to unite an unlimited number of single images. Independent from whatever data structuring a server network may perform internally for efficiency on the millions of hyperspectral images and hundreds of climate simulations, users will always see just a few datacubes they can slice and dice.

EarthServer has established a slate of services for such spatio-temporal datacubes based on the scalable array engine, rasdaman, which enables direct interaction, including 3D visualization, what-if scenarios, common EO data processing, and general analytics. All services strictly rely on the open OGC data and service standards for “Big Geo Data”, the Web Coverage Service (WCS) suite. In particular, the Web Coverage Processing Service (WCPS) geo raster query language has proven instrumental as a client data programming language which can be hidden behind appealing visual interfaces.

EarthServer has advanced these standards based on experience gained. The OGC WCS standards suite in its current, comprehensive state has been largely shaped by EarthServer which provides the Coverages, WCS, and WCPS standards editor and working group chair. The feasibility evidence provided by EarthServer has contributed to the uptake of WCS by open-source and commercial implementers; meanwhile, OGC WCS has been adopted by INSPIRE and has entered the adoption process of ISO.

Phase 1 of EarthServer has ended in 2014 (Baumann et al. 2015a); independent experts characterized the outcome, based on “proven evidence”, that rasdaman will “significantly transform the way that scientists in different areas of Earth Science will be able to access and use data in a way that hitherto was not possible”. And “with no doubt” this work “has been shaping the Big Earth Data landscape through the standardization activities within OGC, ISO and beyond”. In Phase 2, which started in May 2015, this is being advanced even further: from the 100 TB database-size achieved in Phase 1 over the currently more than 500 TB, the next frontier will be crossed by building Petabyte datacubes for ad-hoc querying and fusion (Fig. 1).

In this contribution we present status and intermediate results of EarthServer and outline its impact on the international standards landscape. Further, we highlight opportunities established through technological advance and how future services can cope better with the Big Data challenge in EO.

The remainder of this contribution is organized as follows. In section “Standards-Based Modelling of Datacubes”, the concepts of the OGC datacube and its service standards are introduced. An initial set of services in the federation is presented in section “Science Data Services”, followed by an introduction to the

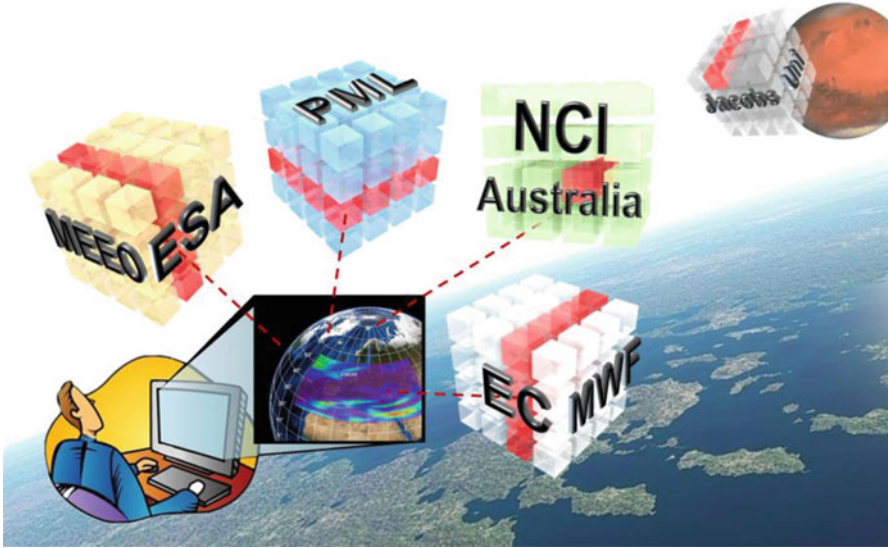


Fig. 1 Intercontinental datacube mix and match in the EarthServer initiative (Source: EarthServer)

underlying technology platform and an evaluation in section “Datacube Analytics Technology”. Section “Conclusion and Outlook concludes the plot with an outlook.

Standards-Based Modelling of Datacubes

EarthServer relies on the OGC “Big Earth Data” standards, WCS and WCPS, as the only client/server protocols for any kind of access and processing; additionally, WMS is offered. In the server, all such requests uniformly get mapped to an array query language which we will introduce later. Advanced visual clients enable point-and-click interfaces effectively hiding the query language syntax, except when experts want to make use of it. Additionally, access through expert tools like python notebooks is provided.

At the heart of the EarthServer conceptual model is the concept of *coverages* as digital representations of space/time varying phenomena as per ISO 19123 (ISO 2004) (which is identical to OGC Abstract Topic 6). Practically speaking, coverages encompass regular and irregular grids, point clouds, and general meshes. The datacube concept, being based on multidimensional arrays, represents a subset of coverages that focuses on regular and irregular spatio-temporal grids (Fig. 2).

The notion of *coverages* (Baumann 2012; Baumann and Hirschorn 2015; Wikipedia 2016) has proven instrumental in unifying spatio-temporal regular and

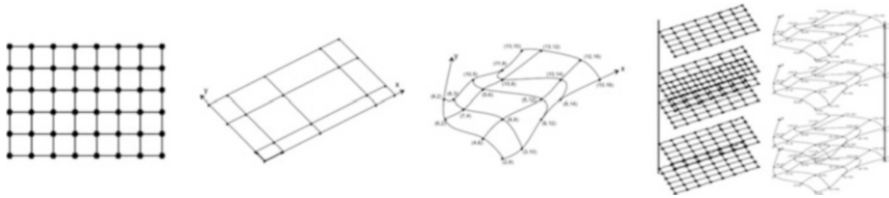


Fig. 2 Sample datacube grid types supported by rasdaman: regular grids (*left*), irregular and warped grids (*center left and right*), multidimensional combinations (*right*) of regular and irregular grid axes (Source: OGC/Jacobs University)

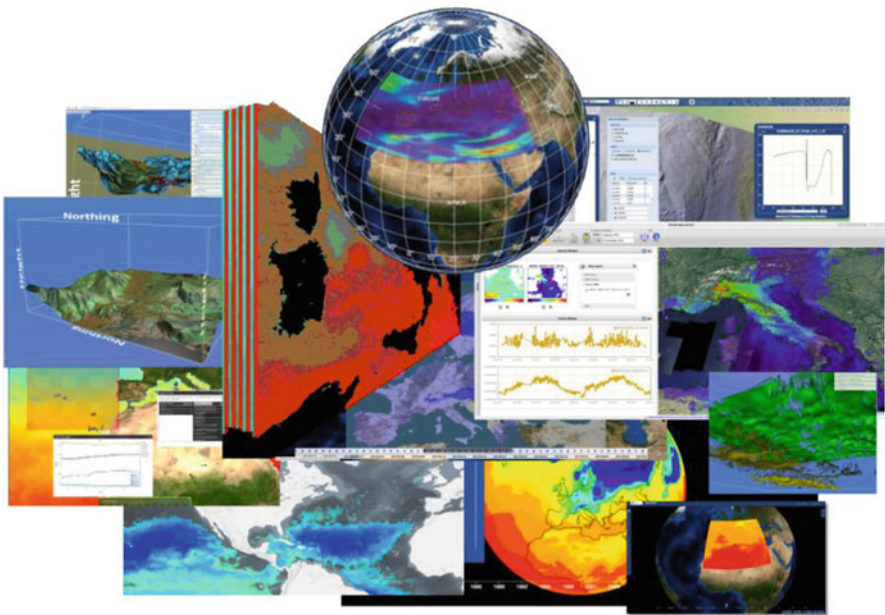


Fig. 3 WCS/WCPS based datacube services utilizing rasdaman (Source: rasdaman/EarthServer)

irregular grids, point clouds, and meshes so that such data can be accessed and processed through a simple, yet flexible and interoperable service paradigm.

By separating coverage data and service model, any service—such as WMS, WFS, SOS and WPS—can provide and consume coverages. That said, the *Web Coverage Service* (WCS) standard offers the most comprehensive, streamlined functionality (OGC 2016a). This modular suite of specifications starts with fundamental data access in WCS Core and has various extensions adding optionally implementable functionality facets, up to server-side analytics based on the *Web Coverage Processing Service* (WCPS) geo datacube language (Fig. 3) (Baumann 2010b).

Below we introduce the OGC coverage data and service model with an emphasis on practical aspects and illustrate how they enable high-performance, scalable implementations.

Coverage Data Model

According to the common geo data model used by OGC, ISO, and others, objects with a spatial (possibly temporal) reference are referred to as *features*. A special type of features are *coverages* whose associated values vary over space and/or time, such as an image where each coordinate leads to an individual color value. Complementing the (abstract) coverage model of ISO 19123 on which it is based, the (concrete) OGC coverage data and service model (Baumann 2012) establishes verifiable interoperability, down to pixel level, through the OGC conformance tests. While concrete, the coverage model still is independent from data format encodings—something which is of particular importance as it allows a uniform handling metadata, and individual mappings to the highly diverse metadata handling of the various data formats.

The OGC coverage model (and likewise WCS) meantime is supported by most of the respective tools, such as open-source MapServer, GeoServer, OPeNDAP and ESRI ArcGIS. In 2015, this successful coverage model has been extended to allow any kind of irregular grids, resulting in the OGC *Coverage Implementation Schema* (CIS) 1.1 (Baumann and Hirschorn 2015) which is in the final stage of adoption at the time of this writing. Different types of axes are made available for composing a multidimensional grid in a simple plug-and-play fashion. This effectively allows to concisely represent coverages ranging from unreferenced over regular grids to irregularly spaced axes (as often occurring in timeseries) and warped grids to ultimately algorithmically determined warpings, such as those defined by SensorML 2.0.

Web Coverage Service

The OGC service definition specifically built for deep functionality on coverages is the Web Coverage Service (WCS) suite of specifications. With WCS Core (Baumann 2010a), spatio-temporal subsetting as well as format encoding is provided; this Core must be supported by all implementations claiming conformance. Figure 4 illustrates WCS subsetting functionality, Fig. 5 shows the overall architecture of the WCS suite. Conformance testing of WCS implementations follows the same modularity approach and involves detailed checks, essentially down to the level of single cell (e.g. “pixel”, “voxel”) values (OGC 2016b). In December 2016, the European legal framework for a common Spatial Data Infrastructure, INSPIRE, has adopted WCS as coverage download service (INSPIRE 2016).

Fig. 4 WCS subsetting: trimming (*left*) and slicing (*right*) (Source: OGC)

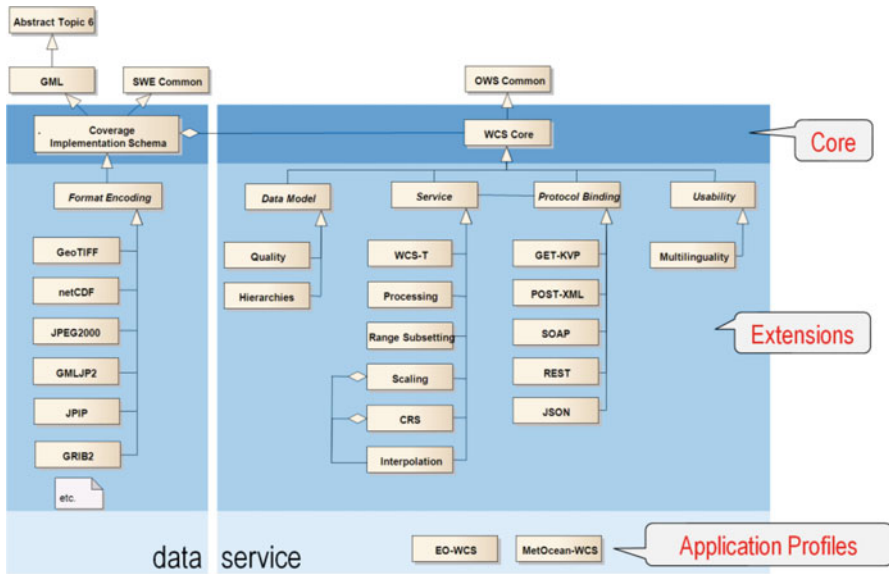
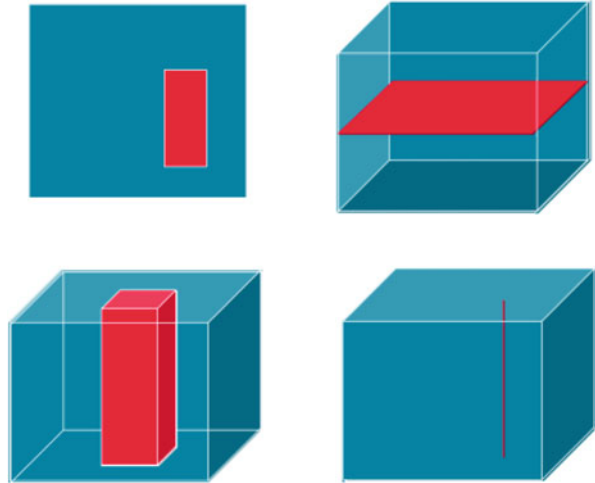


Fig. 5 Overall WCS suite architecture (Source: OGC)

Web Coverage Processing Service

Web Coverage Processing Service (WCPS) is OGC’s geo raster query language, adopted already in 2008 (Baumann 2010b). An example may illustrate the use of WCPS: “From MODIS scenes M1, M2, M3, the difference between red & near-infrared bands, encoded as TIFF—but only those where near-infrared exceeds threshold 127 somewhere.” The corresponding query reads as follows:

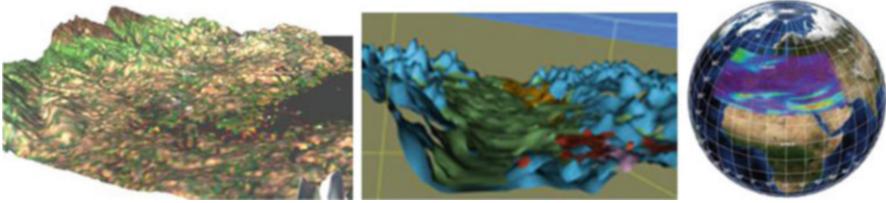


Fig. 6 3D rendering of datacube query results (data and service: BGS, server: rasdaman)

```
for $c in doc("http://acme.com/wcs")//coverage
where some( $c.nir > 127 )
return encode( $c.red - $c.nir, "image/tiff" )
```

Such results can conveniently be rendered through WebGL in a standard Web browser, or through NASA WorldWind (Fig. 6). The syntax is close to SQL/MDA (see below), but with a syntax flavor close to XQuery so as to allow integration with XPath and XQuery, for which a specification draft is being prepared by EarthServer (see section on data/metadata integration further down).

The Role of Standards

As the hype dust settles down over “Big Data” the core contributing data structures and their particularities crystallize. In Earth Science data, these arguably are regular and irregular grids, point clouds, and meshes, reflected by the coverage concept. The unifying notion of coverages appears useful as an abstraction that is independent from data formats and their particularities while still capturing the essentials of spatio-temporal data. With CIS 1.1, description of irregular grids has been simplified by not looking at the grids, but at the axis characteristics. While many services on principle can receive or deliver coverages, the WCS suite is specifically designed to not only work on the level of whole (potentially large) objects, but can address inside objects as well as filter and process them, ranging up to complex analytics with WCPS.

The critical role of flexible, scalable coverage services for spatio-temporal infrastructures is recognized far beyond OGC, as the substantial tool support highlights. This has prompted ISO and INSPIRE to also adopt the OGC coverage and WCS standards. Also, ISO is extending the SQL standard with n -D arrays (ISO 2015; Misev and Baumann 2015). The standards observing group of the US Federal Geographic Data Committee (FGDC) sees coverage processing a la WCS/WCPS as a future “mandatory standard”. In parallel, work is continuing in OGC towards extending coverage world with further data format mappings and to add further relevant functionality, such as flexible XPath-based coverage metadata retrieval. Finally, research is being undertaken on embedding coverages into the Geo Semantic Web (Baumann et al. 2015b), also supporting W3C which has

started studying coverages in the “Spatial Data on the Web” Working Group. A demonstration service for 1D through 5D coverages is available for studying the WCS/WCPS universe (rasdaman 2016a).

Science Data Services

Expertise of the EarthServer project partners covers multiple scientific domains. This ensures that benefits achieved can be made available to the largest possible audience. Partners include Plymouth Marine Laboratory (PML) running the Marine Data Service, Meteorological and Environmental Earth Observation (MEEO) operating the Sentinel Earth Observation Service, National Computing Infrastructure (NCI) Australia running the LandSat Earth Observation Service, European Centre for Medium-Range Weather Forecasts (ECMWF) with its Climate Data Service, and Jacobs University providing the Planetary Data Service. Based on the common EarthServer platform provided by the technology partners Jacobs University, rasdaman GmbH, CITE s.a., and NASA all the aforementioned service partners have set up domain specific clients and data access portals which are continuously advanced and populated over the lifetime of the project so as to cross the Petabyte frontier for single services in 2017. Multiple service synergies will be explored which will allow users to query and analyze data stored at different project partner’s infrastructure from a single entry point. An example of this is the LandSat service being developed jointly by MEEO and NCI. The specific data portals and access options are detailed in the following sections.

Earth Observation Data Services

The use of EO data is getting more and more challenging with the advent of the Sentinel era. The free, full and open data policy adopted for the Copernicus programme foresees access available to all users for the Sentinel data products. Terabytes of data and EO products are already generated every day from Sentinel-1 and Sentinel-2, and with the approaching launch of the Sentinel-3/-4/-5P/-5, the need of advanced access services is crucial to support the increasing data demand from the users.

The Earth Observation Data Service (Alfieri et al. 2013; MEEO 2017) offers dynamic and interactive access functionalities to improve and facilitate the accessibility to massive Earth Science data: key technologies for data exploitation (Multi-sensor Evolution Analysis (2016), (rasdaman 2016b), NASA Web World Wind (NASA 2016)) are used to implement effective geospatial data analysis tools empowered with the OGC standard interfaces for Web Map Service (WMS) (De la Beaujardiere 2016), Web Coverage Service (WCS) (Baumann 2010a), and Web Coverage Processing Service (WCPS) (Baumann 2008)—see Fig. 7.

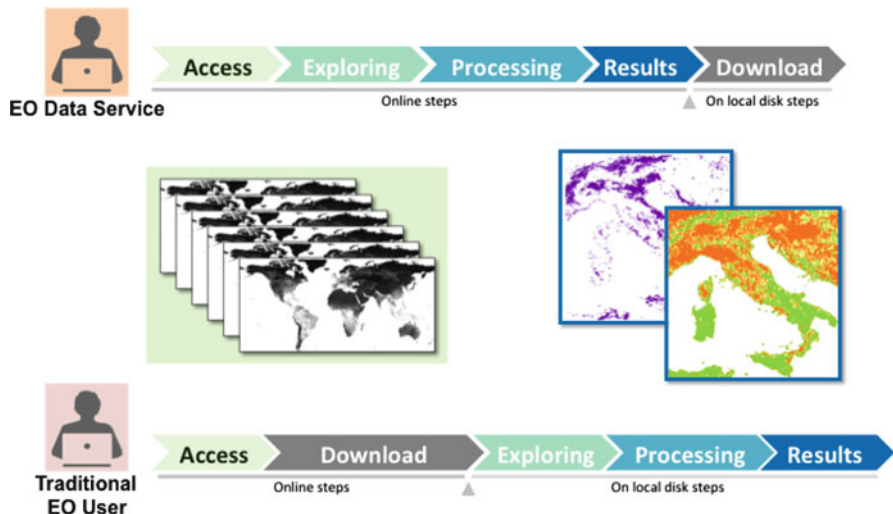


Fig. 7 Data exploitation approaches offered by traditional (*bottom*) and EO Data Service (*top*) approaches (Source: MEEO)

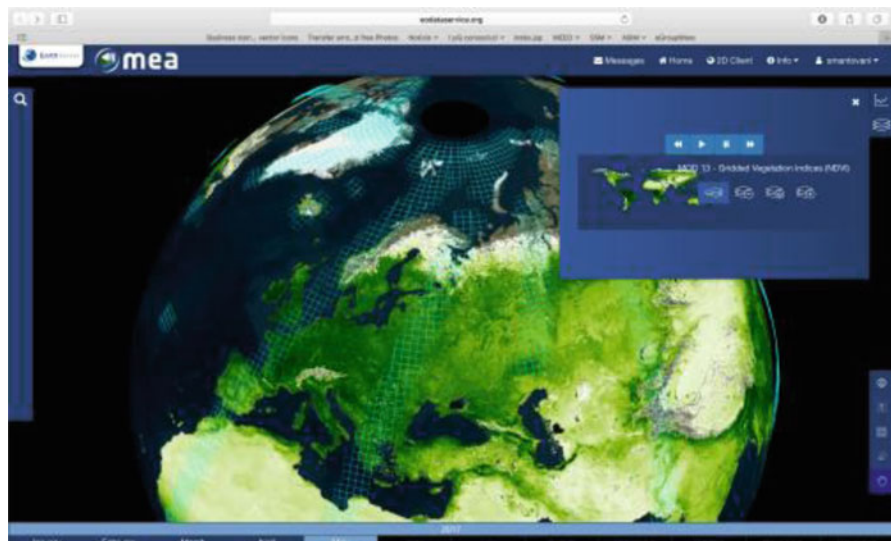


Fig. 8 Availability of global Sentinel 2A data on May 19th, 2017 on top of MODIS normalize difference vegetation index (NDVI) visualized in the ESA/NASA Web World Wind virtual globe (Source: MEEO)

With respect to the traditional data exploitation approaches, the EO Data Service supports on-line data interaction, restructuring the typical steps and moving to the end the download of the real data of interest for the users with a significant reduction of data transfer (Figs. 8 and 9).

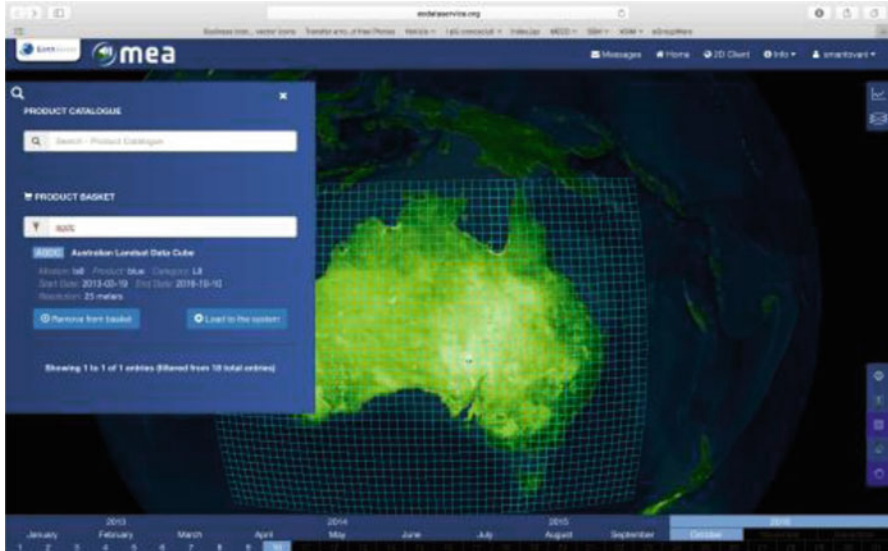


Fig. 9 Australian Landsat Data Cube coverage, presented in the ESA/NASA WebWorldWind virtual globe. Users can select areas of interest and explore Landsat data available at the National Computational Infrastructure (NCI) Australia (Source: MEE0)

The EO Data Service currently provides in excess of one PB of ESA and NASA EO products (e.g. vegetation indexes, land surface temperature, precipitation, soil moisture, etc.) to support Atmosphere, Land and Ocean applications.

In the framework of the EarthServer initiative, the Big Data Analytics tools are being enabled on datacubes of Copernicus Sentinel and Third Party Missions (e.g. Landsat8) data, coming from MEE0 and its federation partner NCI Australia, to support agile analytics and fusion on these new generation sensors through MEE0's service (Fig. 10).

Marine Science Data Service

The marine data service (Marine Data Service) is focused on providing access to remote sensed ocean data. The data available are from ocean colour satellites. The marine research community is well accustomed to using satellite data. Satellite data provides many benefits over in situ observations. The data have a global coverage and provide a consistent and accurate time series of data. The marine research community has recognized the benefit of long time series of data. Time series need to be consistent so that the data are comparable through the whole series. Remote sensed data have helped to provide this consistency.

The ESA OC-CCI project (Sathyendranath et al. 2012) is producing a time series of remote sensed ocean colour parameters and associated uncertainty variables.

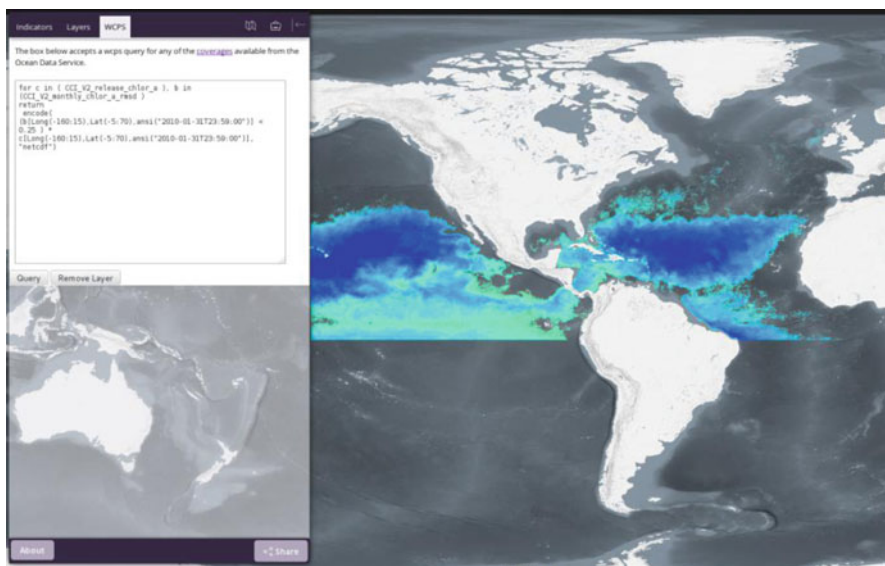


Fig. 10 Screenshot showing GIS client displaying chlorophyll data selected based on the per pixel value of uncertainty criteria, together with corresponding WPCS query (*left*) (Source: PML)

Currently the available time series runs from 1997 to 2015 and represents 1 of 14 subgroups of the overall ESA CCI project. With the creation of these large time series an increasingly technical challenge has emerged, how do users get benefit from these huge data volumes?

The EarthServer project, through the use of a suite of technologies including rasdaman and several OGC standard interfaces, aims to address the issue of users having to transfer and store large data volumes by offering ad-hoc querying over the whole data catalog.

Traditionally a marine researcher would simply select the particular temporal and spatial subset of the dataset they require from a web based catalog and download to their local disk. This system has worked well but is becoming less feasible due to the increases in data volume and the increase in non-specialists wanting access to the data. Take for example a researcher interested in finding the average monthly chlorophyll concentration for the North Sea for the period 2000–2010. Traditional methodologies would require the download of around several gigabytes of data. This represents a large time investment for the actual download as well as a cost associated with storage and processing required (Clements and Walker 2014). By making the same dataset available through the EarthServer project a research can simply write the analysis as a WPCS query and send that to the data service. The analysis is done at the data and only the result is downloaded, in this case around 100 KB. This example outline the clearest cut advantage, however there are more transient benefits that could improve the way that researchers interact with and use data. One example of this would be the testing of novel algorithms that require access to the raw light reflectance data. These data are used through existing

algorithms to calculate derived products such as chlorophyll concentration, primary production and carbon sequestration.

The marine data service currently provides in excess of 70 TB of data. Through the course of the project we will be expanding the data offering to include data from the ESA Sentinel 3 Ocean and Land Colour Instrument (OLCI) (Berger et al. 2012). The aim is to offer as much data from the sensor as is available with the total goal to offer 1 PB of data through the service.

Climate Science Data Service

The Climate Science Data Service is developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). ECMWF hosts the Meteorological Archival and Retrieval System (MARS), the largest archive of meteorological data worldwide with currently more than 170 PB of data (ECMWF 2014). As a Numerical Weather Prediction (NWP) centre, ECMWF primarily supports the meteorological community through well-established services for accessing, retrieving and processing data from the MARS archive. User outside the MetOcean domain, however, often struggle with the climate-specific conventions and formats, e.g. the GRIB data format. This limits the overall uptake of ECMWF data. At the same time, with data volumes in the range of Petabytes, data download for processing on users' local workstations is no longer feasible. ECMWF as a data provider has to find solutions to provide efficient web-based access to the full range of data while at the same time the overall data transport is minimized. Ideally, data access and processing takes place on the server and the user only downloads the data that is really needed.

ECMWF's participation in EarthServer-2 aims at addressing exactly this challenge: to give users access to over 1 PB of meteorological and hydrological data and at the same time providing tools for on-demand data analysis and retrieval. The approach is to connect the rasdaman server technology with ECMWF's MARS archive, thereby enabling access to global reanalysis data via the OGC-based standards Web Coverage Service (WCS) and Web Coverage Processing Service (WCPS). This way, multidimensional gridded meteorological data can be extracted and processed in an interoperable way.

The climate reanalysis service particularly addresses users outside the MetOcean domain, more familiar with common Web and GIS standards. A WC(P)S for climate science data can be of benefit for developers or scientists building Web-applications based on large data volumes, who are unable to store all the data locally. Technical data users for example can integrate a WCS request into their processing routine and further process the data. Companies can easily build customized web-applications with data provided via a WCS. This approach is also strongly promoted by the EU's Copernicus EO programme which generates climate and environmental data as part of its operational services. Companies can use these data for value-added climate services for decision-makers or clients (Fig. 11).

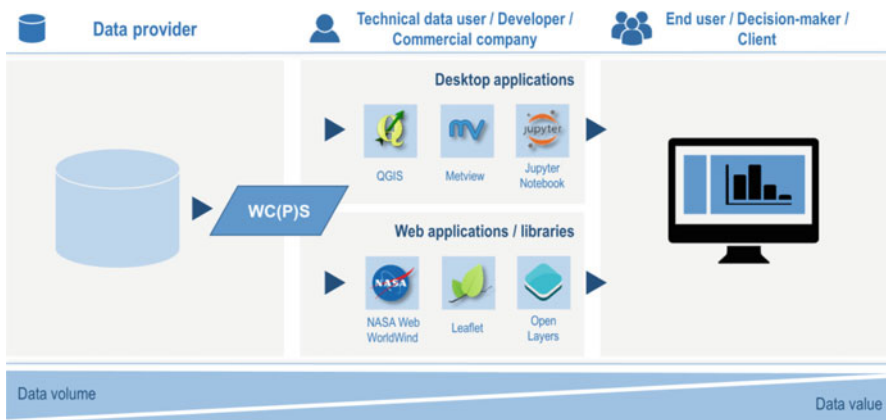


Fig. 11 Example of how a WC(P)S can be integrated into standard processing chains (Source: ECMWF)

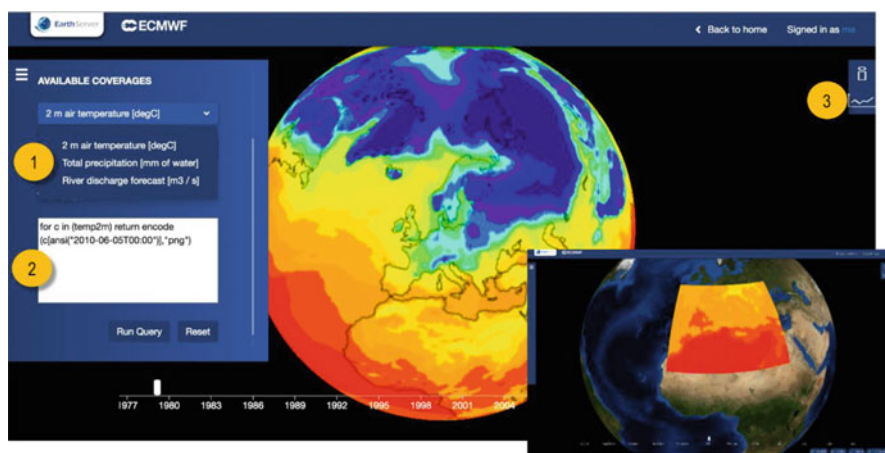


Fig. 12 WebWorldWind client, with three main functionalities: (1) 3D visualization, (2) writing own WCPS queries to choose a coverage subset (compare inlet) and (3) plotting of time series/hydrograph of selected latitude/longitude information (Source: ECMWF)

To showcase how simple it is to build a custom web application with the help of a WC(P)S, a demo web client visualizing ECMWF data with NASA WorldWind has been developed (ECMWF [n.d.](#)) giving access to currently three datasets: ERA-interim 2 m air temperature and total accumulated precipitation (Dee et al. 2011) as well as GloFAS river discharge forecast data (Alfieri et al. 2013). Two-dimensional global datasets can be mapped on the globe (Fig. 12). An additional plotting functionality allows retrieval of data points in time for individual coordinates. This is suitable for ERA-interim time-series data and hydrographs based on river-discharge forecast data (Fig. 13).

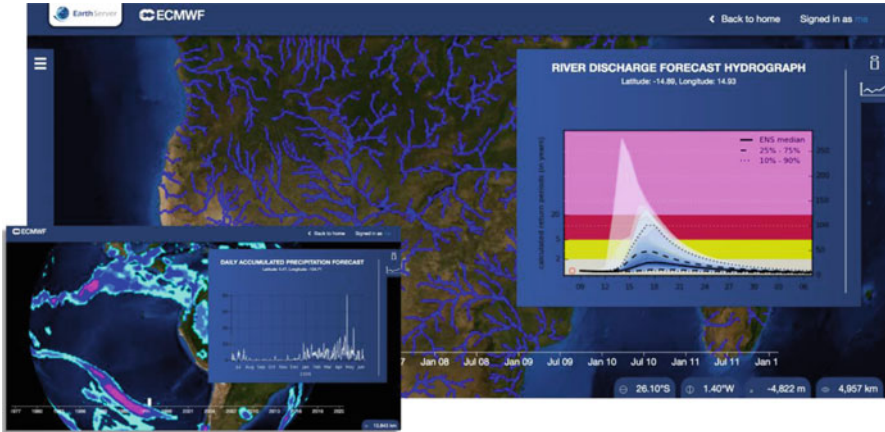


Fig. 13 Sample plotting functionalities. The main image shows a hydrograph plotted based on daily river discharge forecast data. The inlet shows plotting of ERA-interim time series data. The plot shows the total accumulated precipitation for one lat/lon grid point for 1 year (Source: Jacobs University)

In summary, WCS for Climate Data offers a facilitated on-demand access to ECMWF’s climate reanalysis data for researchers, technical data users and commercial companies, within the MetOcean community and beyond.

Planetary Science Data Service

Planetary Science missions are largely based on Remote Sensing experiments, whose data are very much comparable with those from Earth Observation sensors. Data are thus relatively similar in terms of data structure and type: from panchromatic, to multispectral or hyperspectral data, as well as derived datasets such as stereo-derived topography, or laser altimetry, in terms of surface imaging (Oosthoek et al. 2013), in addition to subsurface vertical radar sounding (Cantini et al. 2014), or atmospheric imaging and profiles. The vast majority of these data can be represented with raster models, thus they are suitable for use in array databases.

Planetary raster data have never much suffered from being closed in archives during last decades: all remote sensing imagery returned by spacecrafts is available in the public domain, together with documentation (e.g. (Heather et al. 2013; McMahon 1996)). Nevertheless, archived data are typically lower-level, unprocessed or partially processed images and cubes, not GIS- and science-ready products. In addition, they typically are analyzed as single data granules or with cumbersome processing and analyzing pipelines to be carried out by individual scientists, on own infrastructure.

What is also slightly challenging for the access, integration and analysis of Planetary Science data is the wide range of bodies in terms of surface (or atmosphere) nature, experimental characteristics and Coordinate Reference Systems. The sheer volume of data, counted in few GB for entire missions (such as NASA Viking orbiters) until the 1980s, is now approaching the order of magnitude of tens to hundreds of TB.

All these aspects tend to give a Big Data dignity to Planetary datasets, too. The planetary community expressed the need during the past decade of easier and more effective ways to access and analyze its wealth data (Pondrelli et al. 2011). Most web services to date addressed the availability of maps (e.g. with WMS), but not extensively the deeper access, in terms of analysis and analytics to the complexity and richness of planetary datasets. WCPS demonstrated the capability to address this (Oosthoek et al. 2013; Rossi et al. 2014).

The Planetary Science Data Service (PSDS) of EarthServer, also known as PlanetServer (2016a), focuses on complex multidimensional data, in particular hyperspectral imaging and topographic cubes and imagery. All of those data derive from public archives and are processed to the highest level with publicly available routines.

In addition to Mars data (Rossi et al. 2014), WCPS is offered on diverse datasets on the Moon, as well as Mercury. Other Solar System Bodies are also going to be covered and served. Derived parameters such as hyperspectral summary products and indices can be produced through WCPS queries, as well as derived imagery color combination products.

One of the objectives of PlanetServer is to translate scientific questions into standard queries that can be posed to either a single granule/coverage, or an extremely large number of them, from local to global scale. The planetary and remote sensing and geodata communities at large could benefit from PlanetServer at different levels: from accessing its data and performing analyses with its web services, for research or education purposes; to using and adapting or iterating further the concepts and tools developed within PlanetServer.

PlanetServer in its new iteration is completely based on open source software, and its code available on GitHub (PlanetServer 2016b). The main server component empowering PlanetServer-2 is rasdaman community edition, and its visualization engine is the NASA WorldWind virtual globe (Fig. 14) (Hogan 2011).

A sample, nontrivial WCPS query for returning an RGB combination from MRO CRISM hyperspectral imaging data for the mineral compositional indices *sindex2*, *BD2100_2*, and *BD1900_2* mapped to RGB as described by Viviano-Beck et al. (2014) is the following one, with null values set to transparent:

```
for data in (last_ingestion_2)
return encode(
{
  red:    (int) (255/(max(data.band_233)-min(data.band_233)))
          * (data.band_233 - min(data.band_233));
  green:  (int) (255/(max(data.band_13)-min(data.band_13)))
          * (data.band_13 - min(data.band_13));
  blue:   (int) (255/(max(data.band_78)-min(data.band_78)))
```

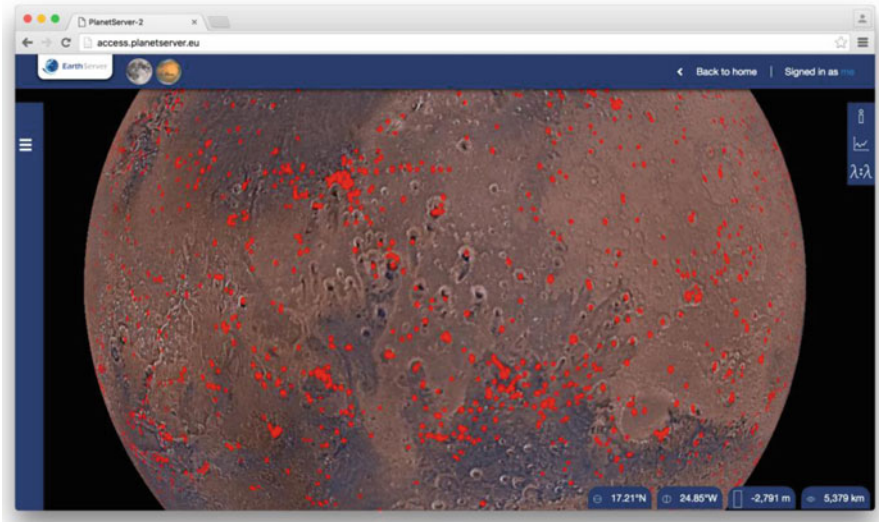


Fig. 14 PlanetServer showing a Mars globe based on Viking Orbiter imagery mosaics produced by the United States Geological Survey (USGS), served from its rasdaman database draped on the WebWorldWind virtual globe using mosaicked NASA LRO mission data (Source: Jacobs University).

```

        * (data.band_78 - min(data.band_78)) ;
    alpha: (int)(data.band_100 > 0) * 255
},
"png", "nodata=65535"
)

```

The result of this query is a map-projected subset of a cube highlighting compositional variations on the Surface (Fig. 15).

Cross-Service Federation Queries

Among the features of the EarthServer platform, consisting of metadata-enhanced rasdaman (see next subsection), is the capability to federate services. Technically, this is only a generalization of the service internal parallelization and distributed processing; externally, it achieves location transparency allowing users to send any query to any data center, regardless of which data are accessed and possibly combined, including across data center boundaries.

A lab prototype of this federation has been demonstrated live at EGU 2015 and AGU 2016 where a nontrivial query required combination of climate data from ECMWF in the UK with LandSat 8 imagery at NCI Australia. This query was alternately sent to ECMWF and NCI; each of the receiving services forked a

Fig. 15 WCPS query result from the RGB combination red: `sindex2`, green: `BD2100_2`, blue: `BD1900_2` (Source: PlanetServer)

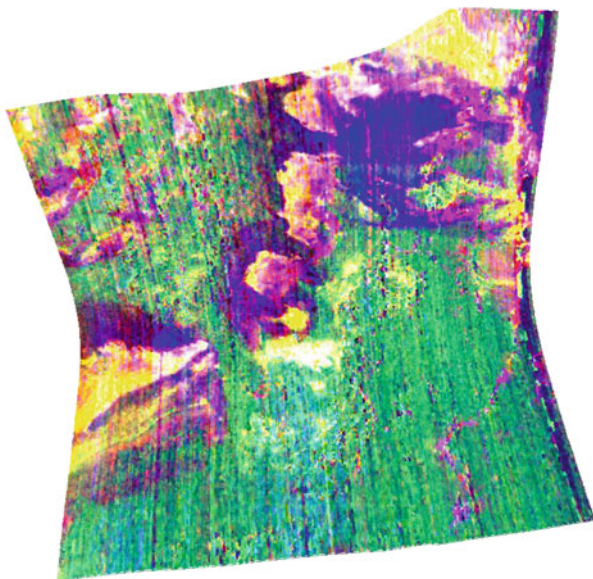


Fig. 16 Visualization of query splitting: original query (*left*), query distribution from Germany to the UK, with subquery spawned to Australia (*center*), query result visualized in NASA WorldWind (Source: EarthServer)

subquery to the service holding the data missing locally. The result was displayed in NASA WebWorldWind, allowing to visually assess equality of the results. Figure 16 shows part of the query, a visualization of the path the query fragments take, and the final result mapped to a virtual globe.

Datacube Analytics Technology

EarthServer uses a combination of Big Data storage, processing, and visualization technologies. In the backend, this is the rasdaman Array Database system which we introduce in the next section. Data/metadata integration plays a crucial role in the

EarthServer data management approach and is presented next. Finally, the central visualization tool, the NASA WorldWind virtual globe, is presented.

Array Databases as Datacube Platform

The common engine underlying EarthServer is the rasdaman Array Database (Baumann et al. 1999). It extends SQL with support for massive multidimensional arrays, together with declarative array operators which are heavily optimized and parallelized (Dumitru et al. 2014) on server side. A separate layer adds geo semantics, such as knowledge about regular and irregular grids and coordinates, by implementing the OGC Web service interfaces. For OGC and INSPIRE WCS, as well as OGC WCPS, rasdaman acts as reference implementation. On storage, arrays get partitioned (“tiled”) into sub-arrays which can be stored in a database or directly in files. Additionally, rasdaman can access preexisting archives by only registering files, without copying them. Figure 1 shows the overall architecture of rasdaman.

Array Storage

Arrays are maintained in either a conventional database (such as PostgreSQL) or its own persistent store directly in any kind of file system. Additionally, rasdaman can tap into “external” files not under its control. Since rasdaman 9.3, an internal tiling of archive files (such as available with TIFF and NetCDF, for example) can be exploited for fine-grain reading. Under work is automated distribution of tiles based on various criteria, optionally including redundancy (Fig. 17).

A core concept of array storage in rasdaman is partitioning or *tiling*. Arrays are split into sub-arrays called *tiles* to achieve fast access. Tiling policy is a tuning parameter which allows adjusting partitions to any given query workload, measured or anticipated. As this mechanism turned out very powerful for users, its generality has been cast into a few strategies available to data designers (Fig. 18).

Array Processing

The rasdaman server (“rasserver”) is the central workhorse. It can access data from various sources for multi-parallel, distributed processing. The rasdaman engine has been crafted from scratch, optimizing every single component for array processing. A series of highly effective optimizations is applied to queries, including:

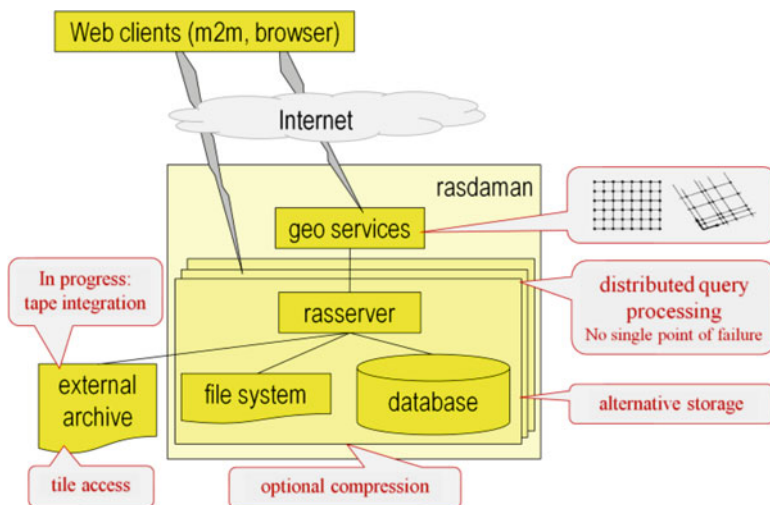


Fig. 17 rasdaman overall architecture (Source: rasdaman)

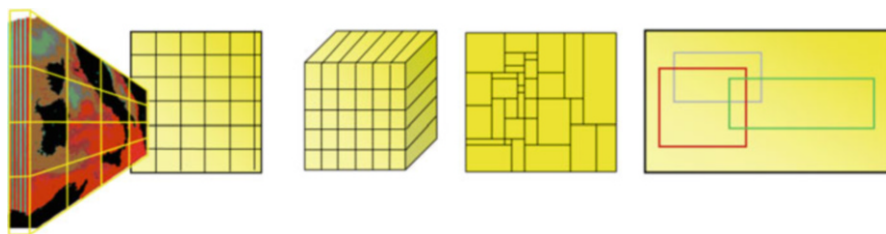


Fig. 18 Sample tiling rasdaman strategies supported (Source: rasdaman)

- Query rewriting to find more efficient expressions of the same query; currently 150 rewriting rules are implemented.
- Query result caching is used to keep complete or partial query results in (shared) memory for reuse by subsequent queries; in particular, geographic or temporal overlap can be exploited.
- Array joins with optimized tile loading so as to minimize multiple loads when combining two arrays (Baumann and Merticariu 2015). This is not only effective in a local situation, but also when tiles have to be transported between compute nodes or even data centers in case of a distributed join.

After query analysis and optimization, the system fetches only the tiles required for answering the given query. Subsequent processing is highly parallelized. Locally, it assigns tiles to different CPUs and threads. In a cluster, query are split and parallelized across the nodes. The same mechanism is also used for distributing processing across data centers, where data transport becomes a particular issue. To maximize efficiency, rasdaman currently optimizes splitting along two criteria

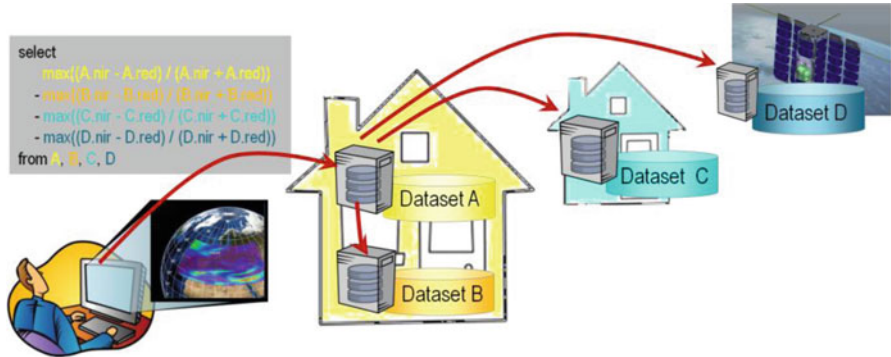


Fig. 19 rasdaman query splitting (Source: rasdaman)



Fig. 20 Visualization workbench for rasdaman distributed query processing (Source: rasdaman)

(Fig. 19): First, send queries to where the data sit (“shipping code to data”); second, generate subqueries that process as much as ever possible locally, minimizing the amount of data to be transported between nodes.

This way, single queries have been successfully split across more than a thousand Amazon cloud nodes (Dumitru et al. 2014). Figure 20 shows an experiment done on the rasdaman distributed query processing visualization workbench where nine Amazon nodes process a query on 1 TB processed in 212 ms.

Tool Integration

Even though the WCS, WCS, and WCPS protocols are open, adopted standards, they are not necessarily appropriate for end users—from WMS we are used to have Web clients like OpenLayers and Leaflet which hide the request syntax, and the same holds for WCS requests and, although high-level and abstract, the WCPS language. In the end, all these interfaces are most useful as client/server

communication protocols where end users are hidden from the syntax through visual point-and-click interfaces (like OpenLayers and NASA WorldWind) or, alternatively, through their own, well-known tools (like QGIS and python).

To this end, rasdaman already supports major GIS Web and programmatic clients, and more are under development. Among this list are MapServer, GDAL, EOxServer, OpenLayers, Leaflet, QGIS, and NASA WorldWind, C++, and Java. Python is in advanced development stage.

The Role and Handling of Metadata

Metadata can be of utmost importance for the utilization of datasets, as apart from textual descriptions and provenance traces, it may provide essential information on how data may be consumed or interpreted (e.g. characteristics of equipment/process, reference systems, error margins). When data management crosses the boundaries of systems, institutions and scientific disciplines, metadata management becomes a complex process on its own. The Earth-Sciences landscape is an ample example where datasets, which are substantially “many”, may be considered from a variety of standpoints, and be produced/consumed by heterogeneous processes in various disciplines with diverse needs and concepts.

Focusing on coverages hosted behind WCS and WCPS services, where metadata heterogeneity is evident due to the liberal approach of the relevant specifications, the EarthServer 2 metadata management system addresses the challenge, by being metadata schema agnostic yet maintaining the ability to host and process composite metadata models. Meanwhile, the system seeks to meet a number of supplementary requirements such as fault-tolerance, efficiency and scalability, looking to a (near) future where hosting billions of datasets will be common case.

The system supports of two modes of operation, with quite distinct characteristics (a) *in situ operation* (metadata are not relocated and services are offered on top of the original store’s metadata retrieval ones) and (b) *federated operation* (metadata are gathered in a distributed store over which the full range of system services may be provided).

The architecture (cf. Fig. 21) consists of loosely coupled distributed services that interoperate through standards, WCS and WCPS being the fundamental ones. XPath is utilized for metadata retrieval/filtering, over NoSQL technologies in order to achieve the desired scalability, performance and functional characteristics. Full text queries are also supported. In federated mode, services are invoked using WCPS or WCS-T standards. Other supported protocols include OpenSearch, OAI-PMH and CSW.

Access to the combined processing and retrieval engine is provided via *xWCPS2.0*, a specification that leverages the agile earth-data analytics layer with effective metadata retrieval and processing facilities, delivering an expressive

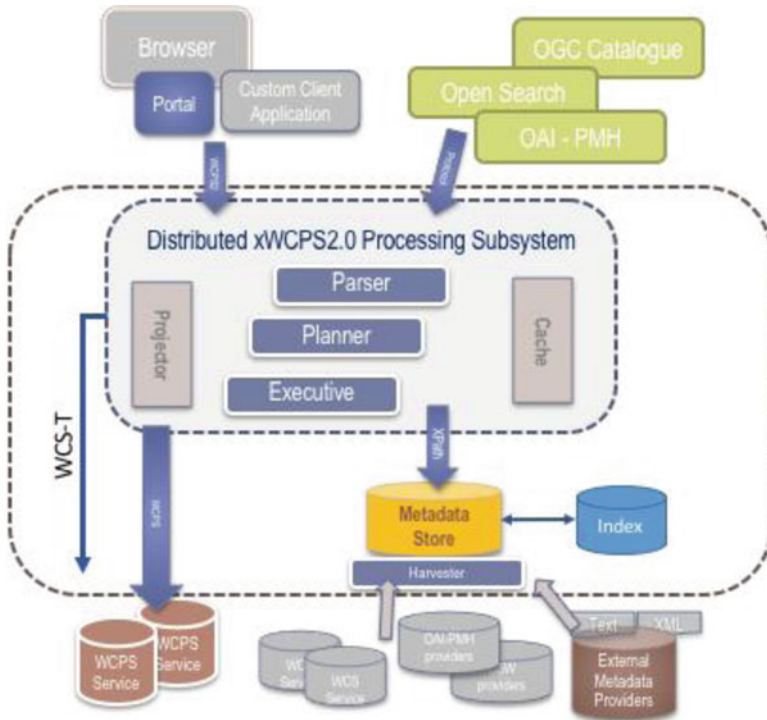


Fig. 21 xWCPS overall architecture (Source: CITE).

querying tool that can interweave data and metadata in composite operations. xWCPS 2.0 builds on xWCPS1.0 (from EarthServer-1) and, apart from an enhanced FLWOR syntax, it delivers features that significantly enhance the ability to issue federated queries.

In the following xWCPS2.0 example, coverages across all federated servers (“*”) are located via their metadata (name of <field> is elevation in where clause) and results consist of xml elements (<result> in return clause), containing the outcome of an XPath expression (metadata) and a WCPS evaluated element (value):

```

for $c in *
where $c:://*[local-name()='field'][@name=elevation]
return
  <result>
    <value>
      $c[Lat(53.08),Long(8.80),ansi("2014-01":"2014-12")]
    </value>
    <metadata>$c:://domainSet</metadata>
  </result>
    
```

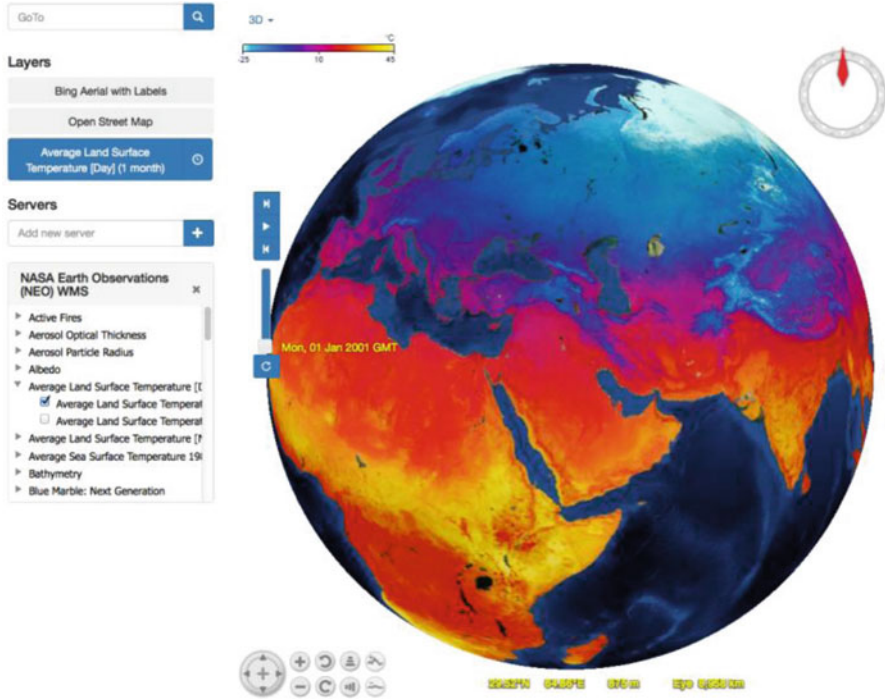


Fig. 22 NASA World Wind with data mapping (Source: NASA).

Virtual Globes as Datacube Interfaces

Visual globes help users experiencing their data visually with the various aspects displayed in their native context. This allows data to be more easily understood and their impacts better appreciated.

NASA is a pioneer in virtual globe technology, substantially preceding tools such as Google Earth. Our primary mission has always been to support the operational needs of the geospatial community through a versatile open source toolkit, versus a closed proprietary product. A particular feature of WorldWind is its modular and extensible architecture. WorldWind as an Application Programming Interface, API-centric Software Development Toolkit (SDK) can be plugged into any application that has spatial data needing to be experienced in the native context of a virtual globe (Fig. 22).

In EarthServer, the virtual globe paradigm is coupled with the flexible query mechanism of databases. Users can query rasdaman flexibly and have the results mapped to the globe. Rasdaman applications can add any 2D, 3D or 4D information to the WorldWind geobrowser for any dynamically generated query result. This enables a direct interaction with massive databases, as the excerpt of interest is prepared in the server while WorldWind accomplishes sophisticated interactive

visualization in the native context of Earth as observed from space, thereby providing access to the various thematic EarthServer databases; with PlanetServer, WorldWind is also used for Mars, Moon and further solar system bodies.

Related Work

A large, growing number of both open-source and proprietary implementations is supporting coverages and WCS (Fig. 3). Specifically, the most recent version (OGC Coverage Implementation Schema 1.0 and WCS 2.0) are known to be implemented by open-source rasdaman (2016b), GDAL, QGIS, OpenLayers, Leaflet, OPeNDAP, MapServer, GeoServer, GMU, NASA WorldWind, EOxServer as well as proprietary Pyxis, ERDAS and ArcGIS. The most comprehensive tool is rasdaman—also OGC WCS Core Reference Implementation—which implements WCS Core and all extensions, including WCPS. This large adoption basis of OGC’s coverage standards promotes interoperability of EarthServer with other services, supporting the GEOSS “system of systems” approach (Christian 2005). Notably, rasdaman is part of the GCI (GEOSS Common Infrastructure) (GEOSS 2016).

Google Earth Engine (Google n.d.) builds on the tradition of Grid systems. Users can submit python code which is executed transparently in a distributed processing environment. However, procedural code does not parallelize easily, therefore—after discussion with the rasdaman team—developers have added a declarative “Map Algebra” interface in addition which resembles a subset of an array query language. Still, many common techniques (like query compilation, heuristic rewriting, cost-based optimizations, adaptive tiling, data compression, etc.) are not being utilized—in the end, a substantial advantage comes from using the massive underlying Google hardware.

SciDB is an Array Database prototype under development (Paradigm4 2016) with no specific geo data support like OGC WCS interfaces. SciQL is a concept study adding arrays to a column store (Zhang et al. 2011). A performance comparison between rasdaman, SciQL, and SciDB shows that rasdaman excels by one, often several orders of magnitude in performance and also conveys better storage efficiency (Merticariu et al. 2015). To the best of our knowledge, only rasdaman has publicly available services deployed (Baumann et al. 2015b). No particular SciDB support for Earth data is known—the only supported ingest format is CSV (comma-separated values), and geo semantics is not available in queries.

Sensor Observation Service (SOS) supports delivery of sensor data (Bröring et al. 2012) which can be imagery. However, there is rather limited functionality, and performance is reported as not entirely satisfactory.

OGC WMTS exposes tiling to clients for maximizing performance (Masó et al. 2010); on the downside, queries are fixed to retrieval of such tiles, i.e. there is no free subsetting and no processing. OGC WPS provides an API for arbitrary processing functionality, however, is not interoperable per se as stated already in the standard (Schut 2007).

In ISO, an extension to SQL is in advanced stage which adds n -D arrays in a domain-independent manner (ISO 2015). SQL/MDA (for “Multidimensional Arrays”) has been initiated by the rasdaman team, which also has submitted the specification; see (Misev and Baumann 2015) for a condensed overview. Adoption is anticipated for summer 2017.

Conclusion and Outlook

Datacubes are a convenient model for presenting users with a simple, consolidated view on the massive amount of data files gathered—“a cube tells more than a million images”. Such a datacube may have spatial and temporal dimensions (such as a satellite image time series) and may unite an unlimited number of individual images. Independently from whatever efficient data structuring a server network may perform internally, users will always see just a few datacubes they can slice and dice.

Following the broadening of minds through the NoSQL wave, database research has responded to the Big Data deluge with new data models and scalability concepts. In the field of gridded data, Array Databases provide a disruptive innovation for flexible, scalable data-centric services on datacubes. EarthServer exploits this by establishing a federation of services of 3D satellite image timeseries and 4D climatological data where each node can answer queries on the whole network, in a federation implementing a “datacube mix and match”. While in Phase 1 of EarthServer the 100 TB barrier has been transcended, in its Phase 2 it is attacking the Petabyte frontier.

Aside from using the OGC “Big Geo Data” standards for its service interfaces, EarthServer keeps on shaping datacube standards in OGC, ISO, and INSPIRE. Current work involves implementation of the OGC coverage model version 1.1, supporting data centers in establishing rasdaman-based services, and enhancing further the data and processing parallelism capabilities of rasdaman.

Acknowledgement The EarthServer initiative is partly funded by the European Commission under grant agreement FP7 286310 and H2020 654367.

References

- Alfieri L et al (2013) GloFAS - global ensemble streamflow forecasting and flood early warning. *Hydrol Earth Syst Sci* 17:1161–1175
- Andrejev A, Baumann P, Misev D, Risch T (2015) Spatio-temporal gridded data processing on the semantic web. *IEEE Intl. Conf. on Data Science and Data Intensive Systems (DSDIS 2015)*, Sydney, Australia, December 11–13, 2015

- Baumann P (2008) OGC Web Coverage Processing Service (WCPS) Language, version 1.0. OGC document 08-068r2. www.openeospatial.org/standards/wcps. Accessed 21 May 2017
- Baumann P (2010a) OGC WCS interface standard – core, version 2.0. OGC document 09-110r4. www.openeospatial.org/standards/wcs. Accessed 21 May 2017
- Baumann P (2010b) The OGC web coverage processing service (WCPS) standard. *Geoinformatica* 14(4):447–479
- Baumann P (2012) OGC Coverage Implementation Schema (formerly: GML 3.2.1 application schema – coverages), version 1.0. OGC document 09-146r2. www.openeospatial.org/standards/wcs. Accessed 21 May 2017
- Baumann P, Hirschorn E (2015) OGC coverage implementation schema, version 1.1, OGC document 09-146r3
- Baumann P, Merticariu V (2015) On the efficient evaluation of array joins. Proc. IEEE Big Data Conf. Workshop Big Data in the Geo Sciences, Santa Clara, US, October 29, 2015
- Baumann P, Dehmel A, Furtado P, Ritsch R, Widmann N (1999) Spatio-temporal retrieval with RasDaMan. Proc. VLDB, Edinburgh, Scotland, UK, September 7–10, 1999
- Baumann P, Mazzetti P, Ungar J, Barbera R, Barboni D, Beccati A, Bigagli L, Boldrini E, Bruno R, Calanducci A, Campalani P, Clement O, Dumitru A, Grant M, Herzig P, Kakaletis G, Laxton J, Koltsida P, Lipskoch K, Mahdiraji A, Mantovani S, Merticariu V, Messina A, Misev D, Natali S, Nativi S, Oosthoek J, Passmore J, Pappalardo M, Rossi A, Rundo F, Sen M, Sorbera V, Sullivan D, Torrisi M, Trovato L, Veratelli M, Wagner S (2015a) Big data analytics for earth sciences: the EarthServer approach. *Int J Dig Earth* :1–27
- Baumann P, Mazzetti P, Ungar J, Barbera R, Barboni D, Beccati A, Bigagli L, Boldrini E, Bruno R, Calanducci A, Campalani P, Clement O, Dumitru A, Grant M, Herzig P, Kakaletis G, Laxton J, Koltsida P, Lipskoch K, Mahdiraji AR, Mantovani S, Merticariu V, Messina A, Misev D, Natali S, Nativi S, Oosthoek J, Passmore J, Pappalardo M, Rossi AP, Rundo F, Sen M, Sorbera V, Sullivan D, Torrisi M, Trovato L, Veratelli MG, Wagner S (2015b) Big data analytics for earth sciences: the EarthServer approach. *Int J Dig Earth* :1–27
- Berger M, Moreno J, Johannessen J, Levelt P (2012) ESA’s sentinel missions in support of Earth system science. *Remote Sens Environ* 120:84–90
- Bröring A, Stasch C, Echterhoff J (2012) OGC® sensor observation service interface standard, version 2.0. OGC document 12-006
- Cantini F, Rossi A, Orosei R, Baumann P, Misev D, Oosthoek J, Beccati A, Campalani P, Unnithan V (2014) MARSIS data and simulation exploited using array databases: Planet-Server/EarthServer for sounding radars. *Geophys Res Abstr* 16:EGU2014-3784
- Christian E (2005) Planning for the global earth observation system of systems (GEOSS). *Space Policy* 21(2):105–109
- Clements O, Walker P (2014) Can EO afford big data – an assessment of the temporal and monetary costs of existing and emerging big data workflows. Proc. EGU, Vienna, Austria, April 27–May 02, 2014
- De la Beaujardiere J (2016) Web Map Service (WMS), version 1.3. OGC document 06-042. www.openeospatial.org/standards/wms. Accessed 21 May 2017
- Dee D et al (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Q J Roy Meteorol Soc* 137(656):553–597. <https://doi.org/10.1002/qj.828/full>
- Dumitru A, Merticariu V, Baumann P (2014) Exploring cloud opportunities from an array database perspective. Proc ACM SIGMOD Workshop on Data Analytics in the Cloud (DanaC), Snowbird, USA, June 22–27, 2014
- EarthServer (2015) EarthServer: big datacubes at your fingertips. www.earthserver.eu. Accessed 21 May 2017
- ECMWF (2014) Data services leaflet 2014 - commercial license. www.ecmwf.int/sites/default/files/DataServices0leaflet020140commercial0licence.pdf. Accessed 1 Aug 2016
- ECMWF (n.d.) Web Coverage Service (WCS) for Climate Data. earthserver.ecmwf.int. Accessed 21 May 2017

- GEOSS (2016) GEOSS resource details: rasdaman. geossregistries.info/geosspub/resource_details_ns.jsp?compId=urn:geoss:csr:resource:urn:uuid:e8d4eb7b-4681-4f71-9e34-0f78cf10ce92. Accessed 21 May 2017
- Google (n.d.) Google Earth Engine. <https://earthengine.google.com>. Accessed 21 May 2017
- Heather D, Barthelemy M, Manaud N, Martinez S, Szumlas M, Vazquez J, Osuna P (2013) ESA's planetary science archive: status, activities and plans. *Eur Planet Sci Cong 8:EPSC2013-626*
- Hogan P (2011) NASA world wind: infrastructure for spatial data. *Proc. 2nd International Conference on Computing for Geospatial Research & Applications*. ACM, New York, NY
- INSPIRE (2016) Technical Guidance for the Implementation of INSPIRE Download Services using Web Coverage Services (WCS). <http://inspire.ec.europa.eu/id/document/tg/download-wcs>. Accessed 21 May 2017
- ISO (2004) Geographic information - schema for coverage geometry and functions
- ISO (2015) ISO 9075 SQL Part 15: multi-dimensional arrays. ISO candidate standard, 2015
- Masó J, Pomakis K, Julià N (2010) OpenGIS Web Map Tile Service Implementation Standard, version 1.0. OGC document 07-057r7
- McMahon S (1996) Overview of the planetary data system. *Planet Space Sci* 44(1):3–12
- MEE0 (2016) MEA multi-sensor evolution analysis. www.meeo.it/wp/products-and-services/mea-multisensor-evolution-analysis/. Accessed 21 May 2017
- MEE0 (2017) EO data service, eodataservice.org. Accessed 21 May 2017
- Merticariu G, Misev D, Baumann P (2015) Measuring storage access performance in array databases. *Proc. 7th Workshop on Big Data Benchmarking (WBDB)*, December 14–15, 2015, New Delhi, India
- Misev D, Baumann P (2015) Enhancing science support in SQL. *Workshop Data and Computational Science Technologies for Earth Science Research (co-located with IEEE Big Data)*, Santa Clara, US, October 29, 2015
- NASA (2016) Web WorldWind. webworldwind.org. Accessed 21 May 2017
- OGC (2016a) Coverages and web coverage service: the big picture. external.opengeospatial.org/twiki_public/pub/CoveragesDWG/CoveragesBigPicture/. Accessed 21 May 2017
- OGC (2016b) WCS 2.0 (OGC 09-110r4) compliance test suite. cite.opengeospatial.org/teamingengine/about/wcs/2.0.1/site/index.html. Accessed 21 May 2017
- Oosthoek J, Flahaut J, Ross A, Baumann P, Misev D, Campalani P, Unnithan V (2013) PlanetServer: innovative approaches for the online analysis of hyperspectral satellite data from Mars. *Adv Space Res* 53(12):1858–1871
- Paradigm4 (2016) SciDB reference manual: community and enterprise editions. Accessed 21 May 2017
- PlanetServer (2016a) Planetary Science Data Service of EarthServer. www.planetserver.eu. Accessed 21 May 2017
- PlanetServer (2016b) PlanetServer GitHub organization and repositories. github.com/planetserver. Accessed 21 May 2017
- PML (2016) Marine data service. earthserver.pml.ac.uk. Accessed 21 May 2017
- Pondrelli M, Tanaka K, Rossi A, Flamini E (eds) (2011) Geological mapping of Mars. *Planet Space Sci.* 59(11-12) 1113
- rasdaman (2016a) Big Earth Data standards. standards.rasdaman.com. Accessed 21 May 2017
- rasdaman (2016b) rasdaman. www.rasdaman.org. Accessed 21 May 2017
- Rossi A, Oosthoek J, Baumann P, Beccati A, Cantini F, Misev D, Orosei R, Flahaut J, Campalani P, Unnithan V (2014) PlanetServer/EarthServer: big data analytics in planetary science. *Geophys Res Abstr* 16:EGU2014-5149
- Sathyendranath S, Brewin R, Mueller D, Brockmann C, Deschamps P-Y, Doerffer R, Fomferra N, Franz BA, Grant M, Hu C, Krasemann H, Lee Z, Maritorena S, Devred E, Mélin F, Peters M, Smyth T, Steinmetz F, Swinton J, Werdell J, Regner P (2012) Ocean colour climate change initiative: approach and initial results. *Geosci and Remote Sens* 2012:2024–2027
- Schut P (2007) Web processing service, version 1.0. OGC document 05-007r7, www.opengeospatial.org/standards/wps. Accessed 21 May 2017

- Viviano-Beck CE et al (2014) Revised CRISM spectral parameters and summary products based on the currently detected mineral diversity on Mars. *J Geophys Res Planet* 119:1403–1431
- Wikipedia (2016) Coverage data. en.wikipedia.org/Coverage_data. Accessed 21 May 2017
- Zhang Y, Kersten M, Ivanova M, Nes N (2011) SciQL, bridging the gap between science and relational DBMS. *Proc. IDEAS. ACM*, New York, NY, pp 124–133

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

