Handling Heterogeneous EO Datasets via the Web Coverage Processing Service

Piero Campalani [a*], Simone Mantovani [b], Alan Beccati [a], Peter Baumann [a]

[a] Jacobs University Bremen (DE) / [b] MEEO Srl (IT)

*{p.campalani@jacobs-university.de

FOSS4G Europe 2014

OsGeo European Conference on Free and Open-Source Software for Geospatial

Bremen (DE), 15-17th July 2014

Credits

- : The *European Union* research fundings

- : *MEEO Srl* for data servers and use-cases

- :
  EarthServer project partners for feedback and for helping improving our service.

- My team at the *L-SIS* research group ▶▶▶

PREFACE: purpose of our investigations

- Evaluate the processing capabilities of the **OGC WCPS** processing query language on time series of EO image products (~TB).

- Towards a different **paradigm** of geospatial web services: away from simple ROI data visualization.

- Demonstrate and benchmark the designed **use cases** on our WCPS service implementation.

Credits: Clements/Walker, EGU 2014
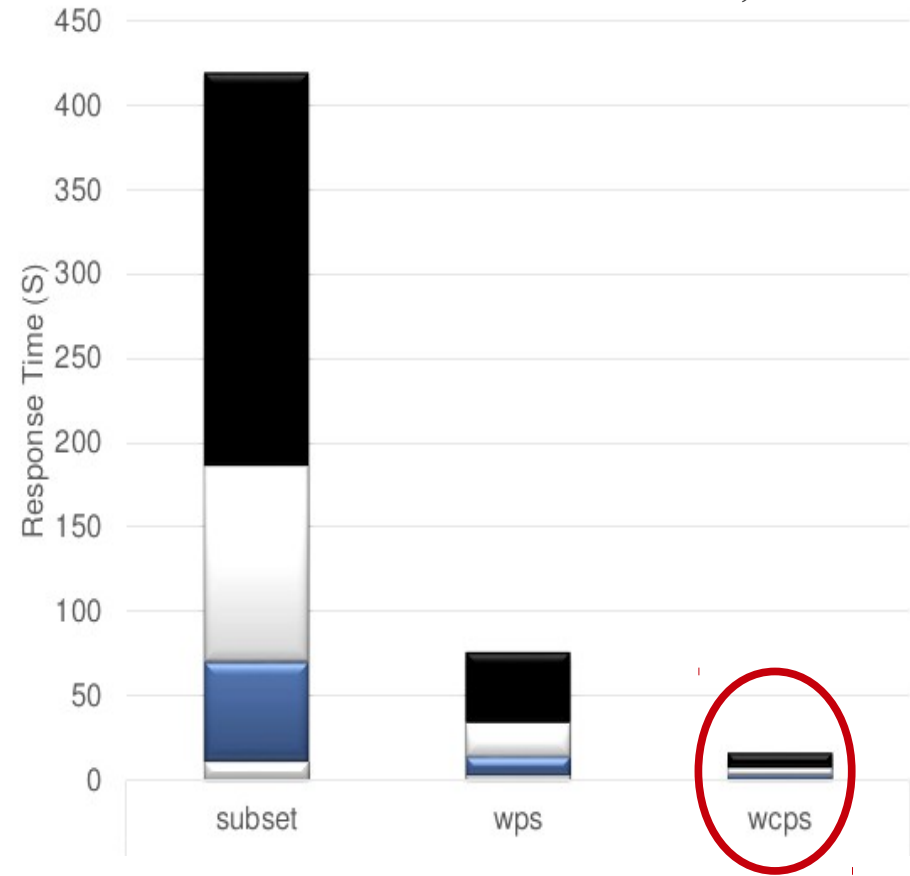
Outline

- **CONTEXT**
  - ▶ `rasdaman` and the *EarthServer* initiative
  - ▶ Multi-sensor Evolution Analysis (MEA) platform

- **OGC OPEN STANDARDS**
  - ▶ Web Coverage Service (WCS)
  - ▶ WCS processing extension (WCPS)

- **USE CASES**
  - ▶ Single-product analysis
  - ▶ Cross-products data comparison
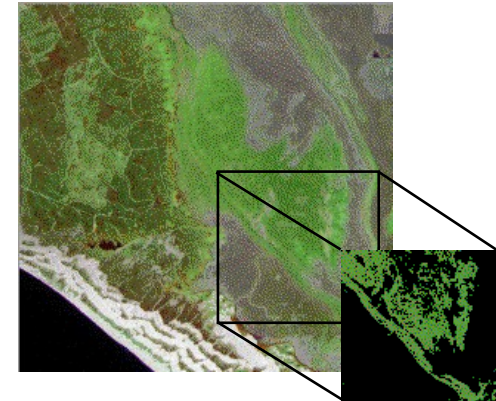 - ▶ Cross-products data merging

- **CONCLUSIONS & OUTLOOK**

context

rasdaman: Agile Array Analytics

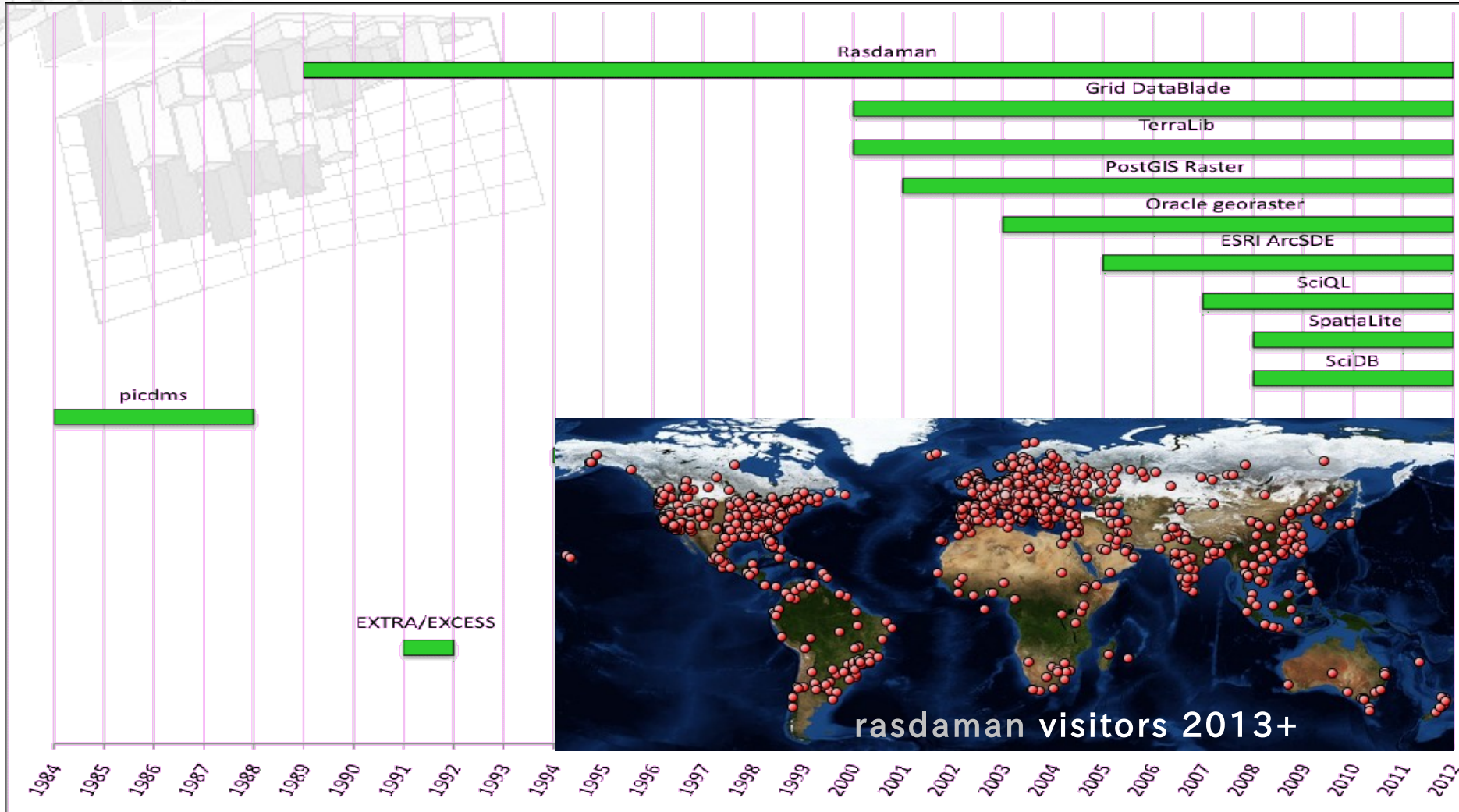- „raster data manager": **SQL + tiled n-D array objects**

```
SELECT img.green[x0:x1,y0:y1] > 130
  FROM LandsatArchive AS img
 WHERE avg_cells( img.nir ) < 17
```

- Scalable parallel "**tile streaming**" architecture.

- In operational use since many years.
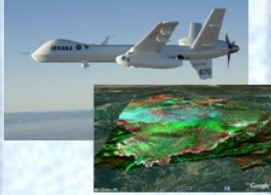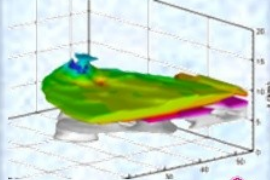
- OGC **WCS Core** Reference Implementation.

2013
Innovation Award

GEOSPATIAL
WORLD
FORUM

CERTIFIED
OGC®
COMPLIANT
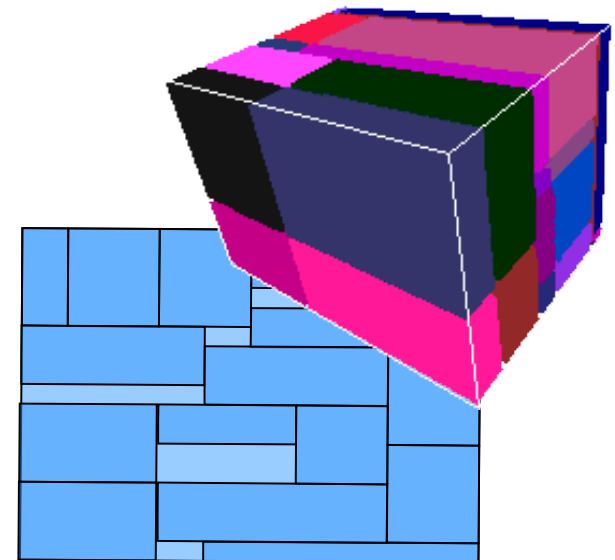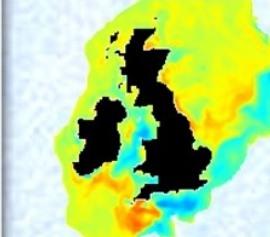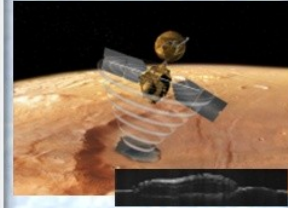
History of Array DBMSs

rasdaman visitors 2013+

EarthServer: *Big Earth Data Analytics*

- Scalable On-Demand Processing for the Earth Sciences
  - EU FP7-INFRA, Sep 2011 – Aug 2014, ~6 mEUR

- Platform: pioneer Array Database technology, `rasdaman`
 - Integrated filtering & processing on metadata, regular/irregular grids, point clouds, ...

- **11** partners (3 SMEs):

Cryospheric Science	Airborne Science	Atmospheric Science	Geology	Oceanography	Planetary Science
landcover mapping	*high-altitude long-endurance drones*	*climate variables*	*geological models*	*marine model runs + in-situ data*	*Mars geology*
EOX	NASA	MEEO Meteorological Environmental Earth Observation	British Geological Survey	PML PLYMOUTH MARINE LABORATORY	JACOBS UNIVERSITY

RTD Activity: Overview

- Big Geo Data engine development
  - Based on `rasdaman` Array Database
  - strictly *open standards* (OGC WCS, WCPS, WMS)

- Regular & irregular grids, point clouds, meshes

- Data/metadata search integration (xWCPS)

- Scalable 2D/3D client toolkits & GUIs
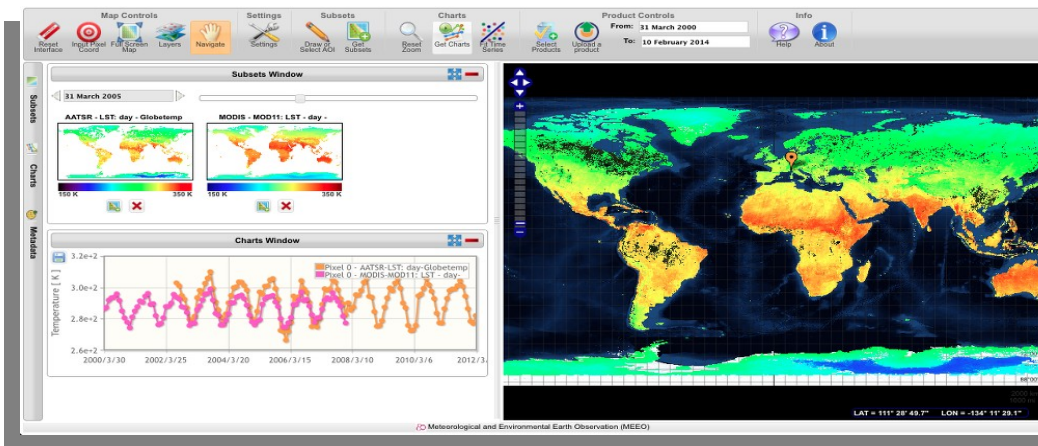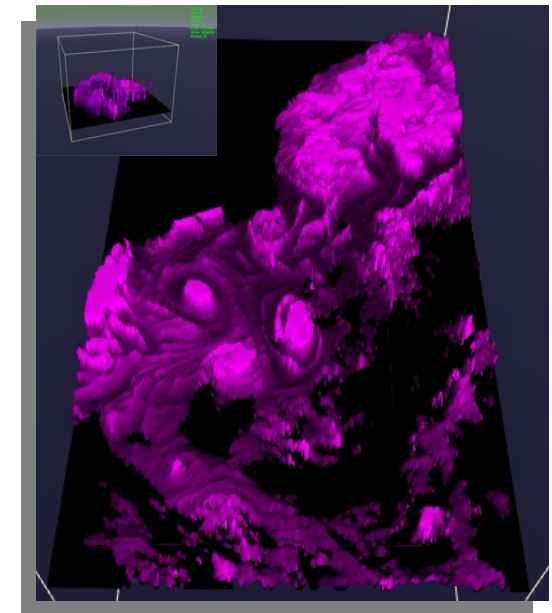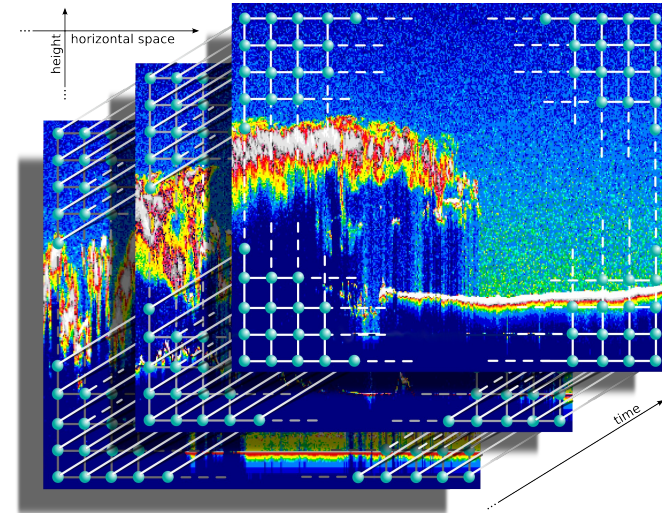
Multi-sensor Evolution Analysis (MEA)

Temporal analysis of atmospheric profiles, aerosols, cloud properties, + other EO products.
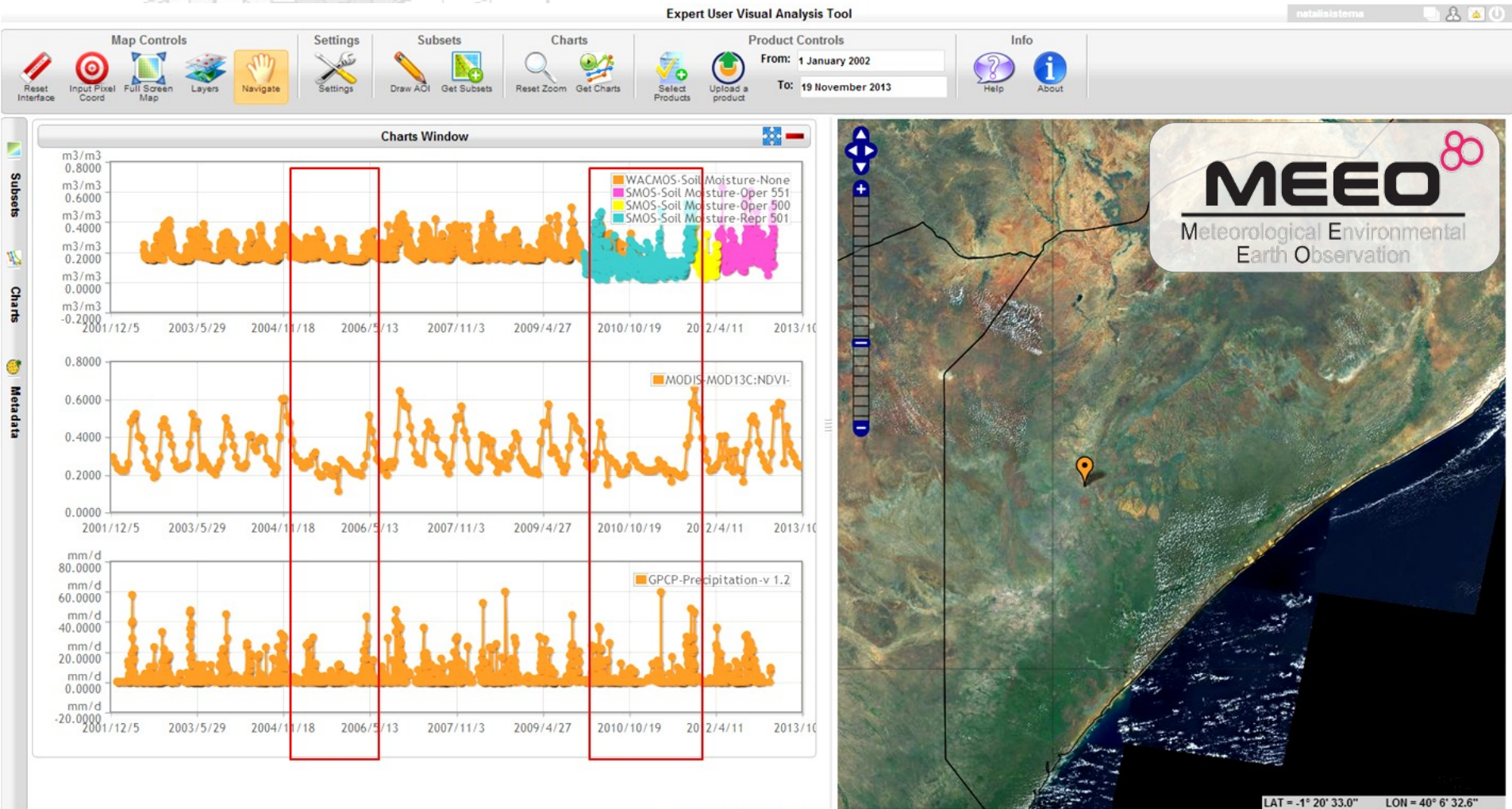
OGC open standards

OGC Web Coverage Service: Big Picture
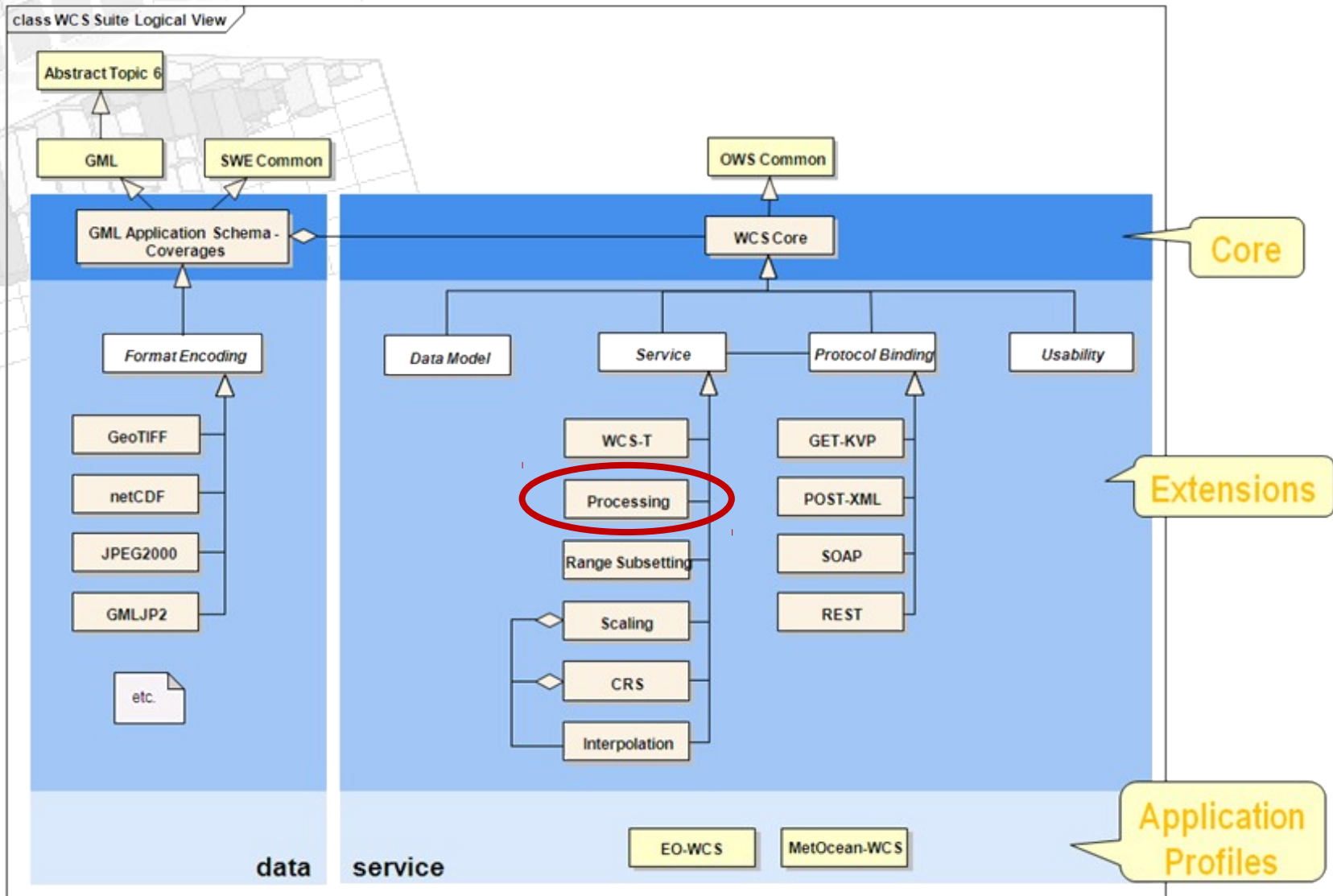
OGC Web Coverage Processing Service

- OGC Web Coverage Processing Service (WCPS) - adopted 2008.
  High-level grid coverage filtering & processing language:
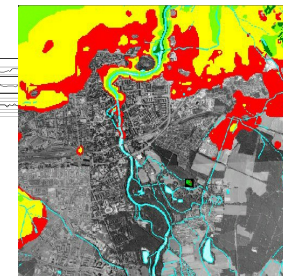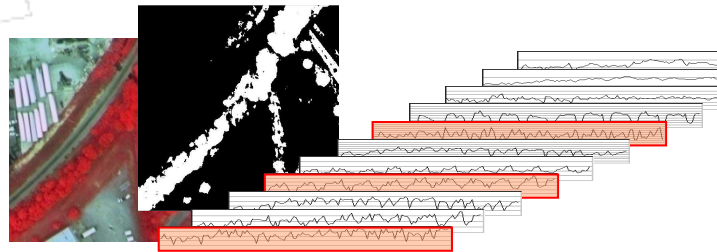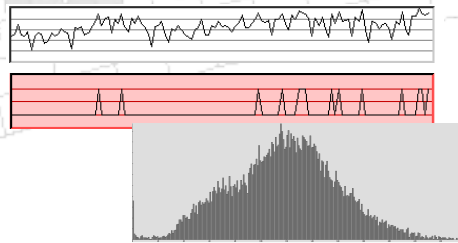
- "From MODIS scenes M1, M2, M3: *difference between red & nir*, as
  JPEG2000, *but only those where nir exceeds 127 somewhere*."

```
for $c in ( M1, M2, M3 )
where some( $c.nir > 127 )
return encode(
    $c.red - $c.nir,
"image/jp2")
```
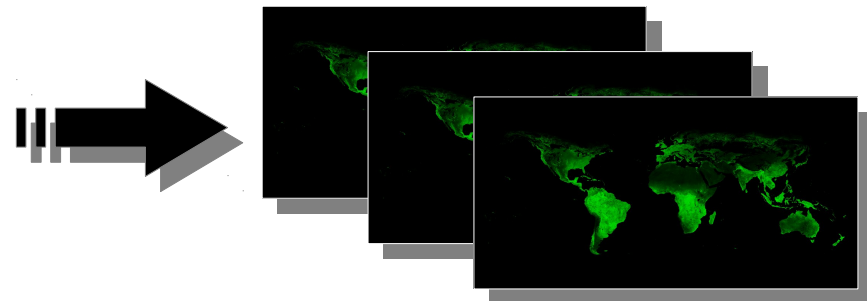
rationale & use cases

Rationale

- Group time-series of rectified EO products into a single geospatial feature: a multi-dimensional **coverage**.

- Combine <u>temporal</u> **Coordinate Reference Systems** (CRSs) with geodetic one to create a single spatio-temporal geometric space.

- **WCPS linear algebra** to provide some first tools of preliminary data analysis for scientists: <u>scaling</u> and <u>condensing</u> operators to enable server-side NODATA-aware grids spatial and temporal co-location/aggregation on datasets with:

  - ▶ Different *temporal resolutions* and support
  - ▶ Different *spatial resolution* and shifted pixel-origins
  - ▶ Either regular or irregular *frequency*

Single product temporal analsyis (I)

- Retrieve **time profiles** (*histories*):
  - ▶ which is the value in time over this point location?
  - ▶ What is the maxima over this ROI in time?
  - ▶ What is the history of weekly averages of this variable?

```
#SINGLE_PXH
coverage single_pixel_history
over     $t t ( imagecrsdomain(c, t) )
values   c[#ROI, t:"CRS:1"($t)
#MERGED_PXH
coverage averaged_pixel_history
over     $t t ( imagecrsdomain(c, t) )
values   avg(c[#ROI, t:"CRS:1"($t)])
#AGGREGATED_PXH
coverage aggregated_pixel_history
over $t t
( 0, (count(c[#SLICE(Lat,Long)])/#AGG_DAYS)-1 )
values  avg(c[#SLICE(Lat,Long),
            t:"CRS:1"($t:$t+(#AGG_DAYS-1))])
```

Single product temporal analsyis (II)

- Simple **statistics** over time:
  - ▶ what are the average aerosols concentrations of last month over Hanoi?
  - ▶ How many times pollution level has exceeded a threshold? PDF?
  - ▶ Where are there highest differences in this product on Tuesday and Wednesday?

```
#AVG
add(((c[#ROI, t("#DATE")]  = #NIL) * (0)
  + (c[#ROI, t("#DATE")] != #NIL) *
    (c[#ROI, t("#DATE")])))
  / count(c[#ROI, t("#DATE")] != #NIL)
#STDEV
sqrt(add(
  pow(
    (c[#ROI, t("#DATE")] = #NIL) * (0) +
    (c[#ROI, t("#DATE")]!= #NIL) * (
     c[#ROI, t("#DATE")] - #AVG(c)), 2))
  / (count(c[#ROI, t("#DATE")] != #NIL) - 1))
```
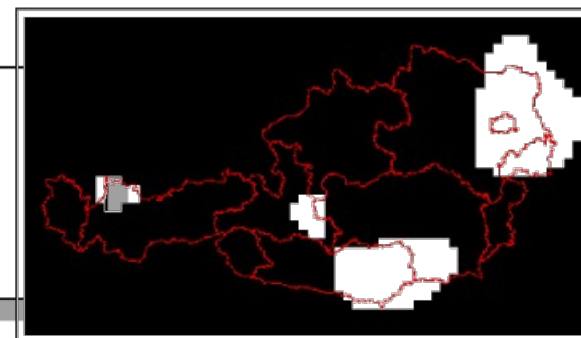
$$\mu$$
$$\sigma$$

Single product temporal analsyis (III)

- What is the black&white 95%-confidence PM **exceedance maps** from statistical forecasts models for tomorrow?
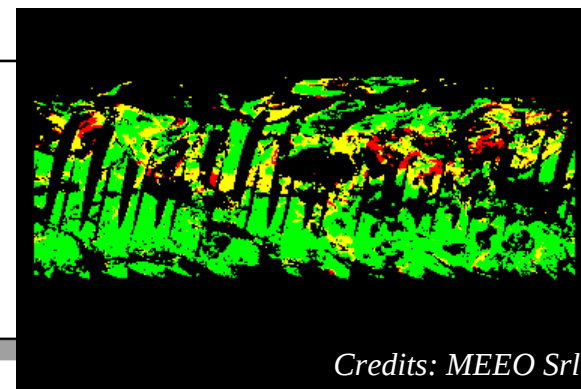
```
#95%_EXCEEDANCE
((c[#ROI, t("#DATE")]).prediction +
 (c[#ROI, t("#DATE")]).variance * 1.644854)
  > #HI_THRESHOLD
```


- What is the air quality traffic-light map on a certain date?

```
#TRAFFIC_LIGHT
{ red:   (char)( c >= LO_THRESHOLD ) * 255;
green: (char)( c <  HI_THRESHOLD ) * 255;
 blue: (char)( c * 0 )
{ [#ROI, t("#DATE")]
```


Credits: MEEO Srl

Time-series cross-comparison (I)

- Retrieve NODATA-aware bivariate statistics
  - ▶ What is the **covariance** of PM values and wind speeds on a near-by location?
  - ▶ What is the time profile of **linear correlation** between two atmospheric variables?

- Time-space <u>delta</u> Δ maps and profiles:
  - ▶ what is the maximum difference in my ROI between weekly data averages?

```
#△_MERGED_PXH
coverage delta_merged_pixel_history
over    $t t ( imagecrsdomain(c1, t) )
values avg(c1[#ROI, t:"CRS:1"($t)]) -
       avg(c2[#ROI, t:"CRS:1"($t:$t+(#M-1))])
#AGGREGATED_△_MAP
scale(c1[#ROI, t("2014-01-01")],
  imagecrsdomain(c2[#ROI, t("2014-01-01")]) -
  coverage delta_pixel_history
  over $x x ( imagecrsdomain(
        c2[#ROI, t("2014-01-01")], Long) ),
       $y y ( imagecrsdomain(
        c2[#ROI, t("2014-01-01")],  Lat) )
  values avg( c2[Lat($y), Long($x),
       t("2014-01-01T00":"2014-01-01T23:59")])
#1:1_△_MAP
scale(c1[#ROI, t("2014-01-01")],
  imagecrsdomain(
    c1[#ROI, t("2014-01-01")]) -
    c2[#ROI, t("2014-01-01T12:00")] ) )
```
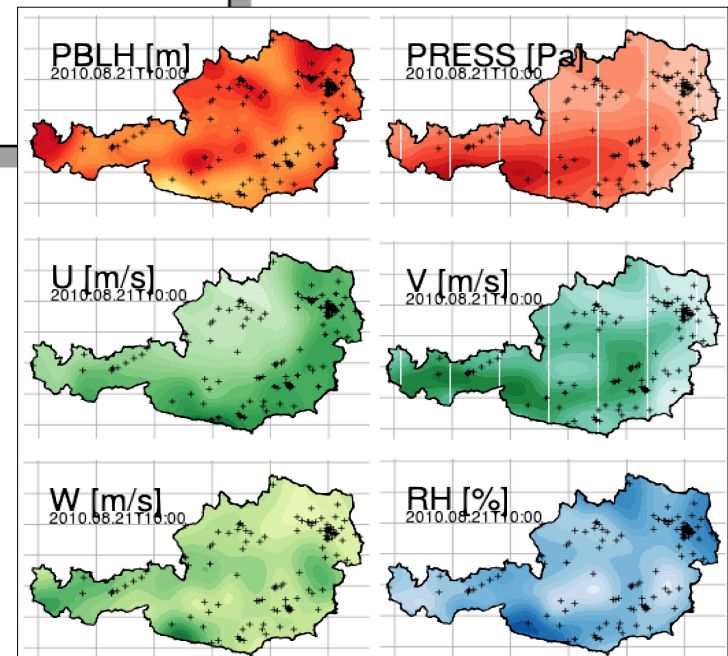

AUT.demnl.proj.tif
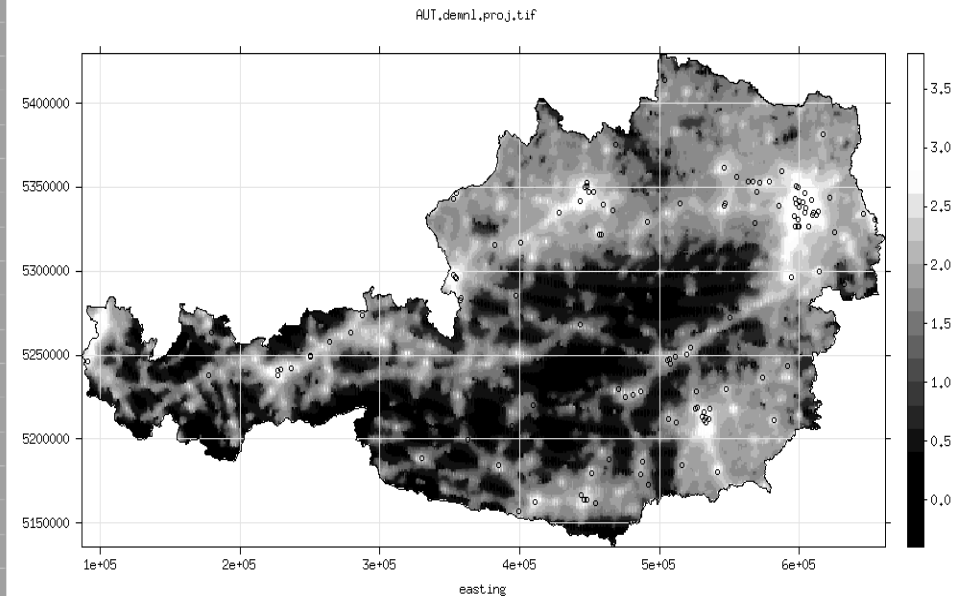
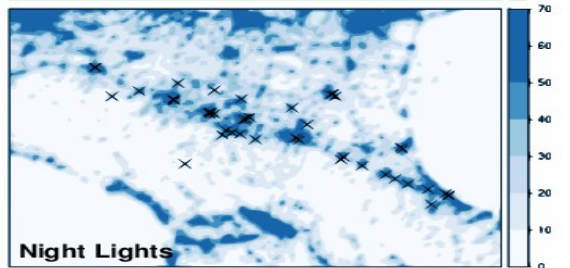# Time-series cross-comparison (II)

- Single predictor **linear model** computations
  - ▶ What is the beta coefficient of my possible predicor?
 - ▶ What are the residuals of an hypothetical linear model applied on this day?

```
#LM_RESIDUALS
c1[#ROI, t("#DATE")] - c2[#ROI, t("#DATE")] * (
    add(c1[#ROI, t("#DATE")] *
        c2[#ROI, t("#DATE")] ) /
    add(pow(c2[#ROI, t("#DATE")], 2)))
```

Spatio-temporal data merging

- Mere averaging: what is the average <u>spatially-enhanced</u> map of AOT obtained by merging MODIS and MISR datasets colelcted on the previous 50 days?

- What are the maxima of daily averages from N different transport <u>models</u>?

- Exploiting statistical estimation <u>error information</u>: what is the MLE-based fusion map of these 3 different kriging models? →
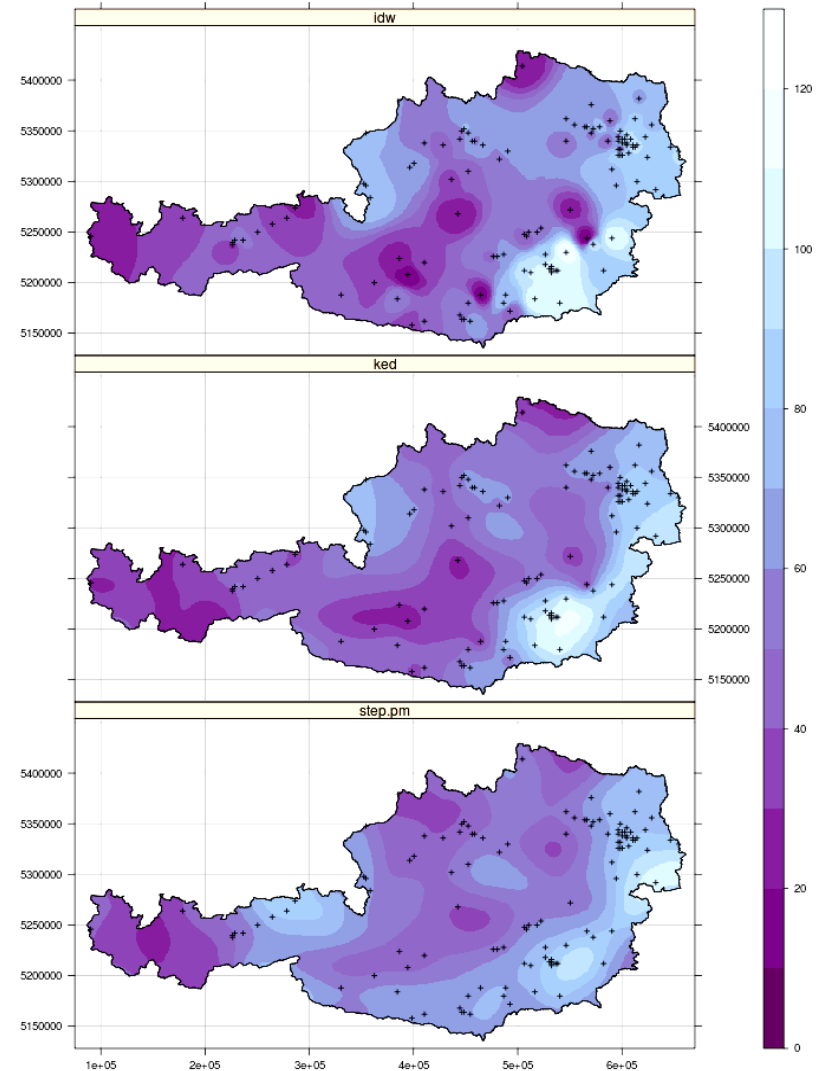
conclusions & outlook

Summary

- The WCPS query language allows for <u>array analytics</u> to be requested directly to your data server.

- WCPS can enhance the capabilities of a web service for <u>value-adders</u> handling remote-sensing imagery data.

- WCPS is an OGC <u>open standard</u>.

- Want to try it out?
  - ► Contact us: `http://rasdaman.org/wiki/MailingLists`
  - ► Check out our *rasdaman* tutorial:
    `http://rasdaman.org/wiki/Workshops/BigDataRasdamanApproach`

WCPS capabilities demonstration app

- Setup OGC WCS endpoint of 1 year EO gridded rectified datasets of *Aerosol Optical Thickness* (AOT) @ 550nm from different sources:

  - ▶ Level-3 24H NASA MOD08 product, 1×1 degrees (+ NODATA)
  - ▶ Level-3 24H ESA CCI AOT product, 1×1 degrees (MEAN/STDEV)(+ NODATA)
  - ▶ 3H ECMWF MACC Reanalysis AOT product, 1.125×1.125 degrees

- Deployment of WCPS queries for:

  - ▶ Single-product data **analysis** and statistics retrieval
 - ▶ Multiple-product data **cross-comparison and merging**

- Project site: *http://worldwind.flanche.net/*

Current limitations

- Only grids which have been pre-<u>aligned</u> to a uniform grid within the coordinate reference system are supported: Level-1 or 2 satellite datasets are left out.

- BBOX-subsetting is the only way to select regions of interest via WCS/WCPS: <u>polygonal clipping</u> – highly common in GIS applications – needs to be designed.

- No CRS <u>reprojection</u> make it harder to combine and cross-compare separate products.
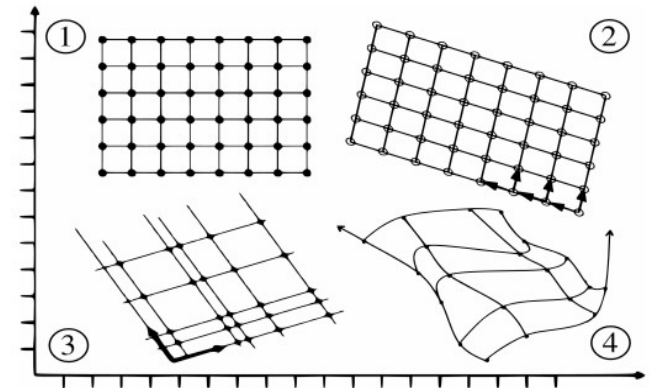
@TODO
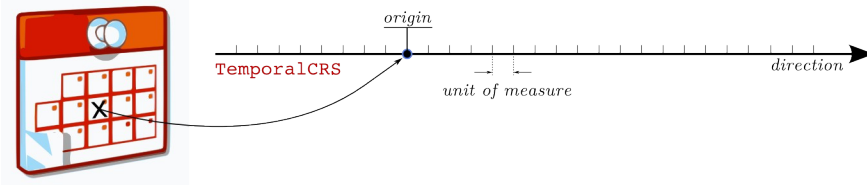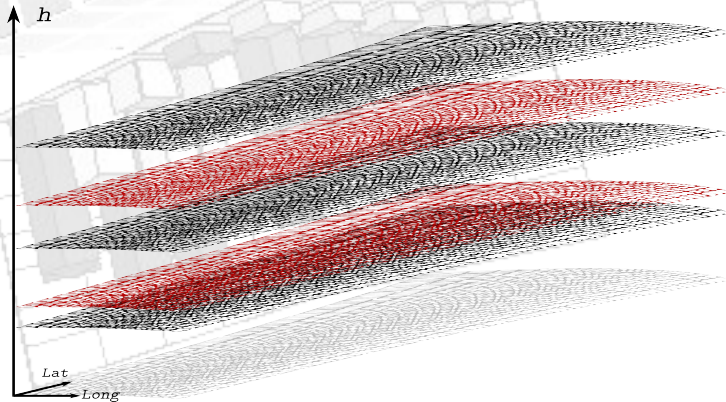
- Extend support to <u>non-uniform</u> („warped")grids for hosting a wider range of EO data;

- Finalize standardization of XML definition for <u>time coordinate reference systems</u> at OGC;

- Extend WCS and WCPS to define <u>intersection operation(s)</u> of grids with general geometries (lines, polygons, etc.)

- Proper handling of <u>data uncertainty</u>;

- More powerful processing with User-Defined Functions and <u>R/Python bindings</u>.

Links

- Big Earth Data Standards: `rasdaman` demonstrations
  http://www.earthlook.org/

- The *EarthServer* initiative
  http://earthserver.eu/

- The open source `rasdaman` project
  http://rasdaman.org/
  http://www.ohloh.net/p/rasdaman

- MEA platform
  http://earthserver.services.meeo.it/

- **OGC** coverage service standards
  http://www.opengeospatial.org/standards/wcs
  http://www.opengeospatial.org/standards/wcps