Funded by European Union

Grant Agreement 952111
H2020 WIDESPREAD-05-2020 (Twinning)

Standard course – Remote Sensing
SC1 – Massive geoprocessing procedures and tools

Rizos Chadoulis, Ioannis Manakos - CERTH

Lluís Pesquer – CREAF

Table of contents

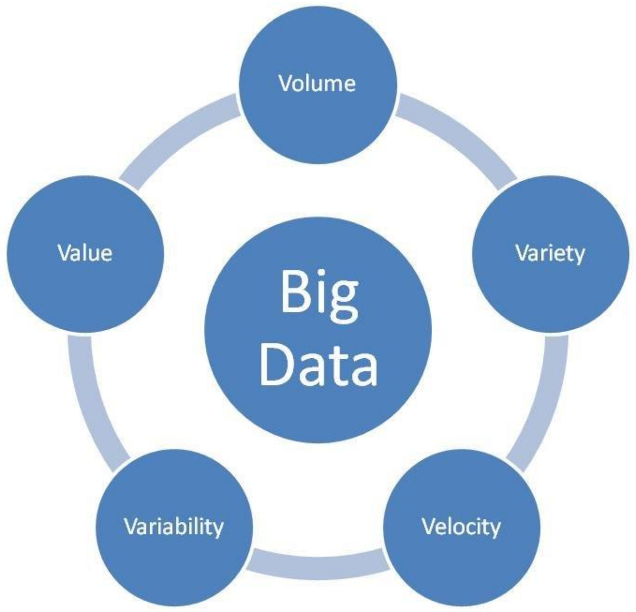
- Introduction and concepts
- Scripts and BATCH processing
- High Performance Computing
- Geospatial information processing
- Datacubes
- Geoprocessing Business Models

Introduction and concepts

CREAF & CERTH

Big data

- The Oxford Dictionary defines *Big Data* as "Datasets that are too large and complex to manipulate or interrogate with standard methods or tools".
- The term was first described by Laney as "the three V's: Volume, Velocity and Variety". Others have extended this list to include validity, veracity, value, and visibility.



Laney, D., 2001. 3D data management: controlling data volume, velocity and variety. META Group Research Note. 6 :p. 70.

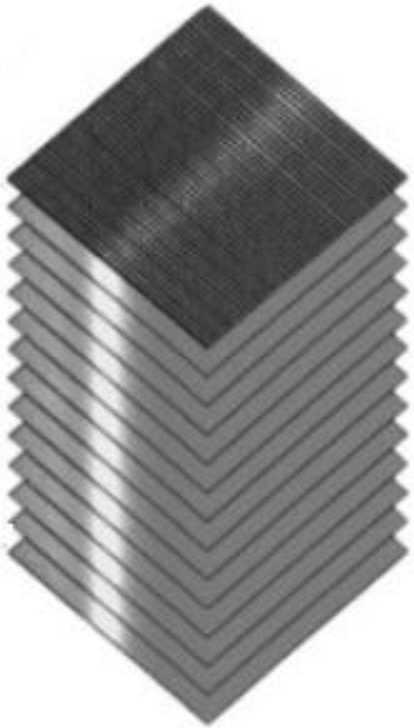
RS Big data

There is a continual global observing data, those are capable of covering the global atmosphere, land surface and oceans.

Currently, more than 200 on-orbit satellite sensors are capturing multi-spatial, multi-temporal data from multi-sensors, we are entering the Big data observation era.

RS Big data

The diversity and high dimensionality of data commonly lead to the complexity of the RS Big Data.



- hyperspectral images
- long time series
- very high resolution
- global coverage products

↓

huge amounts of complex RS data

Geoprocessing

Geoprocessing refers to the processing of spatial or geo-referenced data which relate to conceptual data models and encoding formats from geoscience and neighbouring disciplines (*Muller 2015)*

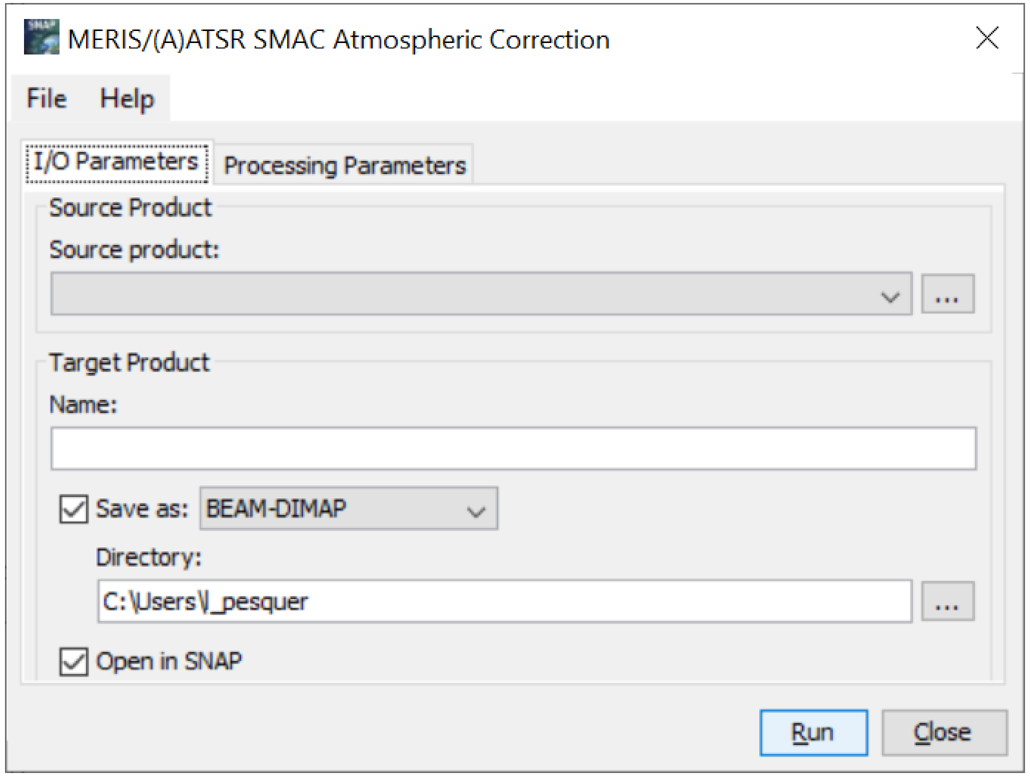
Geoprocessing controls the processing environment and allows you to build custom tools that can further automate your work

It applies tools to create new data from existing data

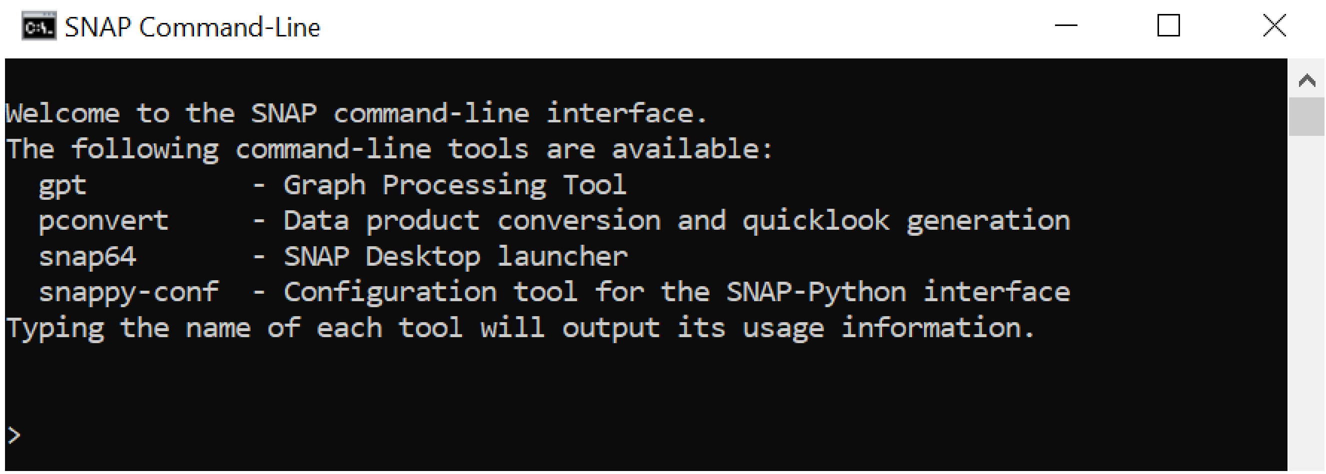
Geoprocessing

Geoprocessing can be done

Interactively (by interface)



By command line



Massive geoprocessing

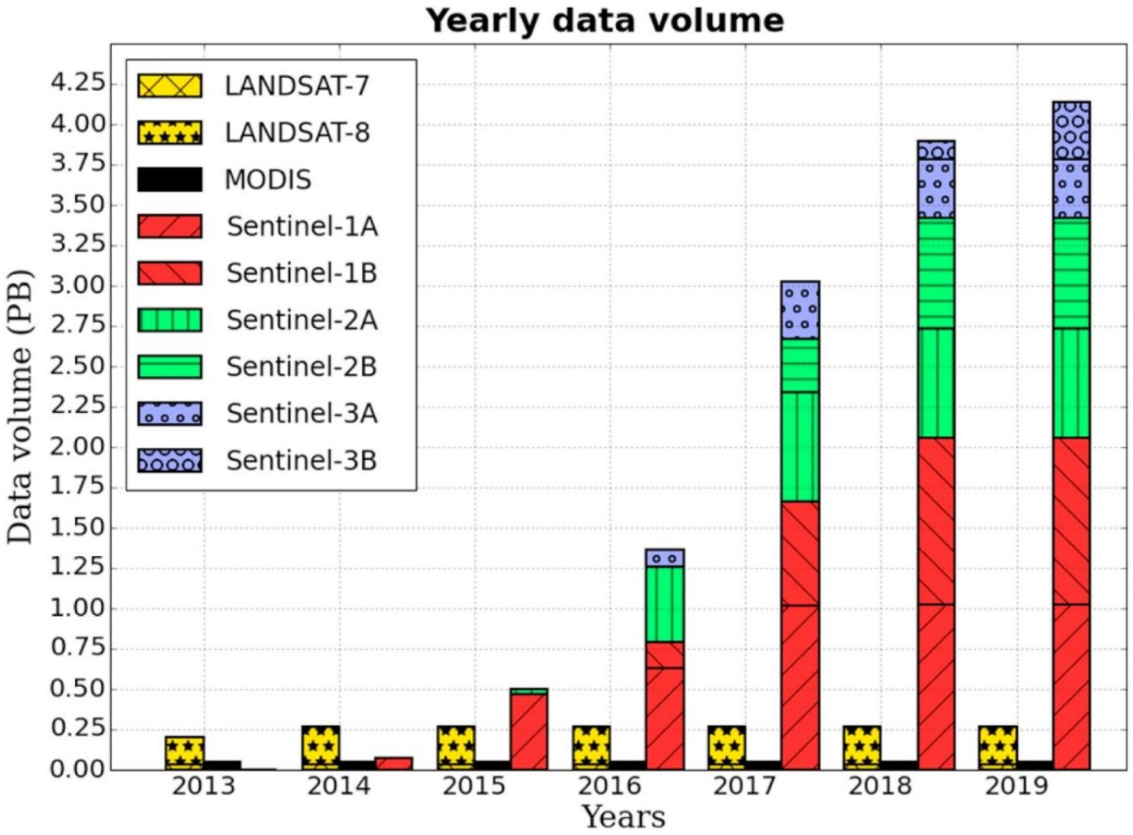
The rapid growth of RS data has increased the complexity of creating responsive and scalable geospatial applications.

Remotely sensed raster datasets can easily have millions or billions of values, and performing operations on them is often a very computationally expensive task.

At least 3 requirements are needed

- Automatization: impossible to handle this amount of data in a interactive way
- Infrastructure: enough processing units and memory resources
- Methodologies and strategies for the efficiently use of these resources

Defining massive



The definition of massive constantly changes because:

- Our computing capabilities increase
- The size of the available data increases

Soille, P., Burger, A., De Marchi, D., Kempeneers, P., Rodriguez, D., Syrris, V., & Vasilev, V. (2018). A versatile data-intensive computing platform for information retrieval from big geospatial data. In Future Generation Computer Systems (Vol. 81, pp. 30–40). Elsevier BV. https://doi.org/10.1016/j.future.2017.11.007

Understanding massive

How much information can be stored in 100 PetaBytes?

- 40 human brains worth of memory storage
- 6.67 years of the Large Hadron Collider’s experimental data
- 190,259 years of continuously playing MP3 encoded music
- 26.6 billion human genomes worth of sequence information
- 16,025 years of data transmissions from the Hubble space telescope

Use of resources

Performance (inversely proportional to execution time) is one of the main operational goals of massive processing. The two next concepts aims to improve the performance with the feasible resources:

- Elasticity is the ability to grow or shrink infrastructure resources dynamically as needed to adapt to workload changes in an autonomic manner, maximizing the use of resources.
- Scalability is the measure of the capability of a system to increase (or decrease), within the existing infrastructure, in performance and cost in response to changes in application and system processing.

Scalability is required for elasticity

Script & Batch Processing

CREAF

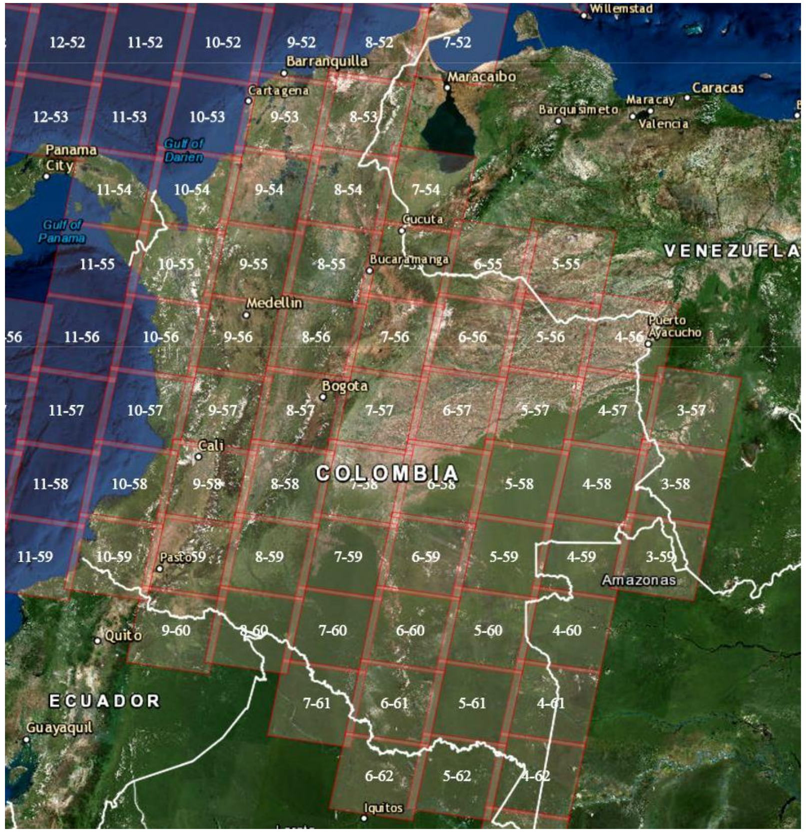
Automatization

The aims of automatic geoprocessing is:

- Applying the same chain of processes to (huge amount of) data distributed on tiles or regions
- Repeat the same process to updated data
- Reproduce the chain of processes

> Reproducibility: the ability of duplicating the results of a prior study using the same materials as were used by the original producer

Reproducibility goes further than the **r**euse of FAI**R** principles



Script processing

Each Remote Sensing software choose their own automatization solution, based on:

- own model and scripts: ModelBuilder ESRI, IDL ENVI, etc..
- existing high level programming language: PyQGIS
- using the Operating System batch

```
FILES=$1/*.$3
for f in $FILES
do echo "Processing $f file..."
  filename=`basename ${f} .${3}`
  echo "Output: ${2}/${filename}.tif"
  gdal_translate -of GTiff ${f} ${2}/${filename}.tif
done
```

GDAL API from Linux console

```
fileName = "/path/to/raster/file.tif"
fileInfo = QFileInfo(fileName)
baseName = fileInfo.baseName()
rlayer = QgsRasterLayer(fileName, baseName)
if not rlayer.isValid():
  print "Layer failed to load!"
```

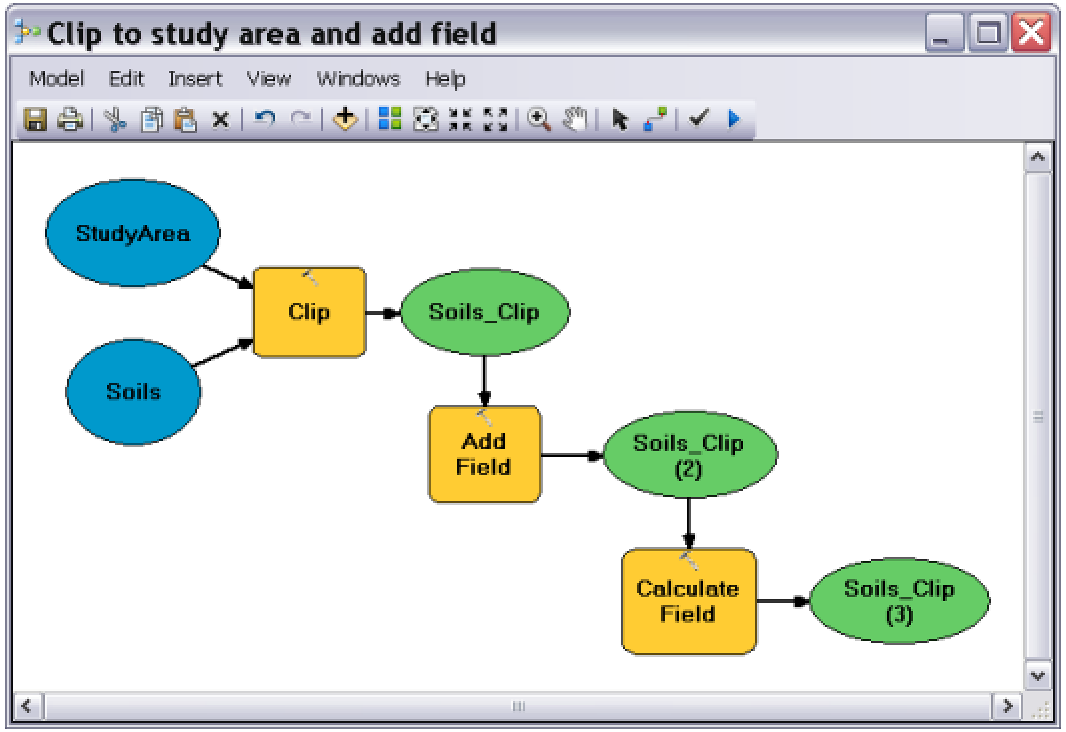
PyQGIS: Python QGIS

Script processing

ModelBuilder (ESRI) is an application you use to create, edit, and manage models.

These models are workflows that string together sequences of geoprocessing tools

ModelBuilder can also be thought of as a visual programming language for building these workflows.

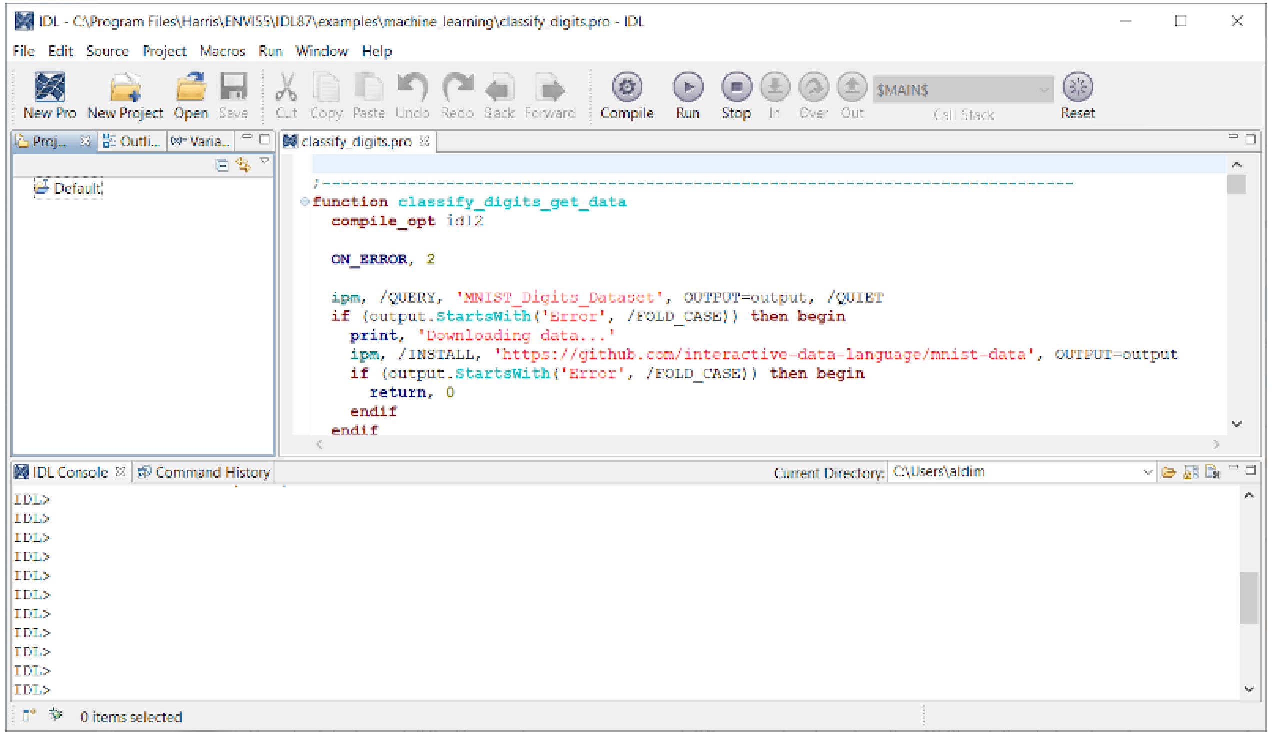


Script processing

IDL (ENVI) (Interactive Data Language) is the programming software used for data analysis, visualization and cross-platform application development.

IDL is based on the Eclipse framework

It is a powerful data manipulation language which contains plotting and mathematical functions with a robust Analysis Engine.

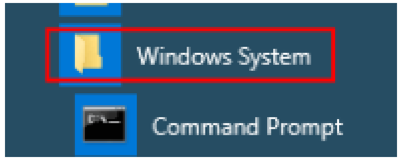


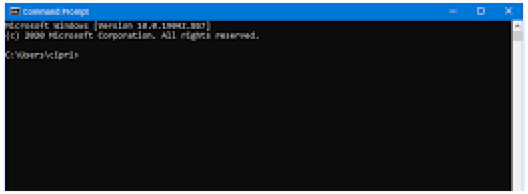
Batch processing

A batch file is a script file in DOS, OS/2 and Microsoft Windows.

It consists of a series of commands to be executed by the command-line interpreter, stored in a plain text file.

cmd.exe is the normal MS Windows command interpreter for batch files (command.com for the first versions)





A batch file can contain any command the interpreter accepts interactively and it can use constructs that enable conditional branching and looping

The filename extension is **.bat** (.cmd for old versions)

PowerShell

PowerShell is a cross-platform task automation solution made up of a command-line shell, a scripting language for task automation and configuration management.

PowerShell runs on Windows, Linux and macOS.

PowerShell operates on objects over text. It is built on the .NET Common Language Runtime (CLR). All inputs and outputs are .NET objects.

Commands in PowerShell can be native executables, cmdlets, functions, scripts, or aliases

The filename extension is **.ps1**

PowerShell

High Performance Computing

CERTH

Improving performance

There are three ways of improving performance:

- Work harder (High Performance Computing)
- Work smarter (Parallel Computing)
- Get help (Distributed computing)

G. Pfister, In Search of Clusters, Prentice Hall PTR, 1998.

Commodity clusters

A commodity cluster is an ensemble of fully independent computing systems integrated by a commodity off-the-shelf interconnection communication network.

Baker, M., & Buyya, R. (1999). Cluster computing: the commodity supercomputer. In Software: Practice and Experience (Vol. 29, Issue 6, pp. 551–576). Wiley. https://doi.org/10.1002/(sici)1097-024x(199905)29:6<551::aid-spe248>3.0.co;2-c

High Performance Computing

IBM Blue Gene/P supercomputer "Intrepid" (Argonne National Laboratory)

HPC integrates systems administration (including network and security knowledge) and parallel programming into a multidisciplinary field that combines digital electronics, computer architecture, system software, programming languages, algorithms and computational techniques.

- Brazell, Jim and Michael Bettersworth. "High Performance Computing". Texas State Technical College, 2005.

HPC platforms for large-scale processing of spatial data

https://hadoop.apache.org/

The Apache Hadoop software library is a framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models.

https://spark.apache.org/

Apache Spark™ is a multi-language engine for executing data engineering, data science, and machine learning on single-node machines or clusters.

https://sedona.apache.org/

Apache Sedona (incubating) is a cluster computing system for processing large-scale spatial data. Sedona extends Apaches Spark and Apache Flink with a set of out-of-the-box distributed Spatial Datasets and Spatial SQL that efficiently load, process, and analyze large-scale spatial data across machines.

HPC platforms for large-scale processing of spatial data

The OpenMP API supports multi-platform shared-memory parallel programming in C/C++ and Fortran. The OpenMP API defines a portable, scalable model with a simple and flexible interface for developing parallel applications on platforms from the desktop to the supercomputer.

Unified Parallel C (UPC) is an extension of the C programming language designed for high performance computing on large-scale parallel machines. The language provides a uniform programming model for both shared and distributed memory hardware.

Message passing interface (MPI) is a standard specification of message-passing interface for parallel computation in distributed-memory systems. MPI isn't a programming language. It's a library of functions that programmers can call from C, C++, or Fortran code to write parallel programs.

HPC platforms for large-scale processing of spatial data

GPGPU

GPGPU, also known as GPGPU computing, refers to the increasingly commonplace, modern trend of using GPUs for non-specialized computations in addition to their traditional purpose of computation for computer graphics. Incorporating GPUs for general purposes enhances CPU architecture by accelerating portions of an application while the rest continues to run on the CPU, ultimately creating an overall faster, high-performance application by combining CPU and GPU processing power.

CUDA® is a parallel computing platform and programming model developed by NVIDIA for general computing on graphical processing units (GPUs). With CUDA, developers are able to dramatically speed up computing applications by harnessing the power of GPUs.

OpenCL™ (Open Computing Language) is an open, royalty-free standard for cross-platform, parallel programming of diverse accelerators found in supercomputers, cloud servers, personal computers, mobile devices and embedded platforms.

HPC platforms for large-scale processing of spatial data

Examples of HPC for spatial statistics include parallelizing the computation of statistical measures (e.g., Moran’s I and Getis-Ord) and geostatistics pattern modelling and interpolation using OpenMP and/or MPI

- Kazar B, Shekhar S, Lilja D et al (2004) A parallel formulation of the spatial auto-regression model formining large geo-spatial datasets. In: SIAM international conference on data mining workshop on high performance and distributed mining (HPDM2004), Florida, 22–24 April 2004
- Pesquer L, Cortés A, Pons X (2011) Parallel ordinary kriging interpolation incorporating automatic variogram fitting. Computers and Geosciences 37 (4): 464-473. DOI:10.1016/j.cageo.2010.10.010
- Wang S, Cowles MK,Armstrong MP(2008) Grid computing of spatial statistics: using the TeraGrid for G(d) analysis. Concurr Comput Pract Exp 20(14):1697–1720

HPC platforms for large-scale processing of spatial data

A parallelization of wavelet transform, which can locate frequency outliers, has been implemented on MPI to scale up outlier detection algorithms.

- Barua S, Alhajj R (2007) Parallel wavelet transform for spatio-temporal outlier detection in large meteorological data. In: International conference on intelligent data engineering and automated learning, Birminghan, UK, 16–19 December 2007. Springer, Heidelber, pp 684–694

HPC platforms for large-scale processing of spatial data

Hadoop and Spark have also been leveraged as platforms to implement Kriging and inverse distance-weighted interpolation algorithms.

- Xu Z, Guan J, Zhou J (2015) A distributed inverse distance weighted interpolation algorithm based on the cloud computing platform of Hadoop and its implementation. In: 2015 12th international conference on fuzzy systems and knowledge discovery (FSKD), Zhangjiajie, China, 15–17 August 2015. IEEE, New Jersey, pp 2412–2416
- Rizki, P. N. M., Eum, J., Lee, H., & Oh, S. (2017). Spark-based in-memory DEM creation from 3D LiDAR point clouds. In Remote Sensing Letters (Vol. 8, Issue 4, pp. 360–369). Informa UK Limited. https://doi.org/10.1080/2150704x.2016.1275053

HPC platforms for large-scale processing of spatial data

GCMF, an end-to-end software system on GPGPU, illustrates the potential of GPGPU as a platform for geospatial information processing, as it can handle spatial joins over non-indexed polygonal datasets containing more than 600,000 polygons on a single GPU within 8 s.

Aghajarian, D., Puri, S., & Prasad, S. (2016). GCMF. In Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. SIGSPATIAL'16: 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM. https://doi.org/10.1145/2996913.2996982

HPC platforms for large-scale processing of spatial data

GPU-based methods have also been introduced to accelerate the computation of change detection

Prasad SK, Shekhar S,McDermottMet al (2013) GPGPU-accelerated interesting interval discovery and other computations on GeoSpatial datasets: a summary of results. In: Proceedings of the 2nd ACM SIGSPATIAL international workshop on analytics for big geospatial data, Orlando, FL, 4 November 2013. ACM, New York, pp 65–72

Challenges

- The ubiquitous existence of spatial autocorrelation makes parallelization not applicable for most geo-related algorithms because the dependence between data partitions requires task interaction, which increases the difficulty of parallelizing serial algorithms in spatial database and spatial data mining functions.
- The load balancing between processing elements is complicated when dealing with sparse data structures for which the pattern of interaction among data elements is data-dependent and highly irregular. Spatial networks (e.g., road networks) are an example of these data structures.

Future Directions

- Generalizing spatial analysis and data mining methods to the network space, such as network spatial interpolation (Kriging), network point density estimation, and linear hotspot detection

Okabe A., Sugihara K. (2012) Spatial analysis along networks: statistical and computational methods. John Wiley & Sons, New Jersey

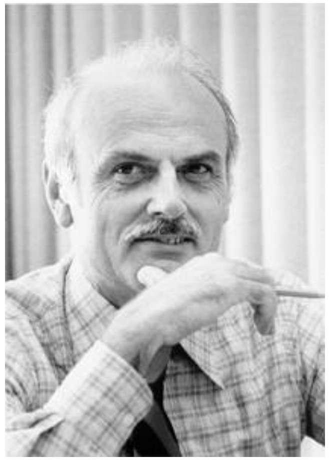
Spatial Database Management Systems

CERTH

The 3 challenges of spatial data mining

- The spatial autocorrelation effect
- spatial data is embedded in a continuous space whereas classical datasets are often discrete.
- spatial context matters

On-Line Analytical Processing (OLAP)

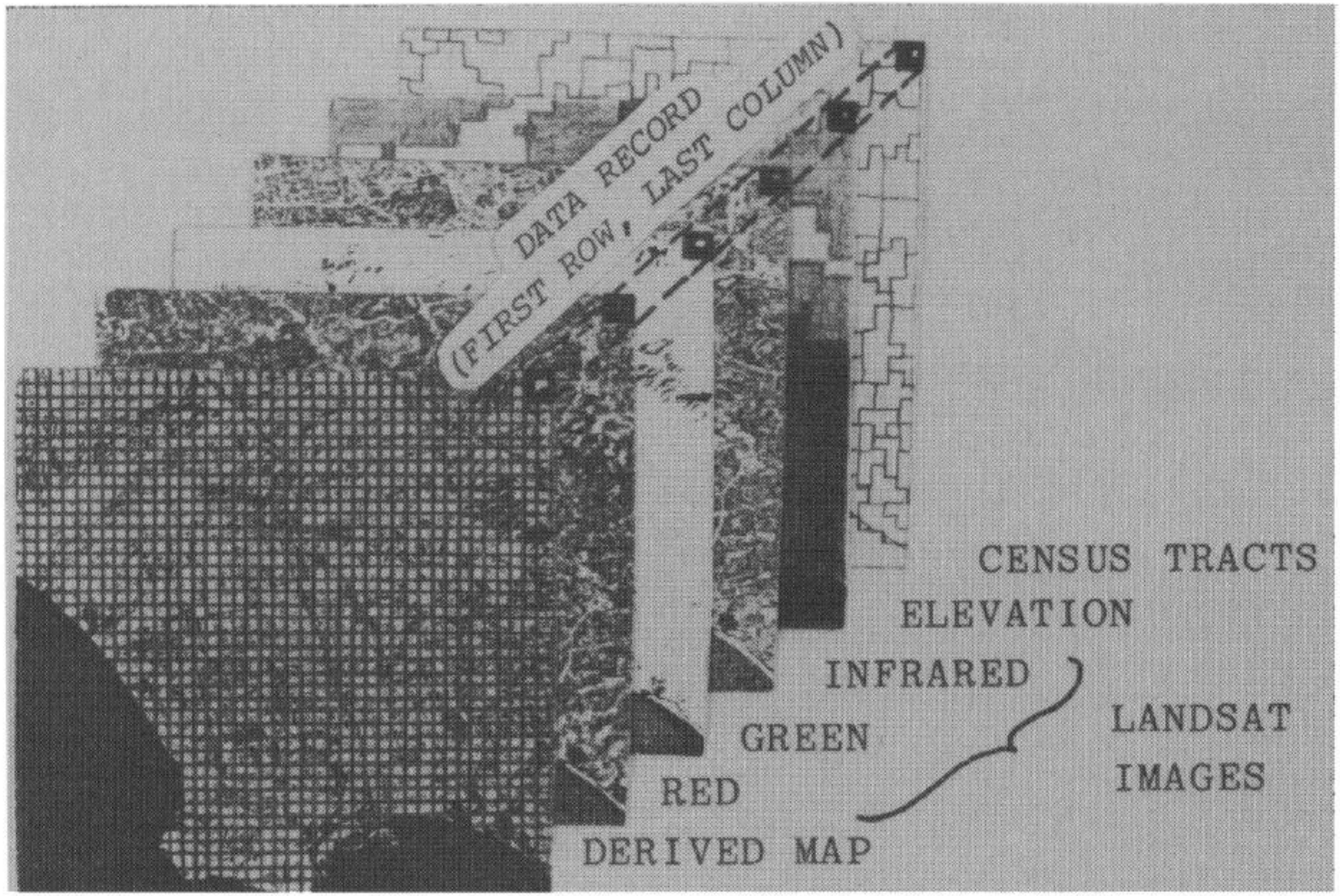


1993

Codd coined the term On-Line Analytical Processing (OLAP) to characterize the requirements for summarizing, consolidating, viewing, applying formulae to, and synthesizing data according to multiple dimensions.

E. F. Codd. Providing OLAP (on-line analytical processing) to user-analysts: An IT mandate. Technical report, E. F. Codd and Associates, 1993.

PICture Database Management System (PICDMS)



- building multiple-variable databases from photographs and other two-dimensional data sources such as maps, drawings, etc.
- manipulating such data using simple logical commands

Chock, M., Cardenas, A. F., & Klinger, A. (1984). Database Structure and Manipulation Capabilities of a Picture Database Management System (PICDMS). IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6(4), 484–492. doi:10.1109/tpami.1984.4767553

Baumann’s contribution

- extended DBMSs with capabilities to manage multidimensional discrete data (MDD)
- proposed a sublanguage for MDD definition and manipulation with a set-algebraic semantics

- Baumann, P.: A Database Array Algebra for Spatio-Temporal Data and Beyond. Proc. NGITS’99, LNCS 1649, Springer 1999, pp.76-93
- Baumann, P.: On the Management of Multidimensional Discrete Data. VLDB Journal 4(3)1994, Special Issue on Spatial Database Systems, pp. 401–444

The Array Manipulation Language

- an Algebra for Arrays
- can be used as a query language for array data, and as a view definition language, to define arrays in terms of existing ones

Marathe, A., Salem, K.: A language for manipulating arrays. Proc. VLDB’97, Athens, Greece, August 1997, pp. 46–55

The Array Query Language (AQL)

- a query language for multidimensional arrays
- treating arrays as functions from index sets to values rather than as collection types

Libkin, L., Machlin, R., Wong, L.: A query language for multidimensional arrays: design, implementation and optimization techniques. Proc. ACM SIGMOD'96, Montreal, Canada, pp. 228–239

Grant Agreement 952111, H2020 WIDESPREAD-05-2020 (Twinning)

Online Geospatial Information Processing

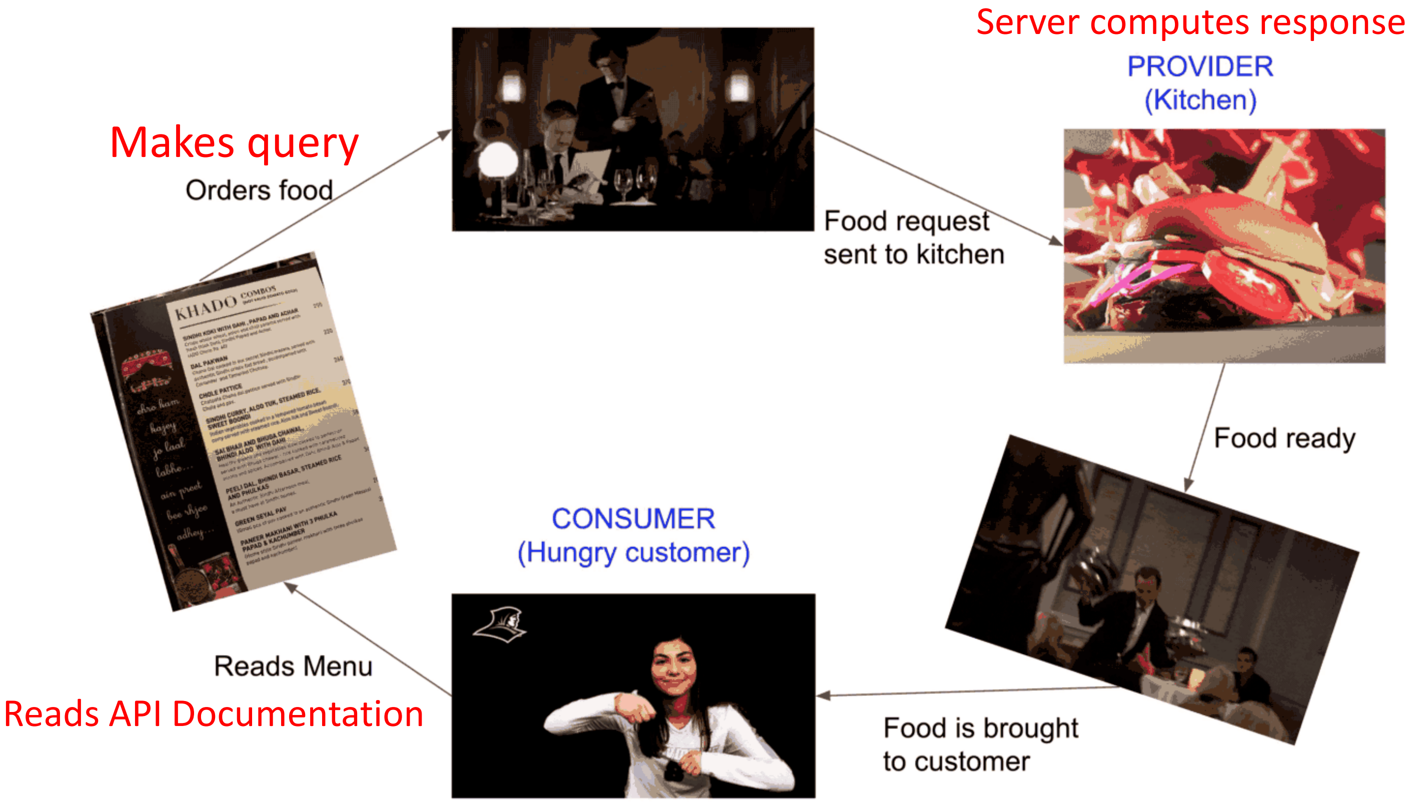
CERTH

Web Service

A software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL). Other systems interact with the web service in a manner prescribed by its description using SOAP-messages, typically conveyed using HTTP with an XML serialization in conjunction with other web-related standards.

"Web Services Glossary § Web service". W3C. 11 February 2004. Retrieved 5 April 2022.

Web Service



Using a Web Service is analogous to ordering and consuming food at a restaurant.

Abhilasha Bhatia, Understanding Web API basics Through Analogies, Link: https://medium.com/@abhilashabhatia/understanding-web-api-basics-through-analogies-8be8da2b5800

Open Geospatial Consortium (OGC)

Open Geospatial Consortium

a worldwide community committed to improving access to geospatial, or location information.

One of its primary functions is developing, coordinating, promulgating, revising, amending, reissuing, interpreting, or otherwise producing technical standards.

- *"Web Services Glossary § Web service". W3C. 11 February 2004. Retrieved 5 April 2022.*
- *Wang Ping (April 2011), A Brief History of Standards and Standardization Organizations: A Chinese Perspective, East–West Center*
- *IEEE Standards Association (IEEE SA), What are Standards? Why are They Important? Link: https://beyondstandards.ieee.org/what-are-standards-why-are-they-important/*

OGC Web Processing Services (WPSs)

The OGC Web Processing Service (WPS) Interface Standard provides a standard interface that simplifies the task of making simple or complex computational processing services accessible via web services.

GetCapabilities

- Provides metadata (XML) documents describing each of the processes offered by a WPS including a short abstract or keywords.

DescribeProcess

- Once a process has been selected, the server responds with the same information provided by the *GetCapabilities* response but with additional detailed information which may include: the required input parameters for the process, allowable formats, and the potential outputs.

Execute

- Requests the server to perform a selected operation or process, using the provided input values and returning the output produced.

- *Web Processing Service*. OGC. Retrieved November 6, 2021, from https://www.ogc.org/standards/wps
- 3.4 Web Processing Services (WPS) | Introduction to OGC Standards | Learning Materials (rspsoc.org.uk) http://learningzone.rspsoc.org.uk/index.php/Learning-Materials/Introduction-to-OGC-Standards/3.4-Web-Processing-Services-WPS

Web Coverage Processing Services

- an extension of the WCS standard with an explicit focus on the processing of coverages
- defines a protocol-independent language for on-demand extraction, processing, and analysis of multi-dimensional gridded coverages (datacubes) representing among others spatio-temporal sensor, image, simulation, or statistics data.

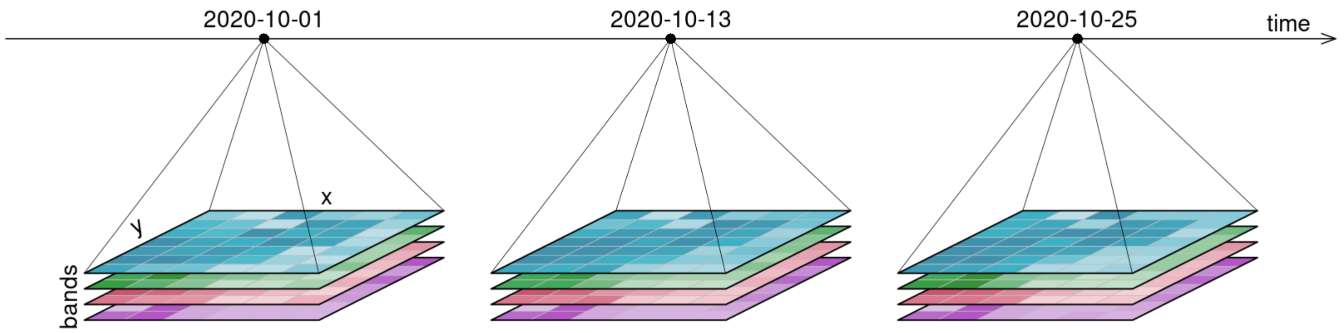
Web Coverage Processing Service (WCPS) Standard, Link: https://www.ogc.org/standards/wcps

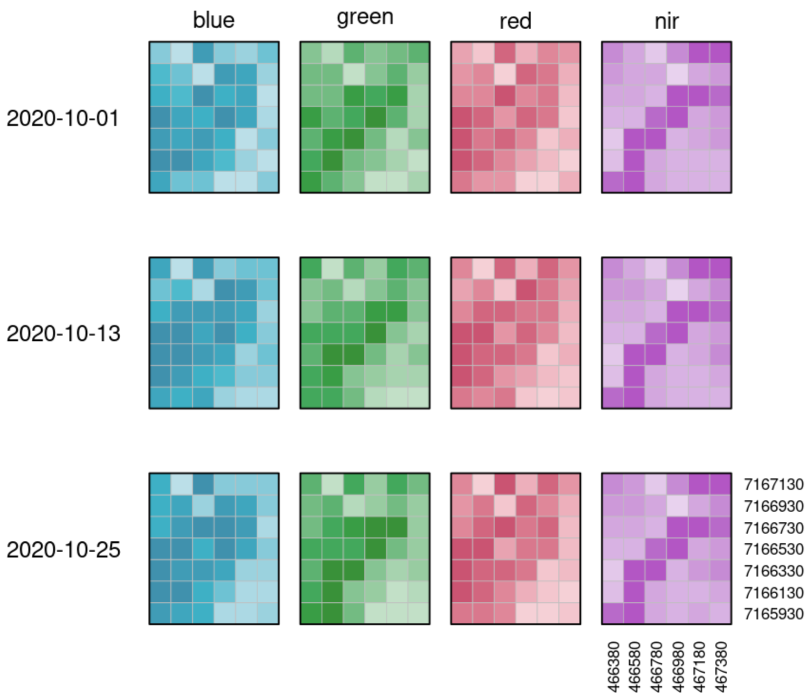
Earth observations data cubes (EODCs)

CERTH

What are datacubes?

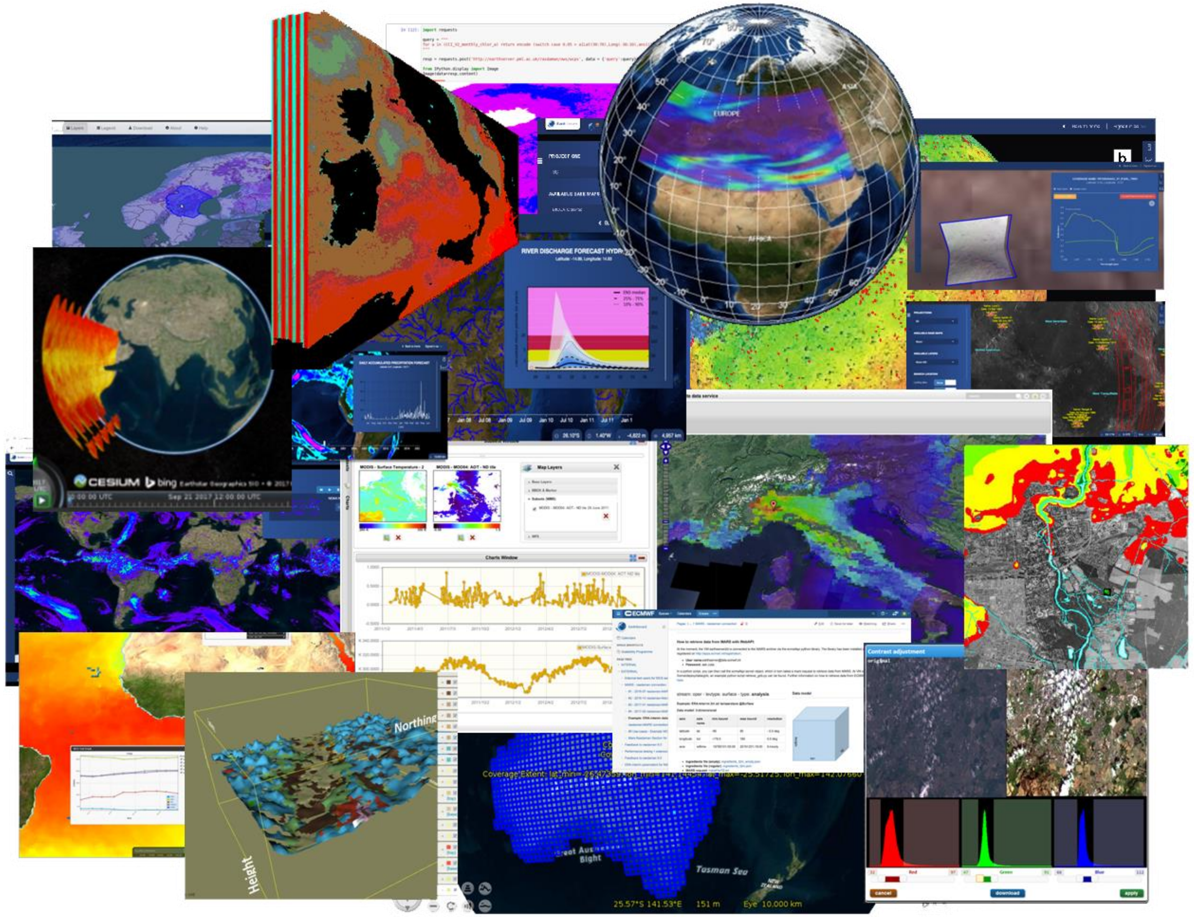
Datacubes are multidimensional arrays with one or more spatial or temporal dimension(s).





https://openeo.org/documentation/1.0/datacubes.html#what-are-datacubes

The datacube manifesto

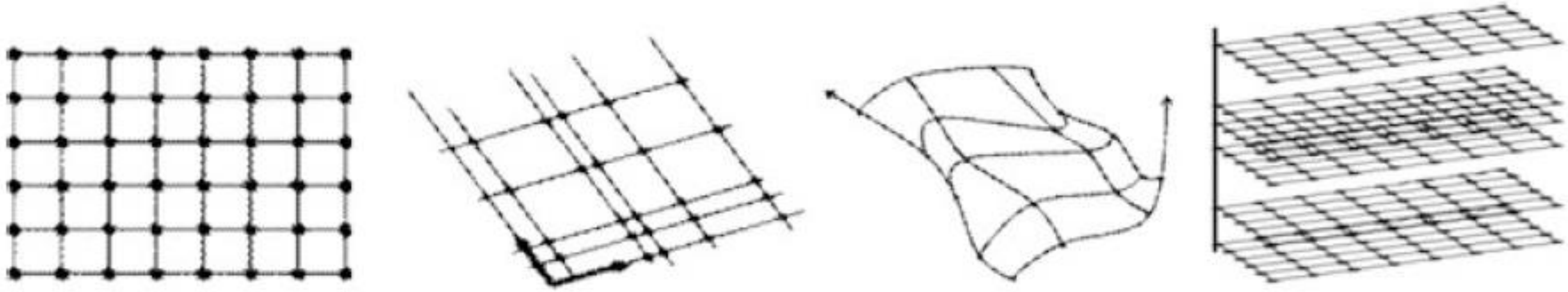


The datacube manifesto, written by P. Baumann, provides a concise, crisp definition of the key concepts of datacubes.

Baumann P., 2017, “The Datacube Manifesto”, http://www.earthserver.eu/tech/datacube-manifesto

Rule #1

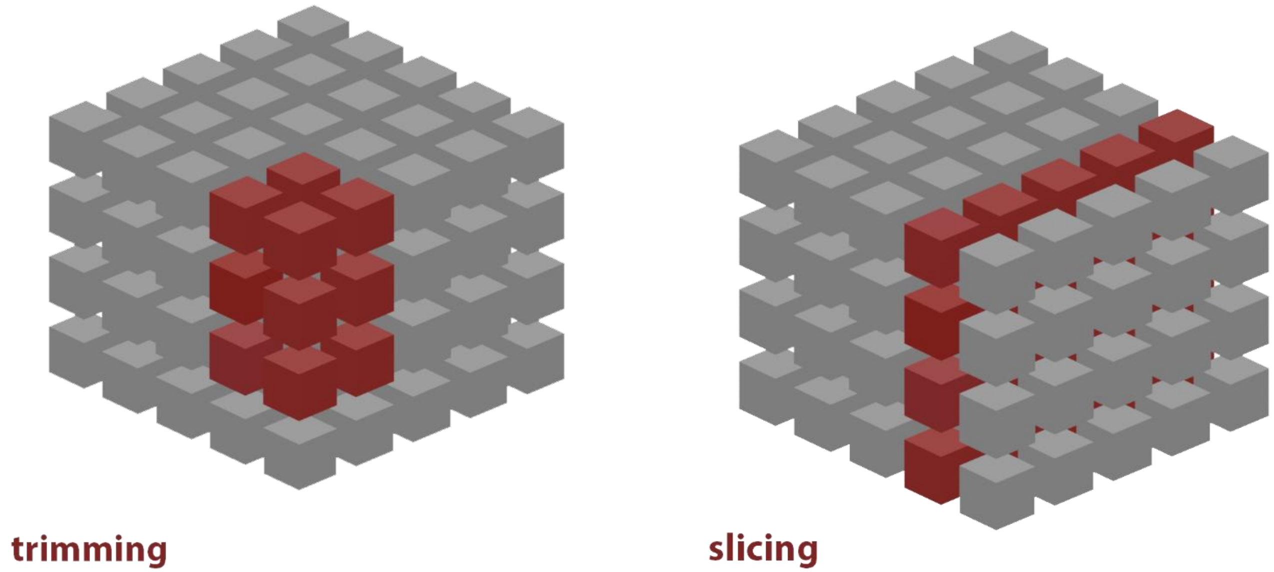
Datacubes shall support gridded data of at least one through four spatial, temporal, or other dimensions.



Baumann P., 2017, “The Datacube Manifesto”, http://www.earthserver.eu/tech/datacube-manifesto

Rule #2

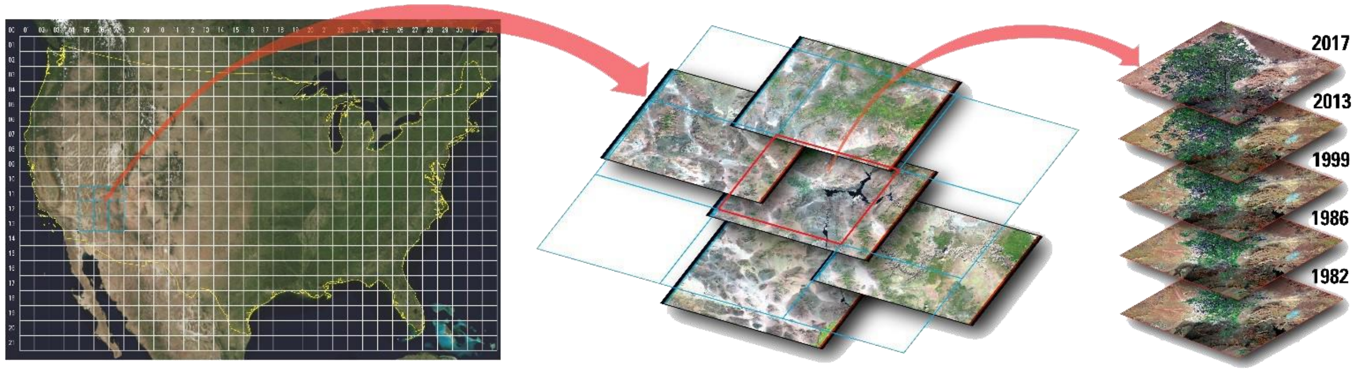
Datacubes shall treat all axes alike, irrespective of an axis having a spatial, temporal, or other semantics. ⇒ Trimming and slicing follow the same syntax along all axes.



Baumann P., 2017, “The Datacube Manifesto”, http://www.earthserver.eu/tech/datacube-manifesto

Rule #3: Efficient trimming and slicing

Datacubes shall allow efficient trimming and slicing along any number of axes from a datacube in a single request.



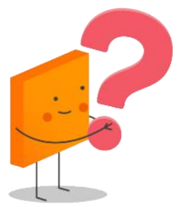
Example visualization of U.S. Landsat Analysis Ready Data tiles

Analysis-ready data means data ready for analysis along all dimensions.

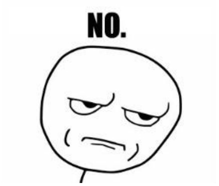
- Dwyer, J.L., Roy, D.P., Sauer, B., Jenkerson, C.B., Zhang, H., and Lymburner, L., 2018, Analysis ready data—Enabling analysis of the Landsat archive: Remote Sensing, v. 10, no. 9, art. no. 1363. doi.org/10.3390/rs10091363
- Baumann P., 2017, “The Datacube Manifesto”, http://www.earthserver.eu/tech/datacube-manifesto

Rule #4: Dimension neutrality

Datacubes shall convey similar extraction performance along any datacube axis (dimension neutrality).



Consider a stack of GeoTIFF scenes representing a timeseries, offered through a datacube API. Does it comply with the aforementioned requirement?



No! It will be fast in extracting image slices, but convey disastrous performance on timeseries analysis.

Baumann P., 2017, “The Datacube Manifesto”, http://www.earthserver.eu/tech/datacube-manifesto

Rule #5: Adaptive Partitioning

Datacubes shall allow adaptive partitioning, invisible to the user when performing access and analysis.

- Partitioning is a must for large datacubes.
- It does not add functionality, but speeds up performance tremendously if done right.
- There is not one single good partitioning; it depends on the individual mix of client access patterns, such as purely horizontal access, purely vertical access, pure timeseries access, or any combination (and variation, over time) of patterns.
- A powerful datacube implementation will offer partitioning as a tuning option

Baumann P., 2017, “The Datacube Manifesto”, http://www.earthserver.eu/tech/datacube-manifesto

Rule #6: Language Support

Datacubes shall support a language allowing clients to submit simple as well as composite extraction, processing, filtering, and fusion tasks in an ad-hoc fashion.

- analysis functionality
- any query, any time, on any size
- some high-level datacube language must be supported where users describe what they want to get, not the detailed algorithm.

Baumann P., 2017, “The Datacube Manifesto”, http://www.earthserver.eu/tech/datacube-manifesto

The 6 faces of the datacube (1/2)

- <u>Parameter Model:</u> The semantics of a cube cell value is described by a parameter model which allows understanding the information stored in each thematic layer of the cube. This includes the parameterization of the property and its quality, as well as the associated metadata that are necessary for the analysis.
- <u>Data Representation:</u> Data representation is the way in which a parameter is discretized and semantically encoded along the different axes or dimensions of the cube such as space, time, and thematic properties.
- <u>Data Organization:</u> The cell values generated by the discretisation of the parameter need to be physically arranged and stored in a machine-readable way. This encompasses issues like file formats, file systems, and database structures.

European Commission. Joint Research Centre. (2017). Proceedings of the 2017 conference on Big Data from Space (BIDS’ 2017): 28th 30th November 2017 Toulouse (France). Publications Office. https://doi.org/10.2760/383579

The 6 faces of the datacube (2/2)

- Infrastructure: The data storage units must be hosted by an IT infrastructure or 'hardware' that also allows their handling. Rapid data access and transfer between storage and processing instances are important criteria.
- Access and Analysis: Within the infrastructure, a wide range of functionalities must be implemented through software to access, manipulate and analyze the stored data (and metadata) and to ingest new products into the data cube. These functionalities must be documented and made available to users by means of APIs and other interactive interfaces (GUIs).
- Interoperability: A large, growing number of open-source and proprietary implementations support WCPS so that interoperable access to data cubes is possible through a wide range of tools

European Commission. Joint Research Centre. (2017). Proceedings of the 2017 conference on Big Data from Space (BIDS' 2017): 28th 30th November 2017 Toulouse (France). Publications Office. https://doi.org/10.2760/383579

Datacube Platforms

https://sepal.io

EURO DATA CUBE

https://eurodatacube.com

Google Earth Engine

https://earthengine.google.com

sentinelhub

https://www.sentinel-hub.com

OPEN DATA CUBE

https://www.opendatacube.org

open EO

https://openeo.org

JEODPP

https://jeodpp.jrc.ec.europa.eu/bdap/

The challenge of interoperability

Interoperability was first defined by the IEEE as “the ability of two or more systems or components to exchange information and to use the information that has been exchanged”.

- Interoperability among EODCs
- Interoperability of EODCs with Other Types of Geospatial Data Cubes
- Interoperability of EODCs with General-Purpose Data Cubes

- Geraci, A. IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries; IEEE Std 610; IEEE Press Piscataway: Piscataway, NJ, USA, 1991; pp. 1–217.
- Giuliani, G., Masó, J., Mazzetti, P., Nativi, S., & Zabala, A. (2019). Paving the Way to Increased Interoperability of Earth Observations Data Cubes. In Data (Vol. 4, Issue 3, p. 113). MDPI AG. https://doi.org/10.3390/data4030113

Cloud Geoprocessing business models

CERTH

Why chose cloud computing?

Cloud scalability in cloud computing refers to the ability to increase or decrease IT resources as needed to meet changing demand. Scalability is one of the hallmarks of the cloud and the primary driver of its exploding popularity with businesses.

Deployment automation allows applications to be deployed across the various environments used in the development process, as well as the final production environments. This results in more efficient, reliable, and predictable deployments. Solutions that automate your deployment processes improve the productivity of both the Dev and Ops teams and enable them and the business to develop faster, accomplish more, and ultimately build better software that is deployed more frequently and functions more reliably for the end-user.

Arup Dasgupta. Cloud computing: The disruptive cloud. https://www.geospatialworld.net/article/cloud-computing-the-disruptive-cloud/

Why chose cloud computing?

Virtualization in cloud computing is defined as a creation of a virtual version of a server, a desktop, a storage device, an operating system, or network resources. It is essentially a technique or method that allows the sharing of a single physical instance of a resource or that of an application amongst multiple organizations or customers. It aids to separate the service from the underlying physical delivery of the service. With the help of this technique, multiple operating systems and applications can run on the same machine and hardware subsequently.

Arup Dasgupta. Cloud computing: The disruptive cloud. https://www.geospatialworld.net/article/cloud-computing-the-disruptive-cloud/

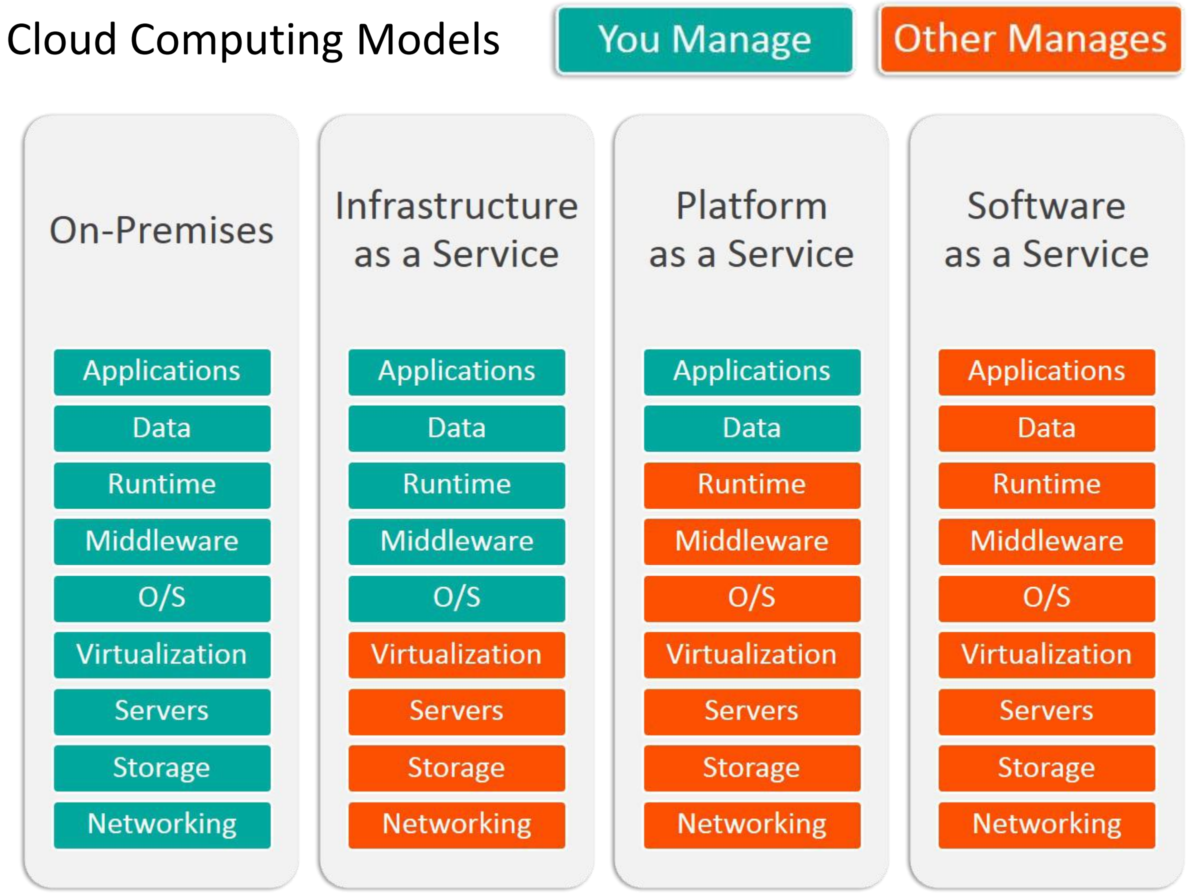
Why chose cloud computing?

Pay-as-you-use (or pay-per-use) is a payment model in cloud computing that charges based on resource usage. The practice is similar to the utility bills (e.g. electricity), where only actually consumed resources are charged. One major benefit of the pay-as-you-use method is that there are no wasted resources (that were reserved, but not consumed), which can be a source of significant losses for the companies. Users only pay for utilized capacities, rather than provisioning a chunk of resources that may or may not be used.

Arup Dasgupta. Cloud computing: The disruptive cloud. https://www.geospatialworld.net/article/cloud-computing-the-disruptive-cloud/

Cloud Geoprocessing business models

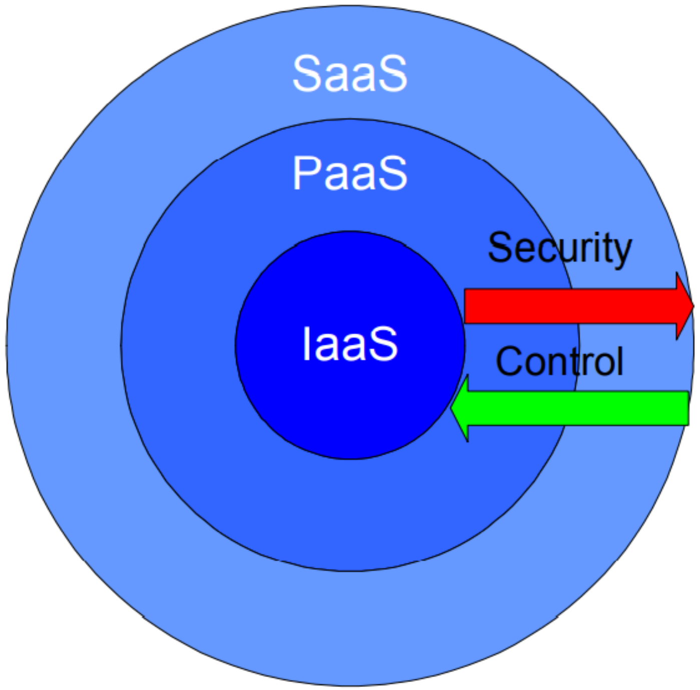
Companies and organizations are experiencing an unprecedented burden on their IT infrastructure as they struggle to meet growing user requirements/expectations for fast, reliable, and secure services. As they try to increase the processing power and storage capabilities of their IT systems, often these companies/organizations find that the development and maintenance of a robust, scalable, and secure IT infrastructure is prohibitively expensive. Cloud computing is a rapidly-growing industry which allows companies to move beyond on-premise IT infrastructure and, instead, rely on internet-based services. Cloud computing is offered in three different service models which each satisfy a unique set of business requirements. These three models are known as Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).



Cloud Computing deployment models

Private	Hybrid	Public
A cloud computing model in which an enterprise uses a proprietary architecture and runs cloud servers within its own data center	A cloud computing model that includes a mix of on-premises, private cloud and third-party public cloud services with orchestration between the two platforms	A cloud computing model in which a third-party provider makes compute resources available to the general public over the internet. With public cloud, enterprises do not have to set up and maintain their own cloud servers in house.
• Single-tenant architecture • On-premises hardware • Direct control of underlying cloud infrastructure	• Cloud bursting capabilities • Benefits of both public and private environments	• Multi-tenant architecture • Pay-as-you-go pricing model

Security vs User-control



Taking into consideration the trade-off between security and control is crucial when choosing which cloud computing business model to use.

While the IaaS model provides the most control over development, network and storage, it is also the most susceptible to risk. Users can upload anything they want onto the machine, or even develop malicious code straight on the machine. This model can be further attacked by taking advantage of the privacy and confidentiality agreements and violating the terms of service agreement. Further up the model layers, less is at risk for the provider and user.

Ludwig, B., Coetzee, S. Implications of security mechanisms and Service Level Agreements (SLAs) of Platform as a Service (PaaS) clouds for geoprocessing services. *Appl Geomat* **5,** 25–32 (2013). https://doi.org/10.1007/s12518-012-0083-3

Service Level Agreements (SLAs)

A cloud SLA (cloud service-level agreement) is an agreement between a cloud service provider and a customer that ensures a minimum level of service is maintained. It guarantees levels of reliability, availability and responsiveness to systems and applications; specifies who governs when there is a service interruption; and describes penalties if service levels are not met.

Example: G632 Sevice Level Agreement ArcGIS Online, ArcGIS Velocity and Site Scan for ArcGIS Mar 14 2022 (esri.com)
https://www.esri.com/content/dam/esrisites/en-us/media/legal/referenced-files/g-632-agol-service-level.pdf