Overview of Webification

Ed Armstrong, Zhangfan Xing
NASA Jet Propulsion Laboratory

2018 ESIP Summer Meeting
Tucson, AZ

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

© 2018 California Institute of Technology.
Government sponsorship acknowledged.

What can webification do for me ?

* Web service to inspect, slice and dice satellite records !
  * Inspect a single variable in a granule
  * Quickly review metadata contents
  * Subsample the data in space dimension
  * Extract a long time series of data for a specific region
  * Filter and extract variables of interest based on quality flags as netCDF files
  * Both machine and human friendly
  * … any piece of information from a science data store can be directly accessed, through a unique and meaningful URL …..

…….One w10n URL is worth a thousand lines of code.

Webification

* Webification
  * AKA: Webification-science or w10-sci (for scientific arrays)
* Web Service protocol for subsetting, data and metadata extraction
* This presentation will focus on the front end capabilities of w10n-sci for satellite data

* ***Philosophy: Everything is a URL… …***
  * A directory branch
  * A granule
  * A variable inside a granule
  * A piece of metadata inside a variable
* Powerful capabilities
  * Subset-by-value: A variable subsetted based on the contents of other variables
  * Outputs in user friendly JSON and netCDF
  * Built on NumPy: fast server side execution
  * Can "proxy" other web services like OPeNDAP
  * Doesn't matter: Level 1, Level 2, Level 3 or Level 4 data

W10n-sci Syntax

* W10n-sci syntax can be long but always resolves to form a URL as follows:
  * http://host:port/path/store/identifier?queryString
  * String /identifier identifies
    * *meta info of a node or leaf* by ending with a forward slash '/', e.g., /sea_surface_temperature/or
    * *data info of a leaf* by ending with an enclosing square-bracket pair '**[...]**', e.g., /sea_surface_temperature**[0:1:1,0:2030:10,0:1354:10]**
  * String ?queryString indicates how the request should be done, such as, in what output format data of a leaf should be returned.

Two server flavors of webifcation

* **Pomegranate** is a reference implementations of w10n-sci in Python and supports data formats commonly used in Earth science: HDF 4 and 5, all versions of netCDF, etc. Pomegranate is open source software.
* **Juneplum** is a special implementation of w10n-sci in Java. It is also open source software

PO.DAAC webification

* All PO.DAAC data has been ”webified”
  * https://podaac-w10n.jpl.nasa.gov/allData/
* The ACCESS developed Virtual Quality Screening Service (VQSS) uses w10n-sci for the backend
  * https://podaac-access.jpl.nasa.gov
* Other NASA w10n servers exists

Future recipe

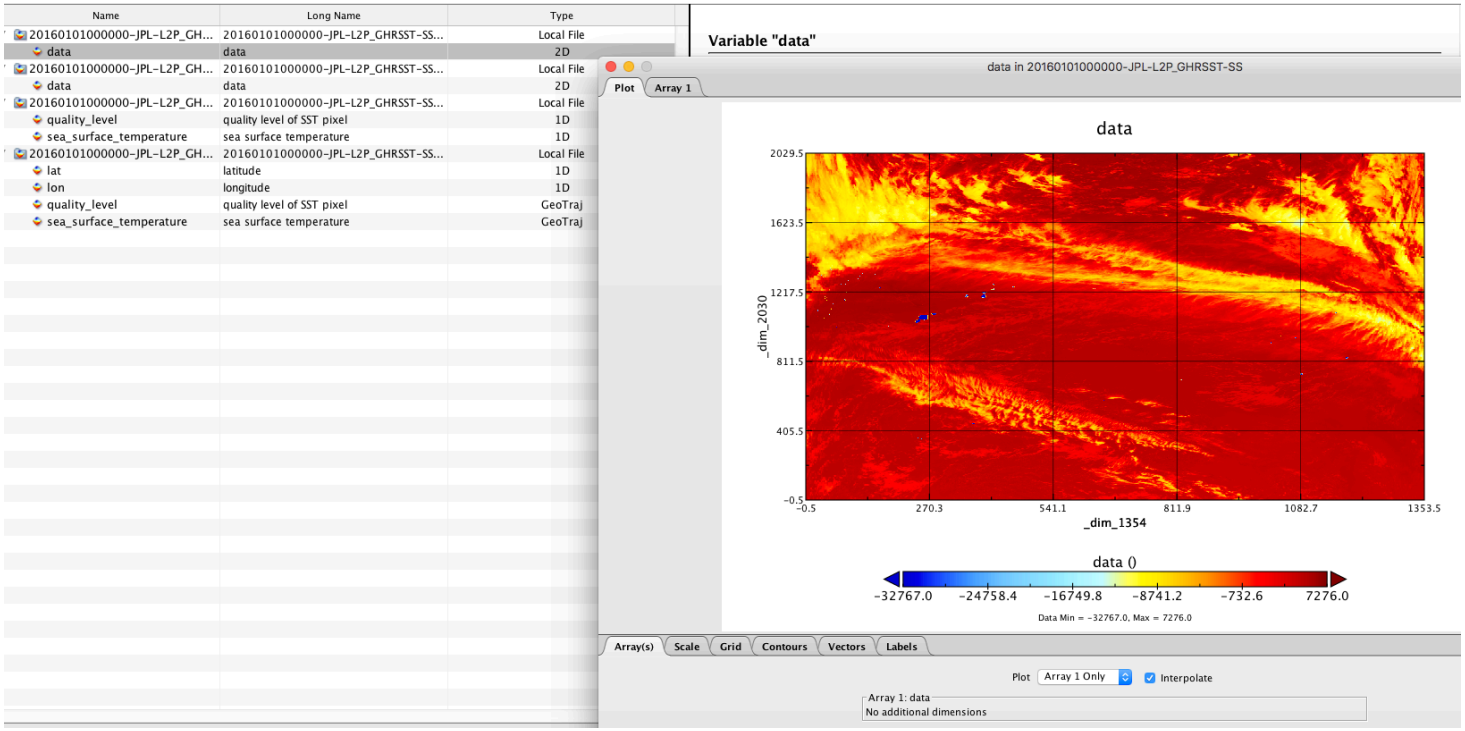
* Python based webification recipe
  * Subset a years worth of GHRSST MODIS Aqua Level 2 SST data over a 0.02x0.02 decimal degree box in Gulf of Mexico screened by highest quality flag
    * 1 km, 100 of granules
    * I know which granules to operate on based on a discovery request

Examples

* Directory and Granule exposure
* Granule metadata
* Granule data extraction
  * Entire array, by stride, by value
  * Different data outputs
* https://podaac-w10n.jpl.nasa.gov/howto/

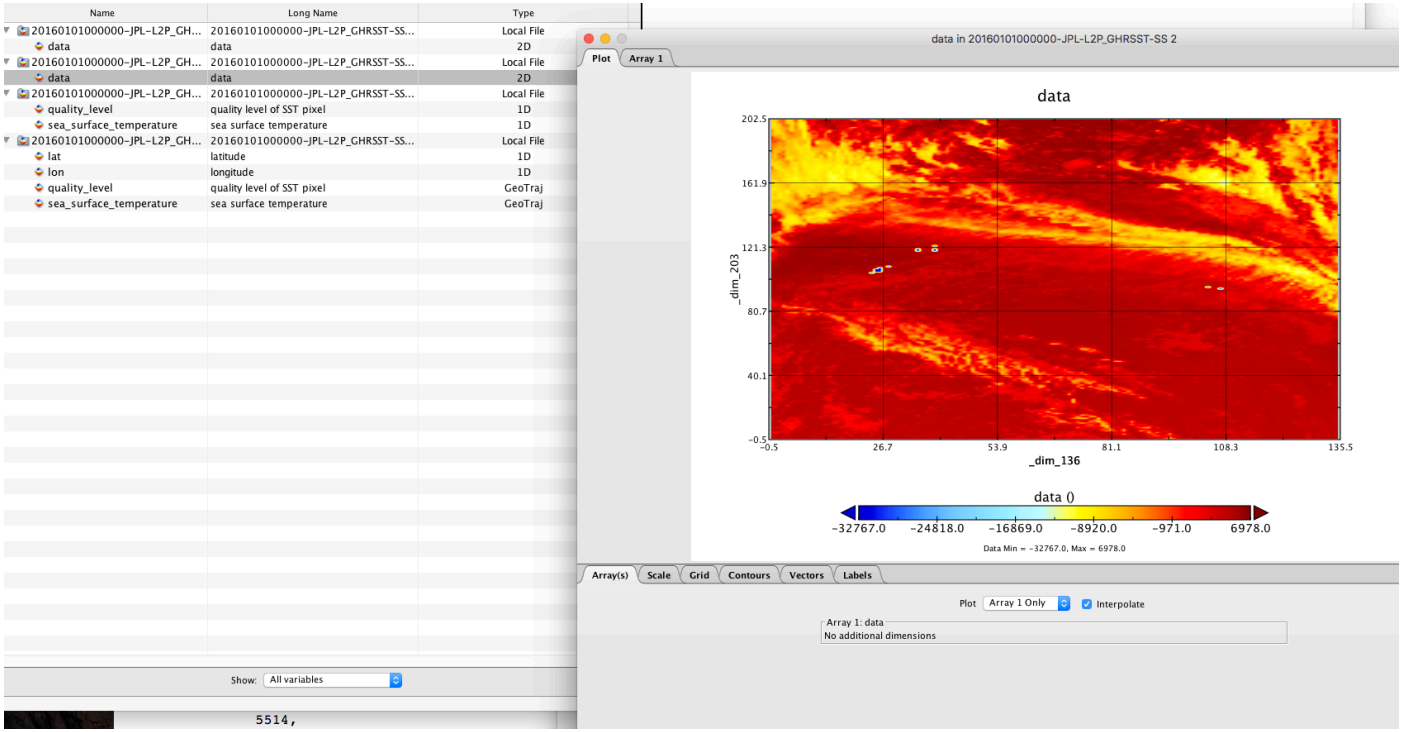
SST MODIS granule

https://.../20160101000000-JPL-L2P_GHRSST-SSTskin-MODIS_A-D-v02.0-fv01.0.nc/**sea_surface_temperature[]?output=nc**



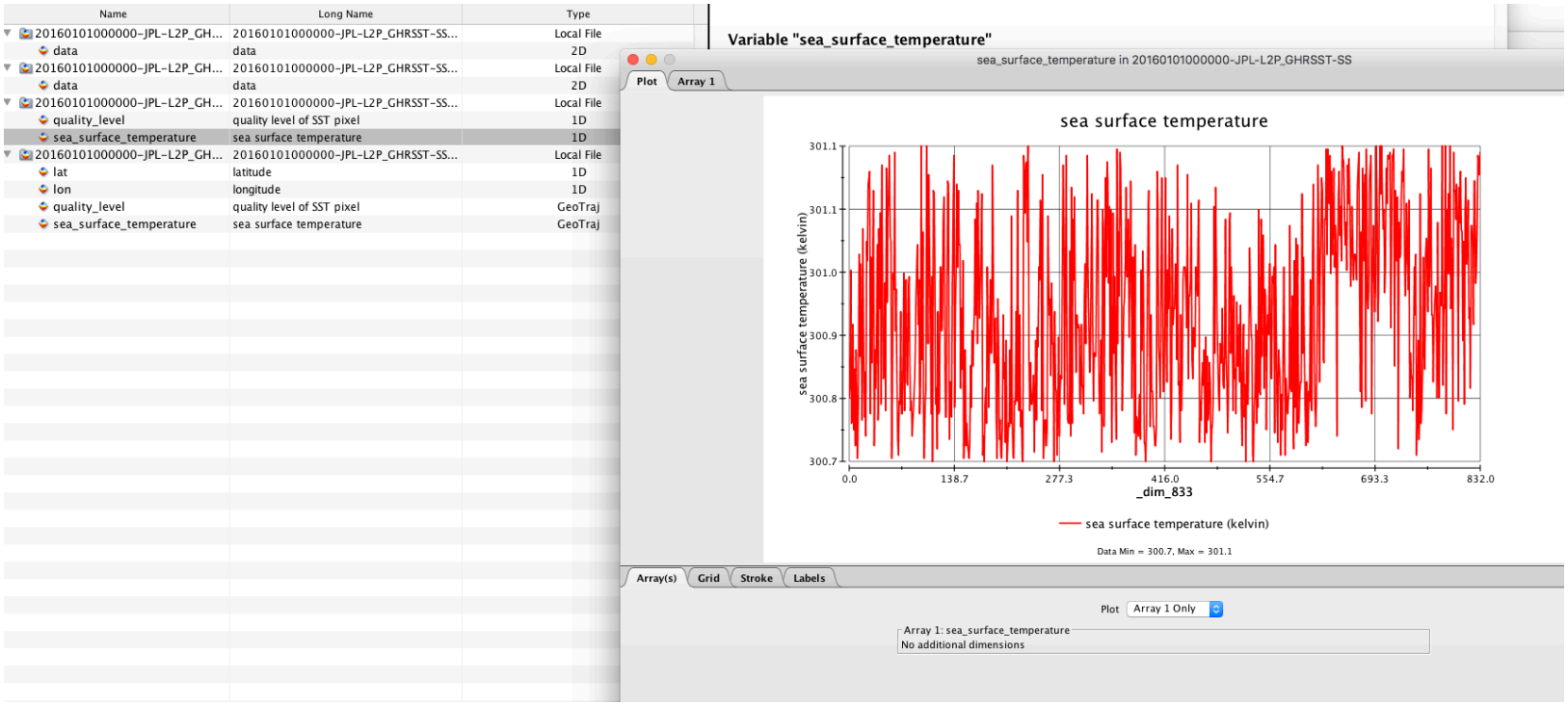
Stride by 10

https://podaac-w10n.jpl.nasa.gov/allData/ghrsst/data/GDS2/L2P/MODIS_A/JPL/v2014.0/2016/001/20160101000000-JPL-L2P_GHRSST-SSTskin-MODIS_A-D-v02.0-fv01.0.nc/**sea_surface_temperature[0:1:1,0:2030:10,0:1354:10]?output=nc**



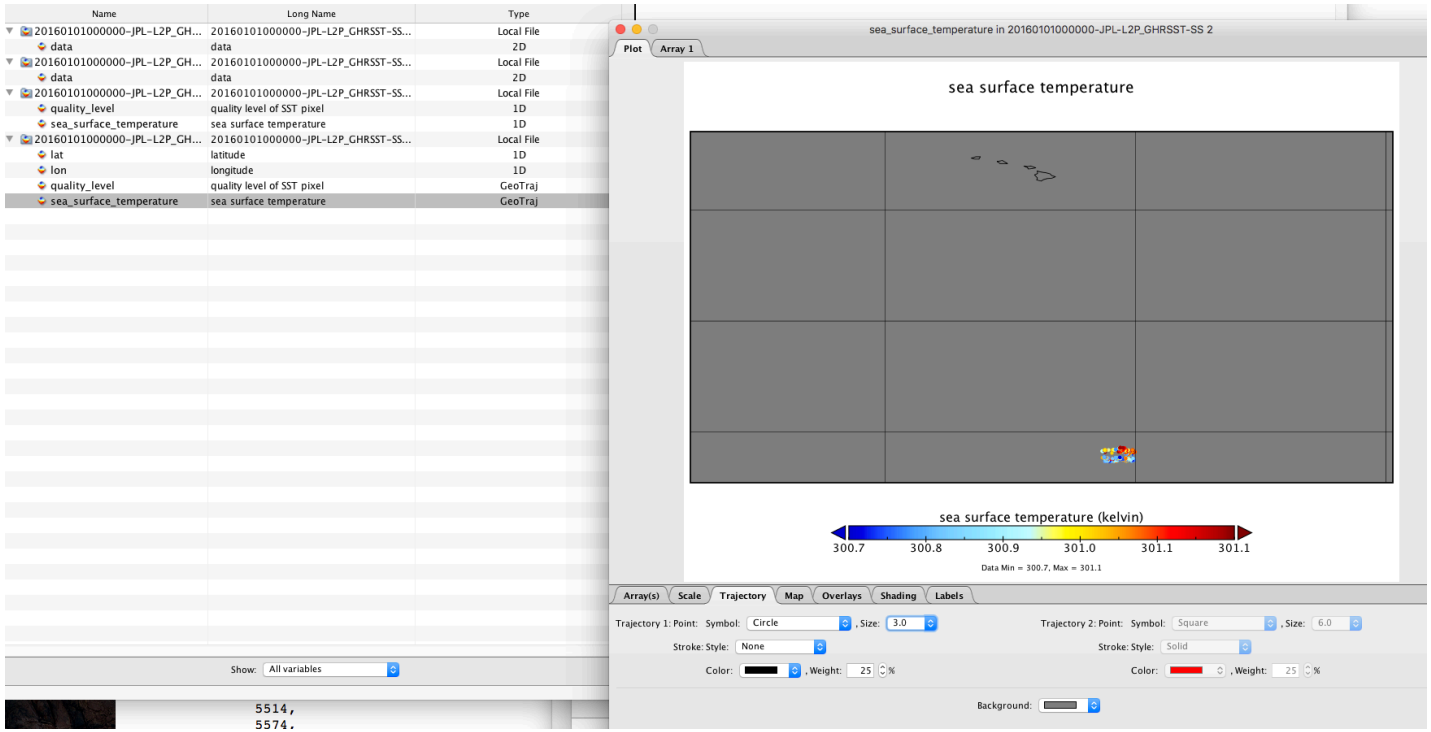
Just quality = 5 (best)

https://.../20160101000000-JPL-L2P_GHRSST-SSTskin-MODIS_A-D-v02.0-fv01.0.nc/{sea_surface_temperature,quality_level}/[0<sea_surface_temperature≤100,-19≤lat≤-17,-152≤lon≤-150,quality_level==5]?output=nc



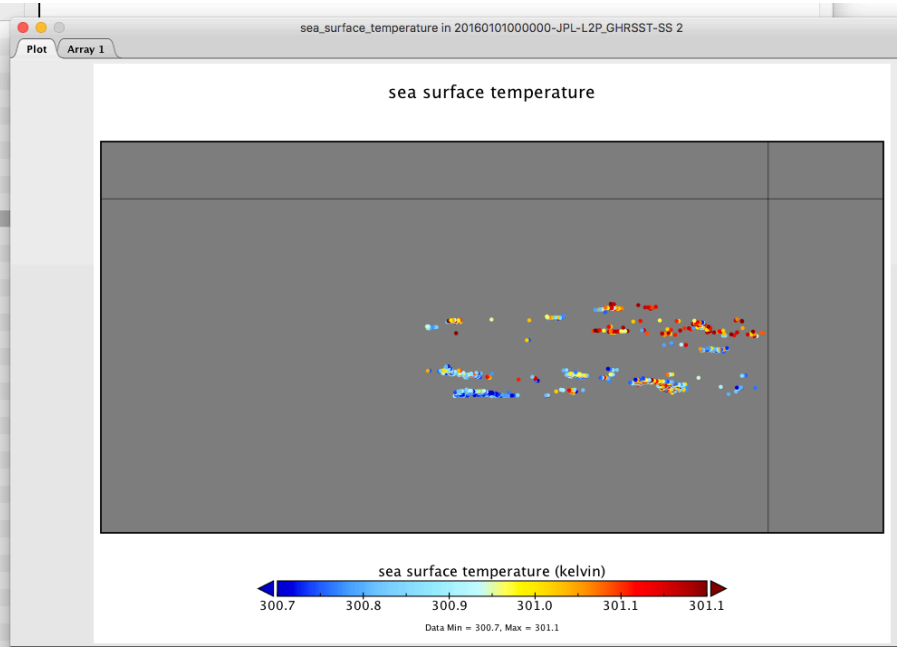
Just quality = 5 (best) as a map

https://.../20160101000000-JPL-L2P_GHRSST-SSTskin-MODIS_A-D-v02.0-fv01.0.nc/{sea_surface_temperature,lat, lon, quality_level}/[0<sea_surface_temperature≤100,-19≤lat≤-17,-152≤lon≤-150,quality_level==5]?output=nc



Just quality = 5 (best) as a map

Name	Long Name	Type
20160101000000-JPL-L2P_GH...	20160101000000-JPL-L2P_GHRSST-SS...	Local File
data	data	2D
20160101000000-JPL-L2P_GH...	20160101000000-JPL-L2P_GHRSST-SS...	Local File
data	data	2D
20160101000000-JPL-L2P_GH...	20160101000000-JPL-L2P_GHRSST-SS...	Local File
quality_level	quality level of SST pixel	1D
sea_surface_temperature	sea surface temperature	1D
20160101000000-JPL-L2P_GH...	20160101000000-JPL-L2P_GHRSST-SS...	Local File
lat	latitude	1D
lon	longitude	1D
quality_level	quality level of SST pixel	GeoTraj
sea_surface_temperature	sea surface temperature	GeoTraj



Additional Resources

* Whitepaper documentation
  * More documentation of specification, examples and comparisons to other services

Backup -- Advanced Examples

* Subset by bit flags values