

Python access to Geo Data Portal

Title

Enabling Geo Data Portal (GDP) functionality for commonly used scientific analysis environments

Primary contacts: Curtis Price, Dave Blodgett, and Rich Signell

Abstract

The Geo Data Portal (GDP) developed by USGS CIDA provides a powerful functionality to process climate and other large remote gridded data, returning summary statistics and areal tabulations over user specified areas. Currently the user interface is a interactive web form, returning a CSV table. This work would allow ArcGIS, Python, R and other common scientific client applications users to access this functionality directly via a GDP Python module.

Description

Project Topic	Tasks	Resources Required	Major Outcomes	Total Funding Needed
Python and Matlab port of GDP client tools (WPS) w/ hooks for ArcGIS Toolbox	Develop python functions to execute GDP web processing services Develop python based ArcGIS Toolbox to return GDP data products	~1 month python developer ~1 month Matlab developer ~1 month ArcGIS developer Travel Funds for two, three-day face to face meetings.	Python scripting tools and ArcGIS toolbox to access GDP web processing services.	Travel \$ money for face to face planning and development ~\$5k-10k Funds for contract developer salary ~3 Months FTE

Additional Criteria

- No ArcGIS or other software required
- Python 2.5, 2.7 etc

Benefit to FSP/Scientists/Mission Areas

The creation of a python module will allow many commonly-used scientific software tools (ArcGIS, Python, R, Matlab, Microsoft Office) to access GDP analysis functionality. The GDP is now accessible through an interactive web page, which is another system to learn and data formats to navigate.

In kind funding and work leveraged

- Opens CIDA's Geo Data Portal functionality to scientific app clients

Interaction with DMWG

Review of software distribution for appropriate metadata, policy implications

Deliverable and its Measurable Benefit

- Python module supported on Windows, Linux (Ubuntu), OS X

Methodology (process)

- Proposed Functionality
1. Input
 - Specified pre-loaded layers on server (by keyword)
 - Allow user-defined OPeNDAP URL as input
 - Validate OPeNDAP URL as valid gridded data
 - Return ranges of lon,lat,time and height (describe grid)
 - Return available statistical techniques
 - Allow specification of l
 - lon/lat range
 - time range
 - depth/height range
 - statistical techniques
 - shapefile or polygon as array (numpy)

- Output
 - Return list of preloaded GDP gridded data sets from GDB server
 - Return available preloaded shapefiles
 - Interface to ArcGIS with validated inputs
 - calculated time and summary statistical data as arrays (numpy)
- Design
 - Co-developed with end users and contractor via agile development approach (1 week iteration for 4 weeks)
- Coding
 - Contractor with experience with Python will write the python module.
 - Curtis will do the ArcGIS interface work.
- Distribution/publishing
 - Include in USGS ArcGIS distribution, available on all USGS ArcGIS Desktops
 - Open Source
 - Ubuntu Distribution (Personal Package Archive)

Partnerships

- NOAA Southwest Fisheries Science Center - Environmental Research Division (SWFSC-ERD)
- Geo Data Portal project (CIDA)
- Interagency Coastal Marine Spatial Planning (CMSP) Group

Original content created at CDI 2011 meeting by Curtis Price and Rich Signell