OpenClimateGIS

A Web Service Providing Climate Model Data in Commonly Used Geospatial Formats

IN41A-1395

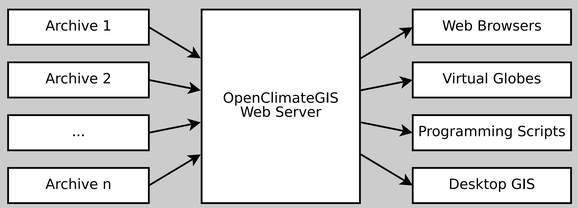
Tyler A Erickson[1], Benjamin W Koziol[1], K Arthur Endsley[1], Richard B Rood[2]

(1) Michigan Tech Research Institute (MTRI) - Michigan Technological University
Ann Arbor, Michigan
(734) 926-8537 - tyler.erickson@mtu.edu - www.mtri.org

(2) Atmospheric, Oceanic and Space Studies - University of Michigan
Ann Arbor, Michigan
(734) 647-3530 - rbrood@umich.edu

Project Overview

Access to climate model simulation data is problematic for many users outside of the climate modeling community because the dataset formats and size of the data typically necessitate specialized software and knowledge. The goal of OpenClimateGIS is to make climate model simulation data more accessible to potential users, by providing the data in commonly used geospatial data formats used by GIS desktop software, browser-based mapping tools, and virtual globes.

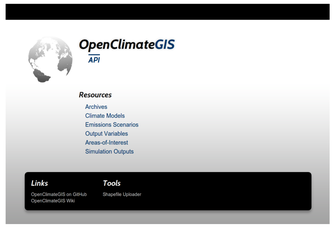


OpenClimateGIS is not an archive of climate model data, but rather is a web service that works with climate model archives that provide their data using the OPeNDAP protocol. OpenClimateGIS requests gridded climate data from the archives, clips the data to user specified areas-of-interest, and provides the data in geospatial formats that can be consumed by a wide range of client applications.

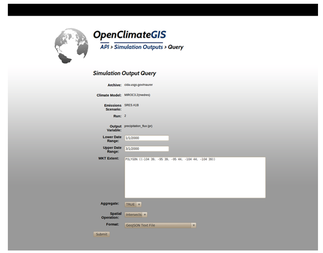
REST-style Architecture

The OpenClimateGIS application programming interface (API) provides access to climate data resources via the standard protocol of the Web: the Hypertext Transfer Protocol (HTTP). This interface has been implemented using a Representational State Transfer (REST) style of architecture. Every client request and response from the server is HTTP-compliant so that it can be understood by a wide range of software applications, including web browsers (i.e. no specialized client is needed). RESTful architectures are a popular choice for web communications, particularly in asynchronous (AJAX) applications. The nature of its architecture makes the OpenClimateGIS data services highly extensible, scalable, decoupled and self-descriptive, important features for an open-source software project with a goal of being re-used and expanded.

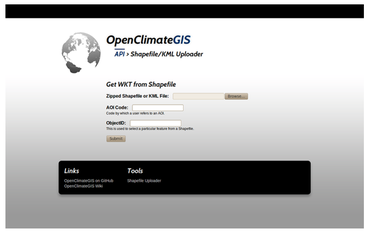
OpenClimateGIS' RESTful web services are made possible through the use of the Django-Piston library which extends the Django open-source web framework.



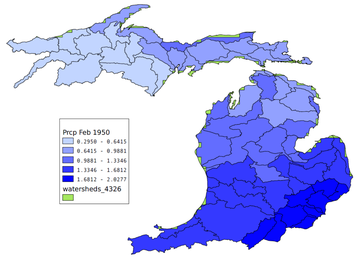
Main API web page that lists links to available data resources



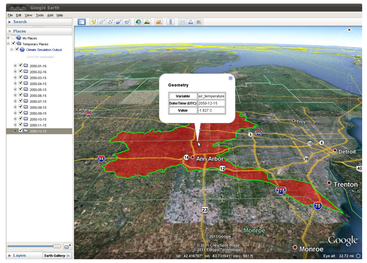
Query web page that assists users in constructing URLs that request data



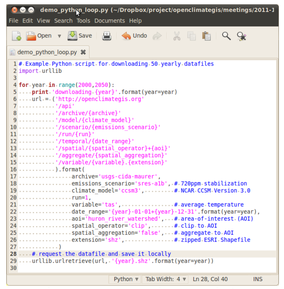
Web page that allows users to upload custom areas-of-interest



ESRI Shapefile output visualized into a desktop GIS package



KML output visualized in a virtual globe application (Google Earth)

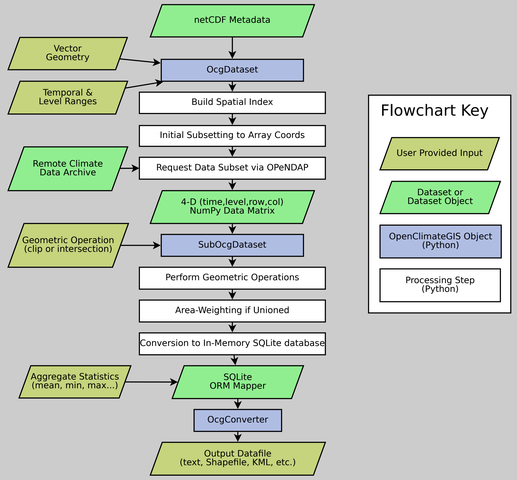


Example programming script used to download a time series of datafiles

Geospatial Processing

The flowchart presented at right gives an overview of the geospatial processing performed by the OpenClimateGIS web service.

Data extraction and conversion proceed through three primary OpenClimateGIS objects. The OcgDataset object connects to a remote netCDF dataset containing gridded climate model data and, using metadata describing the spatial discretization, vertical levels, and time coordinates, generates an initial index to limit the data cube requested from the netCDF dataset. Data retrieved via OPeNDAP is used to populate a SubOcgDataset object which facilitates geometric operations such as clipping and unioning. SubOcgDataset objects are modularly designed to be rapidly split and merged allowing the computationally intensive geometric operations to be performed in parallel using multiple processors. These objects generate an in-memory relational database representation that is passed to an OcgConverter to create various output data formats. Aggregate statistics (currently in development) operate on the relational database and take advantage of the sorting, grouping, and (when applicable) built-in functional capabilities. The modular design of OpenClimateGIS facilitates the rapid addition of new output formats and/or aggregation functions which will be added based on user needs.



Open Development

Fork me on GitHub

The OpenClimateGIS Server has been developed entirely using open source software components. Some of the main components include:

GeoDjango - a Python-based geospatial web framework. GeoDjango forms the core of OpenClimateGIS.
PostGIS - a spatially enabled relational database.
NumPy - a Python package for scientific computing.
netcdf4-python - a Python/NumPy interface for accessing netCDF data files, locally or via the OPeNDAP protocol.
Shapely - a Python package for manipulating geometric objects.
pyKML - a Python package for manipulating KML documents.

In addition the OpenClimateGIS project uses open source development principles. The source code is licensed under the liberal BSD open source license. The project is hosted as a public repository on GitHub, a website that facilitates open collaboration by providing user friendly tools for accessing source code, issue tracking, code review, and documentation. We want to make the development process transparent and as easy as possible for external developers to contribute to the project.

OpenClimateGIS has been developed to run on a standard Linux server. The live website is hosted on a Amazon Web Service virtual machine running the Ubuntu 10.04 operating system.

Project Links:
Public URL: http://OpenClimateGIS.org
Source code: https://github.com/tylere/OpenClimateGIS

www.OpenClimateGIS.org

Acknowledgments

The development of OpenClimateGIS has is supported by funding from the NOAA Climate Program Office through the "*Great Lakes Regional Integrated Sciences and Assessments (GLISA) Center*" project.
http://glisa.umich.edu

OpenClimateGIS has been strongly influenced by participation in the NOAA National Climate Predictions and Projections (NCPP) Platform project.
http://www.esrl.noaa.gov/cog/ncpp/

Thanks to the USGS Center for Integrated Data Analytics (CIDA) for hosting a THREDDS Data Server that provides the OpenClimateGIS system with gridded climate data subsets.
http://cida.usgs.gov/