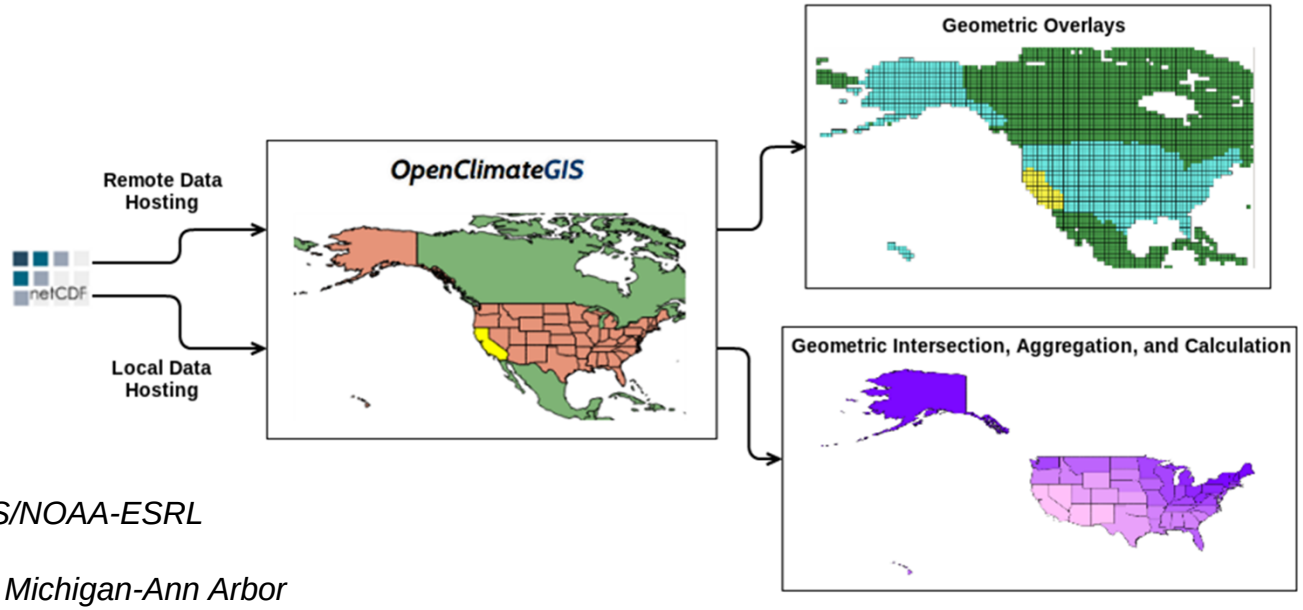
Overview of OpenClimateGIS and ClimateTranslator: A Python Library and Web Interface for Geospatial Manipulations of CF Climate Datasets

ESGF & UV-CDAT F2F Meeting, December 2013

Ben Koziol[1], Luca Cinquini[1,2], Richard Rood[3], Cecelia DeLuca[1], Sylvia Murphy[1]



1. NESII/CIRES/NOAA-ESRL
2. NASA-JPL
3. University of Michigan-Ann Arbor

Presentation Outline

1. Overview of OpenClimateGIS
2. Software Components
   a. Subsetting
   b. Computation
   c. Dataset Bundling
   d. Format Conversion
3. Next Steps

What is OpenClimateGIS?

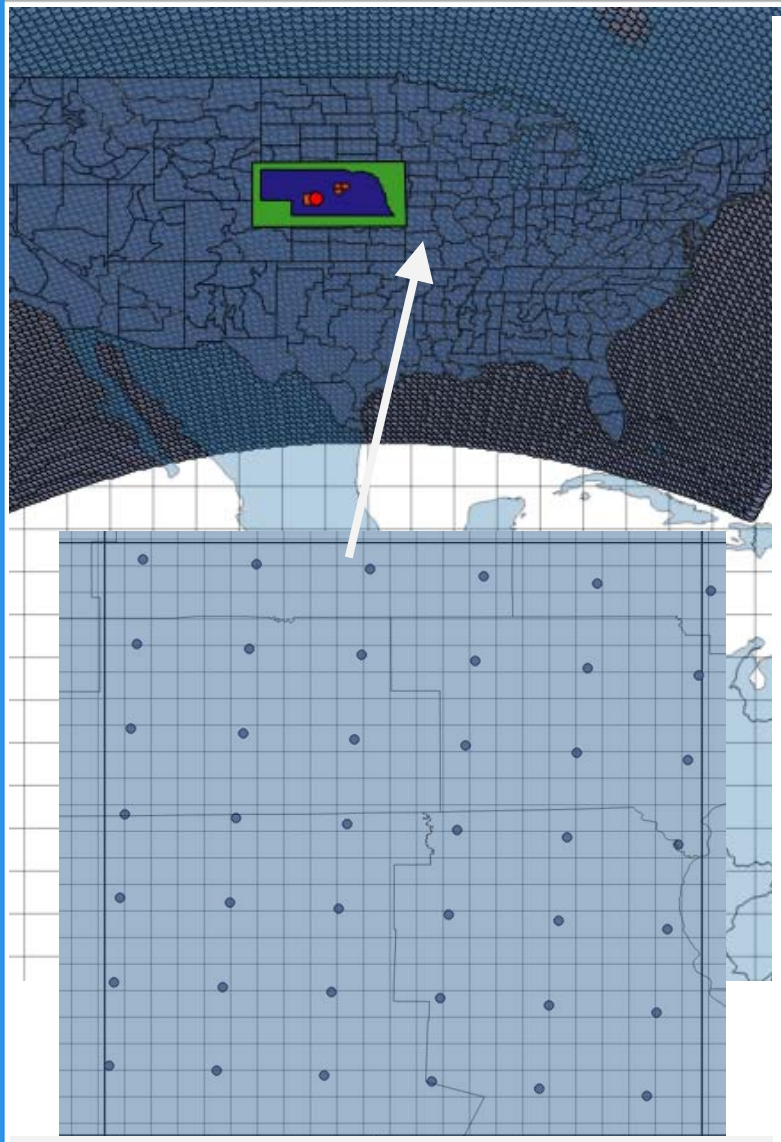
- OpenClimateGIS (OCGIS) is a standalone, Python-based, open source software library enabling dynamic access to and manipulation of climate data

- Software goal is to overcome barriers of usability of climate projections in adaptation planning and resource management
  - Translate out of climate data formats
  - Select geographical regions of interest
  - Select times/levels of interest
  - Compute application-relevant indices
  - Translate into end-user and analysis-ready formats
  - Provide comprehensive metadata

- Builds on numerous open source software libraries:
  - netCDF/HDF, netCDF4-Python, NumPy, Shapely, Fiona, GDAL

- Software has a “request-based” entry point for common operations in addition to a NumPy interface for more advanced analysis/scripting

Subsetting



CanCM4 Decadal Simulation, NARCCAP CRCM-CGCM3, Maurer Observational Grids/Points

- Handles intersects/clip/union spatial operations and many geometry types:
  - Points
  - Polygons
  - Bounding Boxes
  - Collections of Points and Polygons
- Reads geometries directly from ESRI Shapefiles, point/bounding box numeric sequences, or Shapely geometries
- Temporal subsetting - time ranges or “regions” (i.e. arbitrary month and year combinations)
- Level subsetting
- Reads and writes many common coordinate reference systems
- Wrapping and unwrapping for 360 coordinate systems

Computation

- Extensible framework designed to accommodate a variety of climate indices and metrics:
  - *Temporally grouped functions* → monthly means, annual maximums, durations
  - *Simple transforms* → natural logarithm
  - *Multivariate functions* → heat indices
- Goal is to provide a simplified structure for introducing new indices and a straightforward, timely method for documentation (currently works with the Sphinx Python documentation system)

class `ocgis.calc.library.index.heat_index.`**`HeatIndex`**(**args*, ***kwds*) [source]

Bases: `ocgis.calc.base.AbstractMultivariateFunction`, `ocgis.calc.base.AbstractParameterizedFunction`

`calculate`(*tas=None*, *rhs=None*, *units=None*) [source]

```
    dtype = np.float32
    parms_definition = {'units':str}
    required_variables = ['tas','rhs']
    key = 'heat_index'

    def calculate(self,tas=None,rhs=None,units=None):
        if units == 'k':
            tas = 1.8*(tas - 273.15) + 32
        else:
            raise(NotImplementedError)

        c1 = -42.379
        c2 = 2.04901523
        c3 = 10.14333127
        c4 = -0.22475541
        c5 = -6.83783e-3
        c6 = -5.481717e-2
        c7 = 1.22874e-3
        c8 = 8.5282e-4
        c9 = -1.99e-6

        idx = tas < 80
        tas.mask = np.logical_or(idx,tas.mask)
        idx = rhs < 40
        rhs.mask = np.logical_or(idx,rhs.mask)

        tas_sq = np.square(tas)
        rhs_sq = np.square(rhs)

        hi = c1 + c2*tas + c3*rhs + c4*tas*rhs + c5*tas_sq + c6*rhs_sq + \
             c7*tas_sq*rhs + c8*tas*rhs_sq + c9*tas_sq*rhs_sq

        return(hi)
```

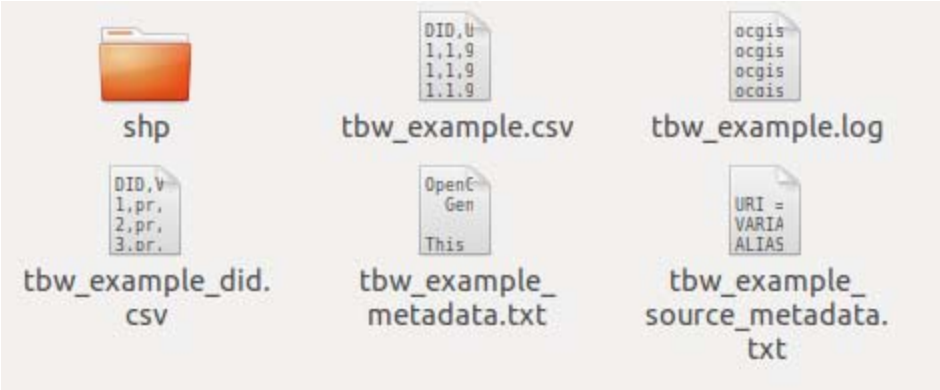
Dataset Bundling

- Bundles or packages are simply groups of data over which to apply a common set of operations
- Example data description:
  a. CMIP5 Decadal Simulation (3 degrees, 360 lat/lon, daily)
  b. NARCCAP CRCM-CGCM3 (50 km, Polar Stereographic, 3-hour)
  c. Maurer Gridded Observational (⅛ degree, 180 lat/lon, daily)
- Example operations description:
  a. Pull out all January dates
  b. Spatially subset and area-weight the values for grid cells intersecting the Nebraska state boundary
  c. Calculate the monthly mean and standard deviation
  d. Write data to CSV

VARIABLE	ALIAS	CALC_KEY	CALC_ALIAS	TIME	YEAR	MONTH	DAY	LEVEL	VALUE
tas	cmip5_tas	mean	mean_cmip5_tas	2001-01-16 00:00:00	2001	1	16		271.785
tas	cmip5_tas	std	stdev_cmip5_tas	2001-01-16 00:00:00	2001	1	16		5.59953
pr	narccap_pr	mean	mean_narccap_pr	1981-01-16 00:00:00	1981	1	16		1.0143e-05
pr	narccap_pr	mean	mean_narccap_pr	1981-02-16 00:00:00	1981	2	16		4.5889e-06
pr	narccap_pr	std	stdev_narccap_pr	1981-01-16 00:00:00	1981	1	16		2.38535e-05
pr	narccap_pr	std	stdev_narccap_pr	1981-02-16 00:00:00	1981	2	16		3.57291e-06
tasmax	maurer_tasmax	mean	mean_maurer_tasmax	1971-01-16 00:00:00	1971	1	16		1.37631
tasmax	maurer_tasmax	std	stdev_maurer_tasmax	1971-01-16 00:00:00	1971	1	16		7.20884

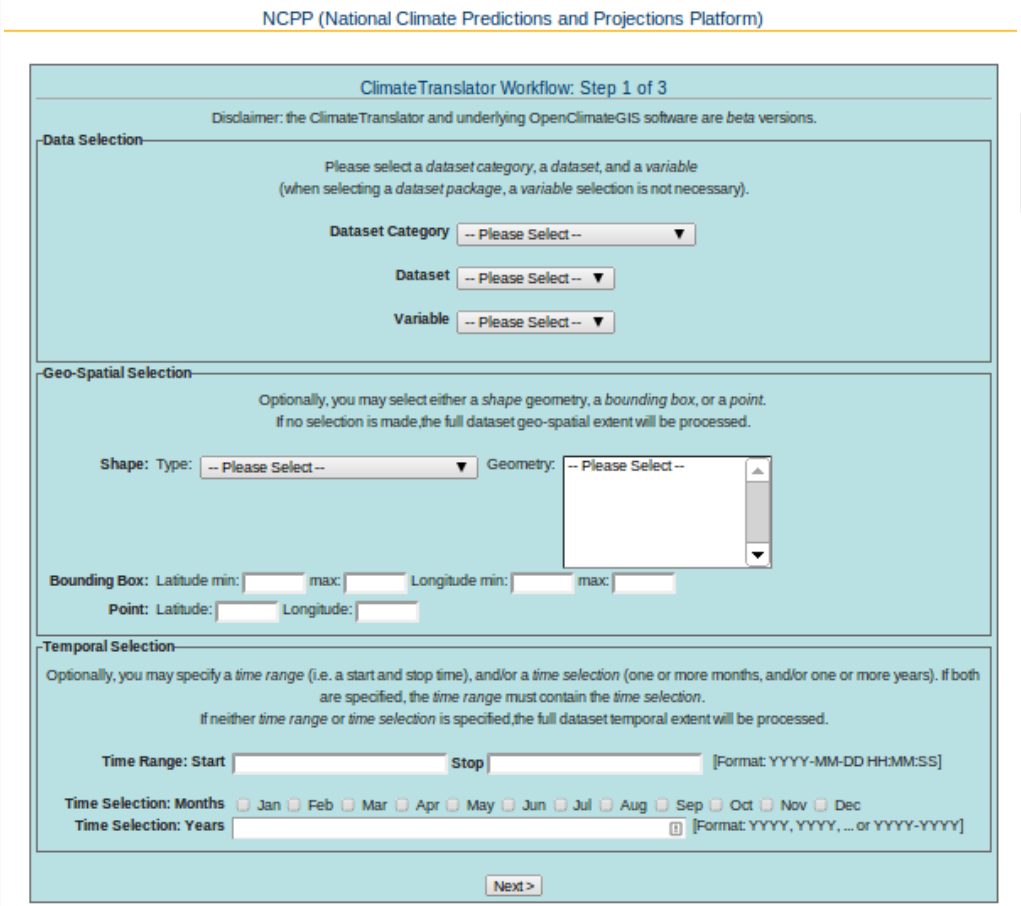
Format Conversion

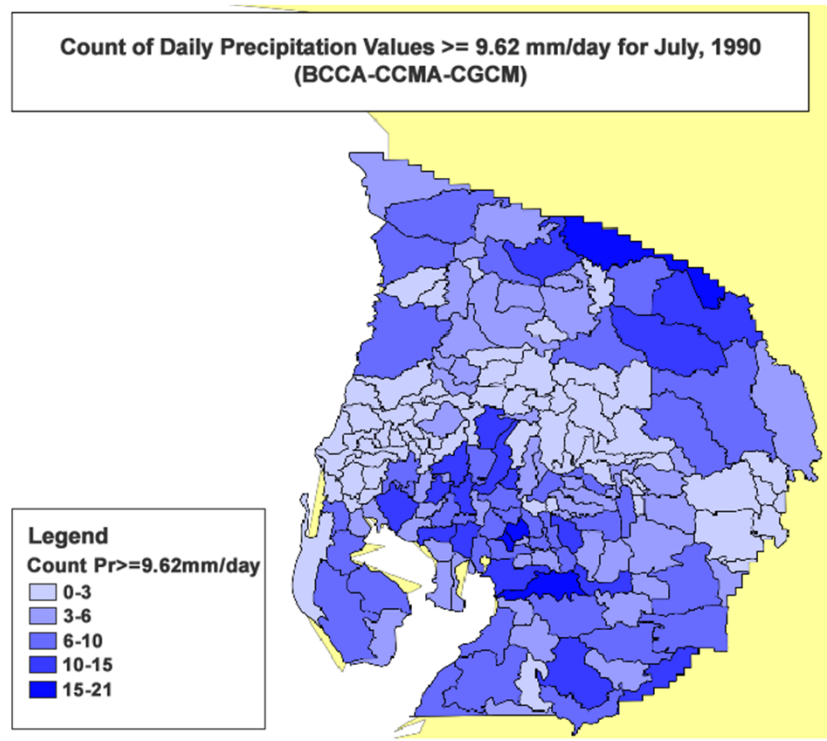
- A general framework for data conversion allows data to be streamed to multiple formats (it is not terribly difficult to add new output formats)
- Common set of headers for output files that may be adjusted to suit a user's needs (i.e. a user may only be interested in a timestamp and associated data value)
- OCGIS takes advantage of Fiona to right to OGC-compliant vector formats (i.e. ESRI Shapefile)
- Metadata dumped for the user to track source data files and applied operations
- Currently supported formats: CSV, Keyed-CSV Shapefile (CSV+), GeoJSON, netCDF, ESRI Shapefile, "NumPy"



ClimateTranslator Web Interface

- NCPP has released an initial version of the ClimateTranslator which uses OCGIS as the processing backend (Django-based)
- The workflow-based web interface exposes much of the OCGIS functionality for data manipulations





Next Steps

- **OpenClimateGIS**
  a. Unit conversion and management
  b. Performance improvements
  c. Connect with ESMPy for regridding
  d. Expand indices catalog
  e. Bring in HDF format
  f. Collaborations with other software projects!
- **ClimateTranslator**
  a. Improve user guidance
  b. Web mapping and visualization
  c. Web Processing Service (pyWPS) as communication between interface and OCGIS
  d. Integrate with CoG-based search to allow subsetting functionality on indexed datasets

Collaboration and Feedback

Questions, comments, suggestions, or “hidden features”:

ClimateTranslator Interface: ct_support@list.woc.noaa.gov

OpenClimateGIS: ocgis_support@list.woc.noaa.gov

All feedback is useful, so please do not hesitate to critique!

Mailing lists and releases:

OpenClimateGIS: ocgis_info@list.woc.noaa.gov

CoG Links:

http://www.earthsystemcog.org/projects/downscaling-2013/climatetranslator

http://www.earthsystemcog.org/projects/openclimategis/

http://www.earthsystemcog.org/projects/esmpy/