NESII Modular, High Performance Infrastructure for Earth System Modeling

Cecelia DeLuca, NOAA/CIRES
Peggy Li, JPL/NASA
Gerhard Theurich, NRL-SAIC

iEMSs Meeting 2014
June 16, 2014

Outline

- Overview
- Part 1: Modeling Tools
  - Earth System Modeling Framework (ESMF) and Grid Remapping
  - ESMF Python (ESMPy)
  - National Unified Operational Prediction Capability (NUOPC) and the Earth System Prediction Suite
  - ESMF Web Services

- Part 2: Metadata and Data Infrastructure
  - Earth System Documentation (ES-DOC)
  - CoG Collaboration Environment
  - OpenClimateGIS

The Basics

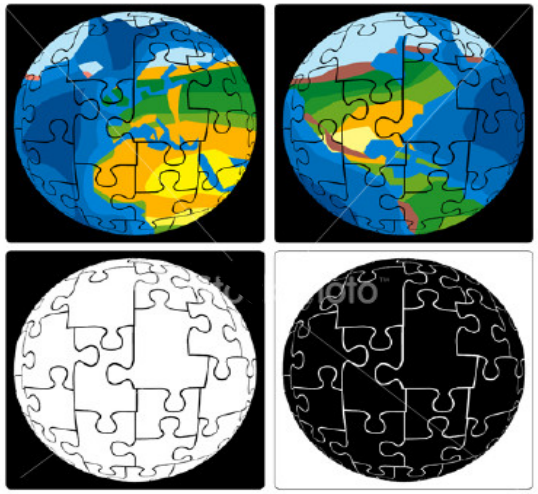
- NOAA Environmental Software Infrastructure and Interoperability Group (NESII) focuses on development of software infrastructure for Earth system modeling.
- Partners and customers are from research and operational centers, weather and climate, across U.S. agencies and international organizations.
- NESII started with the Earth System Modeling Framework (ESMF) project, a comprehensive, high performance community framework for building and coupling climate, ocean and weather models.
- As NESII has grown, its tools have become more diverse and modular.

Key Strategies

- Commitment to standards (data, metadata, component interfaces, services) to maximize interoperability
- Community-driven development and community ownership
  - Formal governance processes in which customers set priorities
  - Frequent public design reviews and demonstrations
  - Expertise shared among partners
- Openness of project metrics, code and information
- Commitment to a globally distributed and diverse development and customer base

The Earth System Modeling Framework

The Earth System Modeling Framework (ESMF) was initiated in 2002 as a multi-agency response to calls for common modeling infrastructure.



ESMF delivered:

- Standard interfaces for model components
- High performance libraries and tools for time management, data communications, metadata and I/O, and parallel grid remapping

Metrics:

~5500 downloads

~3000 individuals on info mailing list

~40 platform/compilers regression tested

~6400 regression tests

~830,000 SLOC

http://www.earthsystemmodeling.org

Standard Interfaces

- All ESMF components have the same three standard methods:
  - Initialize
  - Run
  - Finalize
- Each standard method has the same simple interface:

Steps to adopting ESMF
- Divide the application into components (without ESMF)
- Copy or reference component input and output data into ESMF data structures
- Register components with ESMF
- Set up ESMF couplers for data exchange

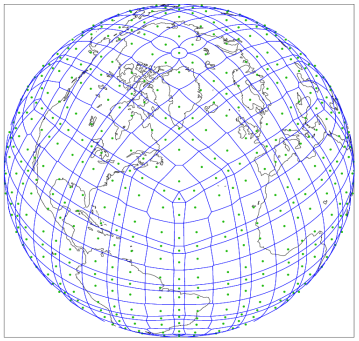
```
call ESMF_GridCompRun (myComp, importState, exportState,
clock, …)

Where:
myComp points to the component
importState is a structure containing input fields
exportState is a structure containing output fields
clock contains timestepping information
```

- *Interfaces are wrappers and can often be set up in a non-intrusive way*

ESMF Grid Remapping

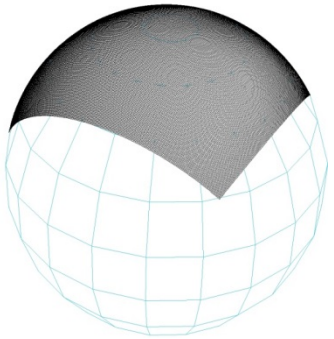
- Uniquely fast, reliable, and general – interpolation weights computed in parallel in 3D space
- Supported grids:
- Logically rectangular and unstructured grids
- Global and regional grids
- 2D and 3D grids
- Supported interpolation methods:
- Nearest neighbor, higher order patch recovery, bilinear and 1st order conservative methods
- Options for straight or great circle lines, masking, and a variety of pole treatments
- Multiple ways to call ESMF grid remapping:
- Generate and apply weights using the **ESMF API**, within a model
- Generate and apply weights using **ESMPy**, through a Python interface
- Generate weights from grid files using **ESMF_RegridWeightGen**, a command-line utility

Some supported grids …



HOMME Cubed Sphere Grid with Pentagons
Courtesy Mark Taylor of Sandia

FIM Unstructured Grid



Regional Grid

ESMPy: ESMF Python Interface for Grid Remapping

- Enables ESMF regridding to be used with very little effort, in an object oriented way:
  - Regridding applied as a callable Python object
  - Numpy array access to distributed data
  - Parallel regridding has the potential to reduce computation from hours to minutes
- Enables ESMF regridding to be used in other scientific packages with Python-based workflows – current users include:
  - UV-CDAT (PCMDI) – Ultrascale Visualization Climate Data Analysis Tools
  - PyFerret (NOAA) – Python based interactive visualization and analysis environment
  - Community Surface Dynamics Modeling System (CSDMS) – tools for hydrological and other surface modeling processes
  - Plus many individuals …

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

ESMPy Classes

- Manager
  - Initialize and Finalize
  - Logging
  - Virtual Machine (parallel distribution)
- Grid
  - Logically rectangular discretization object
- Mesh
  - Unstructured mesh discretization object
- Field
  - Grid or Mesh plus a data array and metadata
  - Derived type of the Numpy Array
- Regrid
  - Callable object which operates on two Fields to compute and apply interpolation weights

Code Examples

Create a Grid and access object information using Python properties:

```
import ESMF, numpy
grid = ESMF.Grid(max_index, num_peri_dims=1, staggerloc=ESMF.StaggerLoc.CENTER)
```

Create a Mesh from UGRID formatted NetCDF file:

```
mesh = ESMF.Mesh(filename="mesh_file.nc", filetype= ESMF.FileFormat.UGRID,
                 meshname="mymesh")
```

Regridding object acting on two Fields:

```
field1 = ESMF.Field(grid, "srcfield")
Field2 = ESMF.Field(mesh, "dstfield")
rhandle = ESMF.Regrid(field1, field2, regrid_method=ESMF.RegridMethod.CONSERE)
field2 = rhandle(field1, field2)
```

Field is derived from Numpy Array:

```
Field1[i,j] = 2.0+math.cos(gridXcoord[i,j])**2*math.cos(2.0*gridYcoord[i,j])
exact_field = sqrt(field1**5)
exact_slice = exact_field[:,:,5:-1]
```

National Unified Operational Prediction Capability (NUOPC)

- ESMF allows for many levels of components, types of components, and types of connections
- In order to achieve greater interoperability, usage and content conventions and component templates are needed
- A collaboration of NOAA, Navy, Air Force, and their research partners is building a "NUOPC Layer" that constrains how ESMF is used, and introduces metadata and other content standards

https://www.earthsystemcog.org/projects/nuopc/

NUOPC Additions to ESMF

The NUOPC Layer adds to ESMF:

- Provide utilities routines and generic components and a dictionary for standard field metadata
- Four flavors of generic components: driver, model, mediator, and connector – user can specialize the generic components.
- A formalism that defines the rules of engagement between the components using metadata
- A formalism that describes and splits the phases of standard methods, such as initialization and run.
- A formalism for checking and reporting whether component requirements are satisfied during the run sequence – *compliance checker* and *component explorer*
- Example applications showing a variety of model interactions (e.g. explicit, semi-implicit, implicit coupling)

The Earth System Prediction Suite

- The Earth System Prediction Suite (ESPS) is a collection of major U.S. weather and climate modeling codes that use ESMF with the NUOPC conventions.
- **The ESPS makes clear which codes are available as ESMF components and modeling systems.**

https://www.earthsystemcog.org/projects/esps

Model Codes in the ESPS

Currently, components in the ESPS can be of the following types:
coupled system, atmosphere, ocean, wave, sea ice

Target codes include:

- The Community Earth System Model (CESM)
- The NOAA Environmental Modeling System (NEMS) and Climate Forecast System version 3 (CFSv3)
- The MOM5 and HYCOM oceans
- The Navy Global Environmental Model (NavGEM)-HYCOM-CICE coupled system
- The Navy Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) and COAMPS Tropical Cyclone (COAMPS-TC)
- NASA GEOS-5
- NASA ModelE

ESPS Code Status

LEGEND

Compliant | (Completion date) In progress | Candidate

Coupled Modeling Systems							
	NEMS	CFSv3	COAMPS / COAMPS-TC	NavGEM-HYCOM-CICE	GEOS-5	ModelE	CESM
	2014	2014			2015	2015	2014
Atmospheres							
GFS/GSM							
NMMB							
CAM							2014
FIM	2014						
GEOS-5 FV					2015		
ModelE Atm						2014	
COAMPS Atm							
NavGEM							
NEPTUNE							
WRF							2014
Oceans							
MOM5							
HYCOM							
NCOM							
MPAS-O							
POP							2014
Ice							
CICE	2014	2014			2014	2014	2014
Wave							
WW3	2014		2014		2014		2014
SWAN							

Spanning major climate, weather, and ocean codes, ESPS is the most direct response to calls for common modeling infrastructure yet assembled

ESMF Web Services and Climate Impacts Modeling

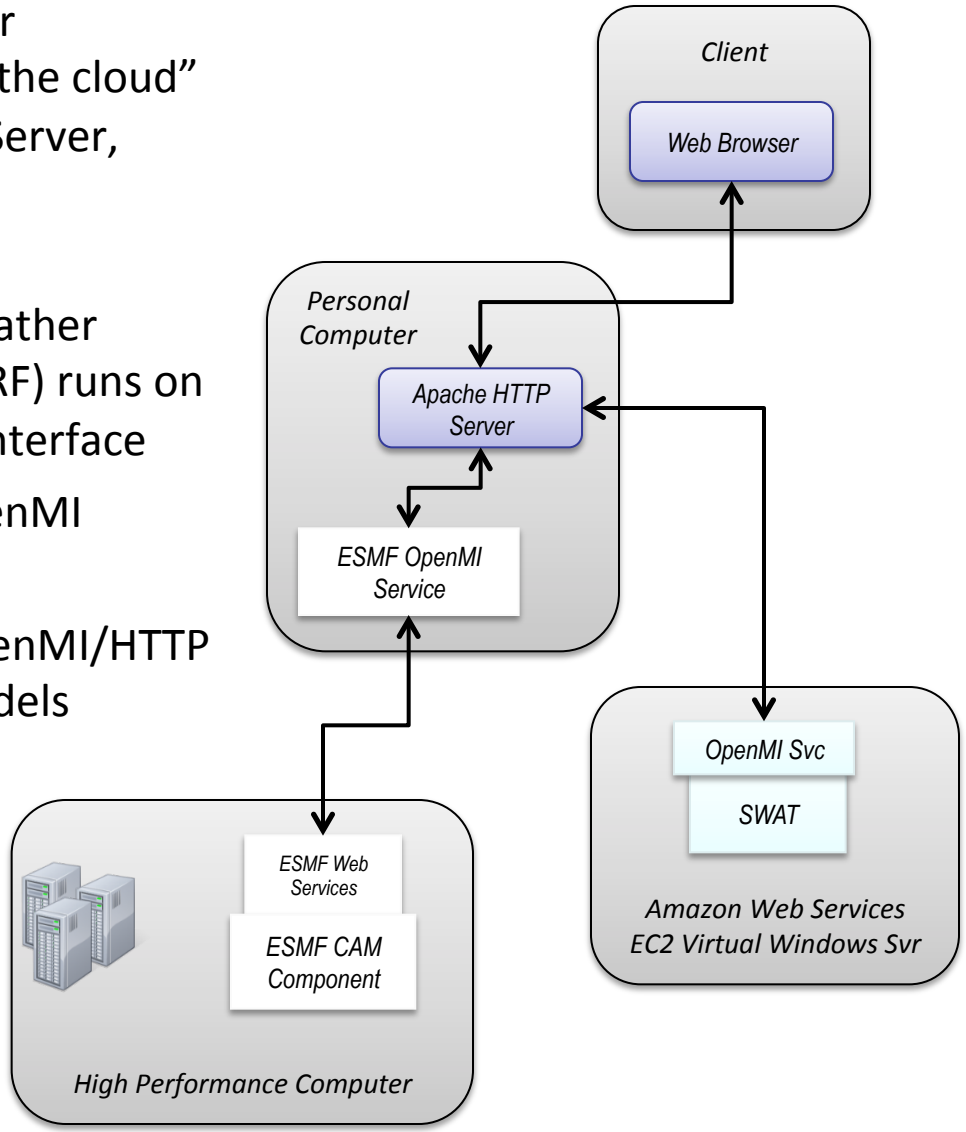
- The ESMF distribution allow any networked ESMF component to be available as a web service.

- Coupling using web services offers a new perspective on climate impacts modeling.
- Instead of what impacts are “put in” the climate model …
- How do we create a linked network of models that links climate to local and regional processes while preserving native infrastructure and specialized information delivery systems?
- Partners include NESII, University of South Carolina, University of Michigan.

Atmosphere-Hydrology Prototype

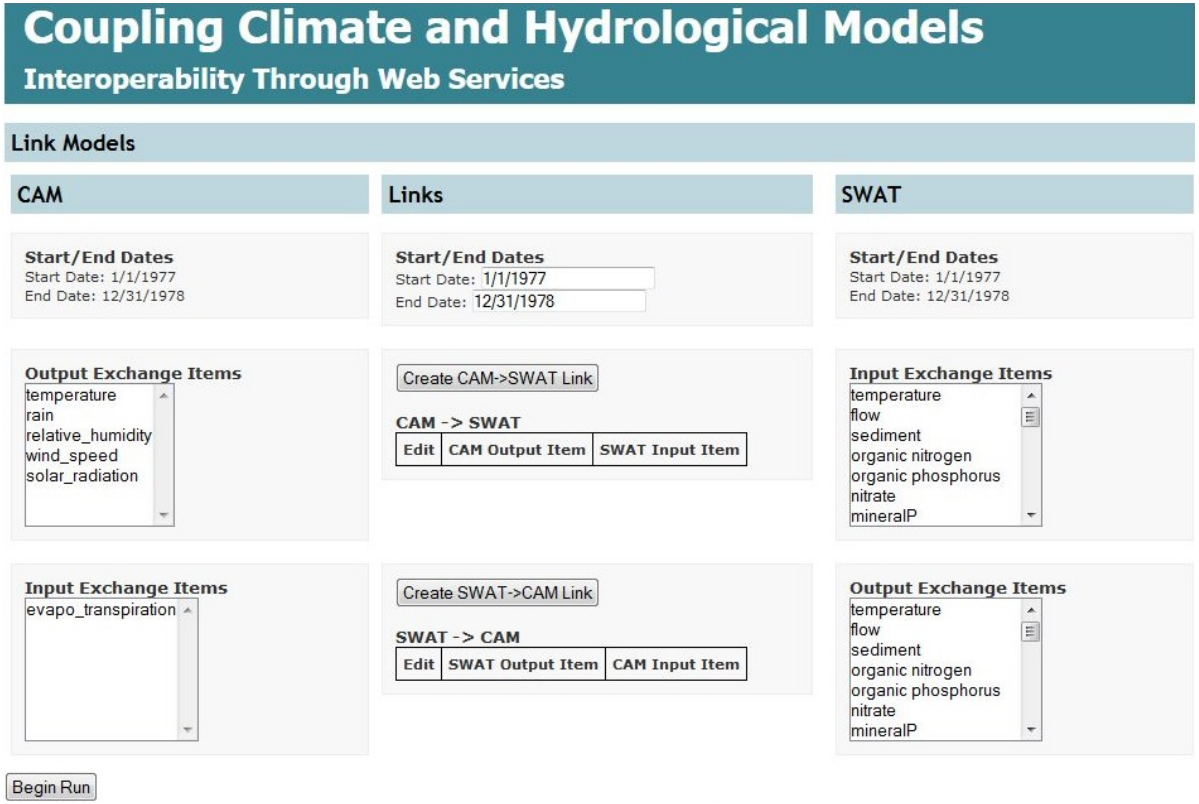
- Hydrologic model: Soil and Water Assessment Tool (SWAT) runs "in the cloud" on Amazon Web Services Virtual Server, with a native (OpenMI) interface
- Atmosphere Model: Community Atmosphere Model (CAM) or Weather Research and Forecast Model (WRF) runs on a supercomputer, with an ESMF interface
- Wrappers for models provide OpenMI RESTful service interfaces
- Driver (Web Application) uses OpenMI/HTTP interface to timestep through models

- Access to the atmosphere model across the network is through ESMF Web Services
- Atmosphere model output data is streamed via ESMF Web Services



Prototype Status

- Technical proof-of-concept is complete, not scientifically validated
- Currently refining a GUI capability which allows for specification of fields exchanged



https://www.earthsystemcog.org/projects/hydroclimatemodeling/

ES-DOC

Earth System Documentation (ES-DOC) is an international effort to develop tools to describe Earth system models in order to better understand and utilize model data. The tools are based on the Common Information Model (CIM) standard.

CIM	A metadata standard used to describe Earth system models. This includes simulations, experiments, and computing resources used by those models. The CIM is now being leveraged by a variety of international model intercomparison projects.
Questionnaire	A customizable package to generate questionnaires that ***CREATE*** model documentation.
Viewer	A browser plugin to ***DISPLAY*** model documentation. It can be embedded directly into web pages but is also available via the Search or Comparator tools.
Comparator	A web-based tool to ***COMPARE*** CIM metadata records currently stored in the CIM archive. Comparison can be output as CSV files or HTML renderings.
Search	A publicly accessible portal to ***SEARCH*** on and ***VIEW*** model documentation.

SEARCH on and VIEW
CMIP5
model documentation

CMIP5 (the Coupled Modeling Intercomparison Project) is the international modeling effort underlying the IPCC assessments. Metadata describing the models used can be viewed at the link below.

https://www.earthsystemcog.org/projects/es-doc-models/

Overview of CoG

- CoG is a collaboration environment and hub to connect projects in the Earth sciences.
- It hosts and links into networks of software development projects, model intercomparison projects (MIPS), events, and workshops.
- It provides projects with a wiki and customizable navigation to wiki content.
- It contains an ontology for the description and management of projects and provides a consolidated look at this content across a project's network.
- It includes a configurable search to data on any Earth System Grid Federation (ESGF) data node.
- It provides services for Earth system model metadata collection and display (through ES-DOC tools)
- It contains a file server for documents and images.

Earth System CoG

Some of the 70+ projects currently hosted on CoG include:

- *NOAA's High Impact Weather Prediction Project (HIWPP)*
- *Atmospheric Dynamical Core Model Intercomparison Project (DCMIP)*
- *Reanalysis Data for CMIP5 (Ana4MIPs)*
- *Observational Data for CMIP5 (Obs4MIPs)*
- *National Unified Operational Prediction Capability (NUOPC)*
- *National Climate Predictions and Projections Platform (NCPP)*
- *Earth System Documentation (ES-DOC)*
- *Earth System Prediction Capability (ESPC)*

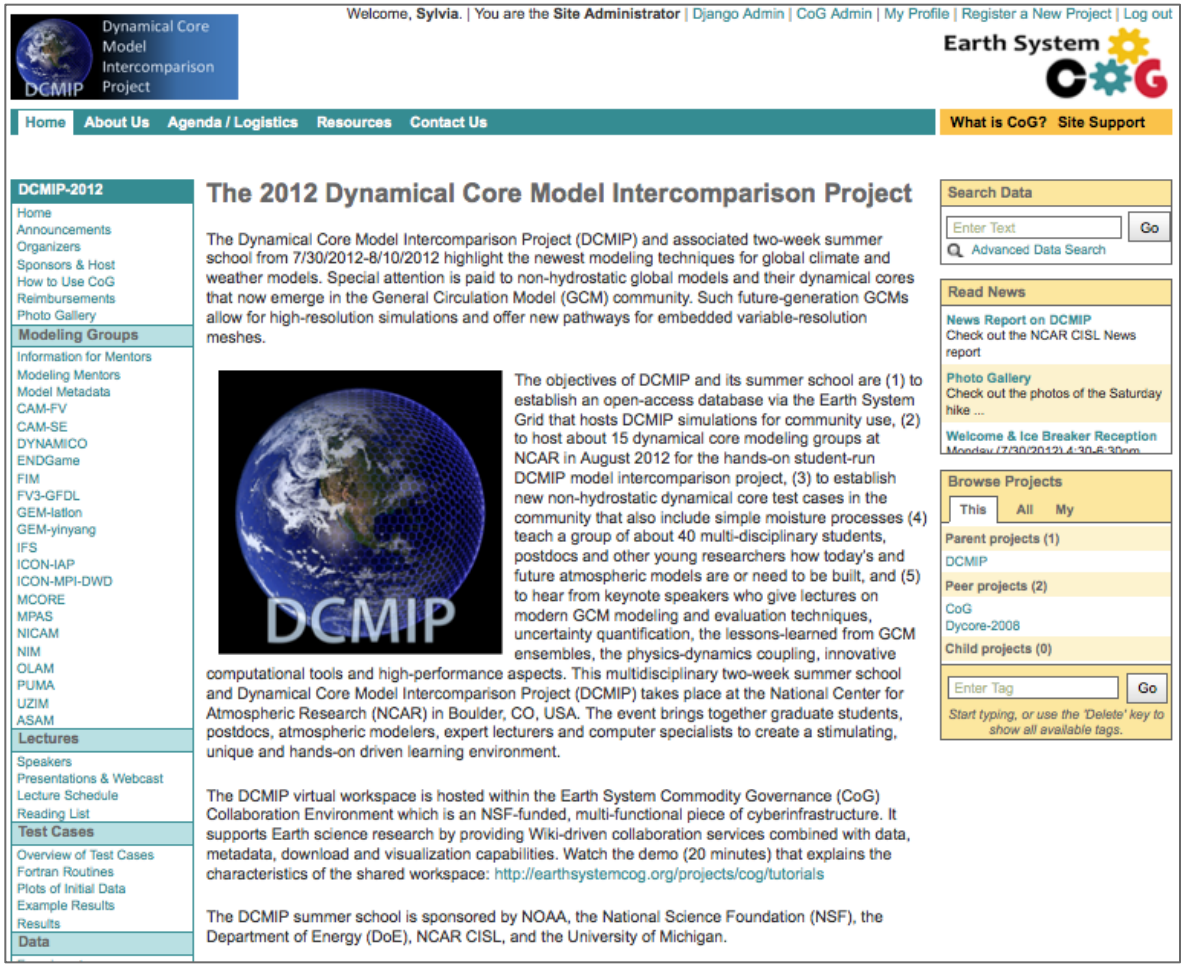
https://earthsystemcog.org/projects/cog/

Wiki and Collaboration Tools

https://www.earthsystemcog.org/projects/dcmip-2012/

The CoG layout is color-coded:

- *The right-hand side (dark yellow) is where services (data, news, project connectivity) are located.*
- *The Upper Navigation bar (dark teal) contains links to project-level metadata.*
- *On the left (light teal) is an auto-generated navigation system created when projects develop freeform content.*
- *The central portion of the site is a wiki that allows projects to create their own content.*



Screenshot of the CoG project workspace for the 2012 Dynamical Core Model Intercomparison (DCMIP) Workshop.

OpenClimateGIS

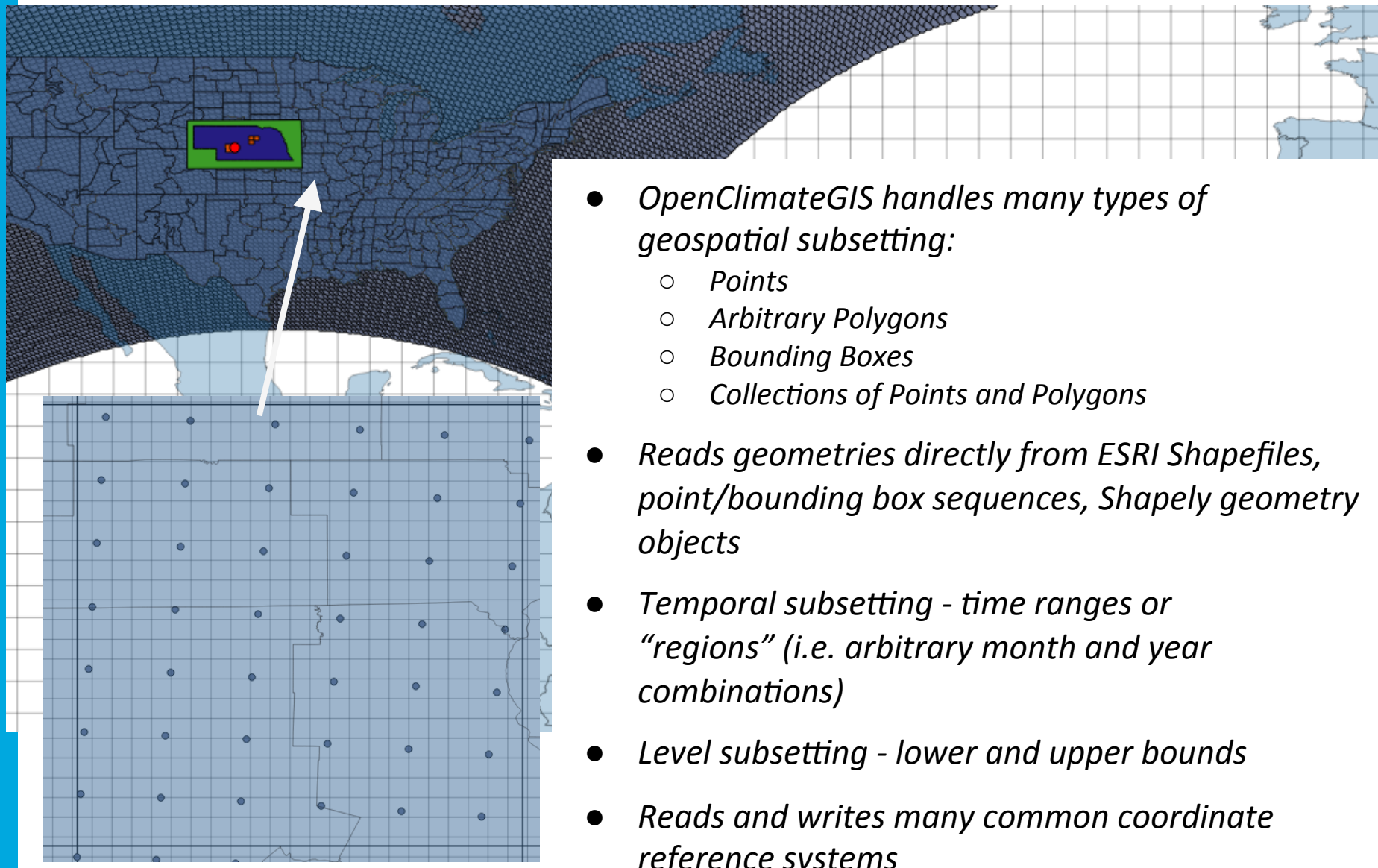
- OpenClimateGIS (OCGIS) is a Python-based, open source software library enabling dynamic access to and manipulation of climate data
- Its goal is to overcome barriers of usability of climate projections in adaptation planning and resource management
  - Translate out of climate data formats to GIS-friendly formats
  - Select geographical regions of interest
  - Select times/levels of interest
  - Compute application-relevant indices
  - Convert to end-user and analysis-ready formats
  - Provide comprehensive metadata
- Developed and used by the NCPP project, the IS-ENES climate4impact project, ClimatePipes, and others

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

Computation and Format Conversion

- OpenClimateGIS is designed to accommodate a variety of climate indices and metrics:
  - **Temporally grouped functions** → monthly means, annual maximums, durations
  - **String-based functions** → 'diff=tasmax-tasmin'
  - **Simple transforms** → natural logarithm
  - **Multivariate functions** → heat indices
- A general framework for data conversion allows data to be streamed to multiple formats: currently, CSV, Key-CSV Shapefile (CSV+), GeoJSON, netCDF, ESRI Shapefile, array-based

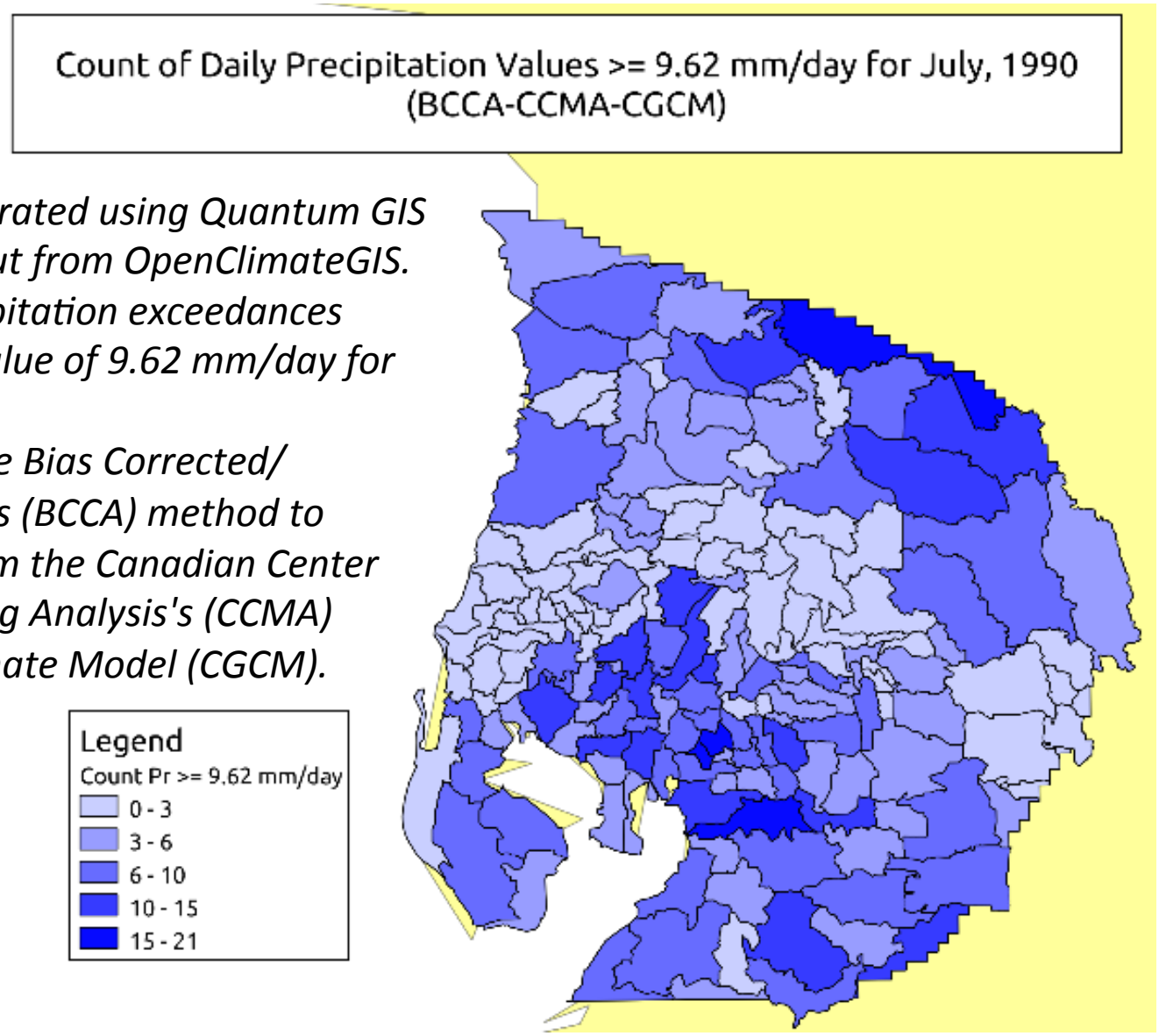
Subsetting



- *OpenClimateGIS handles many types of geospatial subsetting:*
  - *Points*
  - *Arbitrary Polygons*
  - *Bounding Boxes*
  - *Collections of Points and Polygons*
- *Reads geometries directly from ESRI Shapefiles, point/bounding box sequences, Shapely geometry objects*
- *Temporal subsetting - time ranges or "regions" (i.e. arbitrary month and year combinations)*
- *Level subsetting - lower and upper bounds*
- *Reads and writes many common coordinate reference systems*

Precipitation Threshold for Tampa Bay Watershed Basins

- *Figure at right generated using Quantum GIS with shapefile output from OpenClimateGIS.*
- *It shows daily precipitation exceedances using a threshold value of 9.62 mm/day for July 1990.*
- *Source data uses the Bias Corrected/ Constructed Analogs (BCCA) method to downscale data from the Canadian Center for Climate Modeling Analysis's (CCMA) Coupled Global Climate Model (CGCM).*



NESII Collaborators and Customers

- NOAA ESRL GSD/PSD, GFDL, NCDC, PMEL, NCEP Environmental Modeling Center
- NASA JPL, Goddard Space Flight Center, GISS
- DOE PCMDI, Argonne National Laboratory, ORNL, Sandia
- DoD NRL Stennis and Monterey, Naval Oceanography, Army ERDC, Air Force Weather Agency
- NCAR Community Earth System Model, WRF, MPAS, HAO, Unidata
- University of Michigan, Purdue University, University of South Carolina, University of Colorado, Colorado State University, Georgia Institute of Technology
- GO-ESSP, CUAHSI, CSDMS, OpenMI, OGC, CCA, METAFOR
- Delft Hydraulics, British Atmospheric Data Center, CERFACS, IPSL, Univ. Reading, UK Met Office, DKRZ, MPI
- Many more …

Questions?

- For more information, links and references, see our group pages:
  http://esrl.noaa.gov/nesii/
  https://www.earthsystemcog.org/projects/nesii/