

Lab: Getting Started with CyberGIS Gateway

Getting Started with CyberGIS Gateway

CyberGIS is a new-generation geographic information system (GIS) based on advanced computing and information infrastructure. It is designed to provide computationally intensive geospatial analysis and modeling as a service and to allow for collaborative geospatial problem solving. CyberGIS Gateway (<http://gateway.cigi.illinois.edu/home>) is the leading online CyberGIS that provides a suite of scalable geospatial analysis and modeling applications such as TauDEM, CGPySAL, and BioScope. The big computation required by these applications is conducted within a hybrid computing environment of clouds, National Science Foundation's Extreme Science and Engineering Discovery Environment (XSEDE) supercomputers, and the Open Science Grid resources.

1. Introduction to TauDEM

Terrain Analysis Using Digital Elevation Models (TauDEM) is a set of tools used to evaluate and extract hydrologic information from Digital Elevation Models. TauDEM is developed by the Utah State University Hydrology Research Group in collaboration with the US Army Corps of Engineers and the System Wide Water Resources Program. Through the use of USGS National Elevation Dataset data, which has 10 meter resolution, many hydrologic analyses and processes can be run. Capabilities of TauDEM include watershed delineation, flooding techniques to remove pits when conducting watershed analyses, flow direction methodologies, channel network delineation based on spatial variation of drainage, and calculation of slopes and flow directions through different methodologies. Additionally, advanced analyses can be conducted using these DEMs. These analyses include an inverse of the wetness index which essentially is the slope over a catchment area, multiple measurements for flow distances down to streams and up to ridges, slope dependence based on relationship with upslope and downslope pixels, several types of accumulation calculations, range of avalanche debris or "runout", and average slope. Depending on the size of the study area, the TauDEM app on the high-performance CyberGIS platform can save massive amounts of time collecting data, processing data, and ultimately running the analyses on the data.

2. Using the TauDEM App

2.1 Create a new analysis

In the interface of the TauDEM App, you can create a new TauDEM analysis by clicking the "New" or "New from..." button (See Figure 1 below). "New from..." will ask you to select an existing analysis and load the configuration of that analysis to construct another analysis. To save a new analysis, you'll need to click "Submit". To delete an analysis, simply click on the "Delete" button. Please note that you can't start another new analysis before you close the "New Analysis Wizard" tab for your previous analysis. The "Load" button is used for reading the result of an analysis that is completed and ready for visualization. The "Refresh" button enables you to get the latest status of your analysis to see if it's submitted, queued or completed.

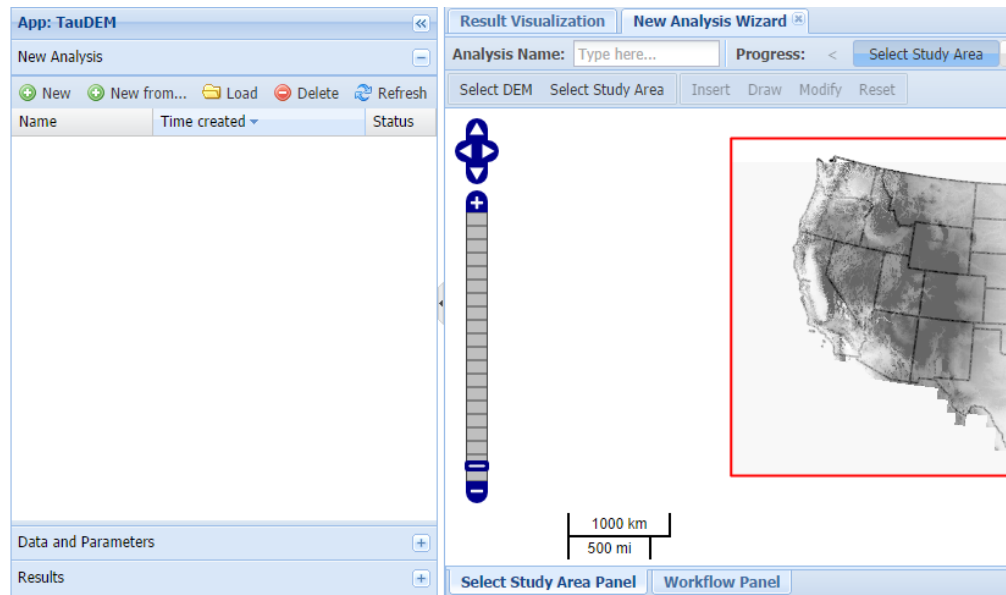


Figure 1. Create a New Analysis in TauDEM App

2.2 Choose a DEM data source

When creating a new analysis, you will have the option to choose elevation data from one of the following three sources.

a) USGS National Elevation Dataset (NED)

Name : USGS National Elevation Dataset (NED) - 1/3 arc (10-meter) resolution

Information : NED is one of the eight layers of map data provided by the National Map project at USGS. 1/3 arc (10-meter) NED dataset is organized as 1x1 degree tiles, covering U.S. terrain. Total dataset size is about 2TB

Data provider : USGS National Map

Data URL : <http://nationalmap.gov>

About Data Provider : The National Map project from the U.S. Geological Survey provides eight layers of map covering the U.S., including elevation, land cover, orthoimagery, structures, boundaries, hydrography, geographic names, and transportation

b) OpenTopography LiDAR Data

Name : OpenTopography LiDAR data (prototype)

Information : This data source dynamically generates LiDAR-derived digital elevation models (DEMs) from the high-resolution LiDAR data repository at the National Science Foundation (NSF) OpenTopography facility. Data resources are integrated into CyberGIS through the service integration of the GISolve Open Service API and OpenTopography Opal2 data services

Data provider : OpenTopography

Data URL : <http://opentopography.org>

About Data Provider : The National Science Foundation (NSF) OpenTopography facility facilitates community access to high-resolution, Earth science-oriented, topography data, and related tools and resources. OpenTopography provides high-resolution LiDAR and LiDAR-derived geospatial data for CyberGIS community

c) Providing your own dataset

Name : I will provide the dataset

Information : Dataset provided by user through uploading or specifying online dataset URL. Currently, GeoTIFF format is supported.

Note: Currently, only USGS NED is supported.

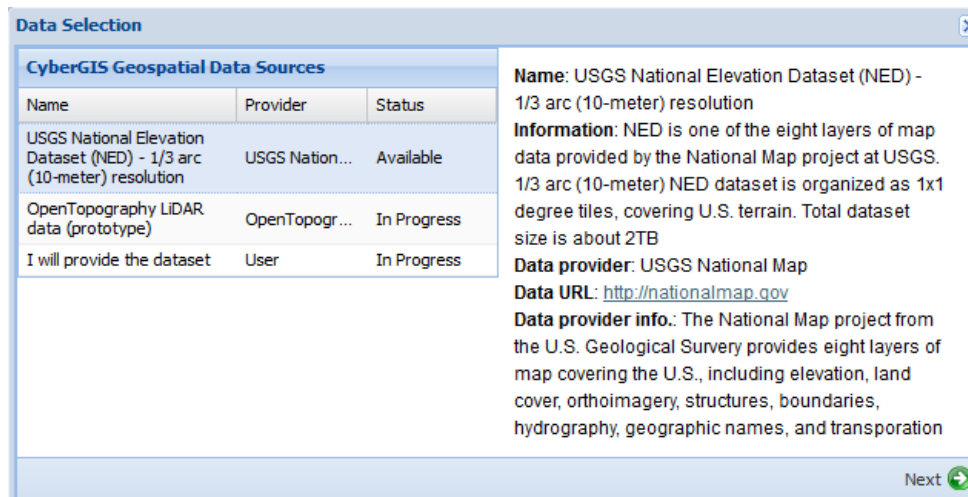


Figure 2. Select DEM data source

2.3 Define a study area

Once a DEM data source is specified, you must define a geographic study area within which to perform the coming analysis. To define a study area, use the zoom and pan functions within the map (on the left hand side of the map viewer) to focus on an area you would like to study. Then, click and drag to create a rectangle defining the bounding box of the study area. The bounding box must be within the geographic area of the underlying dataset and cannot include more than 10 million cells or 10,000 square kilometers. If either issue applies to the bounding box you have selected, a drop-down alert message will appear to request you redraw an acceptable study area. If the study area you have drawn is acceptable, the TauDEM Workflow Wizard will successfully launch to allow the selection of output products.

2.4 Choose the output products

a) Pit removal

First of all, all pits in the DEM will need to be identified; the elevation of each pit will be raised to the average level of the lowest pour cells around it. This step is fundamental to any of the more advanced analyses or processing afterwards.



Figure 3. Hydrologically conditioned elevation grid after pit remove

b) D8 Flow Direction

The D8 flow direction for each grid cell refers to one of either possible directions to its adjacent or diagonal neighbors, depending on which is the steepest descent. The codes used to present the different directions are: 1- East, 2-North East, 3-North, 4-North West, 5 -West, 6-South West, 7-South, 8- South East. Flow direction is reported as “no data” for any grid cell adjacent to the edge of the DEM domain, or adjacent to a “no data” value in the DEM. In flat areas, flow directions are assigned away from higher ground and towards lower ground using the method of Garbrecht and Martz (1997).

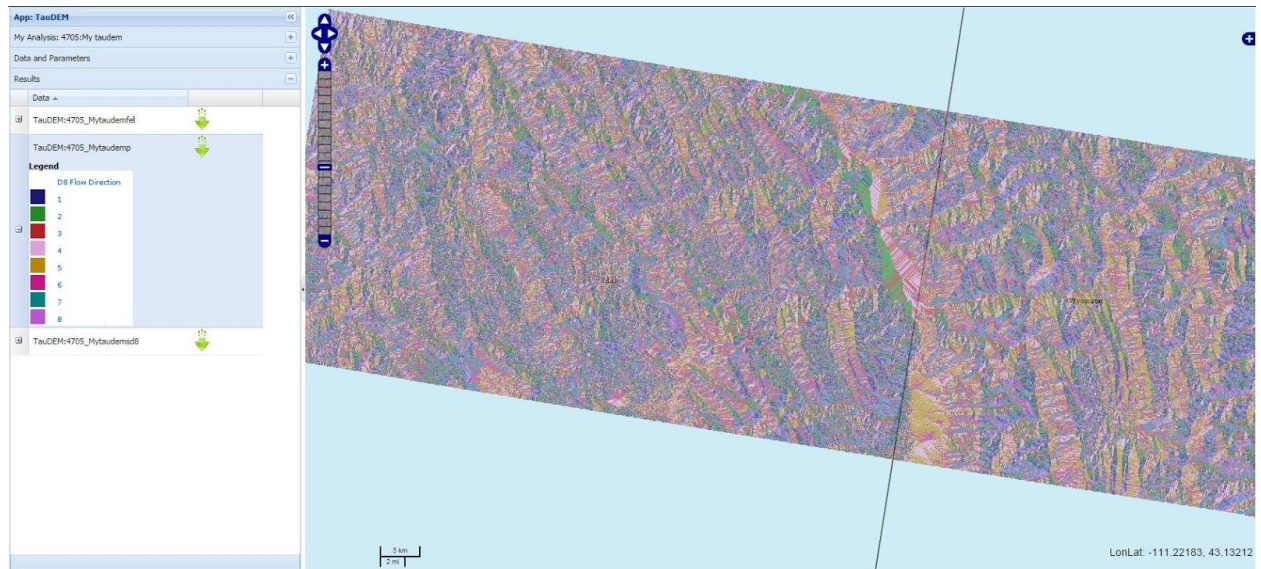


Figure 4. Flow Direction. Flow direction coding: 1-East, 2-North East, 3-North, 4-North West, 5 -West, 6-South West, 7-South, 8- South East.

c) Peuker Douglas Stream Raster

Peuker Douglas Stream Raster is an indicator grid that maps the streams with a value of 1 for stream cells and 0 for others. Certain threshold is applied to a weighted flow accumulation of grid cells to differentiate stream and non-stream. Grid cells with a value greater than or equal to the input threshold are labeled as streams, while the others are treated as non-stream (hillslope) cells. To produce the stream raster, the DEM is first smoothed and then the Peuker and Douglas (1975) method is applied.

2.5 Receiving and visualizing the results

Once the job is submitted, you will need to wait for it to be processed and to receive the results. You can check the job status by clicking the “refresh” button. The icon next to the job indicates the current stage of the job submitted (Figure 5).

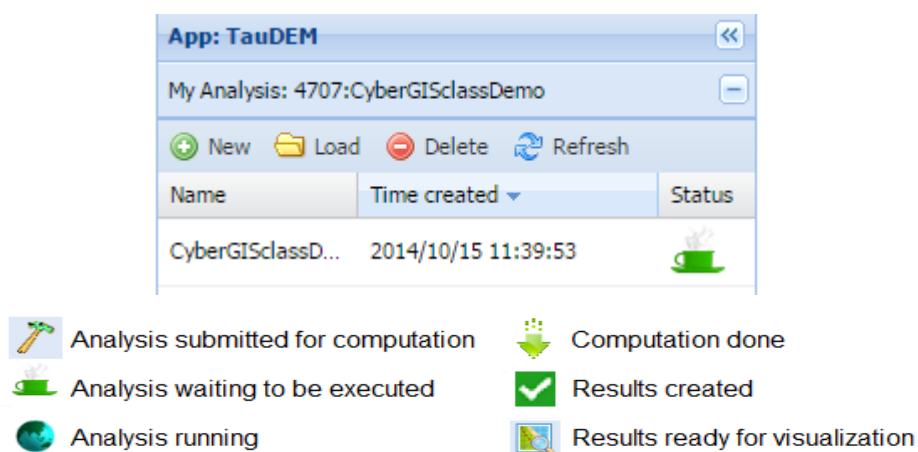


Figure 5. Status icons showing the progress.

When the results are ready for visualization, you can view the outputs of your job in the “Results” tab by double-clicking your analysis under the “My Analysis” tab or by highlighting the analysis and clicking the “Load” button. In the “Results” tab, each output from your analysis is visible as a separate layer with a green arrow next to it (Figure 7). The different layers can be toggled on and off and moved up or down by right clicking each layer and selecting the desired action. The legend for each layer can be seen by clicking on the plus sign next to the layer. Finally, the results can be downloaded as .tif files in a .tar.gz archive by clicking on the green arrow next to the layer name.

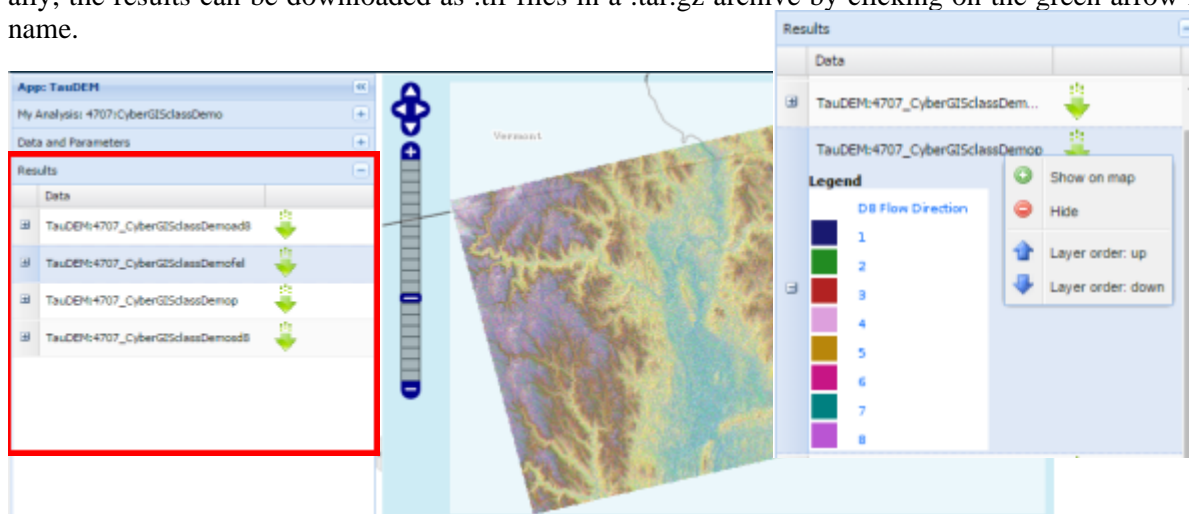


Figure 6. The results tab listing the outputs produced by the analysis and displays them to the right.

References

Garbrecht, J., & Martz, L. W. (1997). The assignment of drainage direction over flat surfaces in raster digital elevation models. *Journal of hydrology*, 193(1-4), 204-213.

Peuker, T. K. and D. H. Douglas, (1975), "Detection of surface-specific points by local parallel processing of discrete terrain elevation data," *Computer Graphics Image Processing*, 4: 375-387.

[Download Word Document Copy of Lab Assignment](#)

Creator Name	Jie Tian
Content Title	Getting Started with CyberGIS Gateway
Content Type	Lab Assignment
Part Of	Module 2: CyberGIS Platform and Software
Learning Objectives	(for all of Module 2) By the end of the semester, the students will be able to: <ul style="list-style-type: none">○ Appreciate the value of CyberGIS in performing computationally intensive analyses on big geospatial data.○ Get familiarized with the CyberGIS software ecosystem and competence using the gateway and the toolkit.○ Re-think the classic geospatial analyses in the context of cyberinfrastructure and high performance computing.
Background Knowledge	(for all of Module 2) <ul style="list-style-type: none">• Concepts and terminology introduced in Module I• Basic computer literacy
Resources Needed	(for all of Module 2) <ul style="list-style-type: none">• Linux Operating System• Cloudera Manager
Work Mode	(for all of Module 2) Watching the lectures and doing lab assignments
Relation to Project	(for all of Module 2) This module (estimated 20-25 hours) is designed to attract students who have taken Module I and would want to learn more practical knowledge and skills of CyberGIS.
Feedback Needed	(for all of Module 2) General feedback