

WRF-Hydro GIS Preprocessing Tools

Kevin Sampson, GISP

Associate Scientist

National Center for Atmospheric Research

Boulder, Colorado

Motivation

To provide a complete set of WRF-Hydro routing grids that have been hydrologically processed according to user specifications.

- Fast, efficient method for producing the 'routing stack'.
- Consistent processing methodology between domains, regions, datasets.
- Remove GIS burden from modelers.

WRF-Hydro & ArcGIS

- Desktop GIS Application Suite
- Site-licenses available at most US academic institutions
- Ecosystem of compatible hydrology tools
 - Spatial Analyst
 - ArcHydro
 - TauDEM
- Extensible using Python API (arcpy)
- Handles everything from projections, to analysis, to mapmaking in one library.



Requirements

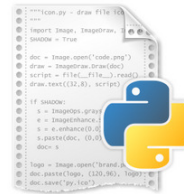
- ArcGIS for Desktop
 - Version 10.2.2 (minimum 10.1 SP1)
 - Basic, Standard, or Advanced license levels
 - Spatial Analyst extension required
 - Python 2.7.5, NumPy 1.7.1
 - Installed with ArcGIS Desktop complete installation
- Optional:
 - TauDEM for ArcGIS 10 (x32 or x64) tools v5.1.2
 - <http://hydrology.usu.edu/taudem/taudem5/index.html>

Python Toolboxes

- Python script wrapped to act as an ArcGIS Toolbox

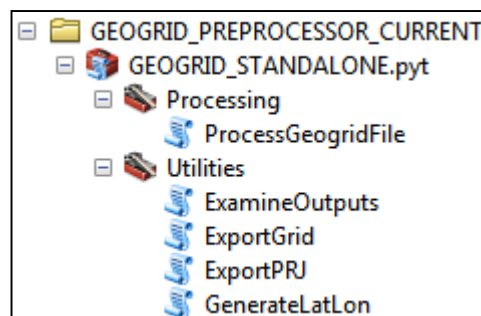


Toolbox
Script (.pyt)



Function
Script (.py)

- PYT file is the toolbox script containing multiple toolboxes
 - Functions called from separate script (wrf_hydro_functions.py)
- Parameter handling and validation



Advantages

Easy to modify

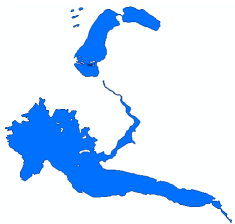
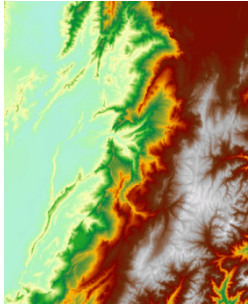
Portable

Many tools organized

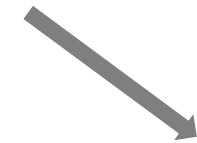
WRF-Hydro preprocessor toolbox as viewed from ArcCatalog.

Pre-Processor

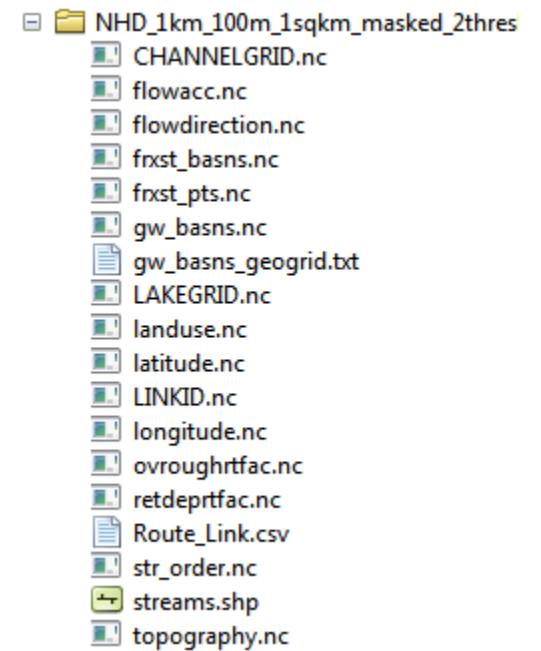
Inputs



Other parameters



Outputs



ProcessGeogridFile

ProcessGeogridFile

Input GEOGRID File

Forecast Points (CSV) (optional)

☐ Mask CHANNELGRID to basins? (optional)

☐ Create reach-based routing files? (optional)

☐ Create lake parameter file? (optional)

Reservoirs Shapefile or Feature Class (optional)

Input Raster

Regridding Factor: 10

Number of pixels to define stream: 200

Output ZIP File: WRF_Hydro_routing_grids.zip

Parameter Values

OVROUGHRTFAC Value: 1.0

RETDEPRTFAC Value: 1.0

ProcessGeogridFile

This tool takes an input WRF Geogrid file in NetCDF format and uses the HGT_M grid and an input high-resolution elevation grid to produce a high-resolution hydrologically processed output.

OK Cancel Environments... << Hide Help Tool Help

Inputs

- Required:
 - WRF GEOGRID file (.nc)
 - High-Resolution Elevation
 - Elevation file (Esri GRID, GeoTIFF, etc.)
 - Mosaic Datasets
- Parameters:
 - Regridding Factor – nesting relationship of routing:land grids
 - Minimum basin size (in routing grid cells)
 - OVROUGHRTFAC - constant
 - RETDEPRTFAC – constsant
- Optional:
 - Station Locations (.csv)
 - Lake Polygons (polygon feature class or .shp)

Input: WRF GEOGRID

The purpose of the GEOGRID file is to define the simulation domain and interpolate various static geographical datasets to the model grid.

- GEOGRID is used in GIS Preprocessor to define the domain's coordinate reference system, extent, resolution, and certain variables:
 - HGT_M (elevation)
 - LU_INDEX (landuse)
- Currently supported GEOGRID coordinate systems
 - MAP_PROJ = 1 (Lambert Conformal Conic)
 - MAP_PROJ = 3 (Mercator)
 - MAP_PROJ = 6 (Cylindrical Equidistant but NOT w/ rotated pole)
 - MAP_PROJ = 2 (Polar Stereographic)

GEOGRID: Projected Coordinate System

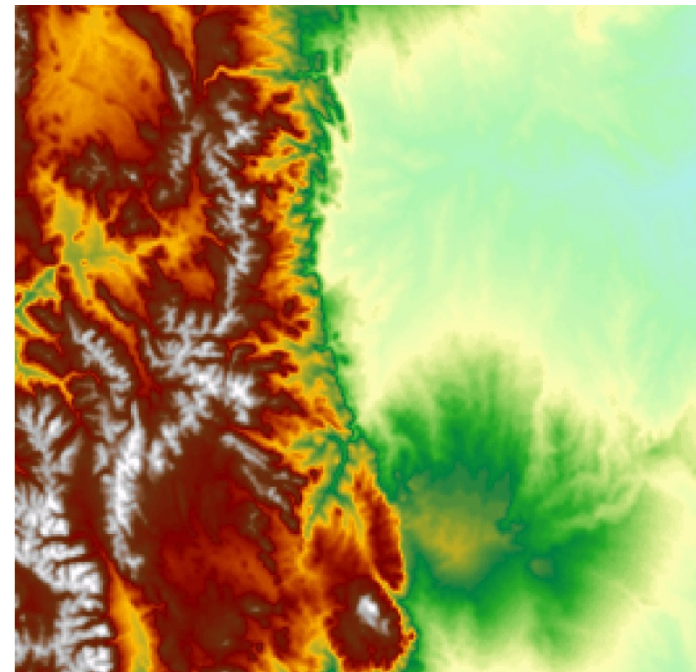
Front_Range_geo_em.d02.nc

```
:TITLE = "OUTPUT FROM GEOGRID V3.5.1";
:SIMULATION_START_DATE = "0000-00-00_00:00:00";
:WEST-EAST_GRID_DIMENSION = 50; // int
:SOUTH-NORTH_GRID_DIMENSION = 36; // int
:BOTTOM-TOP_GRID_DIMENSION = 0; // int
:WEST-EAST_PATCH_START_UNSTAG = 1; // int
:WEST-EAST_PATCH_END_UNSTAG = 49; // int
:WEST-EAST_PATCH_START_STAG = 1; // int
:WEST-EAST_PATCH_END_STAG = 50; // int
:SOUTH-NORTH_PATCH_START_UNSTAG = 1; // int
:SOUTH-NORTH_PATCH_END_UNSTAG = 35; // int
:SOUTH-NORTH_PATCH_START_STAG = 1; // int
:SOUTH-NORTH_PATCH_END_STAG = 36; // int
:GRIDTYPE = "C";
:DX = 1000.0f; // float
:DY = 1000.0f; // float
:DYN OPT = 2; // int
:CEN_LAT = 39.940014f; // float
:CEN_LON = -105.42999f; // float
:TRUELAT1 = 30.0f; // float
:TRUELAT2 = 50.0f; // float
:MOAD_CEN_LAT = 39.940014f; // float
:STAND_LON = -105.0f; // float
:POLE_LAT = 90.0f; // float
:POLE_LON = 0.0f; // float
:corner_lats = 39.783337f, 40.093864f, 40.095993f, 3
:corner_lons = -105.714264f, -105.71753f, -105.14442
:MAP PROJ = 1; // int
:MMINLU = "USGS";
:NUM_LAND_CAT = 24; // int
:ISWATER = 16; // int
:ISLAKE = -1; // int
:ISICE = 24; // int
:ISURBAN = 1; // int
:ISOILWATER = 14; // int
:grid_id = 1; // int
:parent_id = 1; // int
:i_parent_start = 1; // int
:j_parent_start = 1; // int
:i_parent_end = 50; // int
:j_parent_end = 36; // int
:parent_grid_ratio = 1; // int
:sr_x = 1; // int
:sr_y = 1; // int
:FLAG_MF_XY = 1; // int
}
```

esri PE string

```
"PROJCS['Lambert_Conformal_Conic',GEOGCS['GCS_Sphere',DATUM['D_Sph
ere',SPHEROID['Sphere',6370000.0,0.0]],PRIMEM['Greenwich',0.0],UNI
T['Degree',0.0174532925199433]],PROJECTION['Lambert_Conformal_Coni
c'],PARAMETER['false_easting',0.0],PARAMETER['false_northing',0.0]
,PARAMETER['central_meridian',-
105.0],PARAMETER['standard_parallel_1',30.0],PARAMETER['standard_p
arallel_2',50.0],PARAMETER['latitude_of_origin',39.9400138855],UNI
T['Meter',1.0]];-36695400 -29251300 10000;-100000 10000;-100000
10000;0.001;0.001;0.001;IsHighPrecision"
```

geo_em.d01.boulder_creek_1km.prj

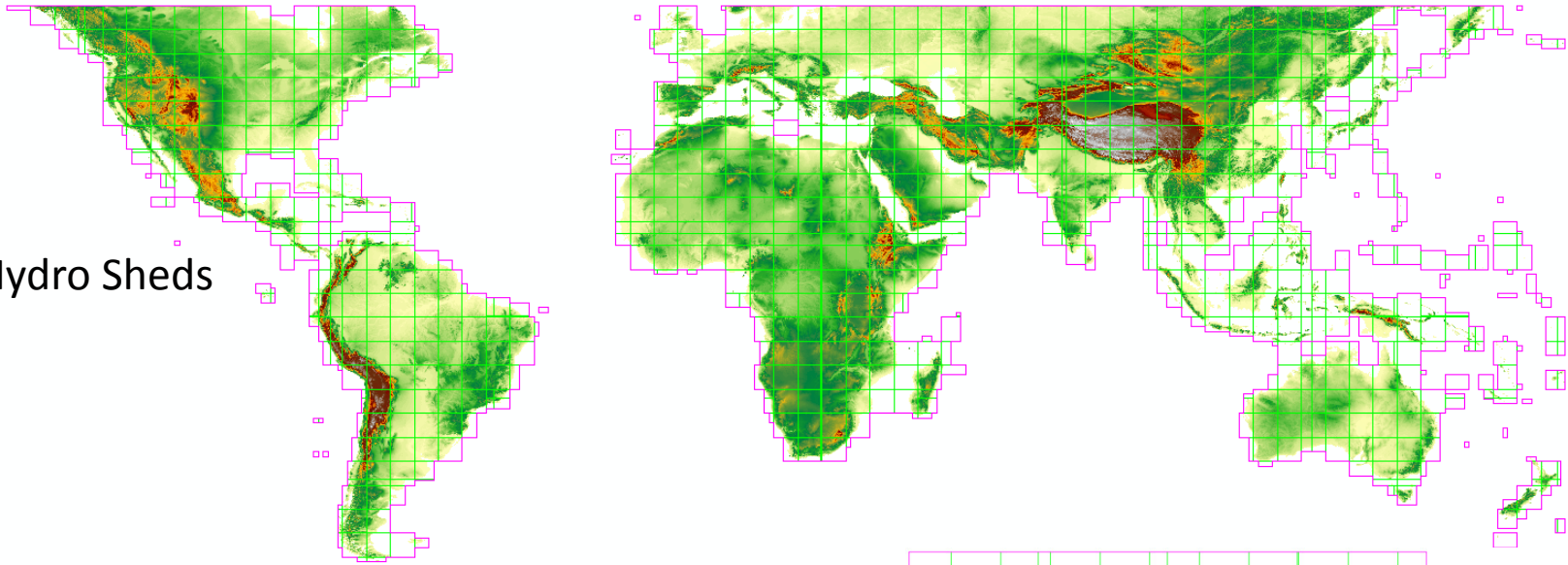


Input: Elevation Raster

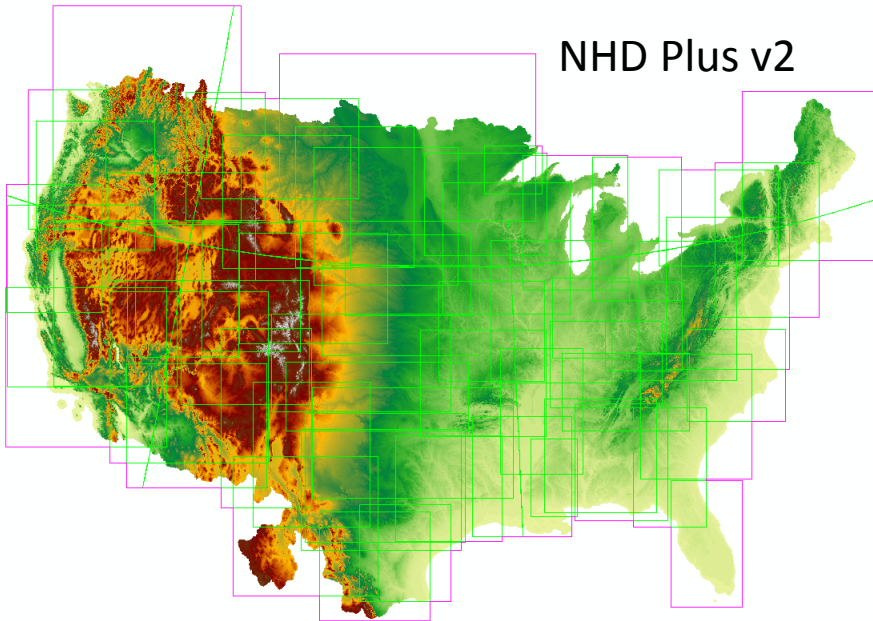
- Must be an ArcGIS readable raster format
- Must contain valid coordinate reference system
- Must cover entire extent of your GEOGRID domain
- Elevation units must be converted to meters (m)
- Should be hydrologically corrected
 - Not necessary but helps with channel placement, hydro enforcement, etc.
- Supported datasets include, but are not limited to:
 - HydroSHEDS
 - NHDPlus (converted from cm to m)
 - EU-DEM

Input: Elevation Mosaics

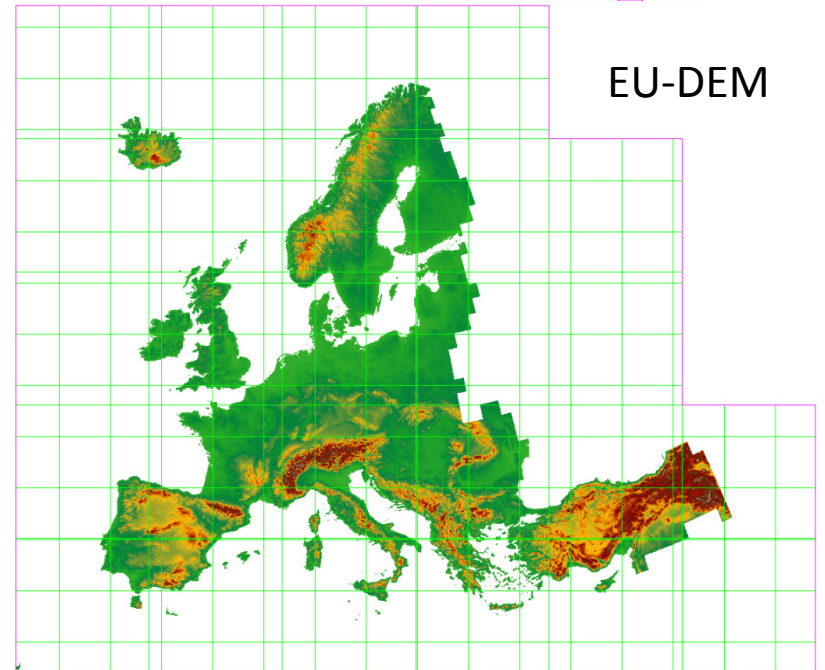
Hydro Sheds



NHD Plus v2



EU-DEM



Input: Station Location File

FID_	LON	LAT	STATION	Name	HOST	ELEV	DRAIN_AREA_SQMI	DRAIN_AREA_SQKM
1	-103.79889	40.26861	S_PLATTE_at_FT_MORGAN	6759500	USGS	4260	14627	37883.93
2	-108.26556	39.23917	COLO_at_CAMEO	9095500	USGS	4813	8050	20849.5
3	-104.39861	38.24806	ARKANSAS_nr_AVONDALE	7109500	USGS	4509	6327	16386.93
4	-105.88002	37.481392	RIO_Grande_nr_ALAMOSA	8223000	USGS	-9999	0	0

FID,LON,LAT,STATION,Name

15,-105.92833,40.08139,Fraser_at_Granby,9033300

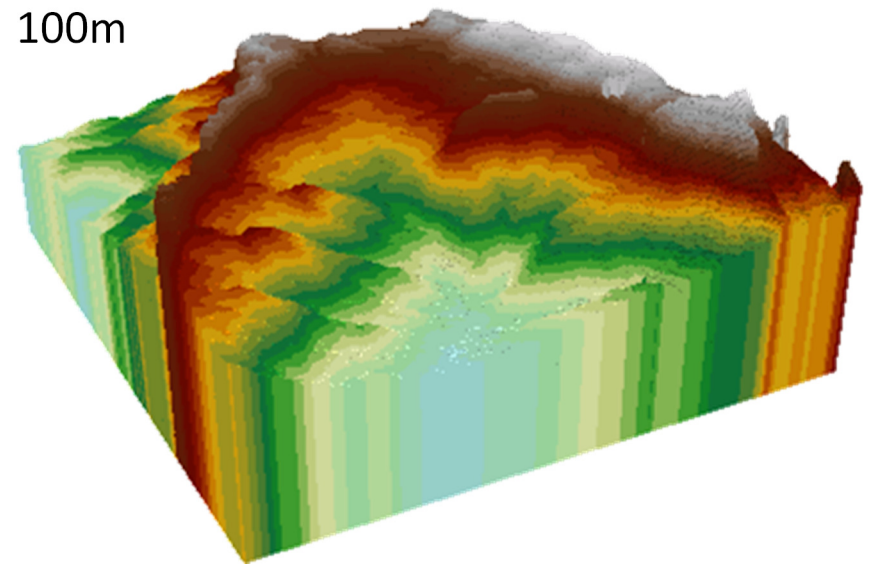
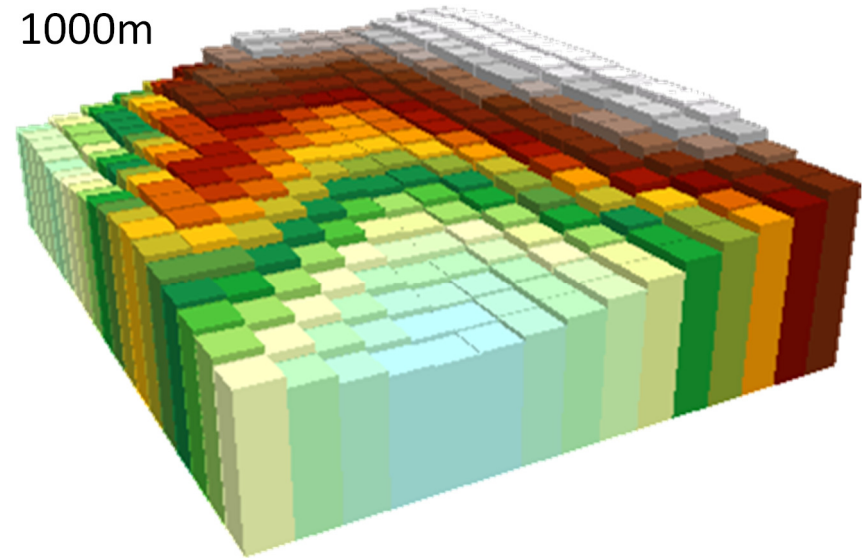
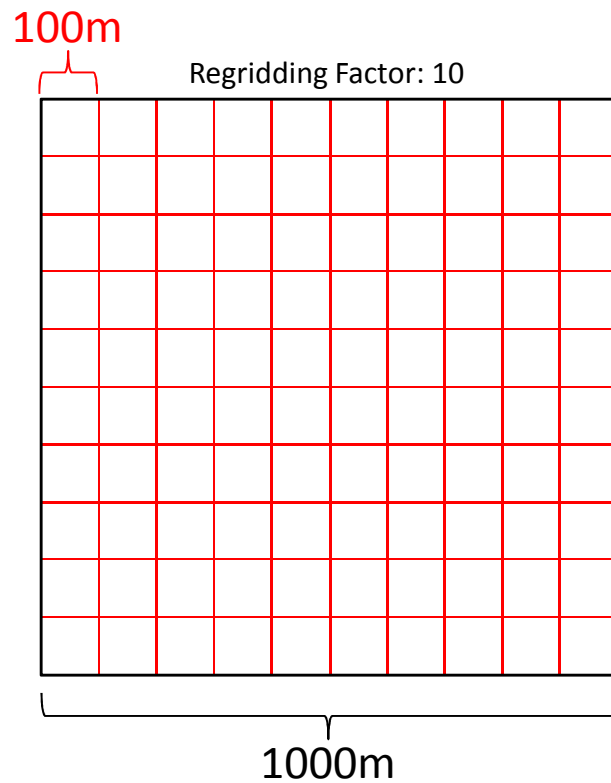
18,-105.9,40.12083,COLO_nr_GRANBY,9019500

20,-106.33333,39.8803,Blue_R_blw_Grn_Mtn,9057500

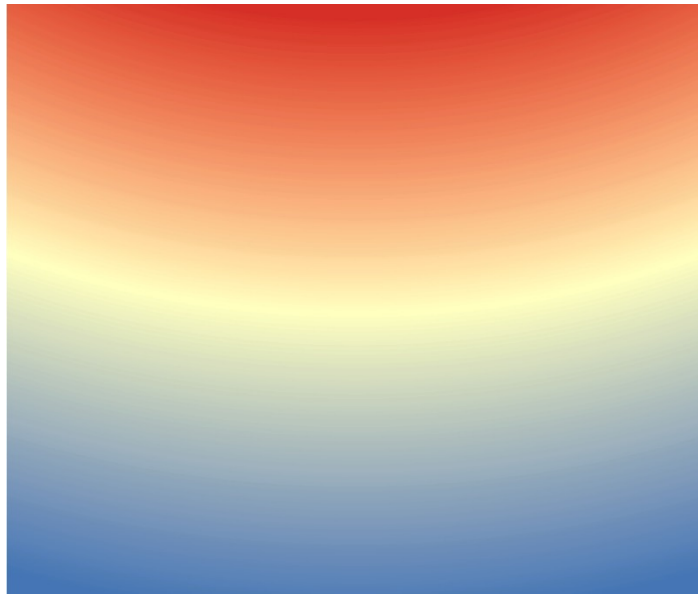
- Create in Xcel, Numbers, Word, etc.
- Direct output of attribute table from shapefile or feature class
- “LON” & “LAT” required
- If present, basins will be delineated using the points provided
 - frxst_basns.nc output file will be created
 - frxst_pts.nc & gw_basns.nc will be populated
- If masked to basins, CHANNELGRID will have values -1, 0, -9999

Input: Regridding Factor

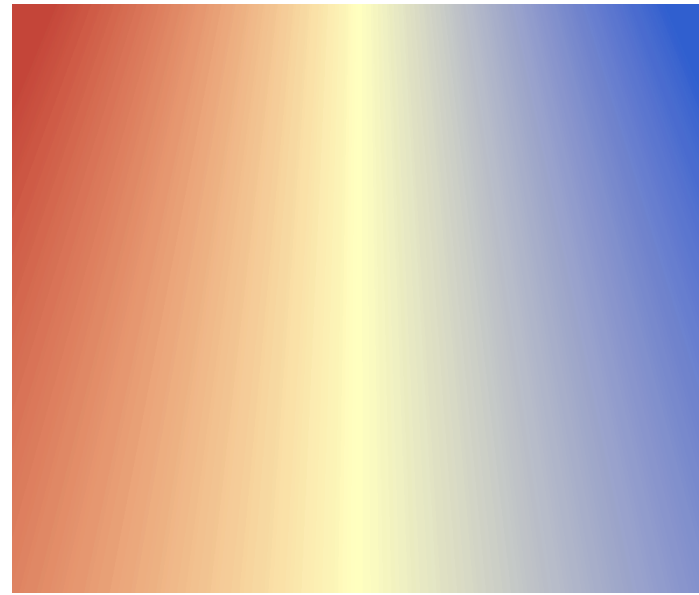
$$\frac{\text{GEOGRID Resolution}}{\text{Regridding Factor}} = \text{Routing Resolution}$$



Process: Latitude & Longitude Grids



Latitude

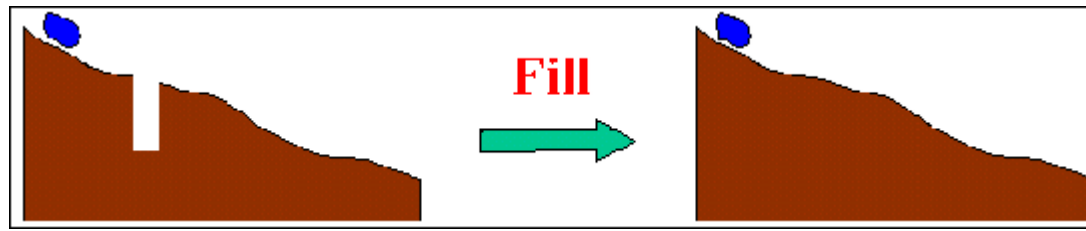


Longitude

- Each pixel on the high-resolution grid will have lat/lon calculated
- Project domain grid to WGS84
- Uses `$$XMAP` & `$$YMAP` functions from deprecated `arcgisscripting` python library.
- Project back to WRF projected coordinate system.

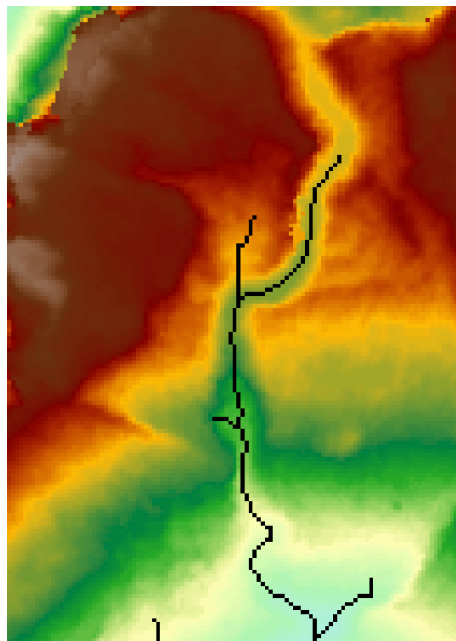
Process: Pit Filling

Spatial Analyst “Fill” Tool

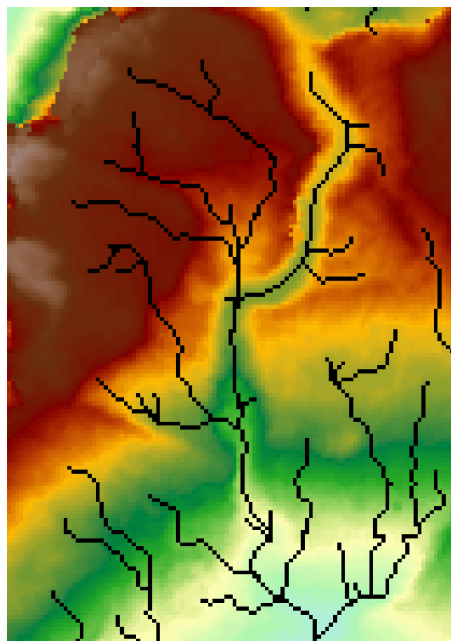


Process: Stream definition

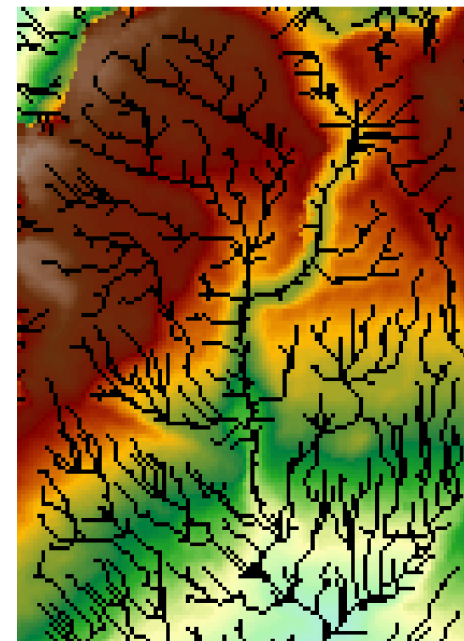
- Input Parameter: Number of pixels to define stream
 - Yields a minimum 'basin' size
 - Given in pixels (unitless), on the high-resolution grid
 - Affects density of generated channel network



1km²

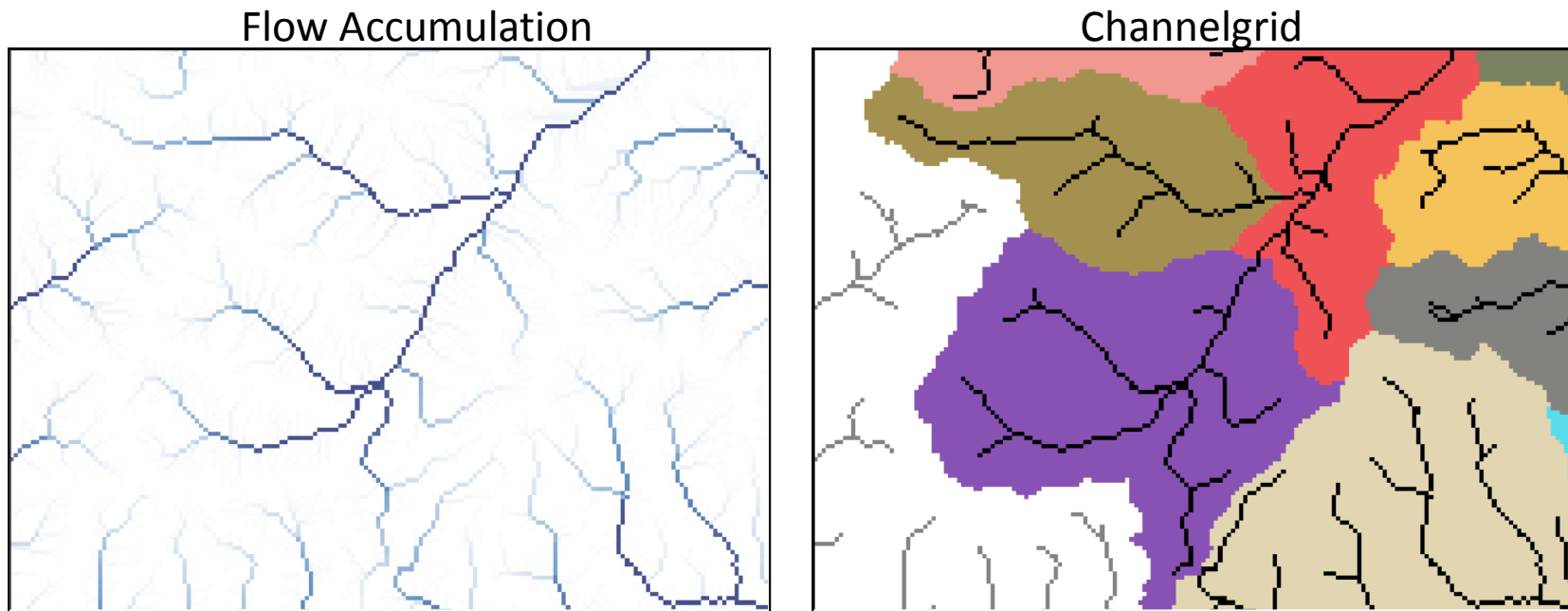


0.1km²



0.01km²

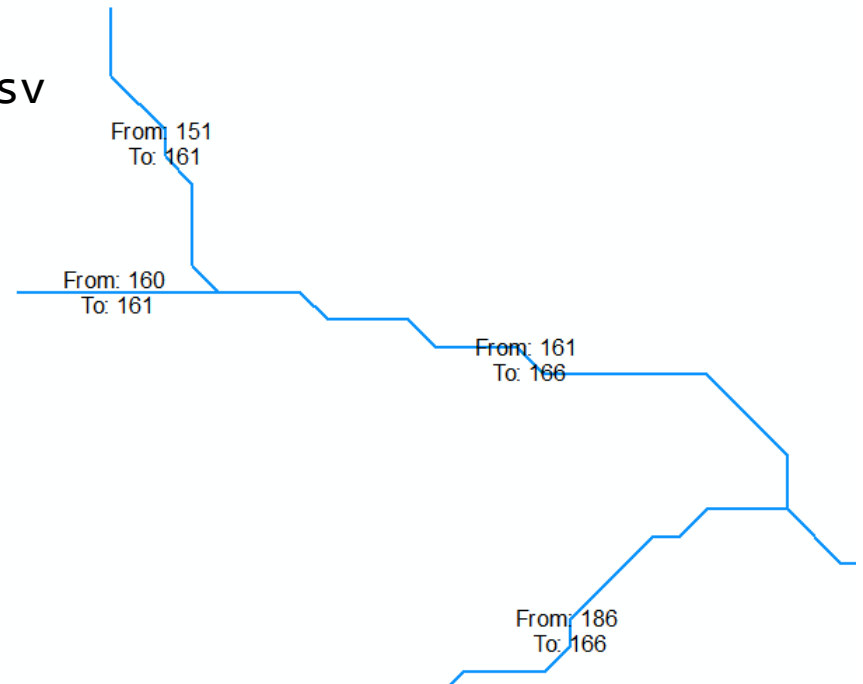
Process: Stream definition



- Use flow accumulation threshold to define channels
- Option: use gaged basins as mask to assign CHANNELGRID values
- If reach-based routing is selected, **Stream to Feature** used to create vector geometry of streams
 - streams.shp shapefile written to output directory

Process: Reach-Based Routing

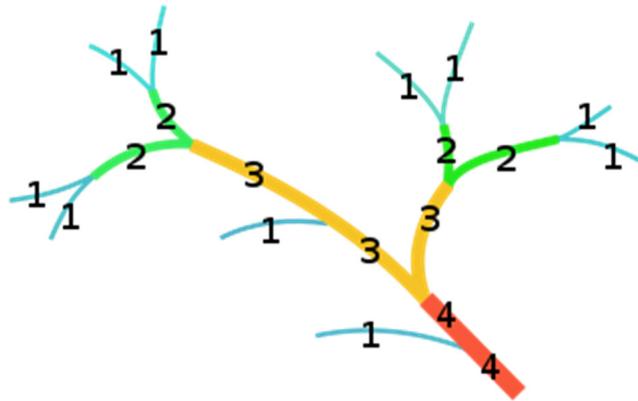
- CHANNELGRID raster is converted to a shapefile (streams.shp)
- Decomposes line geometry to nodes, and gathers elevation, Latitude and Longitude at each node
- Constructs a CSV table with necessary parameters for reach-based routing:
 - Length, Slope, Order, Drop, X/Y, etc.
 - Order-based parameters
 - Writes output file to Route_Link.csv



Process: Stream Order

Stream Order Spatial Analyst tool

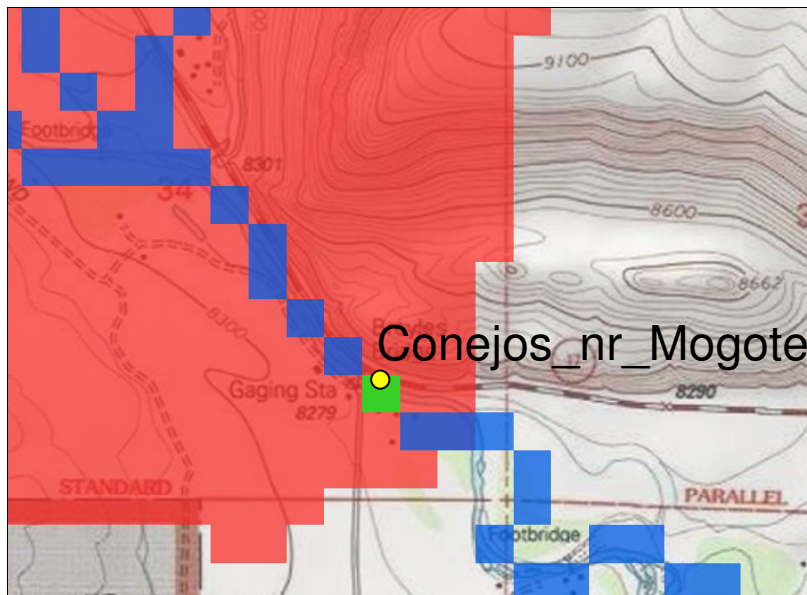
- Strahler stream order



- Writes output file to str_order.nc

Process: Basin Delineation

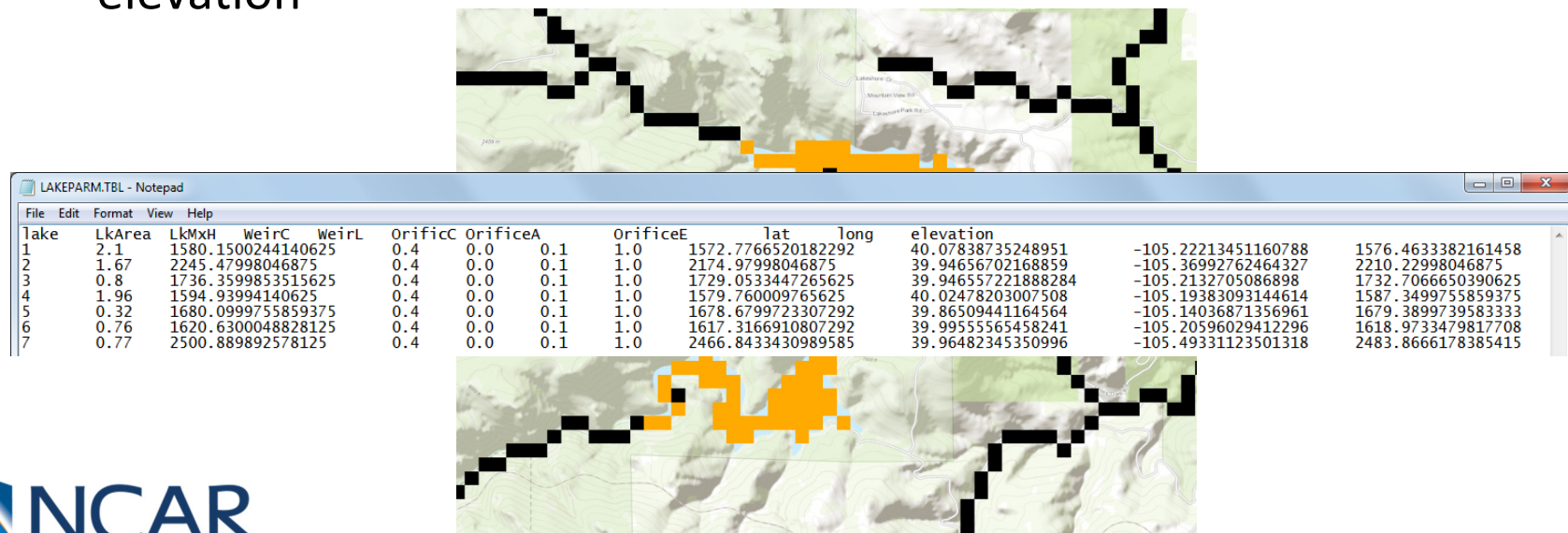
- Snap points to streams
- ‘Walk’ down channel network a specified distance
 - Default = 3 pixel widths
- Delineate basin using ‘Watershed’ Spatial Analyst tool



- Writes output file to
frxst_basns.nc, gw_basns.nc,
gw_basns_geogrid.txt

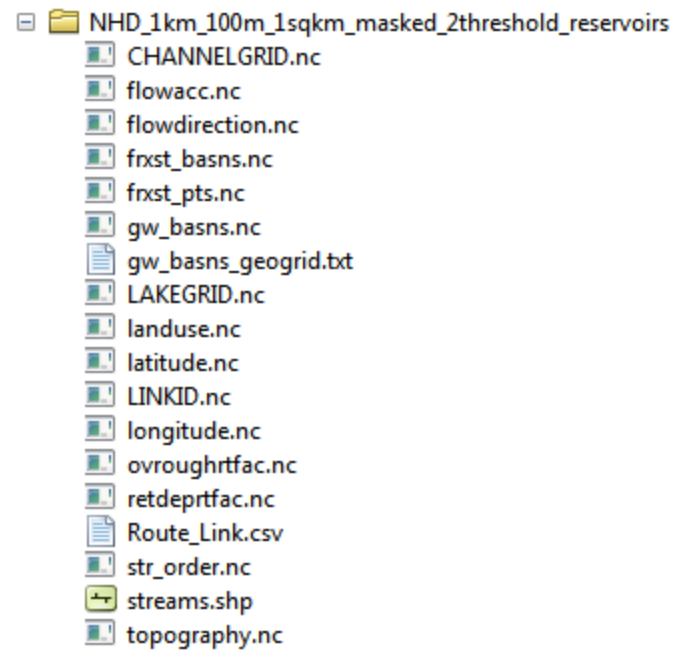
Process: Reservoir Routing

- If option is selected, a polygon shapefile or feature class is required as input.
- Populates LAKEGRID.nc output file
- Assigns lake ID values to pixels where lakes drain into channel.
- Constructs a LAKEPARAM.TBL table with necessary parameters for reach-based routing:
 - Lake area, max elevation, min elevation, base elevation, orifice elevation



Outputs

- Set of netCDF, shapefile, CSV, txt (ascii raster), and .TBL files
 - 13-15 netCDF files
 - 1 ASCII Raster (.txt)
 - 0-1 Shapefiles
 - 0-1 .TBL
 - 0-1 .CSV
 - 1 .log file



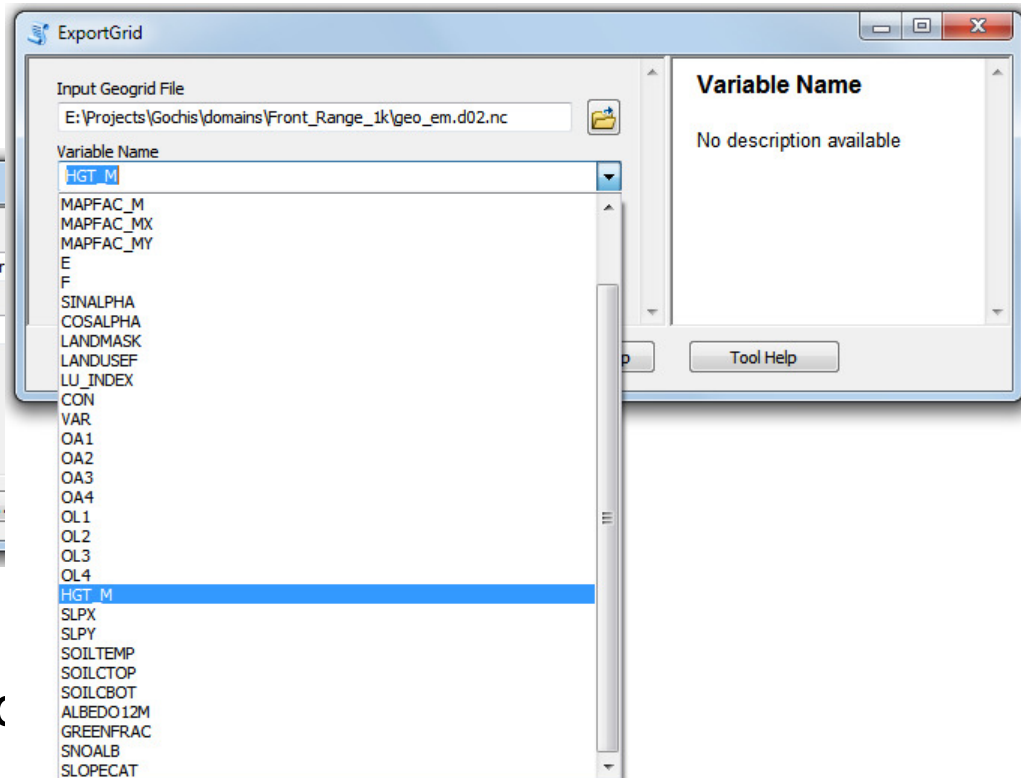
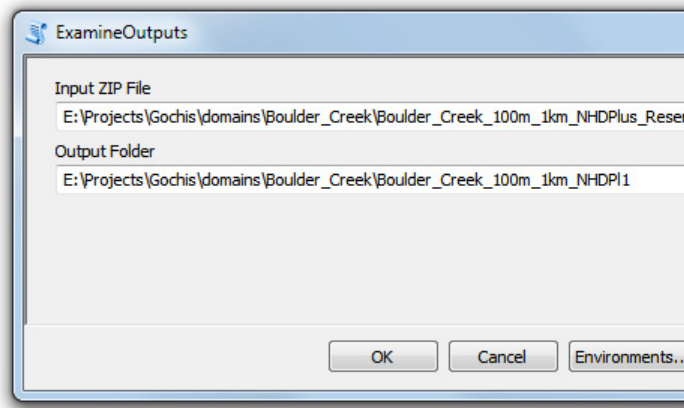
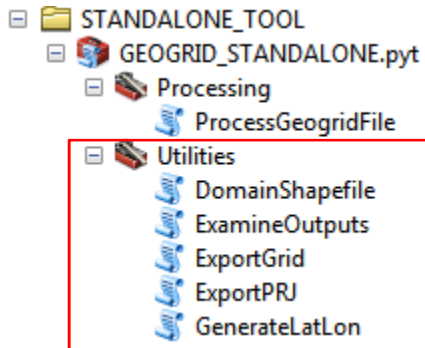
Other Utilities

- ExamineOutputs

- Extracts .zip output file to individual rasters for viewing in Desktop GIS applications.

- ExportGrid

-



- Do

- Creates a polygon shapefile defining the domain boundary.

Customization

- Multiple Toolboxes released for different users
 - UCAR Internal toolbox takes advantage of available services
 - Elevation mosaics and vector services (NHDPlus lakes, flowlines)
 - Modular tool structure allows building rapid new toolsets

```
• import arcpy
• import os, sys, math
• import shutil
• sys.dont_write_bytecode = True
• import wrf_hydro_functions

# Turn functions script and TauDEM script into Project data so it gets copied up to the server directly
• wrfhydro_script = r'E:\Projects\Gochis\TOOLS\GEOGRID_PREPROCESSOR_CURRENT\wrf_hydro_functions.py'

class Toolbox(object):
    def __init__(self):
        """Define the toolbox (the name of the toolbox is the name of the
        .pyt file)."""
        self.label = "GEOGRID_STANDALONE"
        self.alias = ""
        self.description = "This is a standalone processing tool for WRF-HYDRO."

        # List of tool classes associated with this toolbox
        self.tools = [ProcessGeogridFile, ExportGrid, ExamineOutputs, ExportPRJ, GenerateLatLon]

class ProcessGeogridFile(object):
    def __init__(self):
        """Define the tool (tool name is the name of the class)."""
        self.label = "ProcessGeogridFile"
        self.description = "This tool takes an input WRF Geogrid file in NetCDF format + \
        and uses the HGT_M grid and an input high-resolution elevation grid" + \
        "to produce a high-resolution hydrologically processed output."

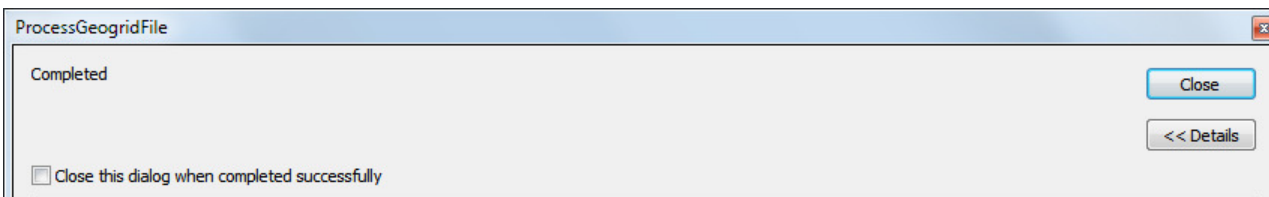
        #self.canRunInBackground = False
        self.canRunInBackground = True
        self.category = "Processing"

    def getParameterInfo(self):
        """Define parameter definitions"""

        # First parameter
        in_nc = arcpy.Parameter(
            displayName="Input Geogrid File",
            name="in_nc",
            datatype="File",
            parameterType="Required",
            direction="Input")

        # Second parameter
        in_csv = arcpy.Parameter(
```

Tool Messages



```
Process: flowdirection.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\flowdir
Process: flowacc.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\flowacc
Process: Flow Accumulation completed without error
Process: retdeprtfac.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\retdeprtfac
Process: ovroughrtfac.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\ovroughrtfac
Process: str_order.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\str_order
Routing table will be created...
Stream to features step complete.
Done reading streams layer.
Done building Nodes layer with adjustments
Done extracting elevations to points.
Done reading elevations.
Done reading Strahler stream orders.
Routing Table:
403 Lines
445 Nodes.
Done writing CSV table to disk.
Routing table created without error.
Process: LINKID.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\linkid
Process: frxst_pts.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\frxst_pts
Process: frxst_pts was empty.
Process: gw_basns.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\gw_basns
Process: gw_basns was empty.
Process: LAKEGRID.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\lakegrid
Process: CHANNELGRID.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\channelgrid
Process: landuse.nc completed without error
Output File: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\landuse
Step 4 completed without error.
Completed without error.
Completed script ProcessGeogridFile...
Succeeded at Mon Nov 17 20:01:54 2014 (Elapsed Time: 2.09 seconds)
```

Results

- Current Session
 - ExamineOutputs [223032_11162014]
 - Output Folder: Boulder_Creek_100m_1km_NHDPI1
 - Inputs
 - Input ZIP File: Boulder_Creek_100m_1km_NHDPlus_Reservoirs.zip
 - Environments
 - Messages
 - Executing: ExamineOutputs E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\ExamineOutputs.bat
 - Start Time: Sun Nov 16 22:30:30 2014
 - Running script ExamineOutputs...
 - Beginning to extract WRF routing grids...
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\CHANNELGRID.nc
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\flowacc
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\flowdir
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\frxst_pts
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\gw_basns
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\LAKEGRID.nc
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\landuse
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\latitudinal
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\longitudinal
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\ovroughrtfac
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\retdeprtfac
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\str_order
 - File Created: E:\Projects\Gochis\domains\Boulder_Creek\Boulder_Creek_100m_1km_NHDPI1\topog
 - Extraction of WRF routing grids completed.
 - Completed script ExamineOutputs...
 - Succeeded at Sun Nov 16 22:30:32 2014 (Elapsed Time: 2.09 seconds)

Documentation & Test Data

- Detailed documentation
 - 40 page PDF, and currently being reviewed
 - Describes tool capabilities, requirements, parameters, and GIS methods used in the tool chain.
- Small GEOGRID domains for testing tool functionality
 - Front Range (Lambert Conformal Conic)
 - India (Mercator)
- Expected Output provided for comparison
- Required Elevation files (.tif) provided
- Optional stream gages & lakes provided (Front Range)

Bottlenecks / Constraints

- Project High-Resolution Dataset for large areas
 - Can be avoided by pre-projecting/resampling high res data
- Flow Accumulation
- Not multi-threaded
 - Processes run on one core
 - Potential to use `python multiprocessing` module to utilize multiple cores.
 - Exception: TauDEM
- `arcpy` cannot build netCDF files with multiple variables
 - Final step is a shell script to combine all output .nc files

Higher Functionality Tools

- TauDEM Integration
 - D-Infinity flow direction algorithm
 - Flow direction apportionment between multiple downstream cells
 - Peucker-Douglas algorithm for stream definition
 - Thins stream network in low-relief areas
 - TauDEM tools allow multiprocessing
 - Specify number of cores to use for each tool
- Burn DEM using existing hydrography

Wishlist

- Better Esri NetCDF functionality (possibly with 10.3 release)
 - Multi-variable netCDF generation
 - Mosaic Datasets ingesting NetCDF multi-file dataset
- netCDF4 python libraries shipped with ArcGIS (10.3 SP1?)
- Pre-processing Web Processing Service (WPS)
- Utilize additional NHDPlus spatial and tabular data
 - Sink points, burn components, etc.

WRF-Hydro GIS Preprocessing Tools Demo

ksampson@ucar.edu

THANK YOU