NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

Building Python Tools in Support of Hydrologic Model Development

Matt Casali, GISP
GIS Program – Research Applications Laboratory
National Center for Atmospheric Research

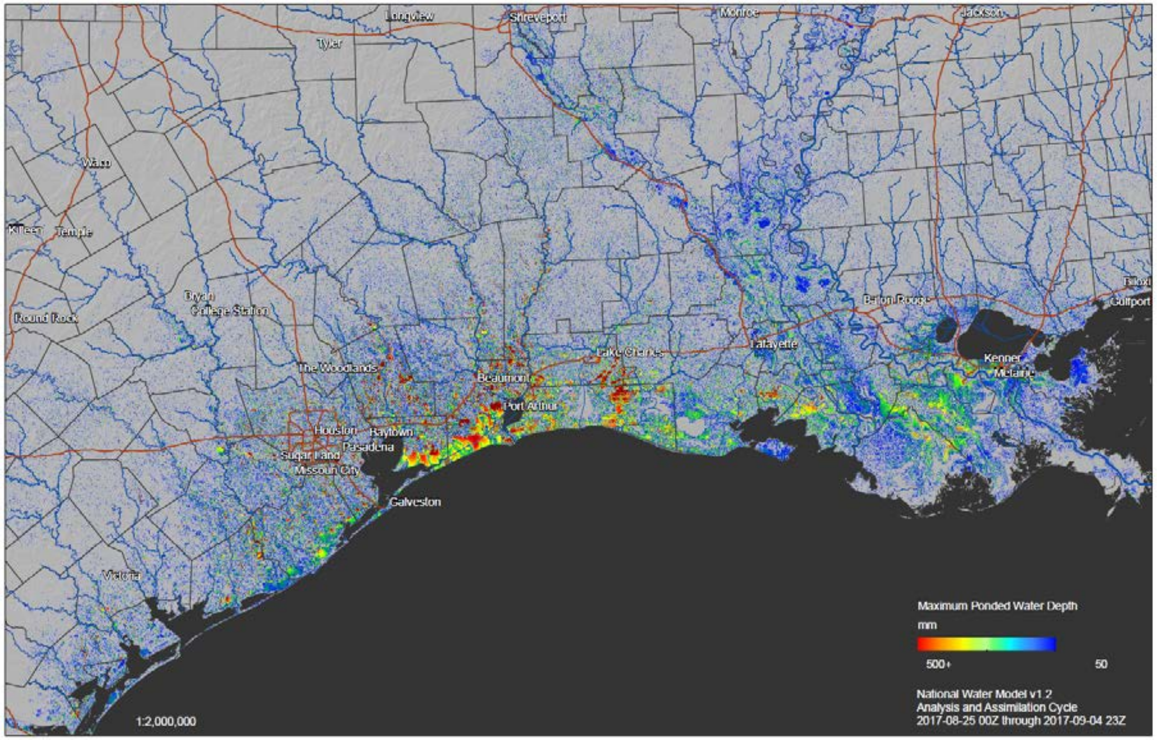
Esri Ocean GIS Forum 2019

November 5, 2019

Redlands, CA

Motivation

- Apply best available geospatial science tools and techniques in support of high resolution atmospheric modeling at NCAR
- Working with atmospheric model data in GIS systems requires specialized data formats, standards, and tools.
- Need for fast, efficient methods of producing model input data, and ways to handle visualizing the model outputs
- Removing the GIS burden from modelers wherever possible



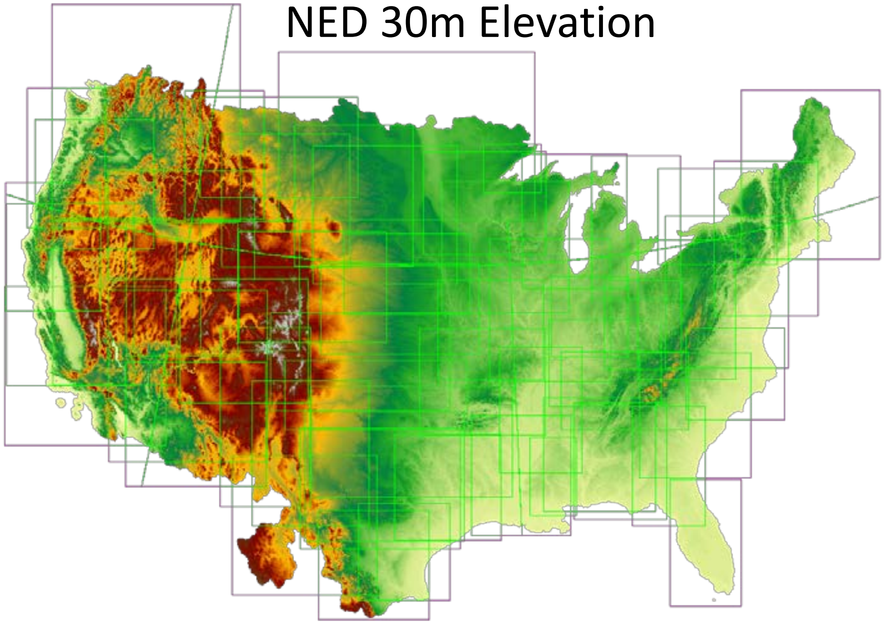
Where GIS plugs-in

- Models are high-level computer code (often Fortran) that operate on gridded data.
  - Modeling architectures typically do not have support for GIS libraries or capability to read common GIS formats
- GIS Pre-processing:
  - Prepare and process GIS data ahead of runtime into model-ready formats
  - Develop re-usable tools for generating model domains/datasets
  - Derive static model parameters from spatial data
- GIS Post-processing:
  - Format conversion, georeferencing model output data
  - Visualization and analysis of model results
  - Serve model data to the community using standard protocols

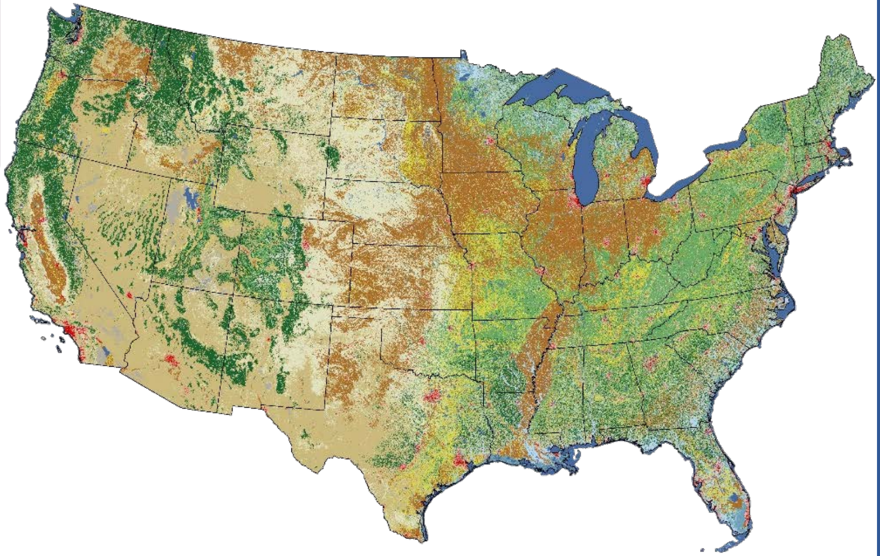
Models are data-hungry

- High-resolution models require very detailed characterizations of the land surface
  - Landuse, vegetation, elevation, soils, etc.
- Remotely-sensed data is key to providing continuous surfaces
  - NLCD 2011, MODIS greenness, SRTM elevation, etc.

NED 30m Elevation

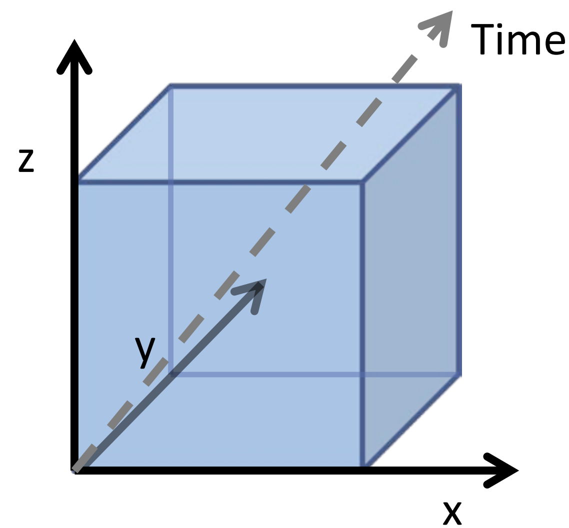


NLCD 30m Landcover



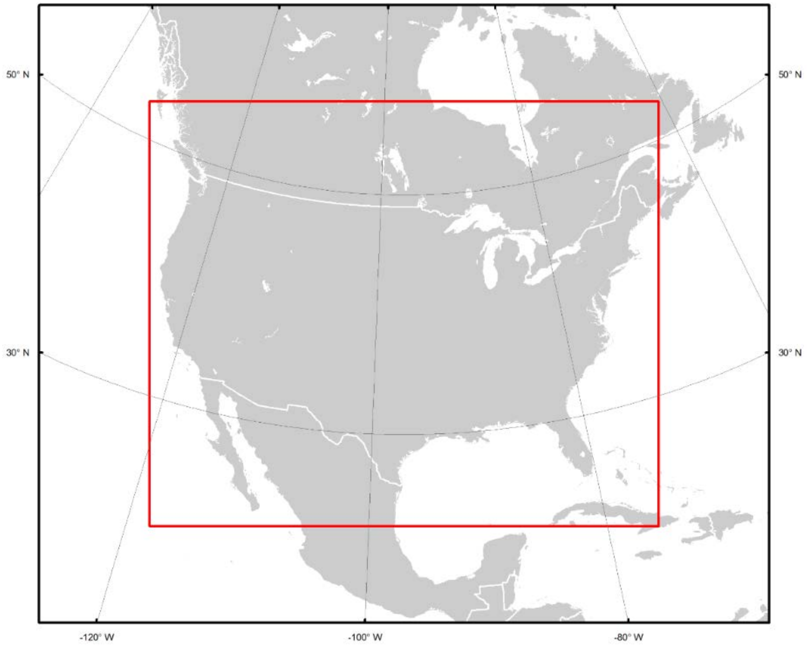
netCDF and CF Conventions

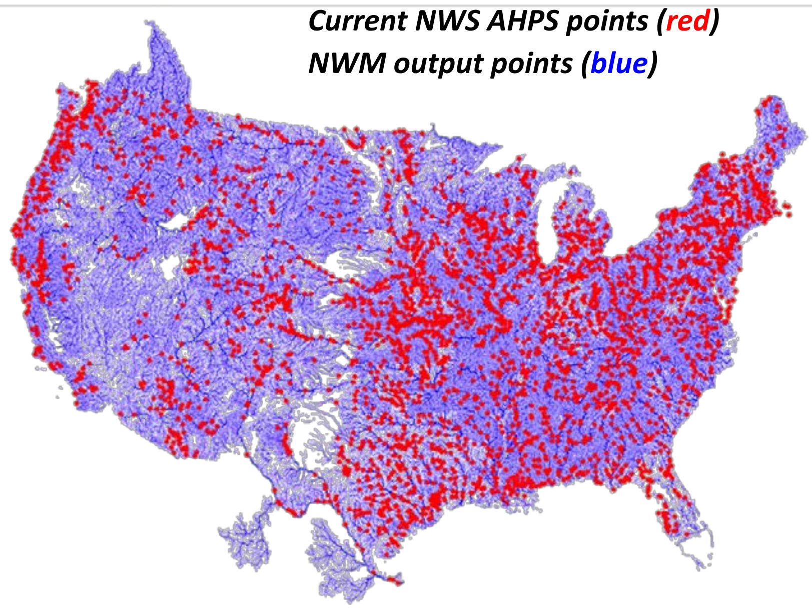
- network Common Data Form
  - ".nc" extension
- Self describing
  - Includes information about the data, coordinate system
- Machine independent
  - Usable in many operating systems
- Used in atmospheric science models
- Multidimensional
  - x, y, z, time
- Climate and Forecast Conventions
  - http://cfconventions.org/latest.html
- CF-compliant netCDF files make them much easier to use
  - Panoply, ArcGIS, QGIS



NOAA National Water Model (NWM)

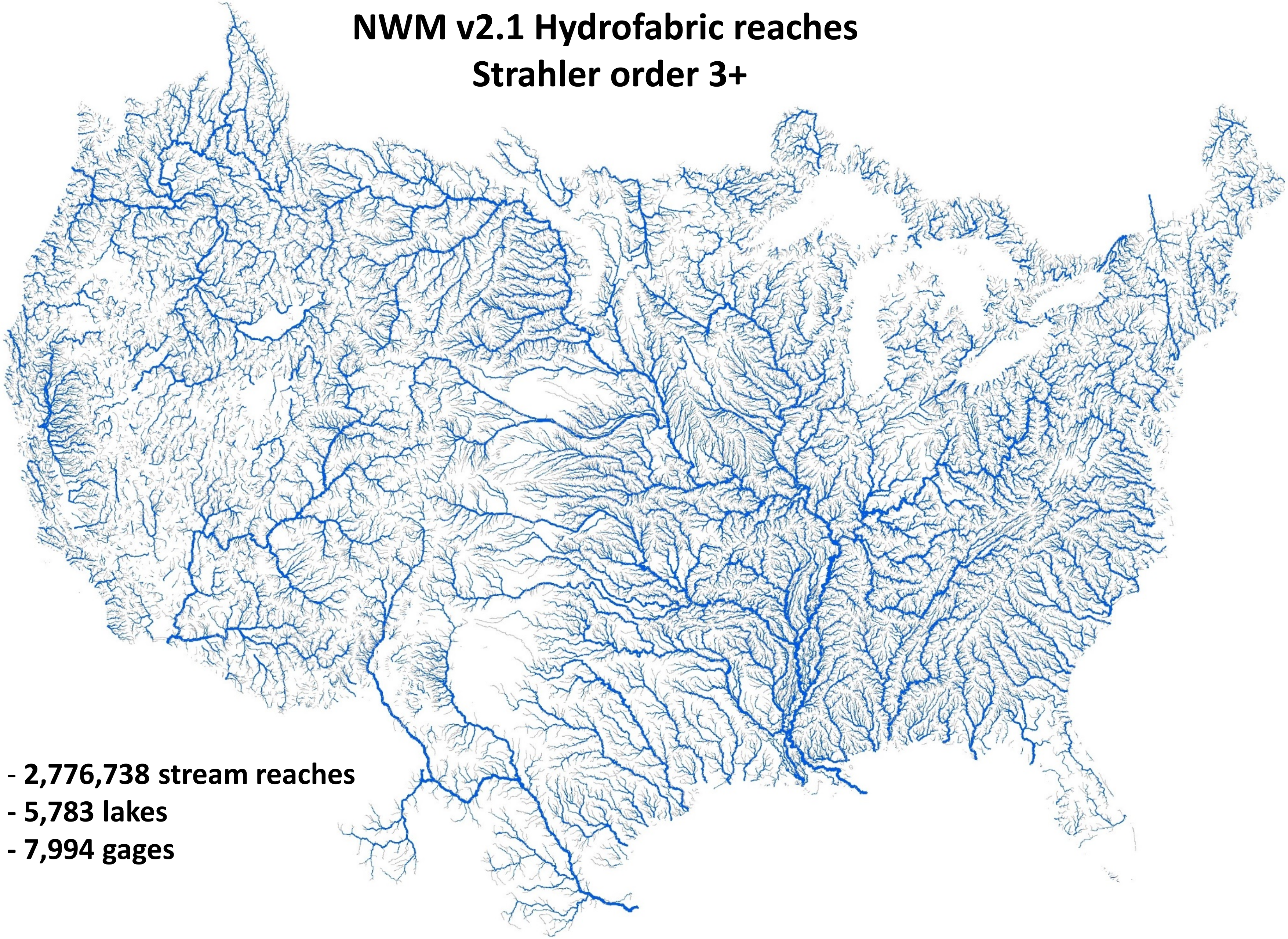
- CONUS-encompassing domain
- 1km Noah-MP Land Surface Model (LSM)
- 250m gridded diffusive-wave overland flow routing
  - Saturated subsurface flow
- NHDPlus v2 Muskingum-Cunge channel routing
- Reservoir routing





http://water.noaa.gov/about/nwm

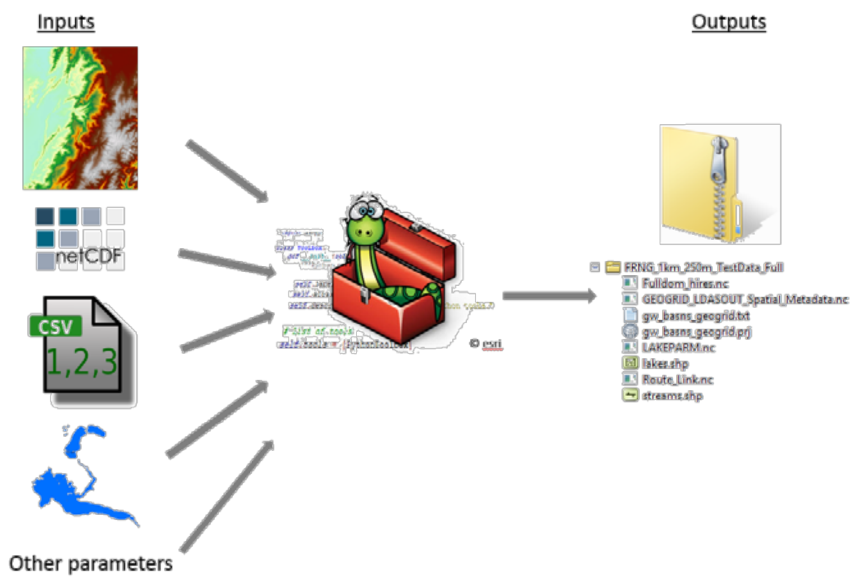
Current hydrography base



- 2,776,738 stream reaches
- 5,783 lakes
- 7,994 gages

Terrain pre-processing tools

- WRF-Hydro ArcGIS Pre-processing Toolset
  - Pre-processing tools, written in Python, using ArcGIS python API (`arcpy`)
  - Variety of WRF-Hydro configuration options supported
  - Consistent processing methodology between domains, regions, datasets
  - Removes the heavy GIS burden from modelers
  - Fast, efficient method for producing the 'routing stack' and other convenience data
  - Output files are model-ready



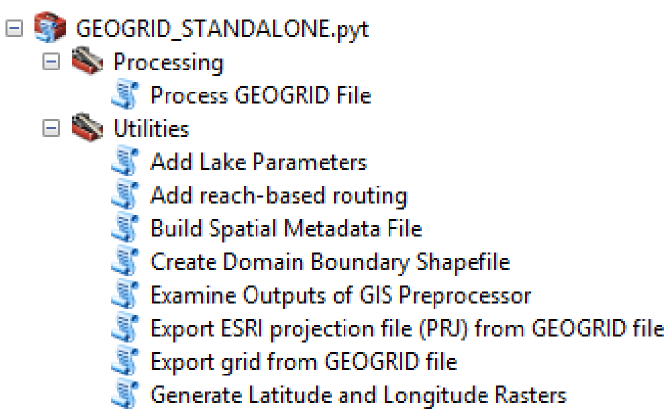
GitHub

/NCAR/wrf_hydro_arcgis_preprocessor.git

Python toolboxes

- Python scripts wrapped to act as an ArcGIS Toolbox
- PYT file is the toolbox script containing multiple toolsets
- Parameter handling and validation
- Requirements:
  - Version 10.3.1 or higher
    - Any license level (Basic, Standard, or Advanced)
    - Spatial Analyst extension required
  - Any Python version installed with ArcGIS Desktop
    - ArcGIS 10.3: Python 2.7.8, NumPy 1.7.1
  - New functionality in ArcGIS Pro / Python 3.x

Toolbox Script (.pyt)



WRF-Hydro preprocessor toolbox as viewed from ArcCatalog.

Advantages

Easy to modify
Portable
Many tools organized

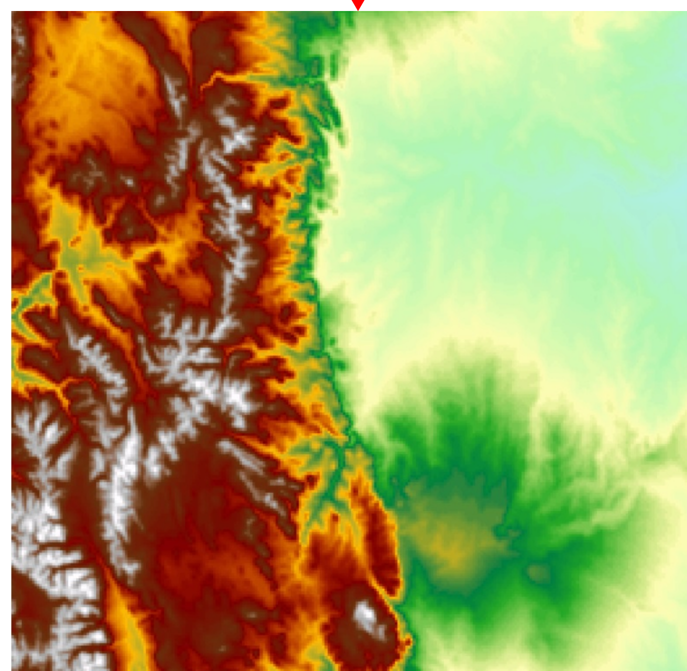
WRF & coordinate systems

```
: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 projection engine string

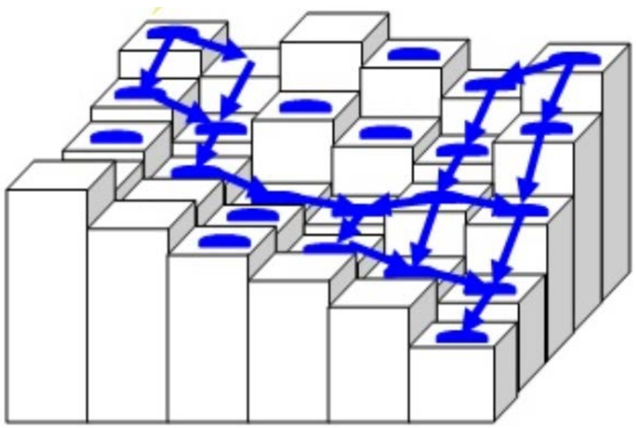
"PROJCS['**Lambert_Conformal_Conic**',GEOGCS['GCS_Sphere',DATUM['D_Sphere',SPHEROID['Sphere',6370000.0,0.0]],PRIMEM['Greenwich',0.0],UNIT['Degree',0.0174532925199433]],PROJECTION['Lambert_Conformal_Conic'],PARAMETER['false_easting',0.0],PARAMETER['false_northing',0.0],PARAMETER['central_meridian',-**105.0**],PARAMETER['standard_parallel_1',**30.0**],PARAMETER['standard_parallel_2',**50.0**],PARAMETER['latitude_of_origin',**39.9400138855**],UNIT['Meter',1.0]];-36695400 -29251300 10000;-100000 10000;-100000 10000;0.001;0.001;0.001;IsHighPrecision"

geo_em.d01.boulder_creek_1km.prj

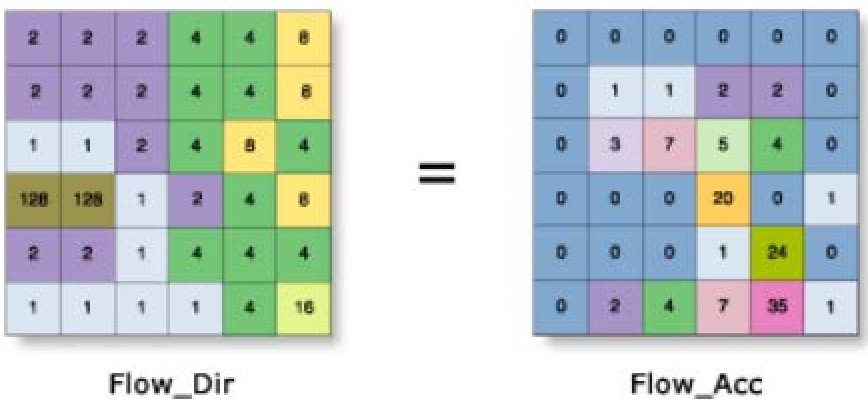


Terrain pre-processing

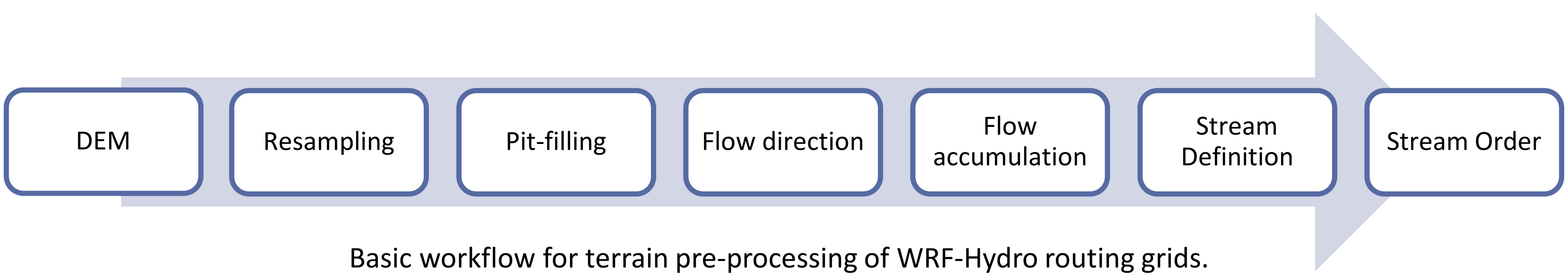
- Because WRF-Hydro is a hydrologic model, treatment of the landscape is required to allow flow paths off of the land surface
- Re-gridding/coarsening ‘breaks’ the hydrologic connectivity of even hydrologically enforced datasets.



Overland flow routing



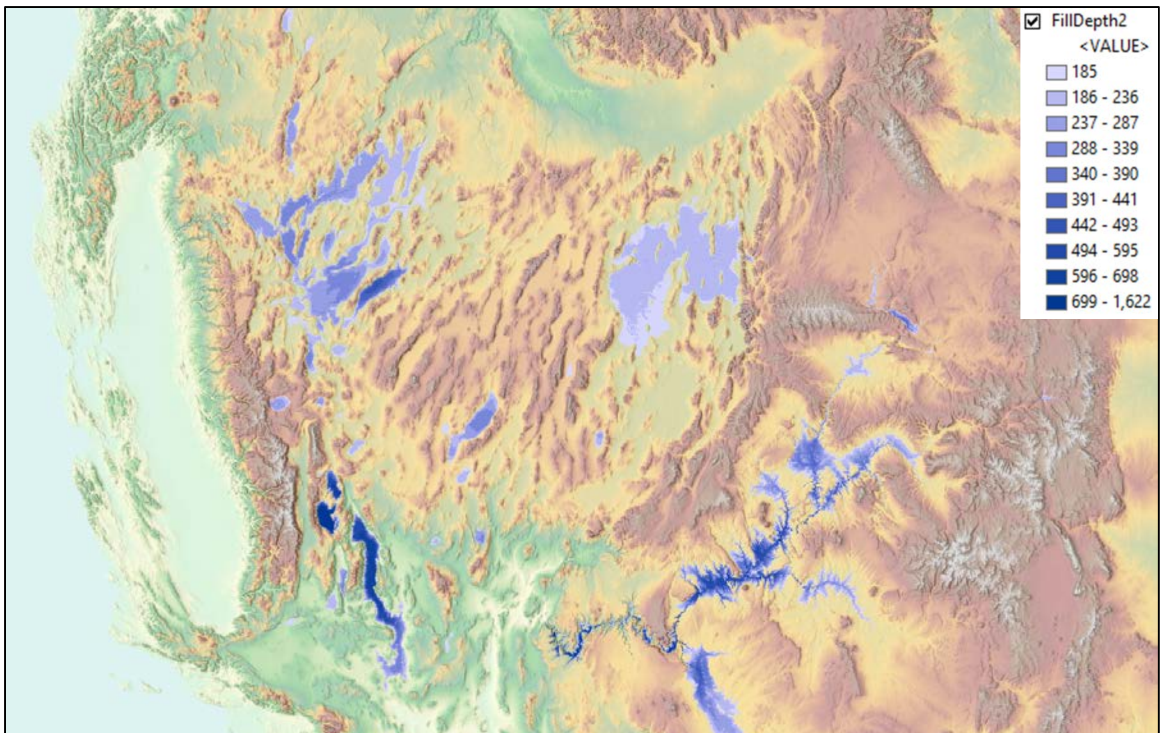
https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/flow-accumulation.htm



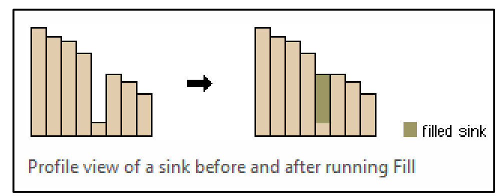
Basic workflow for terrain pre-processing of WRF-Hydro routing grids.

Process: pit filling

- Fill depressions so that water can roll downhill only. This also creates a smoother DEM than you might find in nature.
- This simple hydro-enforcement method can resolve most flow issues in a DEM.
- Optional z-limit (global variable) to limit fill depth.
- Some pits are natural features in the terrain.



Spatial Analyst “Fill“ Tool



© Esri: http://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/how-fill-works.htm

Methods: Tarboton et al (1991)

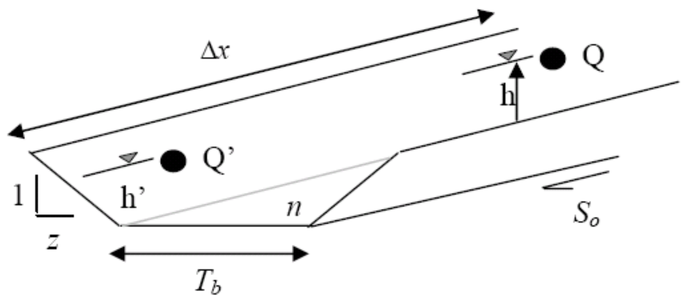
GIS vector data -> model-ready data

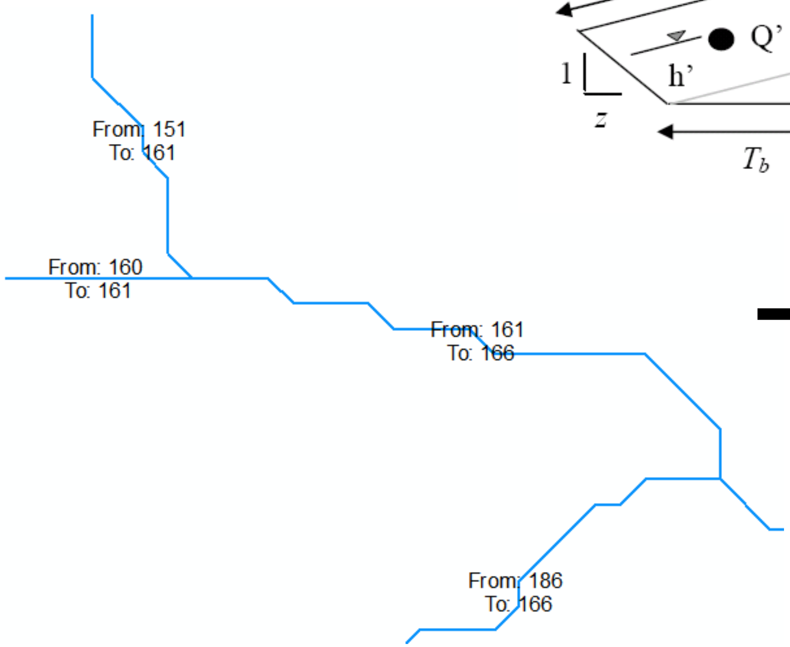
- All features represented as ‘objects’ in the model, collapsed into 1D points with associated parameters
- Chanel parameters are defined for the entire flowline and given the spatial location of the flowline midpoint
- Point data utilize CF H.2.1 standard

```
netcdf RouteLink.nc {
dimensions:
        feature_id = 2729077 ;
variables:
        int link(feature_id) ;
                link:long_name = "Link ID (NHDFlowline_network COMID)" ;
                link:cf_role = "timeseries_id" ;
                link:coordinates = "lat lon" ;
        int from(feature_id) ;
                from:long_name = "From Link ID (PlusFlow table FROMCOMID for every TOCOMID)" ;
                from:coordinates = "lat lon" ;
        int to(feature_id) ;
                to:long_name = "To Link ID (PlusFlow table TOCOMID for every FROMCOMID)" ;
                to:coordinates = "lat lon" ;
        float lon(feature_id) ;
                lon:long_name = "longitude of the segment midpoint" ;
                lon:units = "degrees_east" ;
                lon:standard_name = "longitude" ;
                lon:coordinates = "lat lon" ;
        float lat(feature_id) ;
                lat:long_name = "latitude of the segment midpoint" ;
                lat:units = "degrees_north" ;
                lat:standard_name = "latitude" ;
                lat:coordinates = "lat lon" ;
// global attributes:
                :Convention = "CF-1.6" ;
                :featureType = "timeSeries" ;
                :history = "Created Tue Feb 27 14:45:27 2018" ;
                :processing_notes = "";
```

Deriving channel parameters

- Geospatial features decomposed into Muskingum-Cunge routing parameters for each ‘reach’. Output is netCDF
  - 2D data (stream lines, lake polygons) converted to 1D
    - Geometry & network attributes; Length, width, topology
    - Derived attributes: Slope, elevation, stream order, width





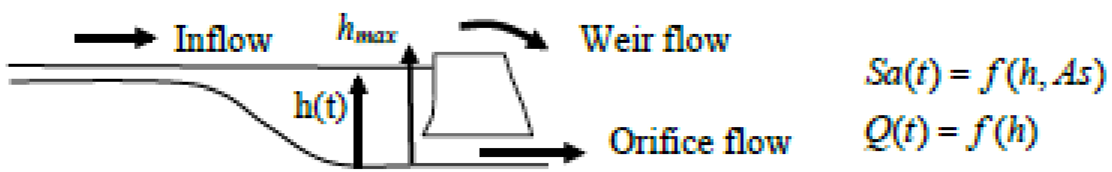
RouteLink

OID	linkDim	link	from	to	lon	lat	alt	order	Length
1	1	663557	0	6635570	-96.5402	46.22878	294.75	1	1070
2	2	663559	0	6635600	-96.53065	46.21349	295.89	1	1117
3	3	663559	0	6635636	-96.50535	46.20151	296.15	1	2303
4	4	663562	0	6635620	-96.61502	46.20052	292.49	1	1119
5	5	663562	0	6635624	-96.63716	46.19552	294.39	1	3171
6	6	663564	0	6635660	-96.41427	46.18661	300.7	1	5338
7	7	663565	0	6635660	-96.39973	46.17369	300.46	1	2344
8	8	663565	0	6635620	-96.61041	46.18512	293.25	1	2558
9	9	663567	0	6635642	-96.53323	46.18099	296.3	1	3251
10	10	663569	0	6635726	-96.30599	46.15612	307.89	1	3510
11	11	663570	0	6635718	-96.13979	46.14558	341.56	1	998
12	12	663570	0	6635718	-96.15008	46.15319	335.79	1	2513
13	13	663571	0	6635716	-96.19209	46.14944	330.73	1	3084

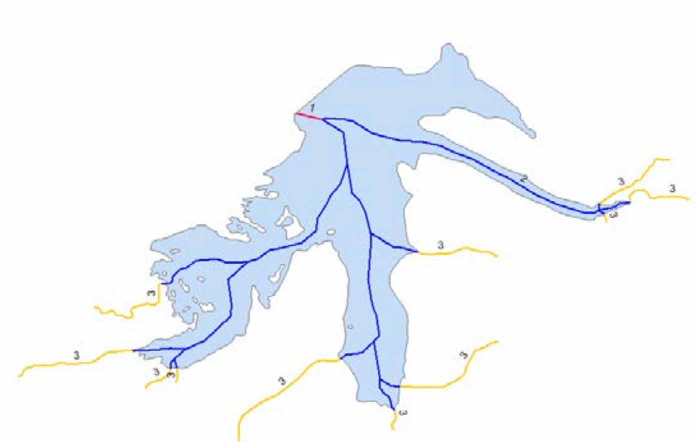
1 (0 out of 2729077 Selected)

Deriving lake parameters

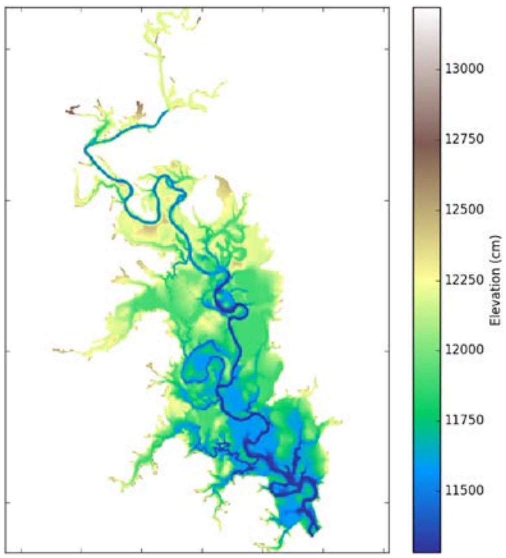
- Parameters largely derived from the hydro-DEM
  - Area, elevation, depth, weir height
- Lakes are represented as objects connected to the outlet flowline
- Lakes “fill and spill” as a function of storage



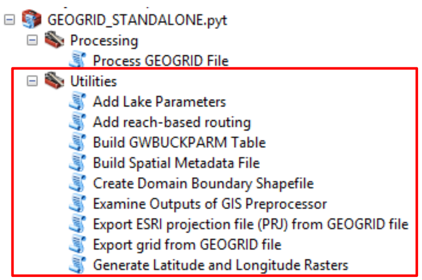
From WRF-Hydro User Guide v3.0, Figure 3.6

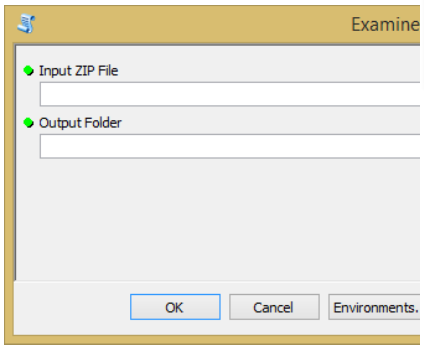


Lake link types: 1 = Outlet, 2 = internal link, 3 = inflow link

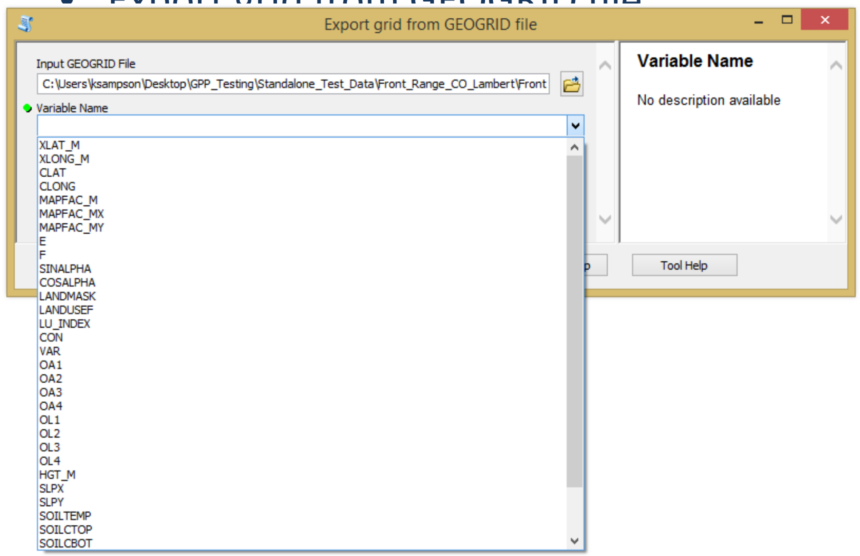


Other tool utilities





- Examine Outputs of GIS Preprocessor
  - Extracts .zip output file to individual rasters for viewing in Desktop GIS applications.
- Export grid from GEOGRID file



file

ster

in boundary from a GEOGRID file

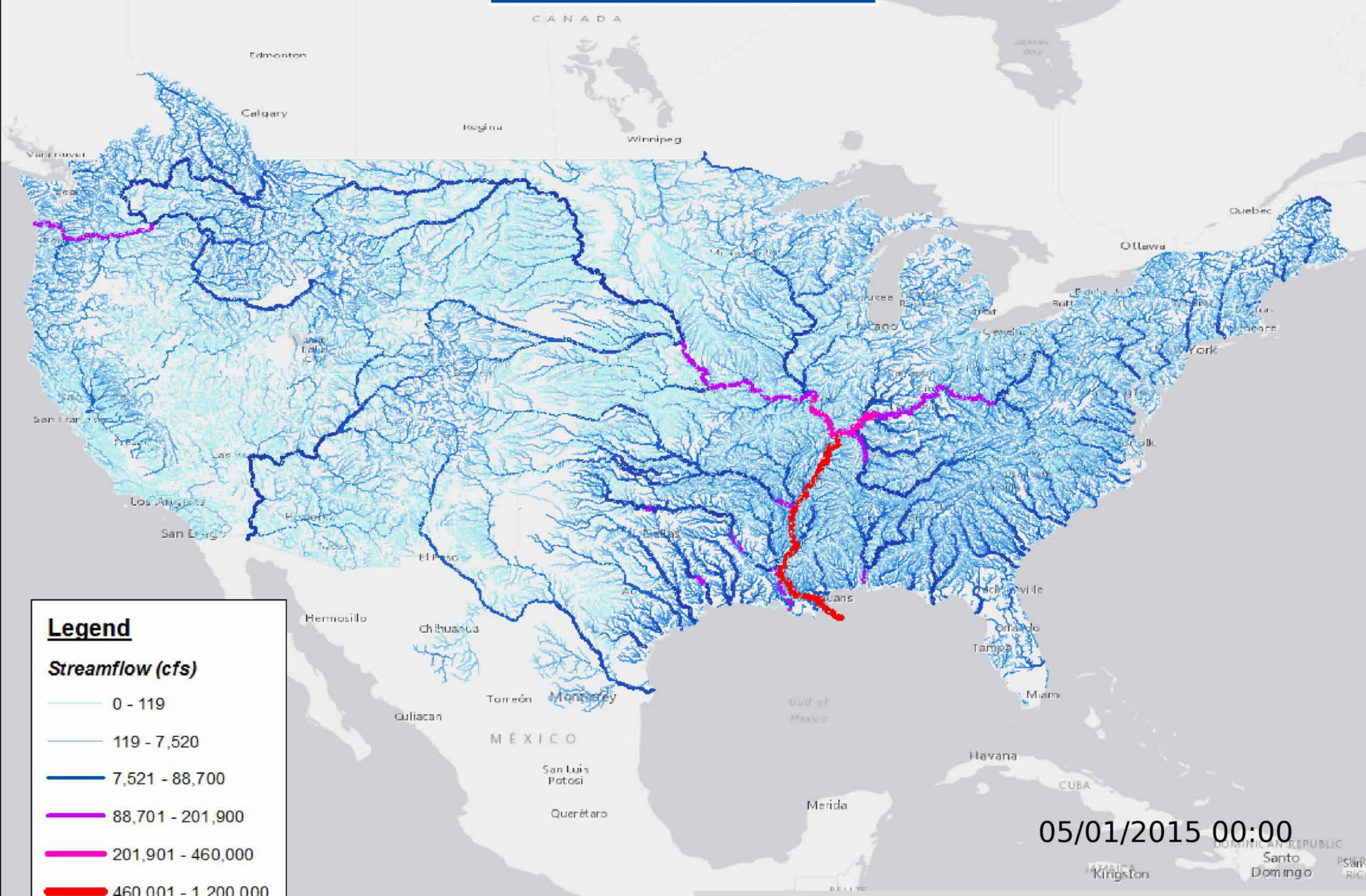
- Build Groundwater Inputs
  - Creates groundwater input files in 3 ways

Bottlenecks and constraints

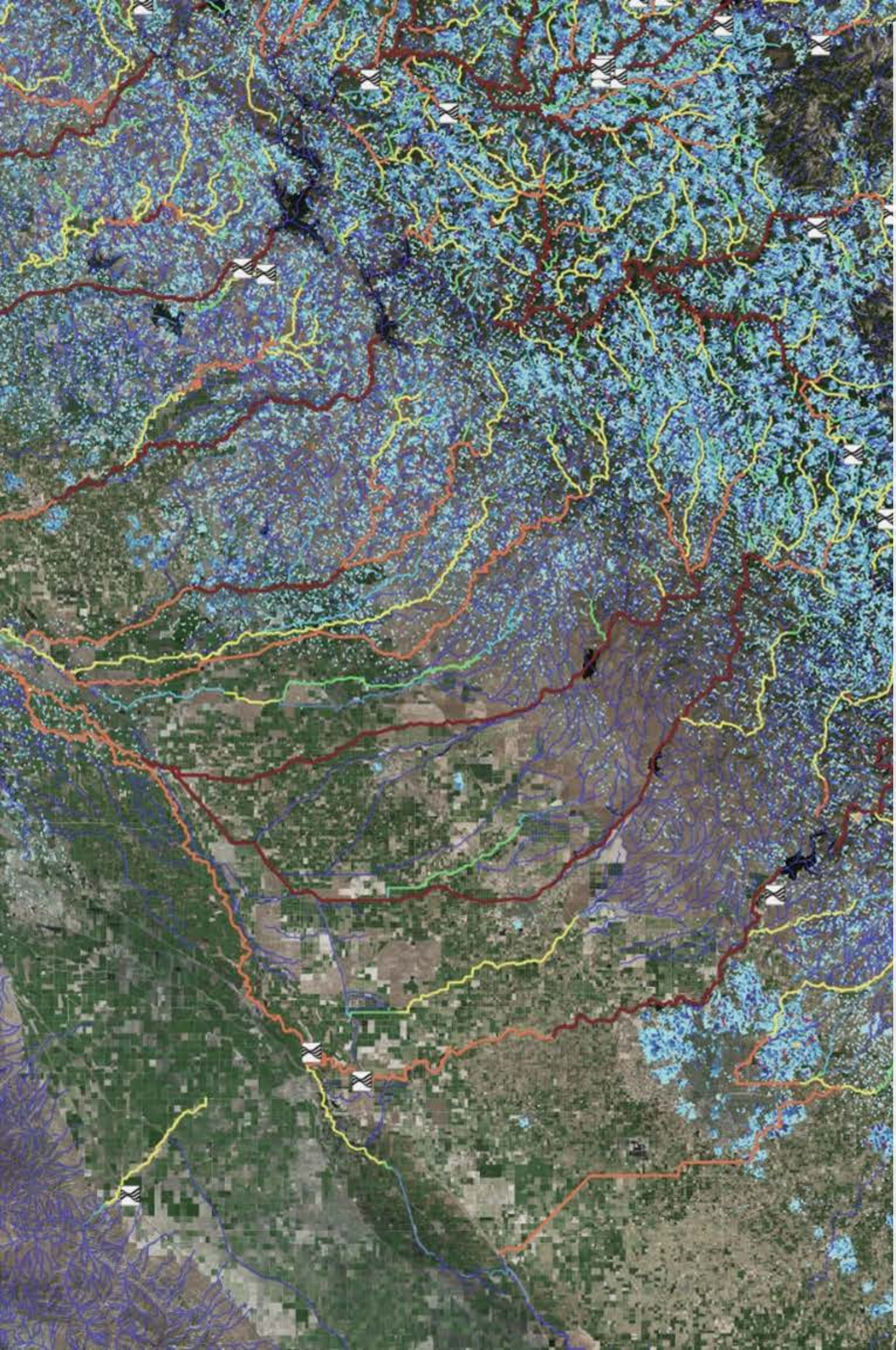
- Project high-resolution dataset for large areas
  - Can be avoided by pre-projecting/resampling high res data before running the GIS pre-processor
- Flow Accumulation – slowest part of the process
- Currently not multi-threaded
  - Process runs on one core
  - Process chain not well suited to parallelization
    - Workflow must be performed in order
- No breaching or hybrid methods available in ArcGIS
  - Pit-filling is only option to deal with depressions
  - Open source GIS such as Whitebox Tools utilizing advanced algorithms for breaching depressions
  - Results in less modification to DEM
- Windows only GIS platform
- Version to version bugs
- Large domains will not process in 32 bit processing
  - 64 bit background geoprocessing is necessary, not installed by default

WhiteboxTools

National Water Model



Credit: Fernando Salas, National Water Center



Thank you!

casali@ucar.edu

NCAR GIS:
gis.ucar.edu

National Water Model:
water.noaa.gov/about/nwm

WRF-Hydro Community Model:
ral.ucar.edu/projects/wrf_hydro

WRF-Hydro GIS Pre-processing Tools:
github.com/NCAR/wrf_hydro_arcgis_preprocessor