The Role of NHDPlus as a Geospatial Framework for USGS SPARROW Modeling

John Brakebill, USGS NAWQA Program

September 24, 2015

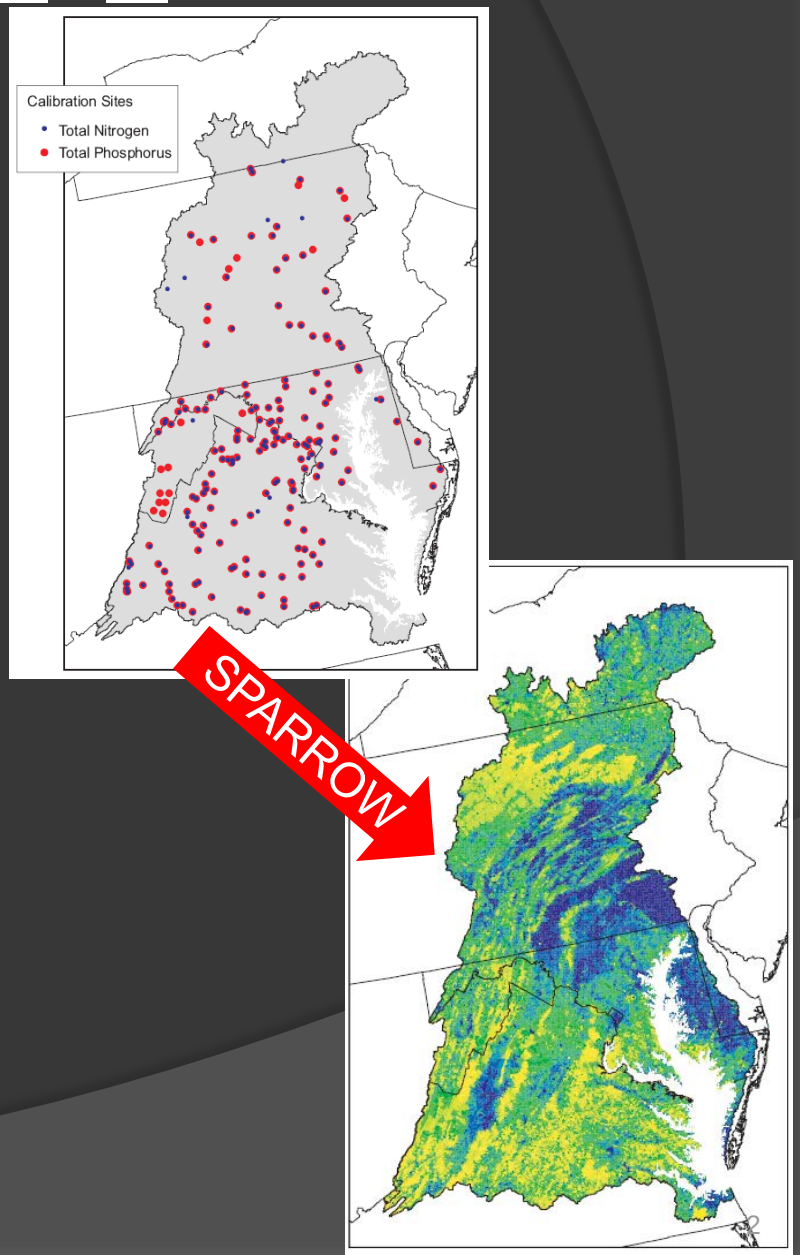
1) SPARROW Background, Required Framework
2) Spatial Data Referencing

Thanks : Greg Schwarz, Steve Preston, Craig Johnston, Laura Hayes, Mike Wieczorek, Jeff Deacon, Tana Haluska, SPARROW Modeling Team

SPARROW

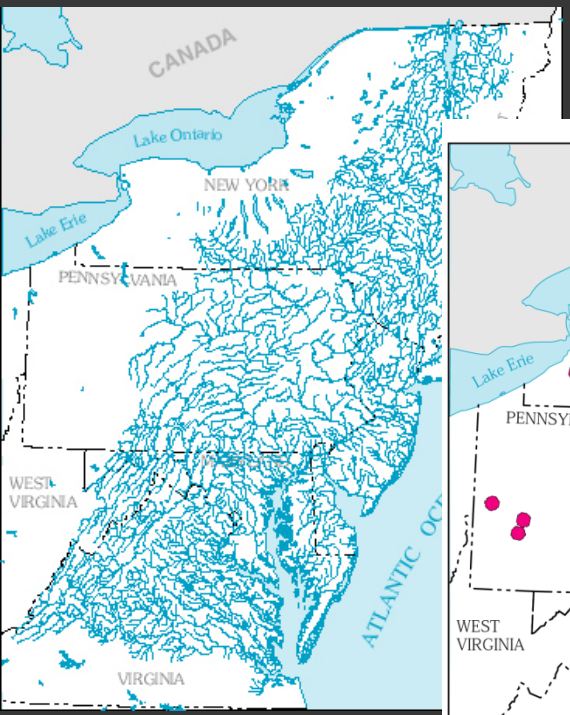
SPAtially Referenced Regressions On Watershed Attributes

- Spatial Statistical Approach that Empirically Relates Contaminant Sources and Transport Factors to Measured Stream Flux
  - Identify the spatial variability and magnitude of contaminant supply
  - Quantify the contributions at various locations
- Tool Provides Spatially Detailed Predictions:
  - Map individual contaminant sources in unmonitored locations
  - Statistical importance and quantification of contaminant sources
  - Provides measures of uncertainty
- Spatial Framework
  - Explicit for evaluating geographic distribution of sources that can be used for WIP's
  - Potential Geographic Targeting

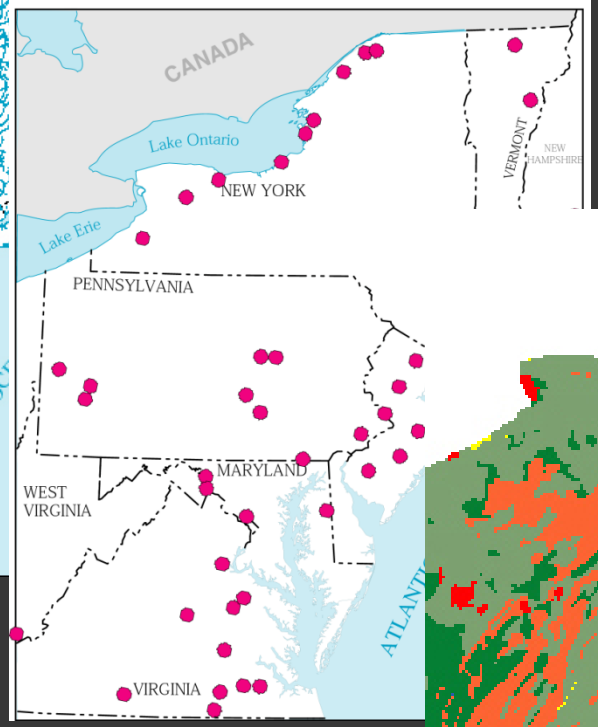


SPARROW Spatially Designed

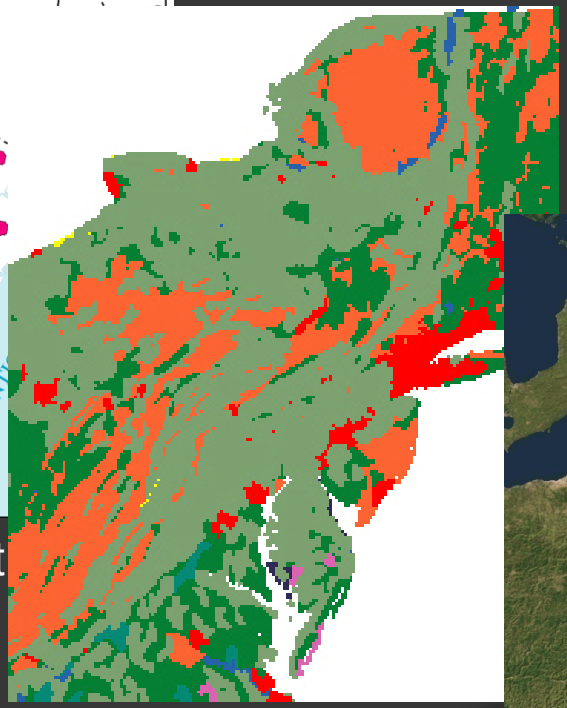
Integrates spatial data over multiple scales to predict origin & fate of contaminants



Network of connected and attributed streams and watersheds

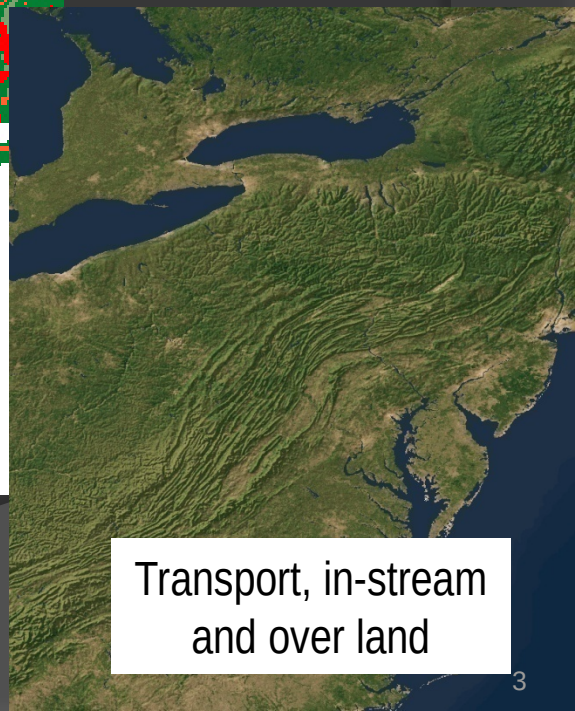


Monitoring Data (Dependent Variable)



Source data

Slope, Physiography, Soil Characteristics, Channel Characteristics and Reservoir Systems



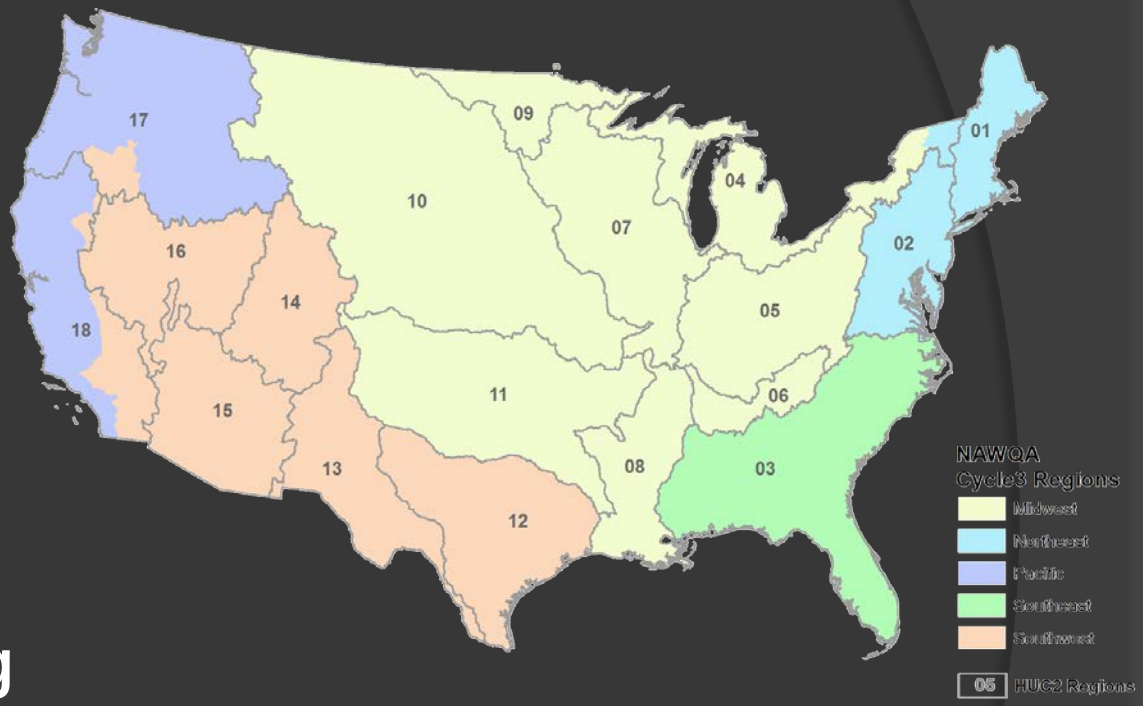
Transport, in-stream and over land

USGS
science for a changing world

SPARROW

National and Regional Modeling

1) Contiguous U.S.
2) USGS NAWQA Major River Basin Studies (modeld 2002 time period)
3) Regional models in the Chesapeake and Potomac Basins, Long Island Sound and New England, Mississippi, and Coastal South Carolina.

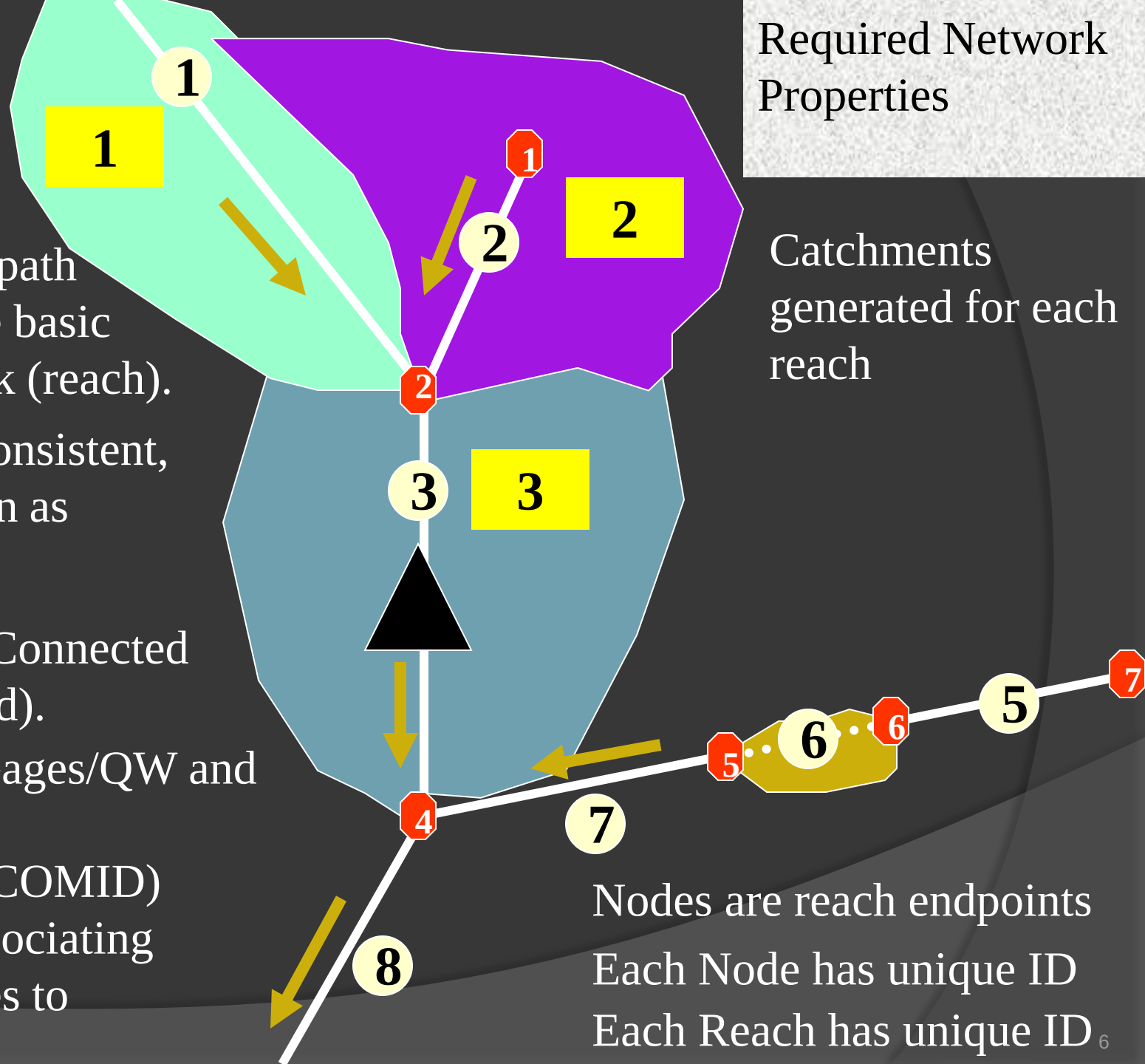


4) Current modeling - Refined to 5 major regions, modeling 2012 time period (NAWQA Cycle III).
5) Water volume, Suspended Sediment, Nitrogen and Phosphorus
6) NHDPlus (medium res) is the primary framework supporting NAWQA Cycle III models

Historical SPARROW Networks

- HUC-08 (~2,294 Nationally)
- RF1 (~66,000 reaches), average 16 km
  - Mid 1980's EPA product, 1:500,000
  - Provided flow and velocity estimates
  - Catchments (1k scale)
  - Median flow ~70 CFS
- Modified RF1
  - Enhanced monitoring network, better catchments, updates to attributes like time of travel and reservoir size estimates
- Enhanced RF1
  - Supported NAWQA Cycle II MRB modeling
  - Additional monitoring
- NHDPlus (> 2.5 million), average 2.2 km
  - Even more monitoring (~10 – 20%)
  - Median Flow ~1 CFS

Required Network Properties



A single flowpath represents the basic building block (reach).

Orientation consistent, same direction as streamflow.

Continuous (Connected and sequenced).

Referenced Gages/QW and Reservoirs

Unique IDs (COMID) allows for associating stream reaches to catchments

Catchments generated for each reach

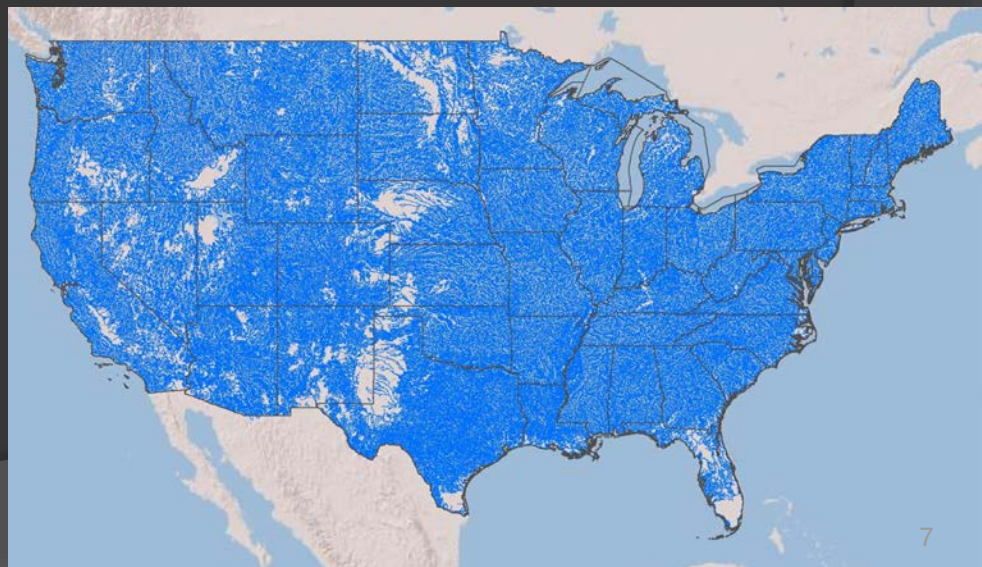
Nodes are reach endpoints

Each Node has unique ID

Each Reach has unique ID

NHDPlus and SPARROW

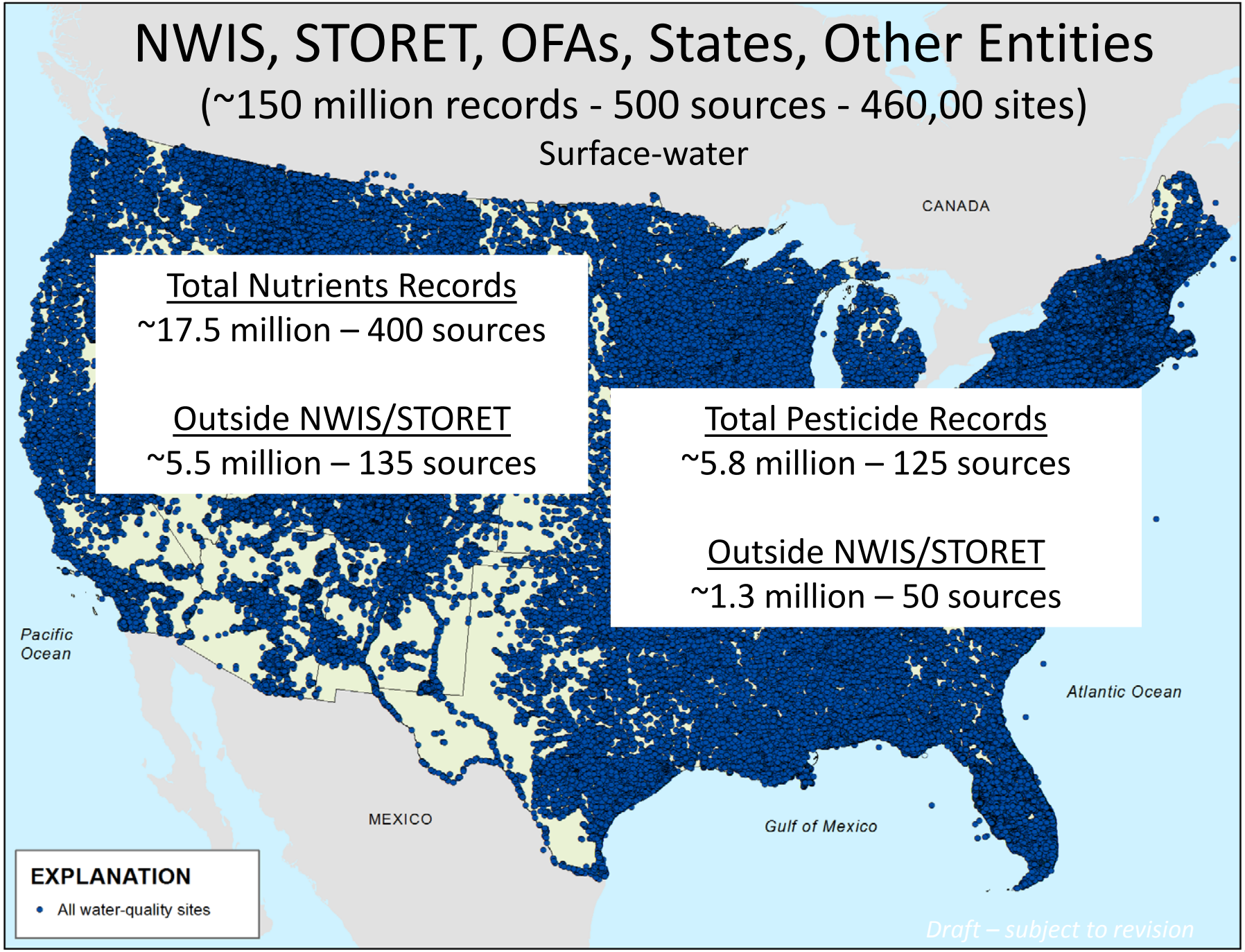
- Spatial Data Aggregation and Analysis:
  - NHDPlus V2 is a primary Framework (SW reach and catchment)
    - Associations (water-quality monitoring, gages, dams/impoundments, municipal point sources)
    - Characterizations (physical, source, climate)
    - Connectivity Improvements
    - Thinned (less reaches) versions will be generated for SPARROW modeling
      - HUC12 designations
  - Value added attributes for other uses



NWIS, STORET, OFAs, States, Other Entities

(~150 million records - 500 sources - 460,00 sites)

Surface-water



Total Nutrients Records

~17.5 million – 400 sources

Outside NWIS/STORET

~5.5 million – 135 sources

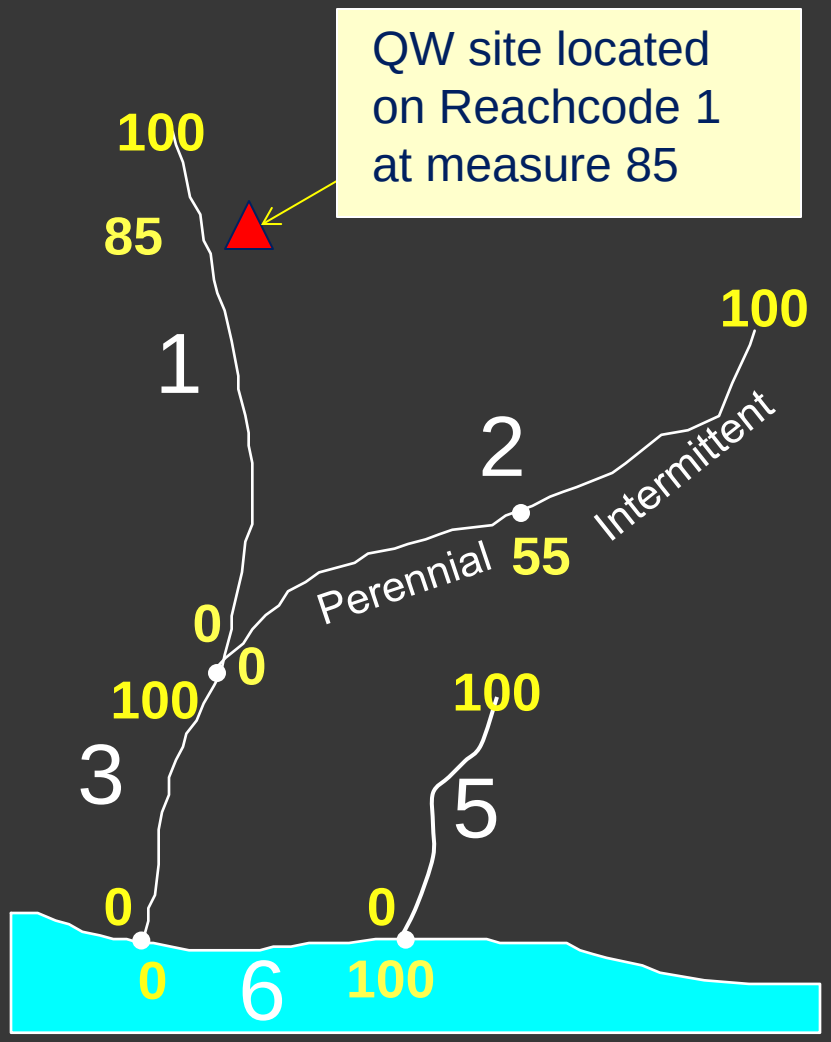
Total Pesticide Records

~5.8 million – 125 sources

Outside NWIS/STORET

~1.3 million – 50 sources

Addresses on Reaches



- An NHD reach is comprised of one or more flowlines
- Addresses on linear reaches are proportional.
  - measure 0-100 from bottom to top

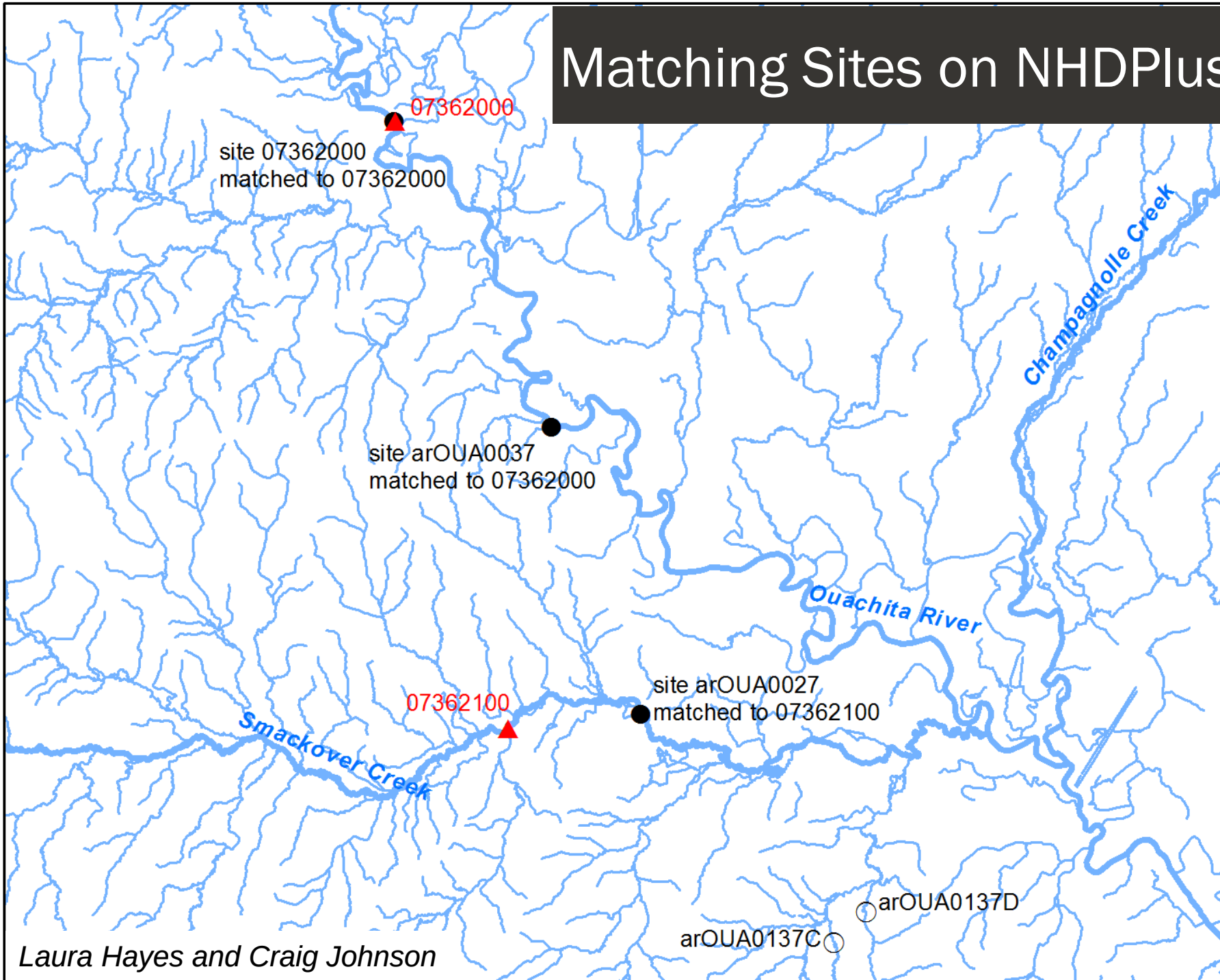
Laura Hayes and Craig Johnson

Scores for QA/QC

qaqc_name	comid_diff	uninit	dist_gt250	minordiv	qa_rivarea	qaqc_score
0	0	0	1	0	0	1
0	1	0	1	0	0	2

- ComID from catchment intersect does not match ComID from linear referencing
- Site referenced to uninitialized flowline
- Site snapped from a far distance away
- Site is on a secondary path divergence
- Site falls within NHDArea and near a confluence
- Total QA/QC “score” to help prioritize

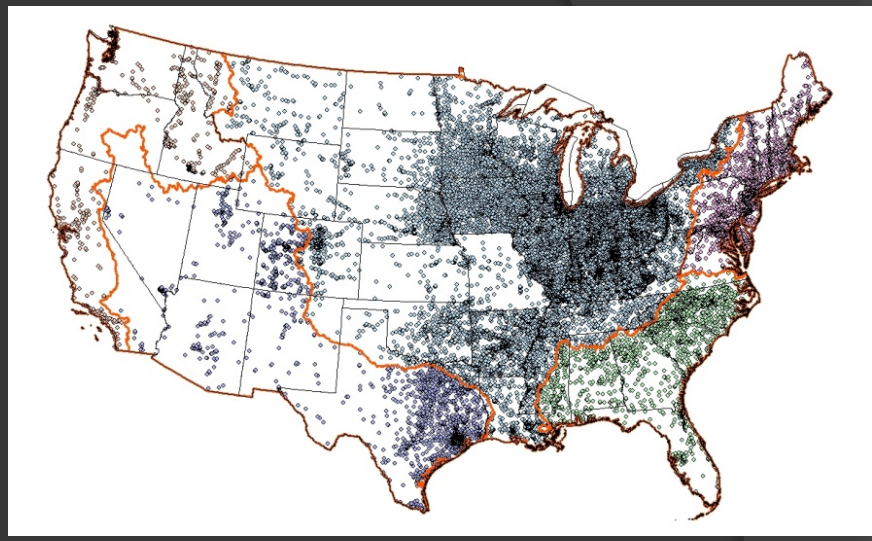
Matching Sites on NHDPlus



Laura Hayes and Craig Johnson

Municipal Point Sources

1. Approach
   - Review locations/associations and reach associations for majors identified in 2002
   - New national pull for 2012 from ICIS
   - Identify new majors since 2002 and review locations and associations
   - Seek discharge and concentration data for all majors missing that data



2. Current Specific Tasks:
   - Each regional SPARROW team working to complete the data review and compilation for majors

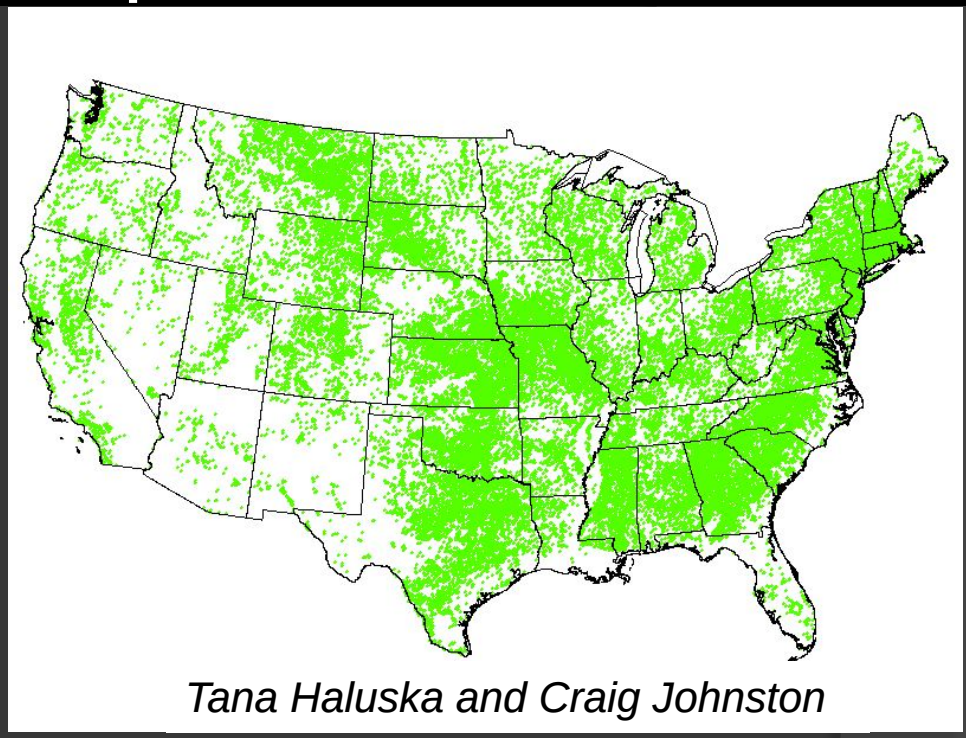
Goal – update the current point source data base to include data through 2012

3. Progress:
   - Most regional teams have nearly completed evaluating locations of majors
   - Some work has been completed to find missing discharge and concentration data and minors

Molly Maupin, Tammy Ivahnenko, Ken Skinner, Lori Sprague, Charlie Crawford, Steve Preston, SW Modeling Team, Others

Reservoirs/Impoundments

- Approximately 59,000 dam locations from several versions of the NID database were snapped to the medium resolution NHDPlus
- Started with the 2009 NID data associated to the High-res NHD , 2011 version associated to Medium Res, and a fresh pull from the 2013 NID.
- Only about 11% have been physically (visually) reviewed
- Script written to properly locate dams on the primary waterbody outlet flowlines.



Tana Haluska and Craig Johnston

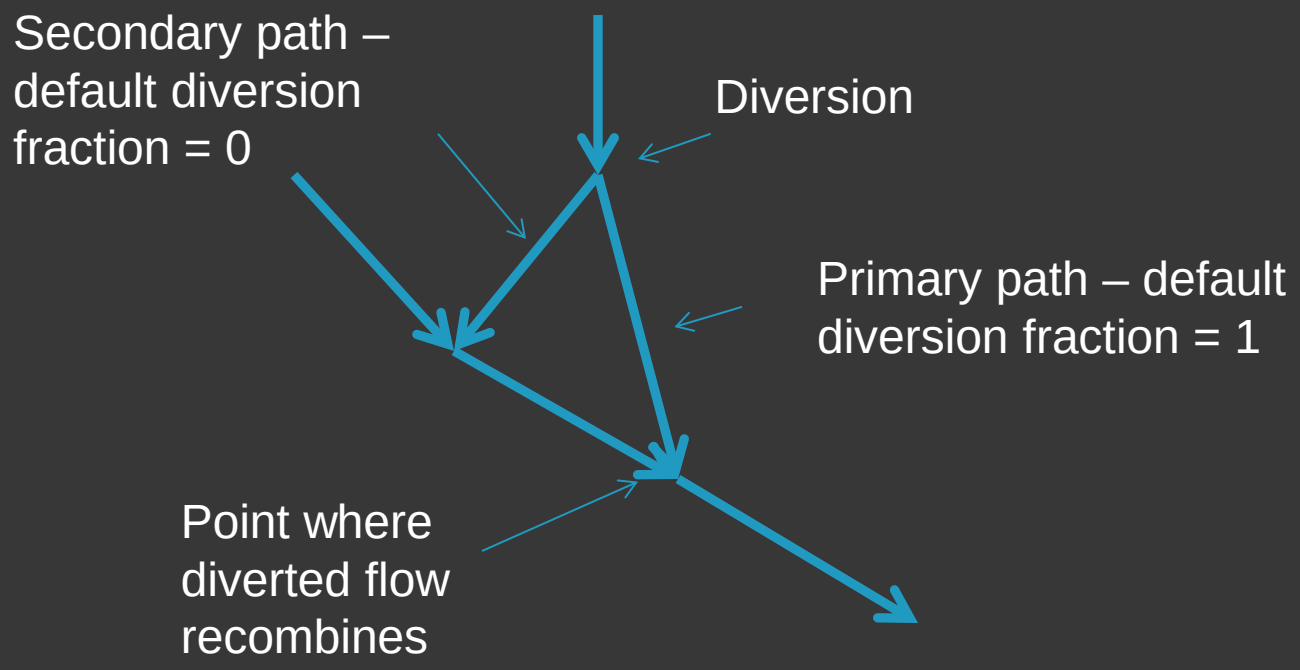
- Screening process being developed for unreviewed dam locations with snapping distance greater than 300 meters (11,000 plus)

Challenges

- Snapped and should not have been
- Some dams were snapped to an incorrect flowline.
- Location issues of lakes and or NID lat longs
- Some of the flowlines are pointed in the wrong direction.
- A few missing NIDID's
- Duplicate Dam locations
- Some known dams not in NID (2013) but in earlier versions
- 351 dams have multiple records in the NID database, these have been matched up with the correct NID record.
- Several thousand NID records are missing some key attributes
- We are aware of the sensitive nature of the NID locations and associated data.

	Drainage Area	Spillway Type	Surface Area	NID Storage	Normal Storage	Max Storage	Max Discharge	Year Completed	Purposes	Dam Type	River
	15936	21656	15179	130	2483	1374	21130	6162	1883	2186	1558
Percentage missing:	**27.58%**	**37.48%**	**26.27%**	**0.22%**	**4.30%**	**2.38%**	**36.57%**	**10.66%**	**3.26%**	**3.78%**	**2.70%**

NHDPlusV2 Diversions – Basic Concepts



All NHDPlusV2 diversions are assigned a primary path and a secondary path

For almost all diversions, NHDPlusV2 accumulations assign a diversion fraction of 1 to the primary pathway and a diversion fraction of 0 to all secondary pathways

Results of Diversion Evaluation

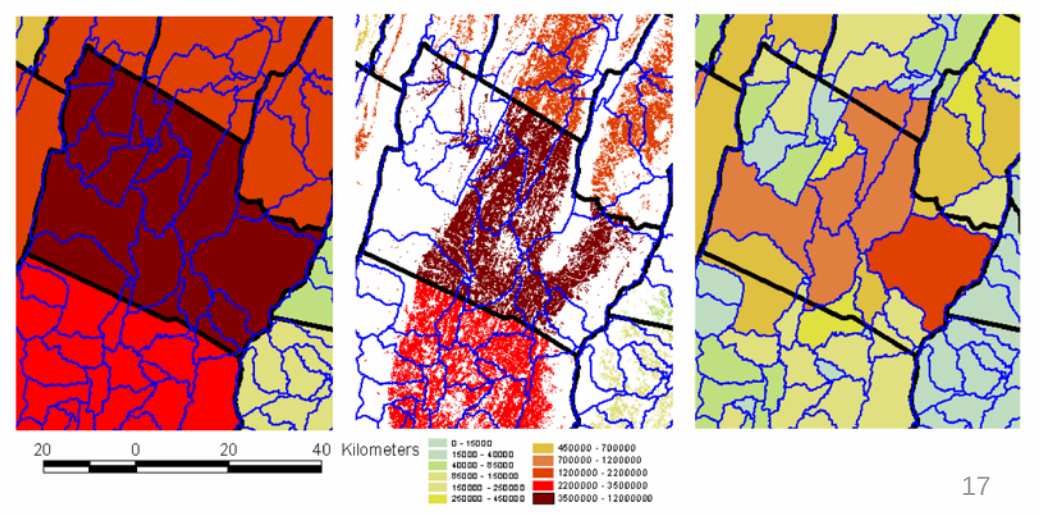
- Approximately 400 gages were identified for inconsistencies between NWIS and NHDPlusV2 drainage area
  - 65 gages were found to be improperly located
  - Remaining gages (80%) required modification of diversion routing
- 2,750 diversions were evaluated and 1,689 (61%) were found to contain routing errors (either primary path was miss-identified or the diversion was incorrectly placed on the network requiring changes in flow direction for some flowlines)
- New attributes added that include:
  - Flowlines that terminate into the ground
  - Flowlines with unresolved diverted flow
  - Groups (more than 2) of diversion-affected flowlines
- Google Earth used to verify path, fraction, and flow direction
- SAS Code incorporates routing modifications, checks that modifications are logically consistent (no broken flowpaths), and rebuild most network variables to be consistent with the revised routing.

Examples of Other Data Associated

- Fertilizer, Manure
- Atmospheric Deposition
- Nutrient content from Septic
- Natural Sources of P
- Land Use (NLCD, CDL)
- Soils
- Tile Drainage
- Geology/Lithology
- Irrigation/Water Use
- Climate variables

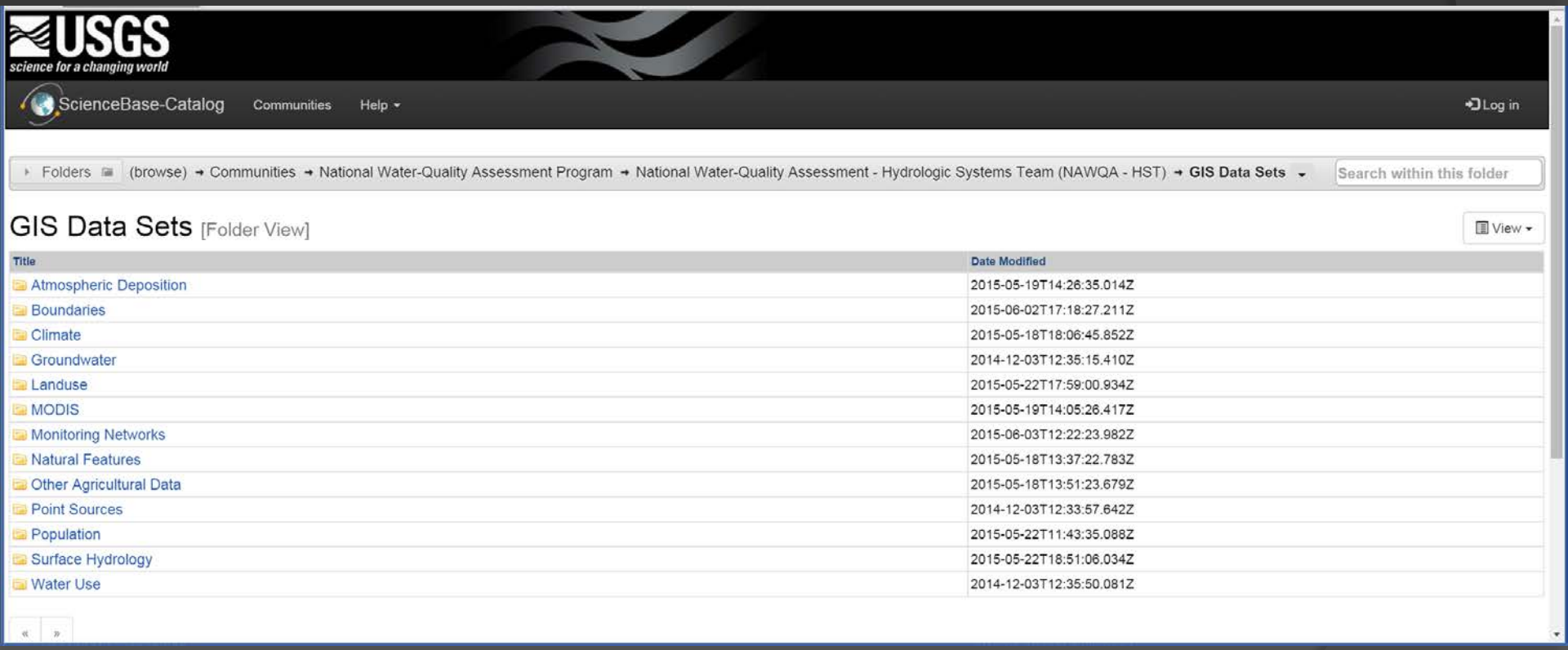
Tool Development

1) Apportion values to catchments
2) Accumulate upstream contributions



Geospatial Data Storage and Access

- ScienceBase – Online storage and access, raw data
  - Meets new USGS open access data requirements



Med-res NHDPlus, Closing Comments

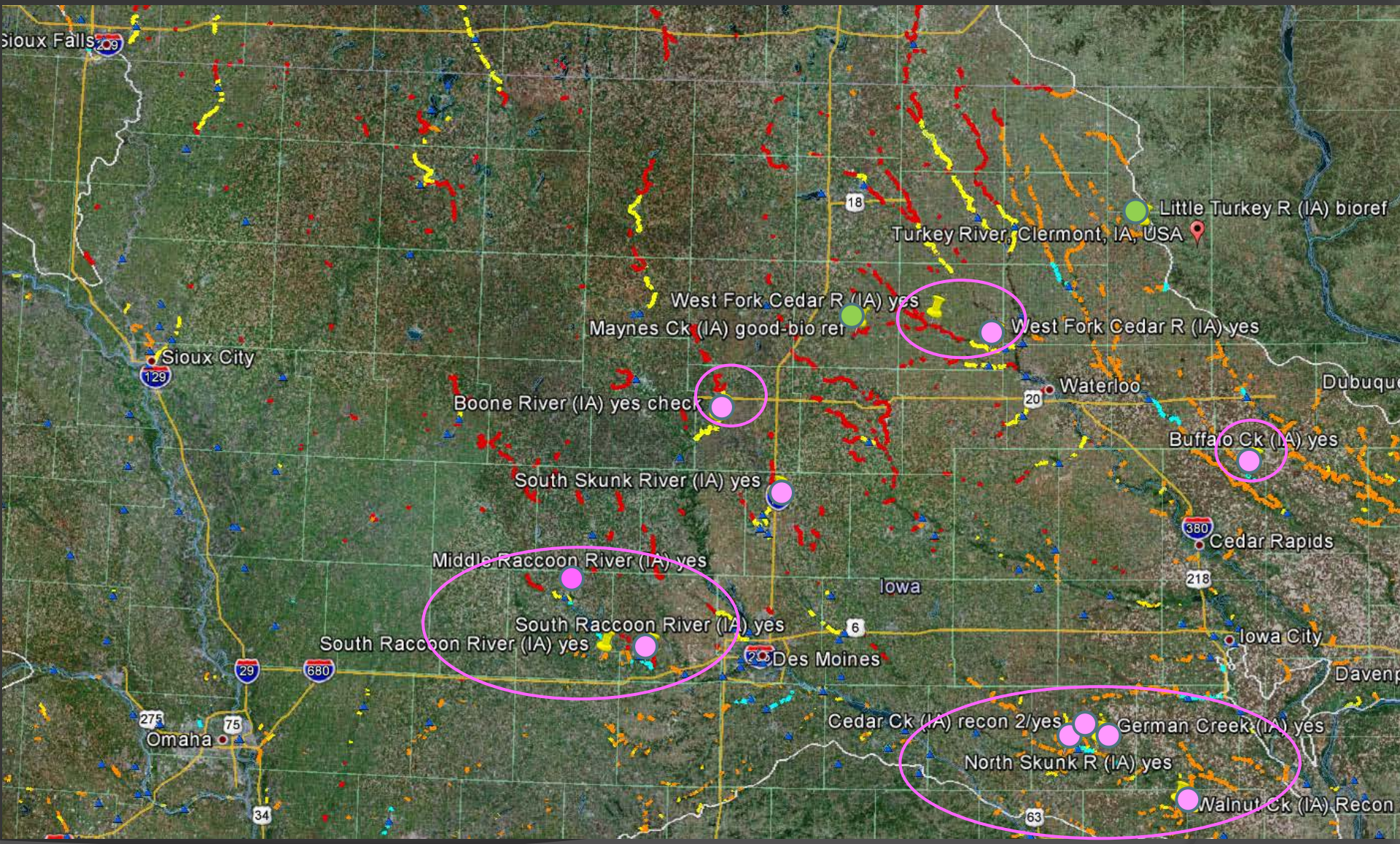
- Pro's
  - Manageable at National Scales
  - Widely accepted and available Nationally
  - Spatial improvement from previous networks
  - Enhancements to date have made a "better" dataset
  - Framework for "additional and/or updated attributes" – ALL can benefit
- Con's
  - Topography out of date (location of stream)
  - Attributes need updating (and new ones added, coordinated, maintained)
  - Crosswalk lacking from finer scales (poor scalability)
  - Tools developed in ArcGIS always need updating
  - Difficult to apportion data explicitly at monitoring locations
  - COMID's are less stable -long term
  - Taken quite a bit of effort (and $$) to get here

Thanks

Questions?

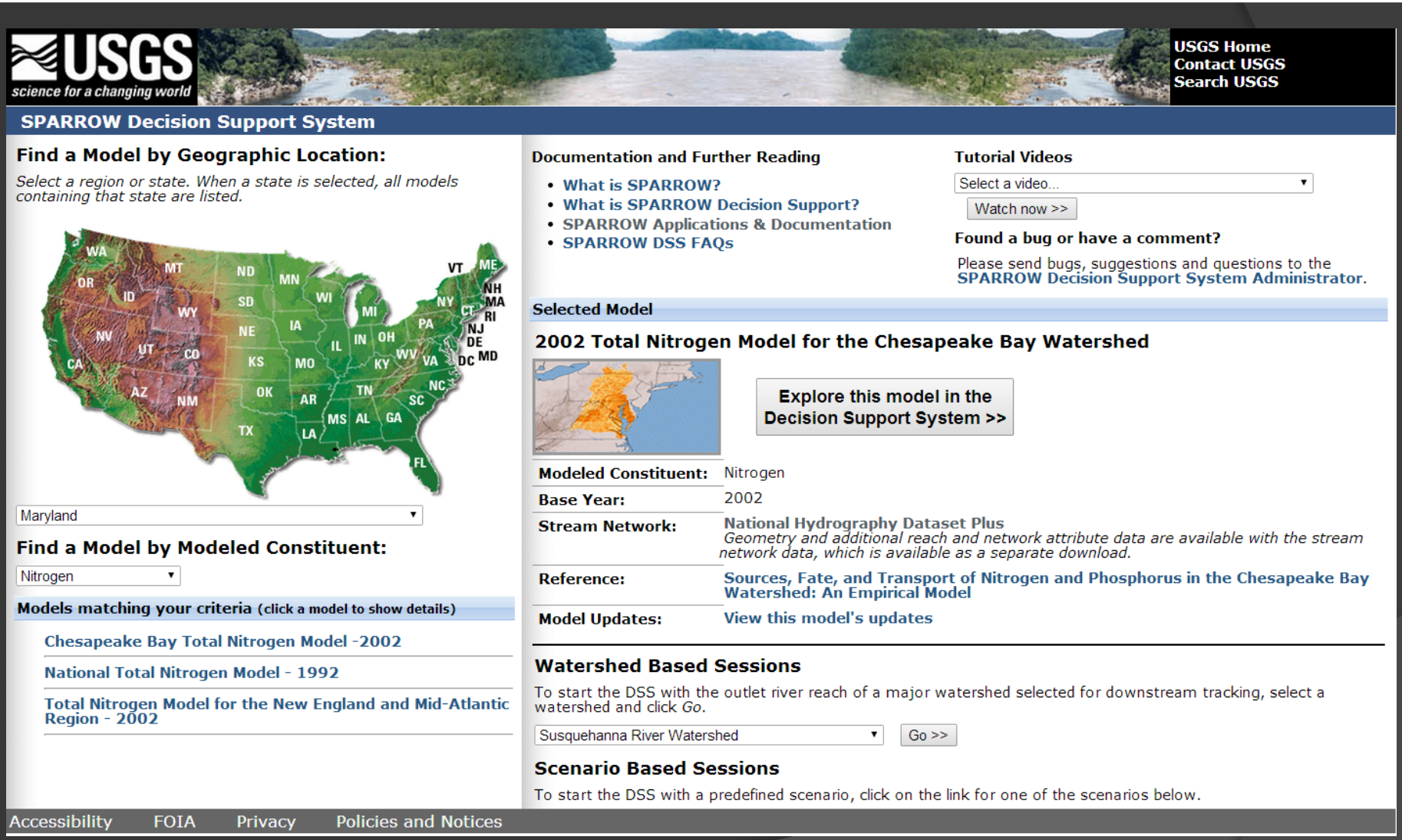
Applications, Iowa – RSQA

Mike Wieczorek and Naomi Nakagaki

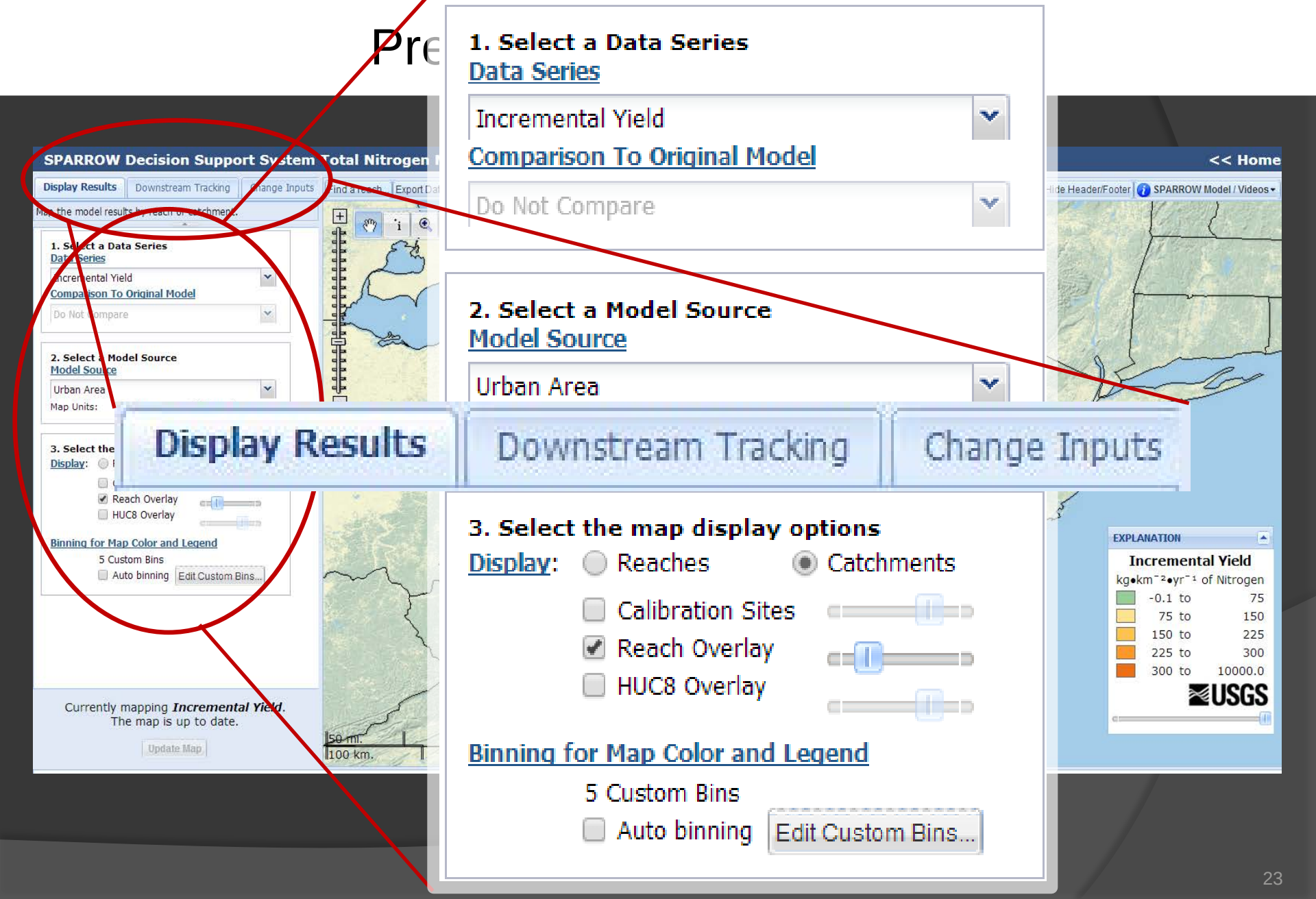


- Good-bio reference sites
- Geospatial reference site candidates

SPARROW Decision Support System

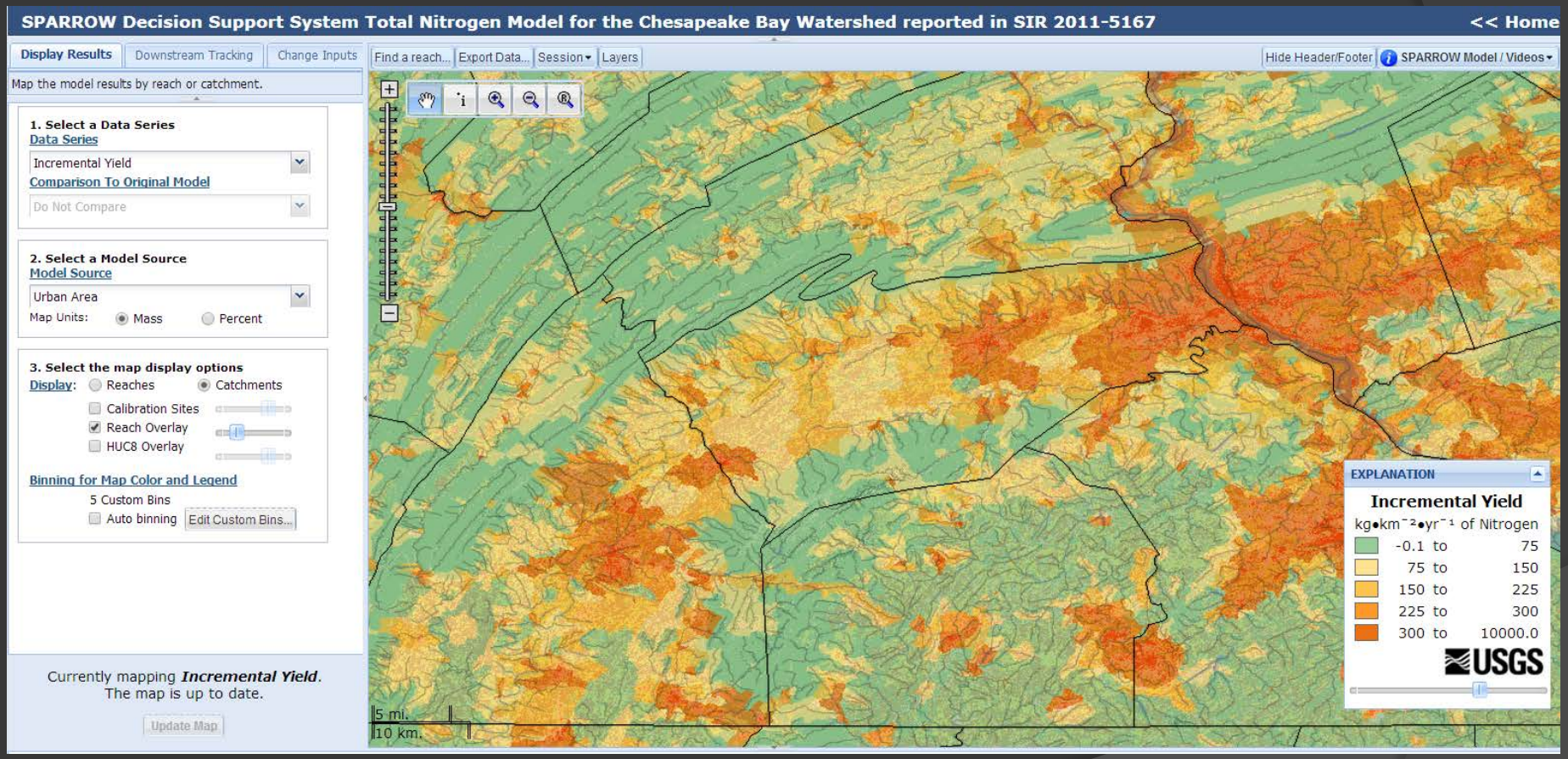


SPARROW DSS – “Display Results”



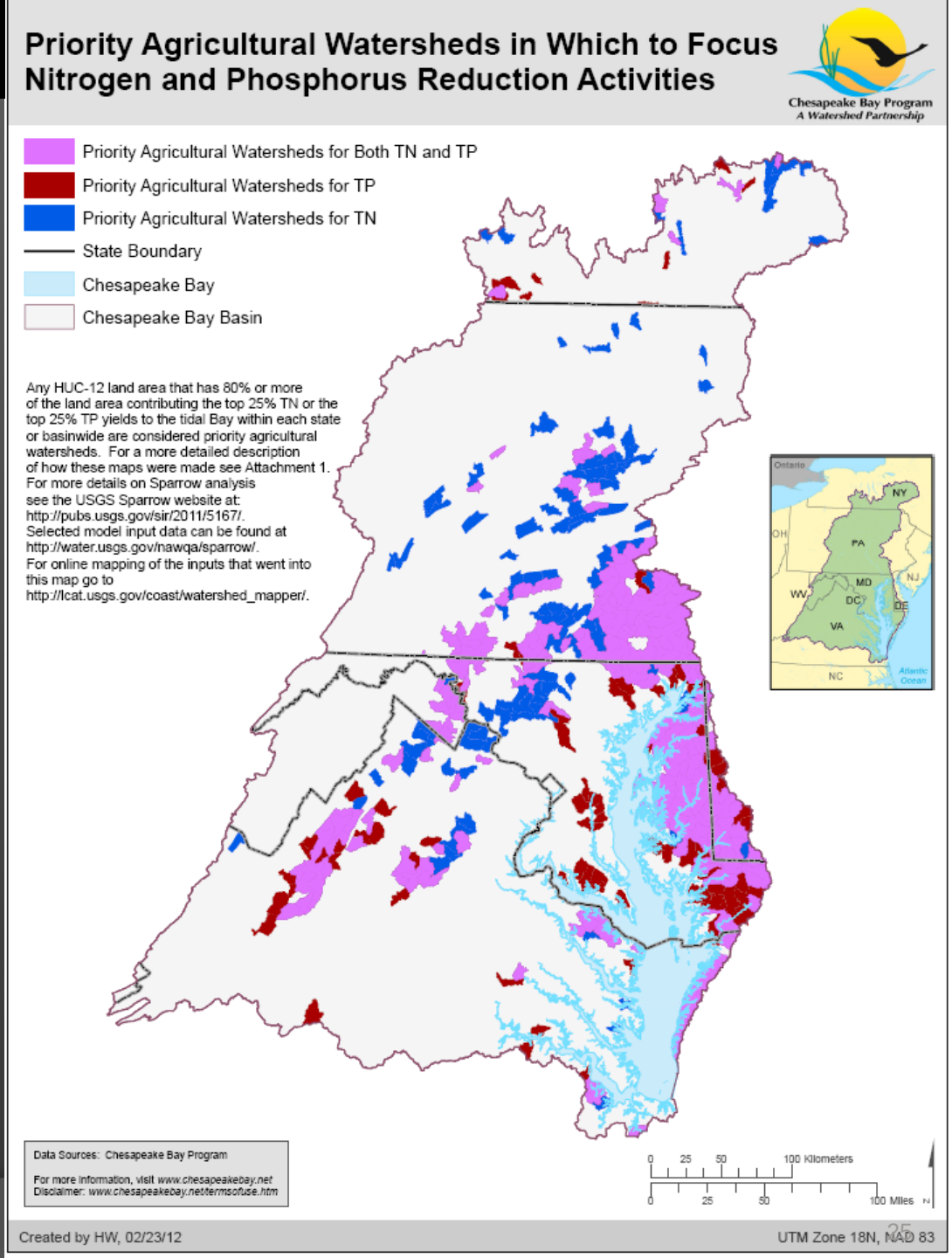
SPARROW DSS – “Display Results”

Predicted Nitrogen Yields from Urban Areas



USGS
science for a changing world

PRIORITY WATERSHEDS BASED ON SPARROW RESULTS HELPED FOCUS CBSF GRANTS TO AREAS YIELDING THE GREATEST NITROGEN OR PHOSPHORUS TO CHESAPEAKE BAY (NATIONAL FISH AND WILDLIFE FOUNDATION, 2012). MAPS SUCH AS THESE HELP GOVERNMENT AGENCIES AND NONPROFIT GROUPS FOCUS RESOURCES WHERE THEY HAVE THE GREATEST IMPACT.



SPARROW Mass-Balance Model

Nonlinear regression

$$\text{Load}_i = \left\{ \sum_{j \in J(i)} \left[\sum_{n=1}^{N} S_{n,j}\, \underbrace{\beta_n \exp(-\alpha' Z_j)} \right] \underbrace{\exp(-\delta' T_{i,j})} \right\} \underbrace{\exp(\epsilon_i)}$$

Load leaving the reach	=	Load generated within upstream reaches and transported to the reach via the stream network	+	Load originating within the reach's incremental watershed and delivered to the reach segment

Source

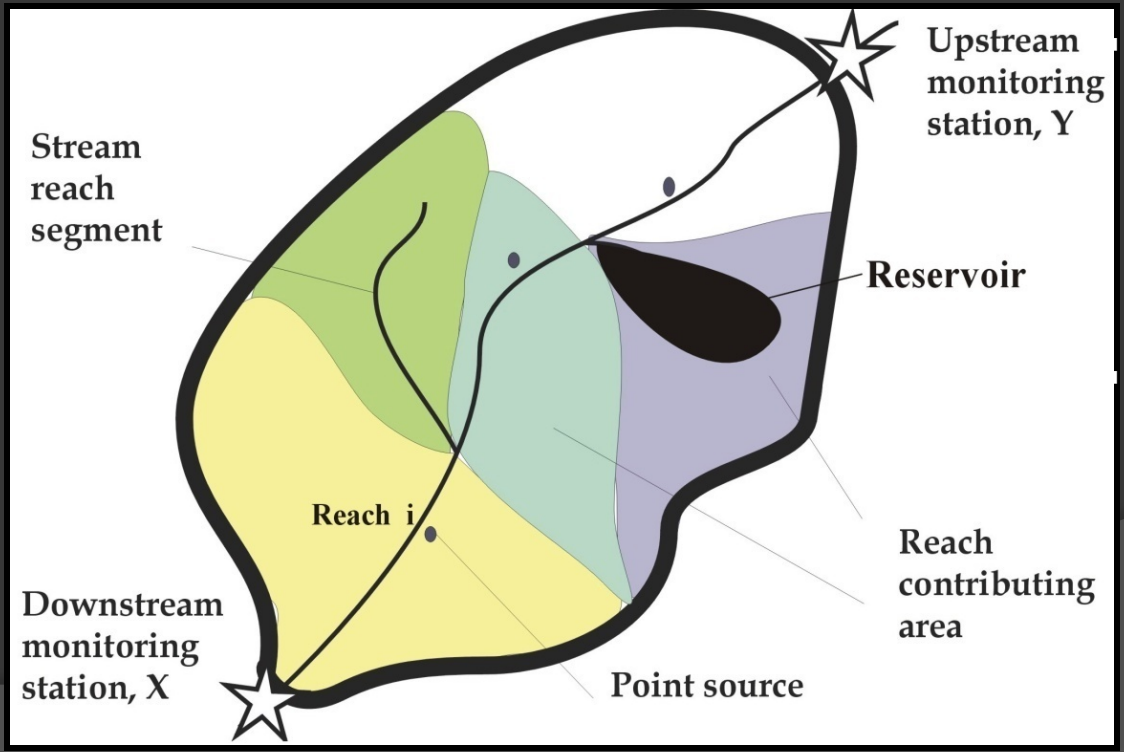
Delivery

Decay/storage

in-stream

reservoir

Monitoring



Nonlinear model structure includes topography and water routing; provides separation of land and water processes

Steady-state, mass-balance structure gives improved interpretability of the model coefficients and predictions

Schwarz et al., 2006

USGS
science for a changing world