



SIFN Data-Sprint: Update & Next Steps

Finishing Imports of Delivered
Data, Getting Next Sources in
the Queue, and Flow-
Hypothesis Exploration

Goal: What We Set Out to Do

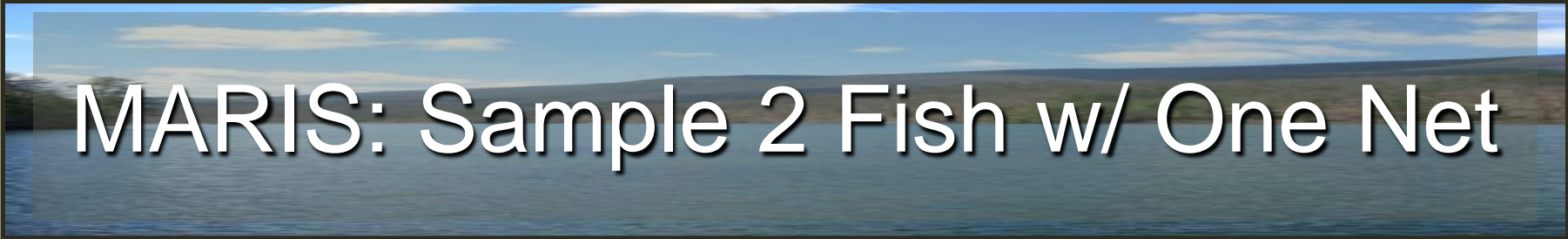
Give Scientists Access to a Broader Range of Data!!!

Specifically:

- By working together, we can:
 - ◆ Minimize workload
 - ◆ Maximize consistency
 - ◆ Get to the end point more swiftly
- Do this by:
 - ◆ Develop Data Standards
 - ◆ Provide access to Database Technical support
 - ◆ Provide access to Scientific Technical Support
 - ◆ Facilitate Data Access

We Felt that doing this required:

- Flexible Data Storage Format
 - ◆ Wide variety of data requires flexible framework
- Data Storage Physical Infrastructure
 - ◆ Someplace to house this data
- Locate and Import Data
 - ◆ Large State Databases Became Primary
 - ◆ Technical Support / Scientific Support
 - ◆ Develop "portal" to other large, online data sources



MARIS: Sample 2 Fish w/ One Net

MARIS Already Had/Was Developing:

- Flexible Data Storage Format
 - ◆ MARIS data framework built to accommodate varied data sources / spatial resolutions
- Data Storage Physical Infrastructure
 - ◆ MARIS transitioning to USGS Infrastructure
- Online query interface to access all data

So that leaves us with Sprinting...

- Locate and Import Data
 - ◆ Large State Databases Became Primary
 - ◆ Technical Support / Scientific Support



Networks and Partnerships

- **Using Networks – they work.**
 - ◆ Networks leverage the abilities of partners with **common** interests, **diverse** assets
 - ◆ Networks – sharing access to network members – the crucial piece in locating **diverse** assets & **common** interests
- **SARP/SIFN** –access to the network of PEOPLE, organized events, constant contact, nuts & bolts, funding
- **TNC** – Structure, motivation, funding, eyes on the prize (x – y plots), connected with MARIS
- **MARIS** – infrastructure, long-term commitment to housing the data, some funding to assist states
- **States** – provided data, staff time, enthusiasm



Why Do We Need This?

Isn't it good enough if a single state gets their data organized for their OWN internal use?

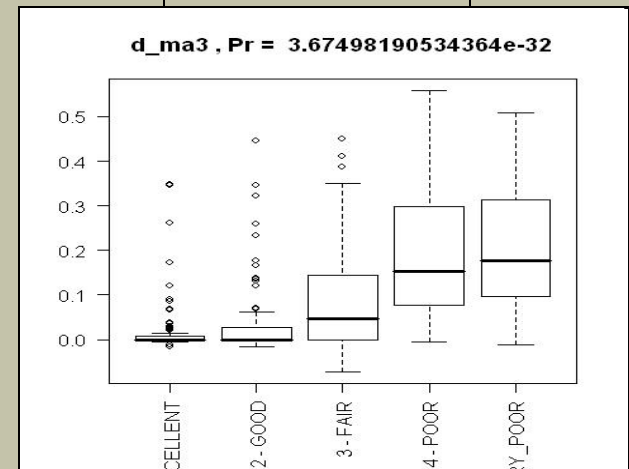
- **More is More** – common data frameworks provide everyone with more data to explore, greater chance for robust hypotheses
- **Hydrologic Connectivity** – Aquatic organisms don't always honor political boundaries

Panning for Gold: More Is Less

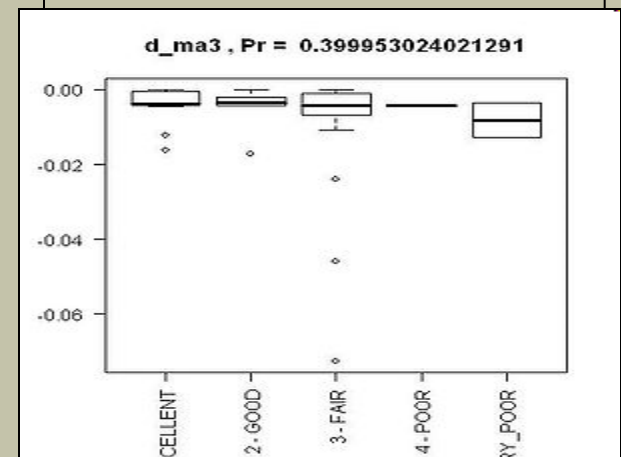
Flow-ecology relationship project in Potomac River basin

- Benthic health scores (IBI)
- Developed baseline and current flows model for 20+ year period to assess flow alteration
- 500 sub-watersheds ~ 90% have INCREASING FLASHINESS (VI+)
- $\Delta VI+$ Data is VERY NOISY (the urban signal over-rides everything)
- $\Delta VI-$ Data shows **opposite trend** for ΔVI versus IBI as compared to VI+ sites

VI – All Sites

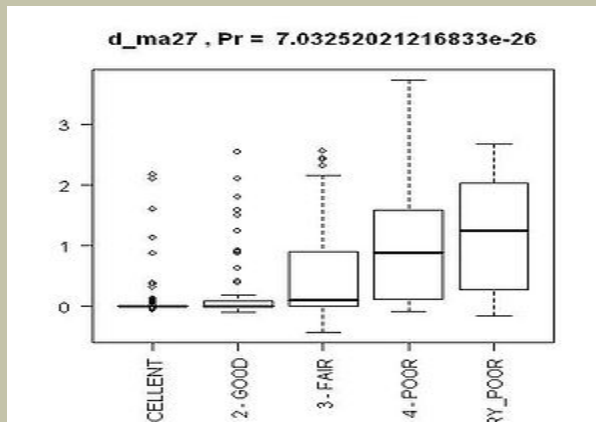


VI – Sites w/decreasing VI

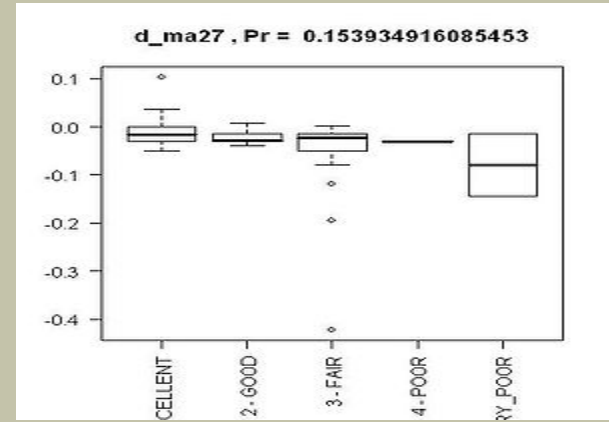


More is Less – April/May VI

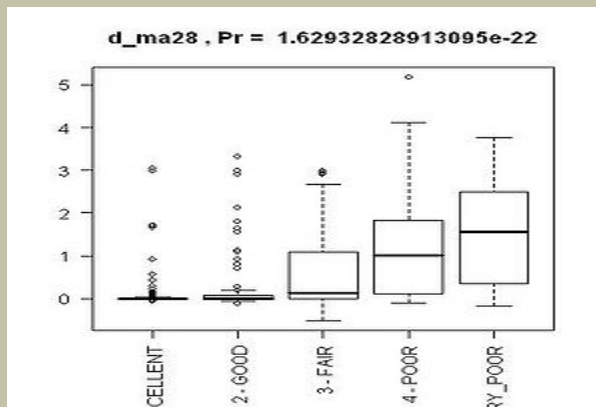
VI April – All Sites



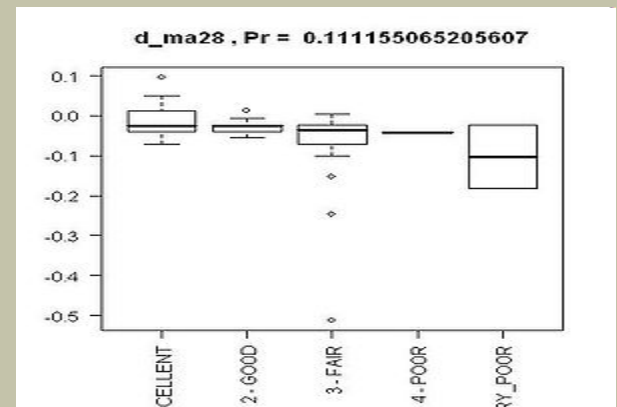
VI April – Sites w/decreasing VI



VI May – All Sites



VI May – Sites w/decreasing VI



Why: Fish Cross Boundaries

Clinch River Basin

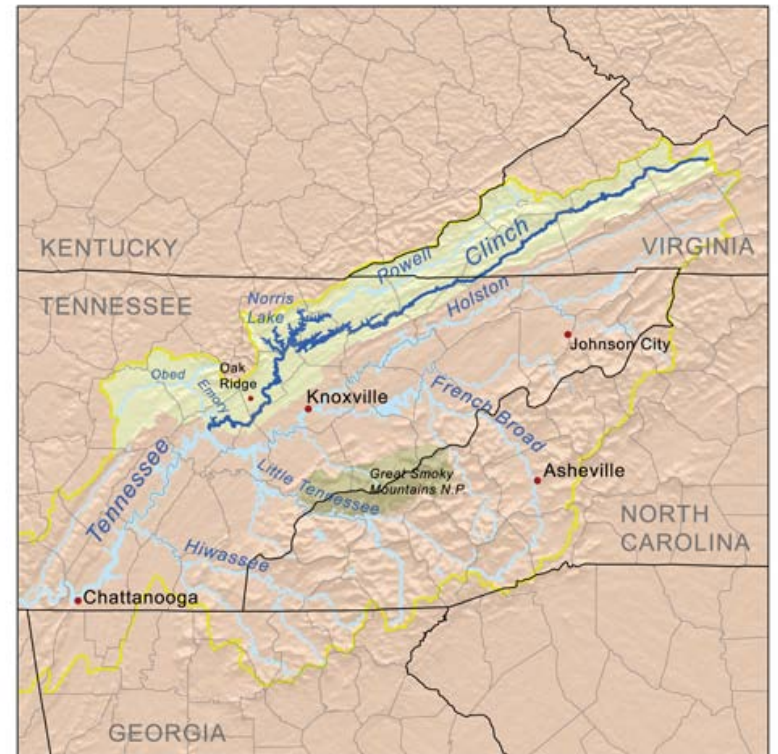
- Spans 4 states: Georgia, North Carolina, Tennessee, and Virginia
- Bio-diversity hot-spot

Potomac River

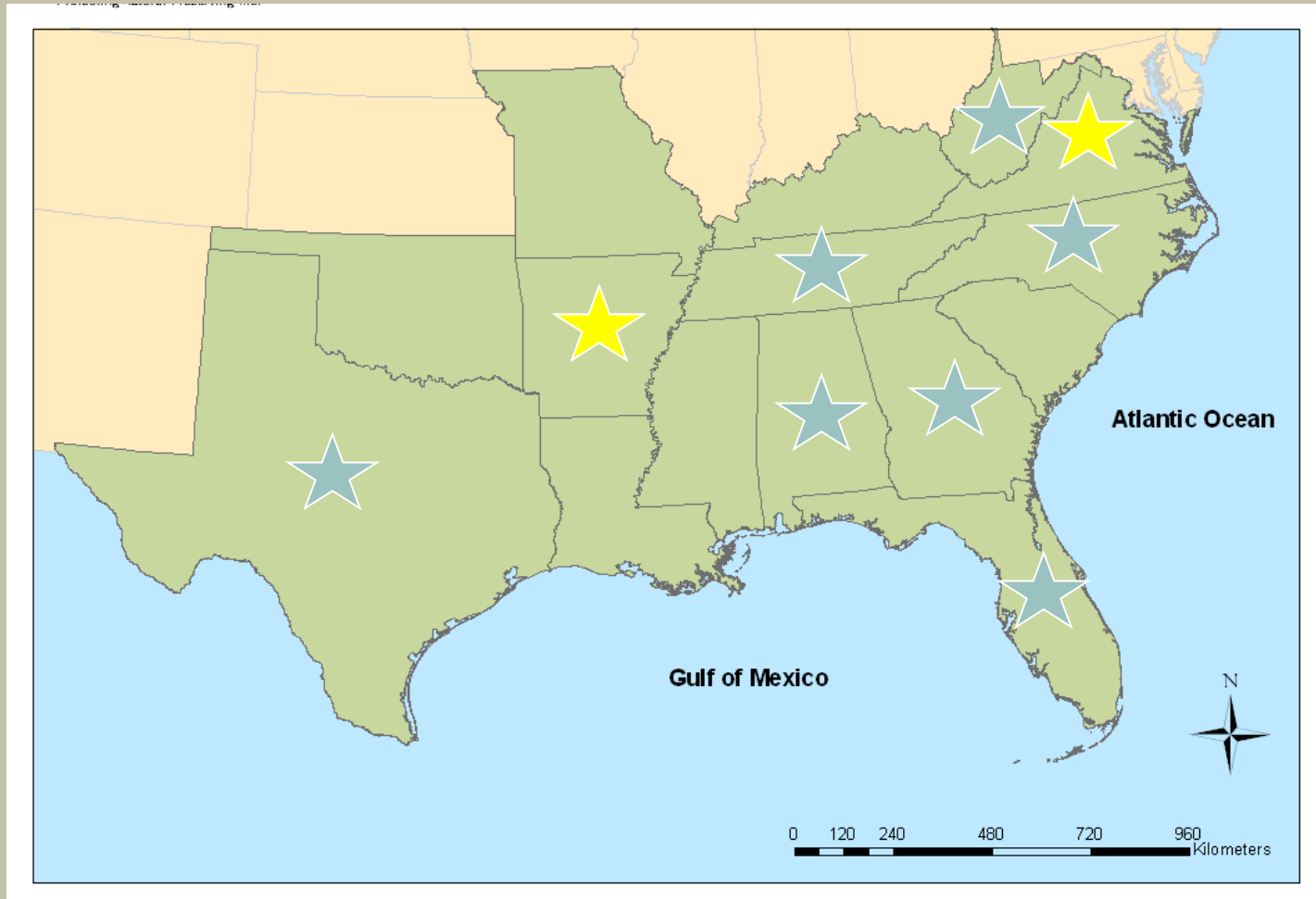
- MD, VA, WV

Roanoke River

- VA, NC



States Who Have Committed





At the Starting Line

The data-sprint experience thus far:

- A sprint can seem like a marathon

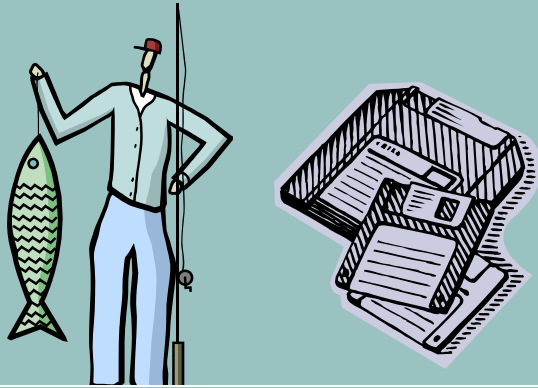
Activities:

- Preparation to import data
 - ◆ Used WebEx to explore data tables, GIS layers together
 - ◆ Documenting table conversion columns

Takeaways:

- ◆ Having means to update previous imports primary (species elevation, error-correction)
- ◆ Data-holders need repeatable queries for future exports
- ◆ Need a means of tracking their imports (in order to avoid redundant imports, and make corrections easy)

Example: VA Fish Data



Data owners/gatherers



Tech Support

Year	Loc	RKM	Spp	Catch	Effort (h)	CPUE (n/h)
2009	Howardsville	322	AME	21	1.00	20.832
2009	Howardsville	322	BLC	0	1.00	0
2009	Howardsville	322	BLG	4	1.00	3.968

MARIS Field Name	Data Value	Data Value
STATE	VA	VA
ORIGINATOR	DGIF	DGIF
WATER_ID		
STATION_ID	Howardsville	Howardsville
SAMPLE_BEGIN_DATE	1/1/2009	1/1/2009
SAMPLE_END_DATE	12/31/2009	12/31/2009
TARGET	ALL	ALL
TARGET_STD	ALL	ALL
STATE_SPECIES	AME	BLC
EFFORT_TIME	1.0	1.0
TOTAL_CATCH	21	21
TOTAL_WEIGHT	NULL	NULL
CPUE_TIME	NULL	NULL



Going Forward

Going Forward:

- Finishing this round of data imports
- Technical assistance for next round of data
 - ◆ Get states to use the TA already in place
 - ◆ Get more TA volunteers
- Using MARIS:
 - ◆ Data extraction
 - ◆ Flow hypothesis formulation
- Flow-hypothesis workshops?
 - ◆ WebEx, work sessions to begin using the data we have
 - ◆ Using the Wiki to guide and document our process
- Size and age table

Size and Age Table: SIFN Recommended

SIFN Proposed Field Name	Caption	SIFN Proposed Data Type	SIFN Proposed Description
sample_type	Type of Sample Event	Text 2	Description of the record type, Q - Quantitative, i.e., population fields are meaningful, or O - Occurrence only.
originator	Data Originator	Text 8	SAME AS MARIS CODE FOR DATA
size_min	Minimum Size	Number (double)	Minimum size specimen of population sampled
size_max	Maximum Size	Number (double)	Maximum size specimen of population sampled
size_mean	MeanSize	Number (double)	Mean size of population sampled
size_median	Median Size	Number (double)	Population Median size
age_min	Minimum Size	Number (double)	Minimum age specimen of population sampled
age_max	Maximum Size	Number (double)	Maximum age specimen of population sampled
age_mean	Mean Size	Number (double)	Mean age of population sampled
age_median	Median Size	Number (double)	Population Median age



Next Steps

Meet with Data-Sprint Technical Advisors while you are at this conference to discuss how we can help you to move forward

- Room 524 – Data Sprint Headquarters

Contact info:

Robert Burgholzer

robert.burgholzer@deq.virginia.gov

805-698-4405

<http://sifn.bse.vt.edu/sifnwiki/>