An approach to monitoring cyanobacteria blooms at surface drinking water intakes using satellite imagery

USIALE 2016 - Asheville

John Clark (ORISE@USEPA)
Blake A Schaeffer (USEPA)
John A Darling (USEPA)
Erin A Urquhart (USEPA)

Image: MODIS Aqua - 10/9/2011 - NASA

Cyanobacteria Blooms

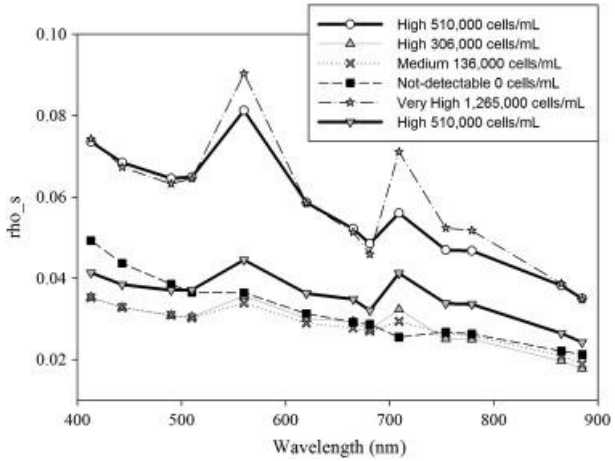
- Photosynthetic freshwater bacteria
- Impacts
  - Human health
  - Economic costs
- Remote sensing
  - Uniform and systematic approach for identifying cyanobacteria blooms
  - Support ground monitoring efforts

Image: MODIS Aqua - 10/9/2011 - NASA

Remote Sensing of HABs

Lunetta et al, 2015:
Cyanobacteria concentration estimated from MERIS imagery

- Cyanobacteria Index (CI-multi) = multiple spectral shape algorithm
- *In situ* validation
- ~3 day return interval
- 300-m spatial resolution



R.S. Lunetta et al. / Remote Sensing of Environment (2015)

CYAN

USGS science for a changing world

Image: MODIS Aqua - 10/9/2011 - NASA

- MERIS processed for three regions for Jan 2008 through Dec 2011
  - Florida: 843 scenes
  - New England: 1155
  - Ohio: 1024
- CONUS coverage ~ summer 2016
- Transition to OLCI in near future

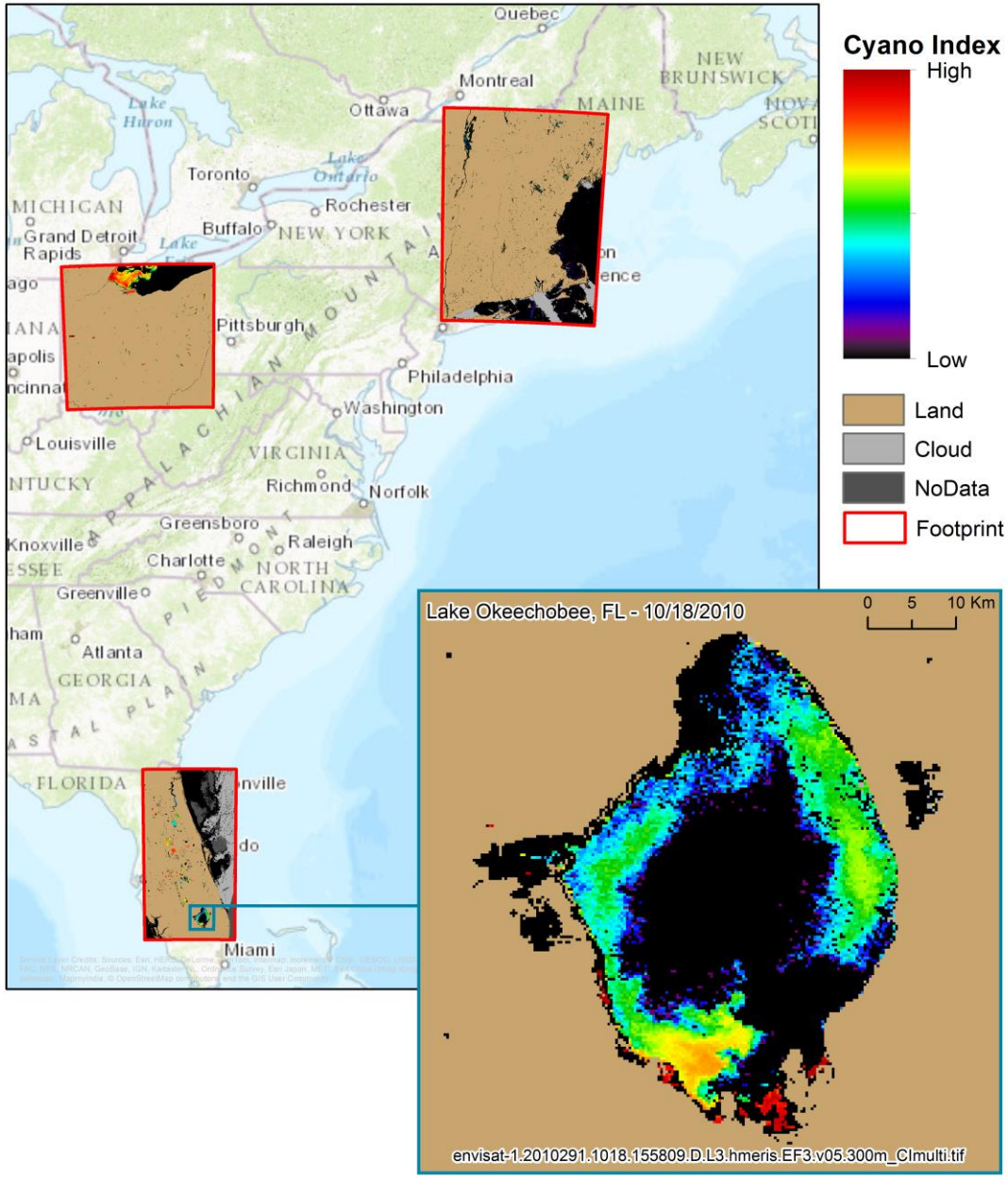


Image: MODIS Aqua - 10/9/2011 - NASA

Objectives

- National scale
  - Estimate spatially resolvable features
- Regional scale
  - Calculate temporal coverage
  - Calculate bloom frequency
- Local scale
  - Relate features to nearby observations
  - Extract time series
  - Summarize across space and time

Image: MODIS Aqua - 10/9/2011 - NASA

Waterbodies

- NHD waterbody features
- Subset following EPA 2012 National Lakes Assessment (NLA) Site Evaluation Guidelines
  - Exclude ephemeral or brackish
  - Exclude surface area less than 1 ha

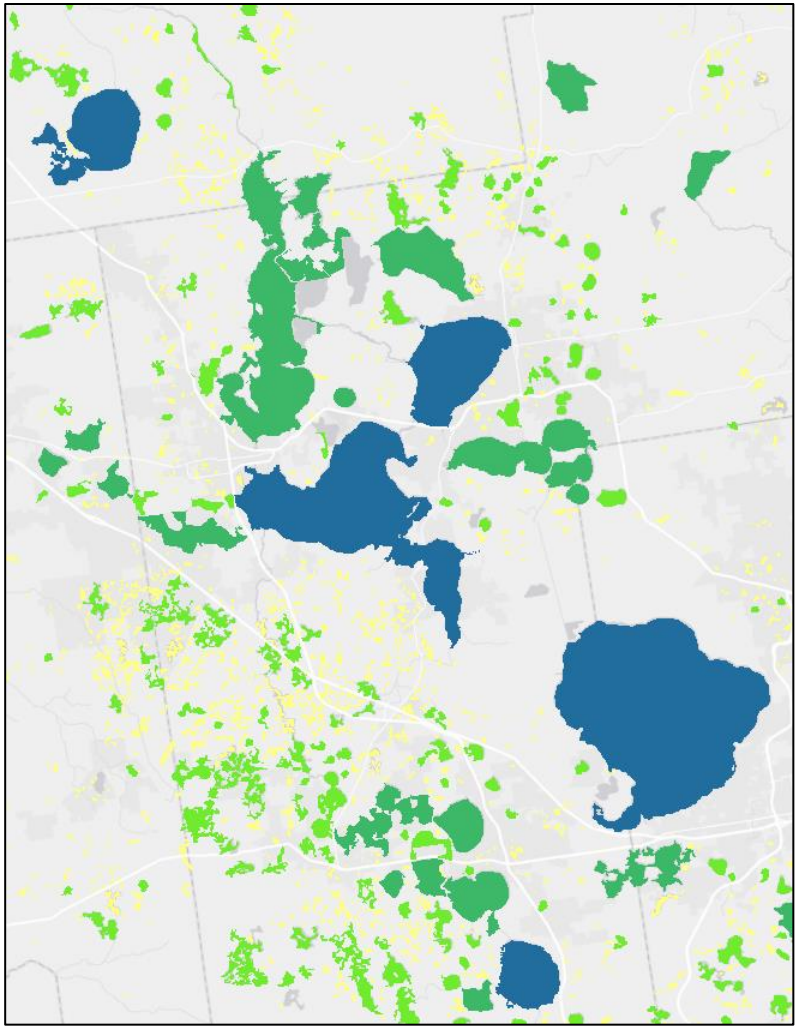


Image: MODIS Aqua - 10/9/2011 - NASA

Public Water Systems

- USEPA Office of Ground Water & Drinking Water (OGWDW) Public Water Systems
  - Subset to PWS surface intakes < 100 m from NHD waterbody
- Sensitive information, restricted access to data
  - NO LOCATIONS REPORTED/PRESENTED

Image: MODIS Aqua - 10/9/2011 - NASA

Estimating Resolvable Waterbodies

- Target spatial resolutions reflect potential remote sensing products
  - Focal widths: 30 m, 90 m, 300 m, 900 m
- Max focal width ∝ max distance to shore

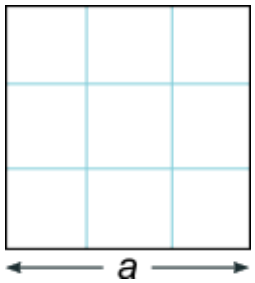
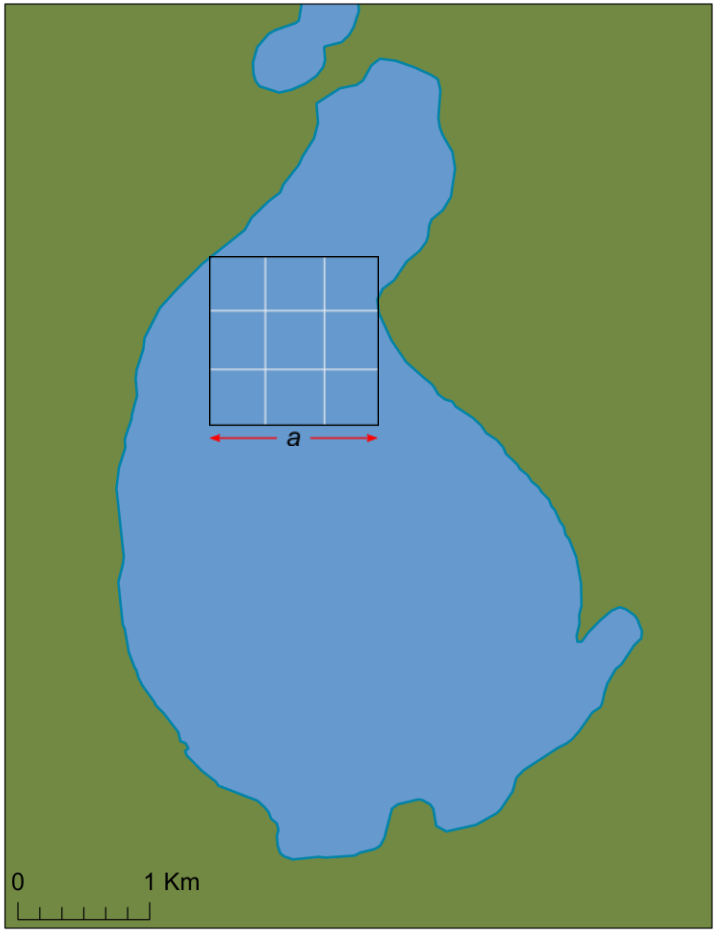


Image: MODIS Aqua - 10/9/2011 - NASA

Estimating Resolvable Waterbodies



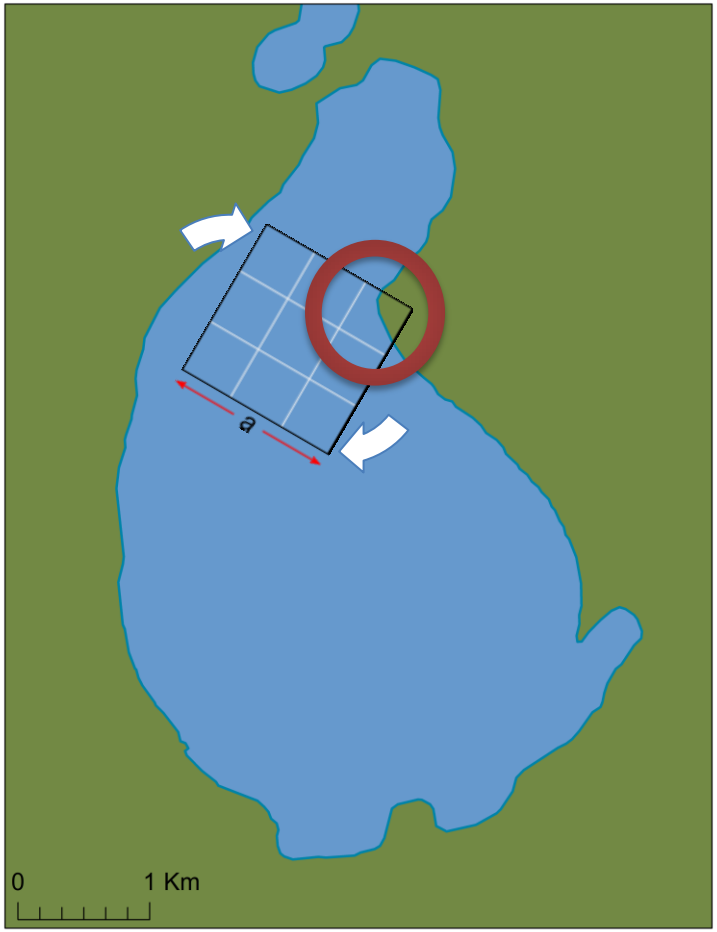
a = Window Width

What is the minimum distance to shore (*R*) that will accommodate a focal window of width *a*?

$$R = \frac{a}{2}?$$

Image: MODIS Aqua - 10/9/2011 - NASA

Estimating Resolvable Waterbodies



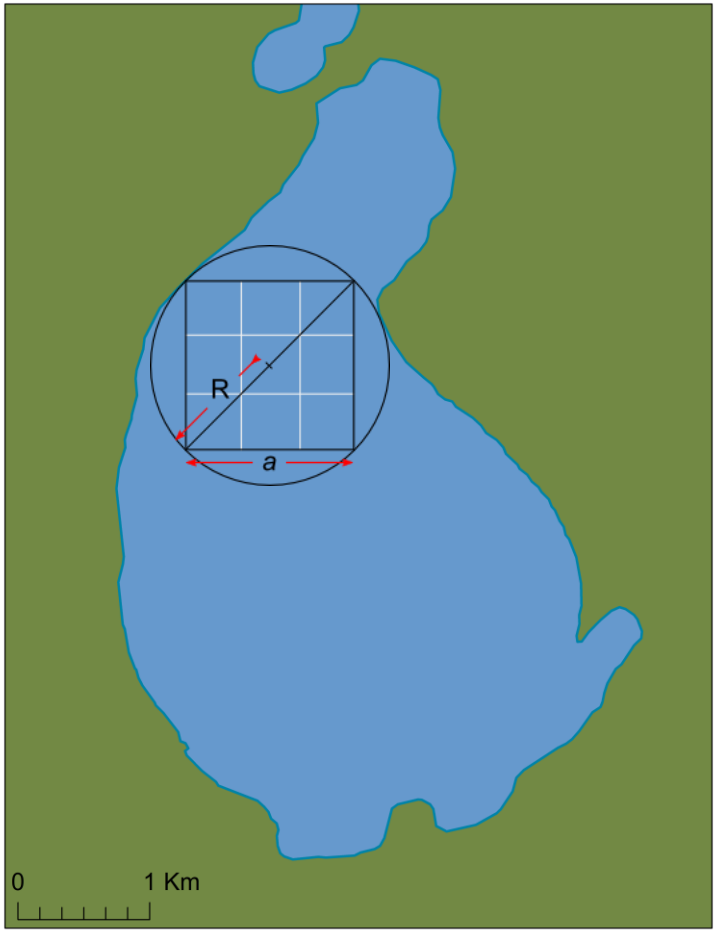
a = Window Width

What is the minimum distance to shore (*R*) that will accommodate a focal window of width *a*?

$$R = \frac{a}{2}?$$

Image: MODIS Aqua - 10/9/2011 - NASA

Estimating Resolvable Waterbodies

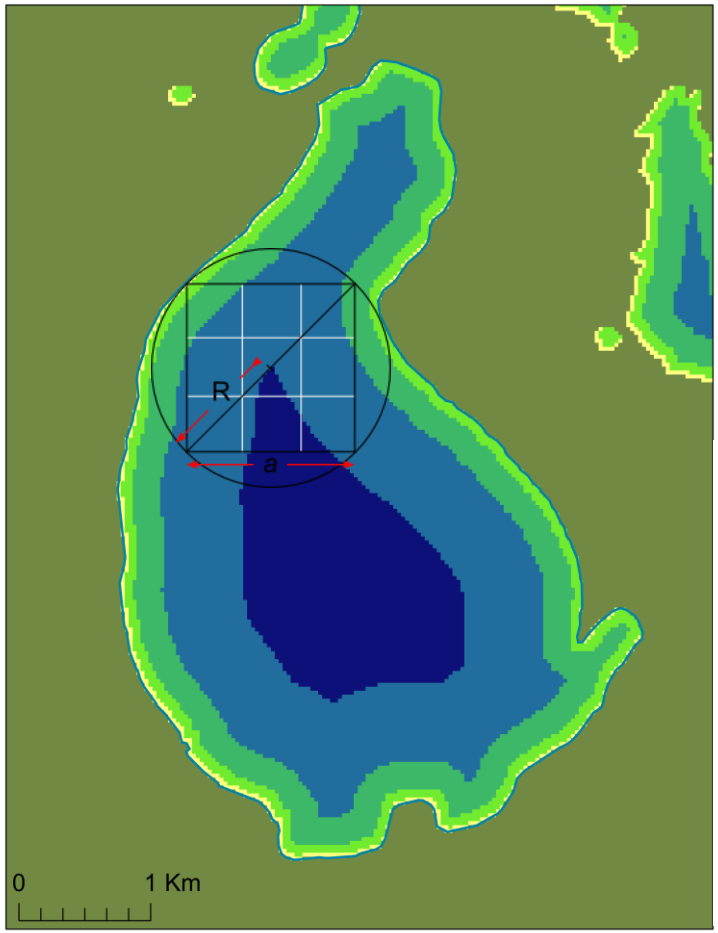


What is the minimum distance to shore (*R*) that will accommodate a focal window of width *a*?

$$R = a\frac{\sqrt{2}}{2}$$

Image: MODIS Aqua - 10/9/2011 - NASA

Estimating Resolvable Waterbodies



What is the minimum distance to shore (*R*) that will accommodate a focal window of width *a*?

$$R = a\frac{\sqrt{2}}{2}$$

$$R(900) = 636\,m$$

Max Window Width

- < 30 m
- 30 m
- 90 m
- 300 m
- >= 900 m

Image: MODIS Aqua - 10/9/2011 - NASA

Estimated Resolvable Waterbodies

	Window Size							
	30x30 m (Landsat)		90x90 m (3x3)		300x300 m (MERIS/OLCI)		900x900 m (3x3)	
Feature	n	%	n	%	n	%	n	%
Waterbody	275897	100	170240	61.7	15545	5.6	1862	0.7
PWS	1991	100	1849	92.9	860	43.2	300	15.1

Image: MODIS Aqua - 10/9/2011 - NASA

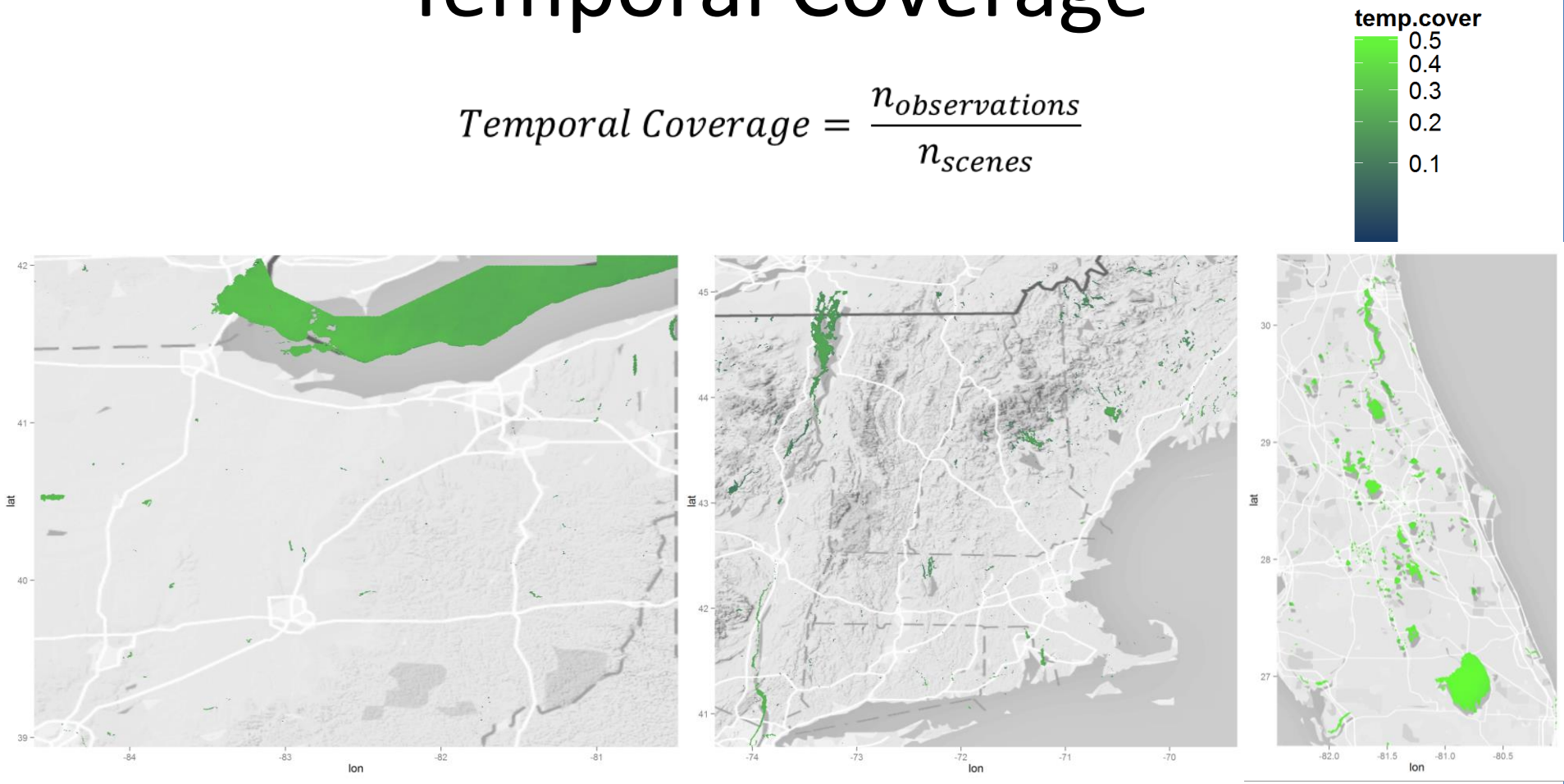
Temporal Coverage

$$Temporal\ Coverage = \frac{n_{observations}}{n_{scenes}}$$

Image: MODIS Aqua - 10/9/2011 - NASA

Temporal Coverage

$$Temporal\ Coverage = \frac{n_{observations}}{n_{scenes}}$$



Bloom Frequency

$$Bloom\ Frequency = \frac{n_{observations > threshold}}{n_{observations}}$$

Image: MODIS Aqua - 10/9/2011 - NASA

Bloom Frequency

$$Bloom\ Frequency = \frac{n_{observations > threshold}}{n_{observations}}$$

Table 1. World Health Organization guidancevalues for the relative probability of acute health effects during recreational exposure to cyanobacteria and microcystins, based on information presented in Chorus and Bartram 1999.

Relative Probability of Acute Health Effects	Cyanobacteria[1] (cells/mL)	Microcystin-LR[2] (μg/L)	Chlorophyll-*a*[3] (μg/L)
Low	< 20,000	< 10	< 10
Moderate	20,000-100,000	10-20	10-50
High	100,000-10,000,000	20-2,000	50-5,000
Very High	>10,000,000	>2,000	>5,000

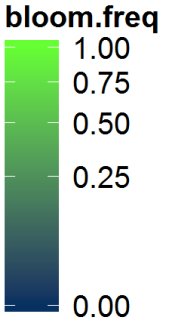
[1] The WHO guidelines were developed for *Microcystis* dominated samples with an assumed toxin content of 0.2 picograms of microcystin per *Microcystis* cell or 0.4 micrograms of microcystin per microgram of chlorophyll-a with a minimum criteria of at least cyanobacterial dominance.

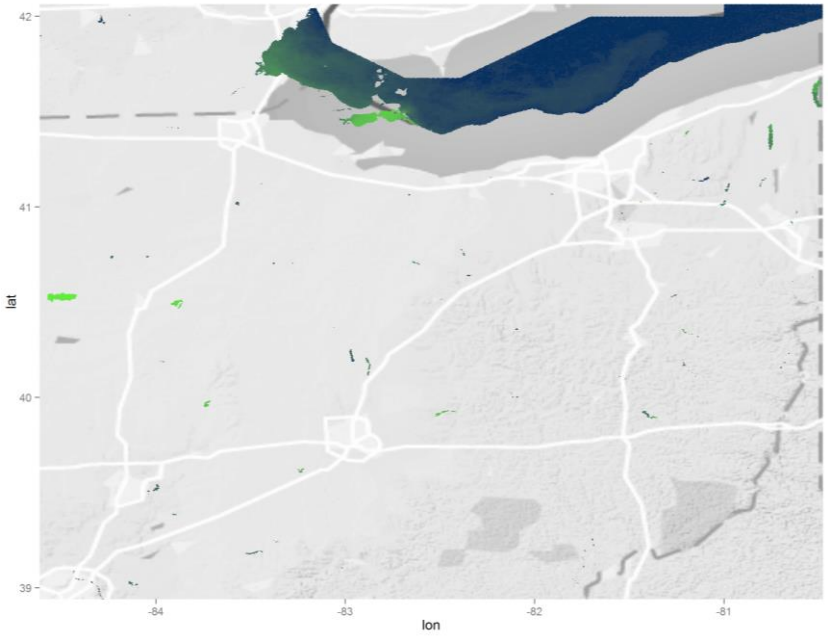
Graham et al., Lake Line, 2009

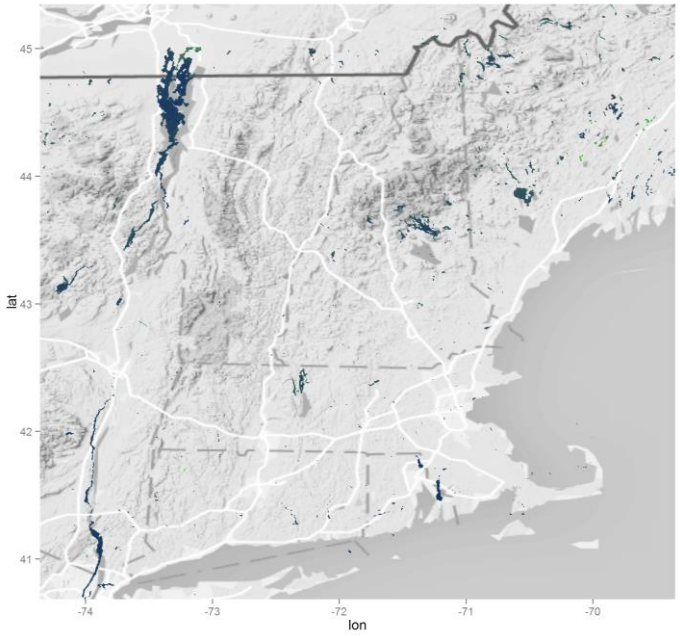
Image: MODIS Aqua - 10/9/2011 - NASA

Bloom Frequency

$$Bloom\ Frequency = \frac{n_{observations > 100{,}000\ cells/mL}}{n_{observations}}$$







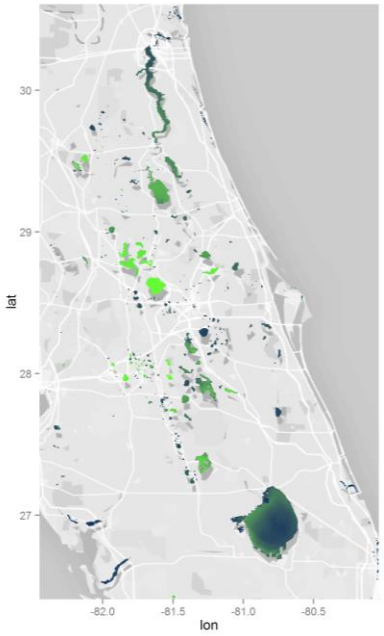


Image: MODIS Aqua - 10/9/2011 - NASA

Local Scale: Reporting

- Spatially relate intakes to nearby CI-multi observations
- Extract time-series data
- Aggregate observations across space and time
- Summarize and report

Preliminary data exploration

Focal Filter

- Target contiguous data
- Reduce mixed pixel effects (land/water)

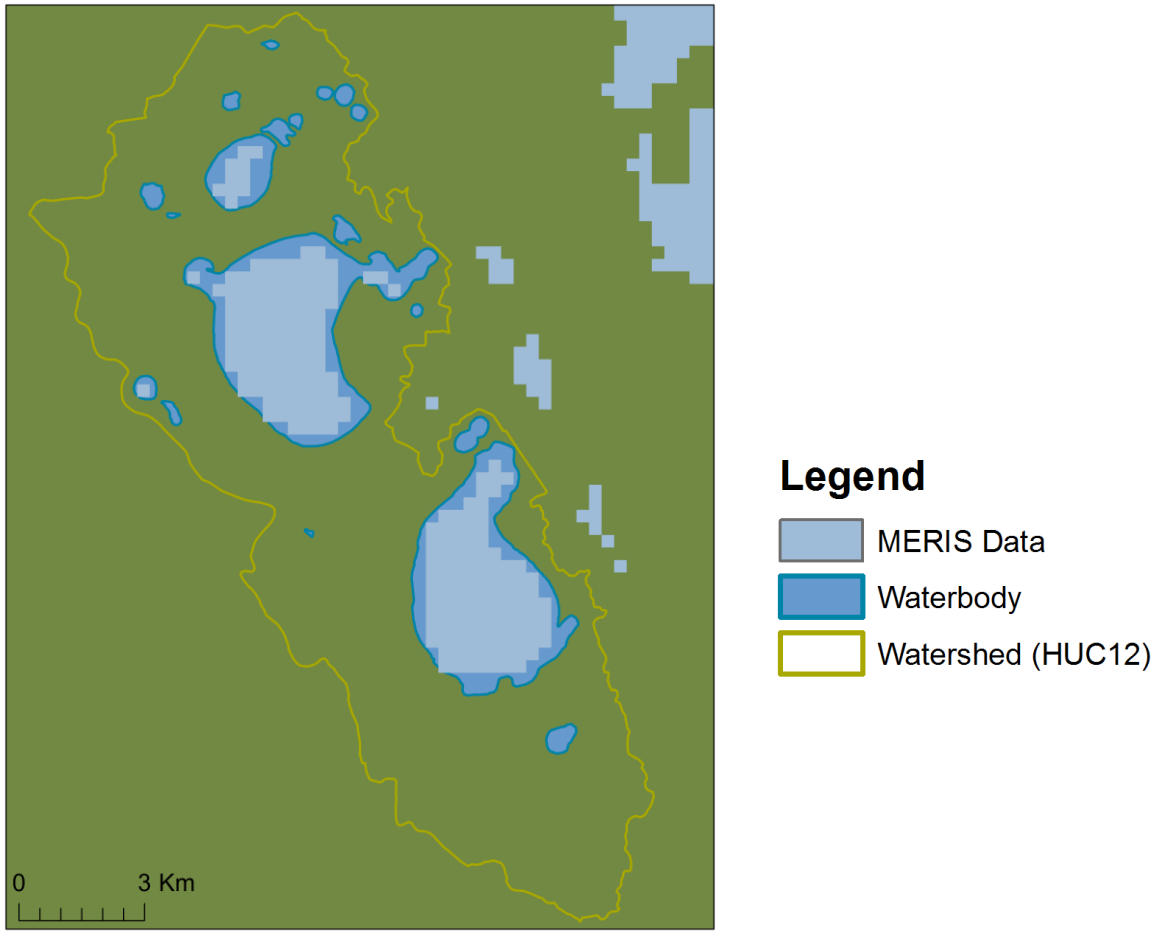
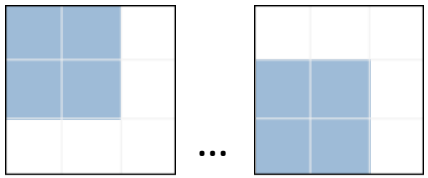


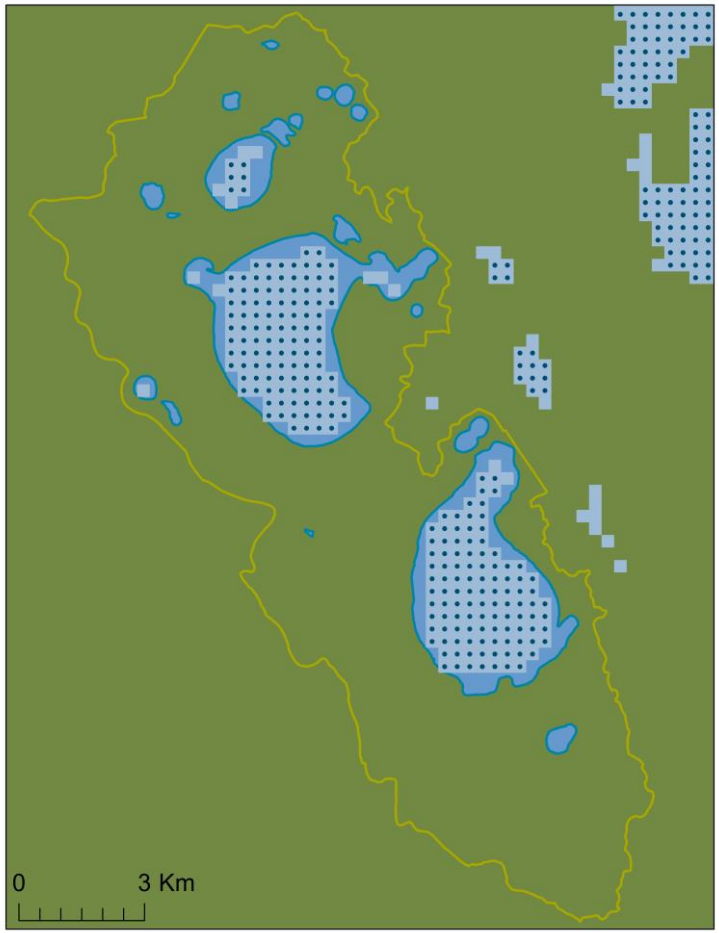
Image: MODIS Aqua - 10/9/2011 - NASA

Focal Filter

- Target contiguous data
- Reduce mixed pixel effects (land/water)

- 4-pass 2x2 focal filter removes isolated pixels





Legend

- MERIS Data
- Waterbody
- Watershed (HUC12)
- Candidate Pixel

Image: MODIS Aqua - 10/9/2011 - NASA

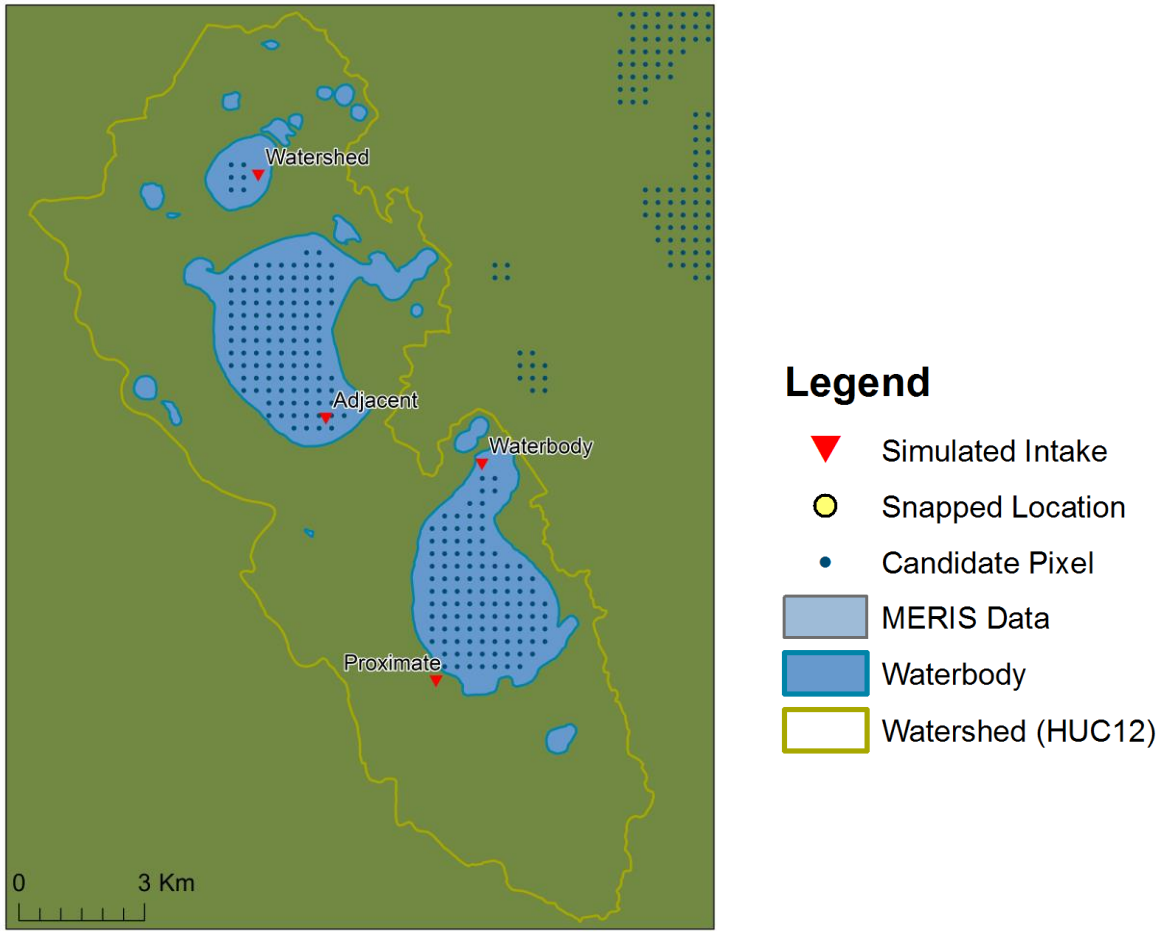
Location Snapping

- Reported PWS locations may not fall within waterbody
  - Likely to exhibit characteristics and bloom events similar to open water pixels in close proximity
- Maximize coverage while accounting for increasing uncertainty

Image: MODIS Aqua - 10/9/2011 - NASA

Location Snapping

1. ‘Adjacent’ case: Nearest 3x3 pixel array within 300 m of reported PWS location
2. ‘Proximate’ case: Nearest 3x3 pixel array within 900 m of reported PWS location
3. ‘Waterbody’ case: All pixels of nearest waterbody within 900 m of the reported PWS location containing >= 9 pixels
4. ‘Watershed’ case: All pixels within the same HUC12 watershed as the reported PWS location if watershed contains >= 9

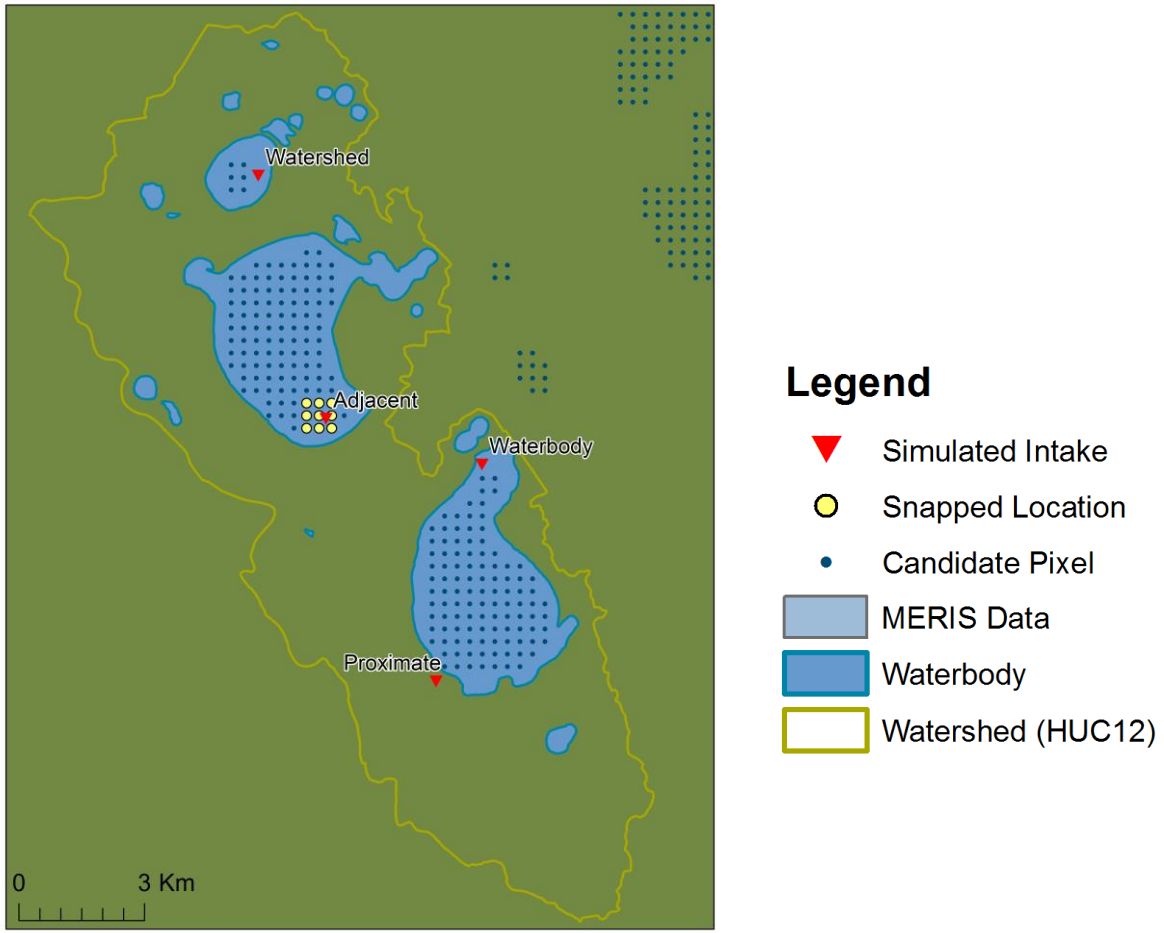


Simulated data, does not represent true intake locations.

Image: MODIS Aqua - 10/9/2011 - NASA

Location Snapping

1. **'Adjacent' case: Nearest 3x3 pixel array within 300 m of reported PWS location**
2. 'Proximate' case: Nearest 3x3 pixel array within 900 m of reported PWS location
3. 'Waterbody' case: All pixels of nearest waterbody within 900 m of the reported PWS location containing >= 9 pixels
4. 'Watershed' case: All pixels within the same HUC12 watershed as the reported PWS location if watershed contains >= 9



Simulated data, does not represent true intake locations.

Image: MODIS Aqua - 10/9/2011 - NASA

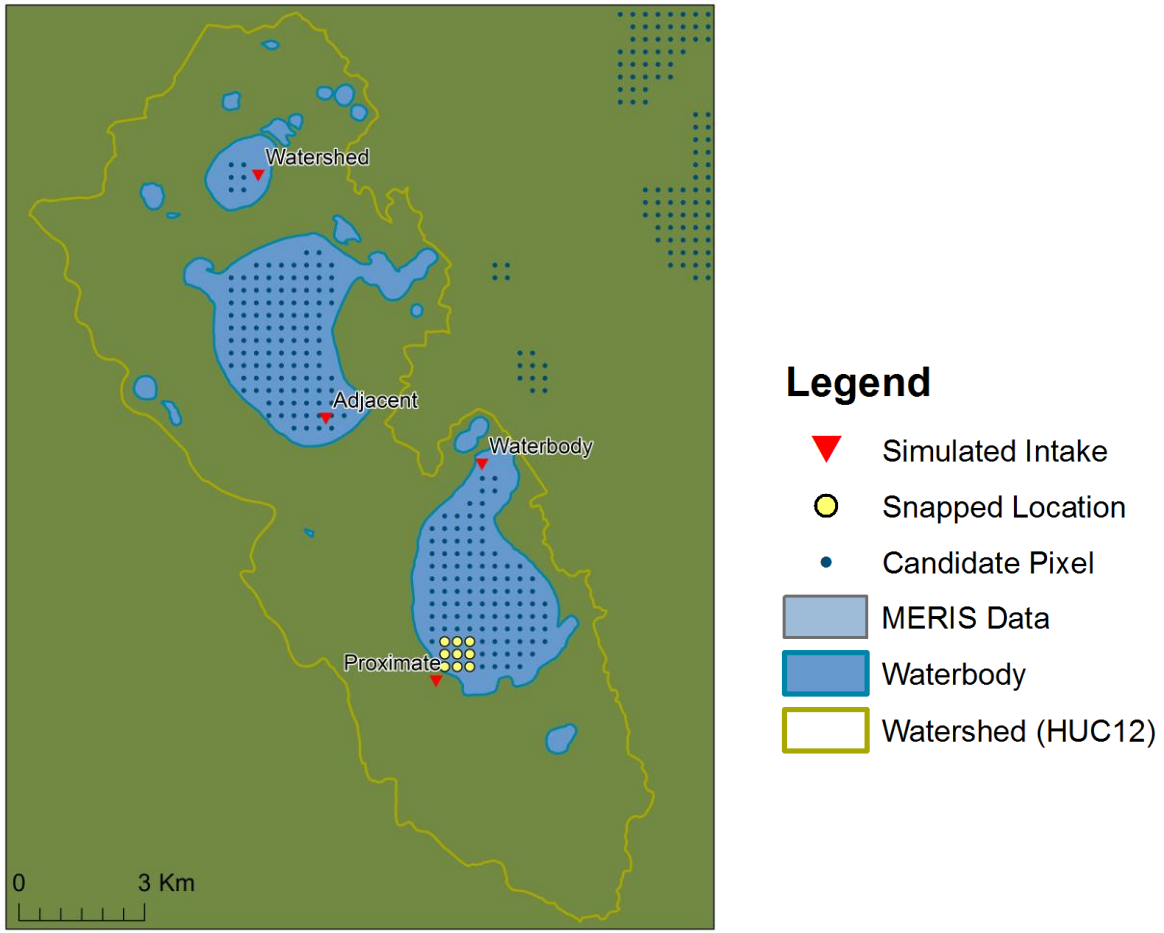
Location Snapping

1. 'Adjacent' case: Nearest 3x3 pixel array within 300 m of reported PWS location

2. **'Proximate' case: Nearest 3x3 pixel array within 900 m of reported PWS location**

3. 'Waterbody' case: All pixels of nearest waterbody within 900 m of the reported PWS location containing >= 9 pixels

4. 'Watershed' case: All pixels within the same HUC12 watershed as the reported PWS location if watershed contains >= 9

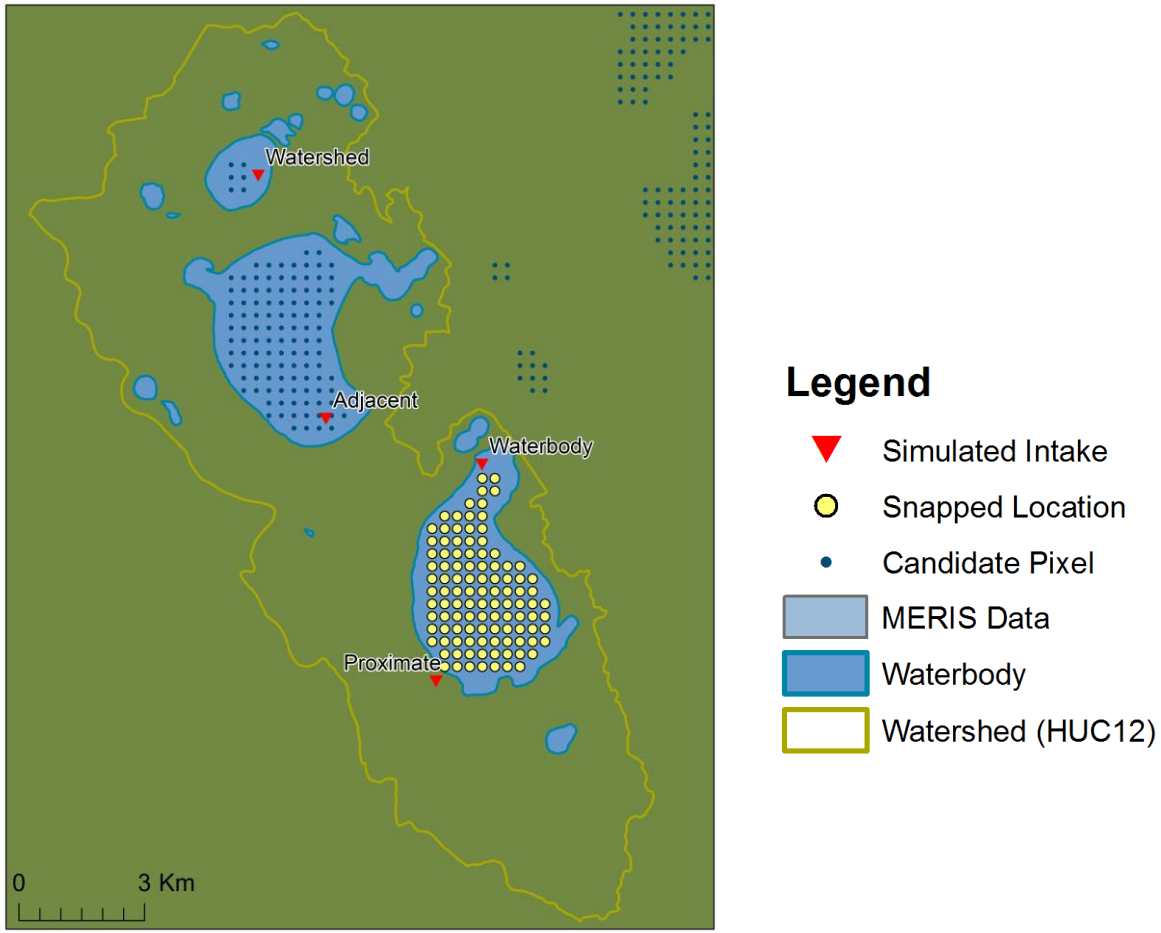


Simulated data, does not represent true intake locations.

Image: MODIS Aqua - 10/9/2011 - NASA

Location Snapping

1. 'Adjacent' case: Nearest 3x3 pixel array within 300 m of reported PWS location
2. 'Proximate' case: Nearest 3x3 pixel array within 900 m of reported PWS location
3. **'Waterbody' case: All pixels of nearest waterbody within 900 m of the reported PWS location containing >= 9 pixels**
4. 'Watershed' case: All pixels within the same HUC12 watershed as the reported PWS location if watershed contains >= 9

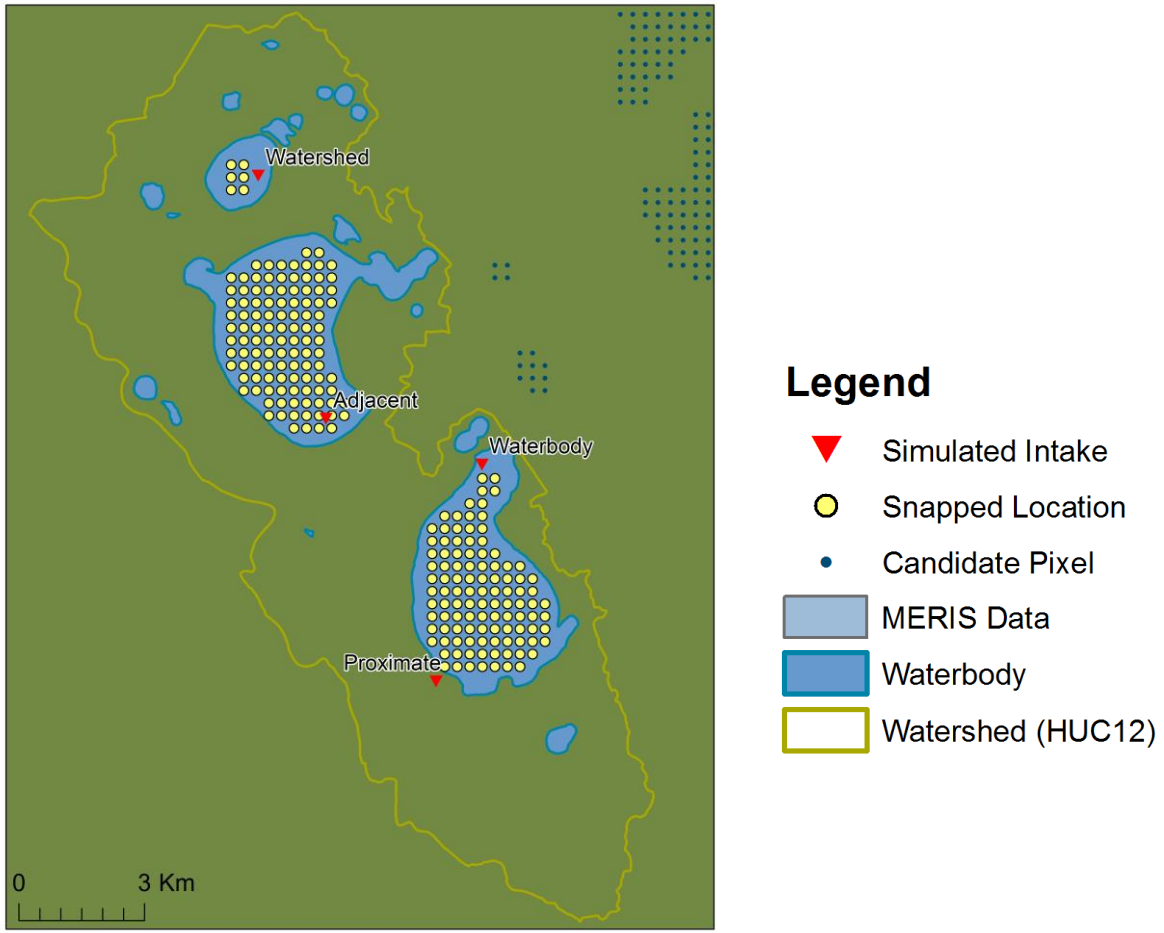


Simulated data, does not represent true intake locations.

Image: MODIS Aqua - 10/9/2011 - NASA

Location Snapping

1. ‘Adjacent’ case: Nearest 3x3 pixel array within 300 m of reported PWS location
2. ‘Proximate’ case: Nearest 3x3 pixel array within 900 m of reported PWS location
3. ‘Waterbody’ case: All pixels of nearest waterbody within 900 m of the reported PWS location containing >= 9 pixels
4. **‘Watershed’ case: All pixels within the same HUC12 watershed as the reported PWS location if watershed contains >= 9**



Simulated data, does not represent true intake locations.

Image: MODIS Aqua - 10/9/2011 - NASA

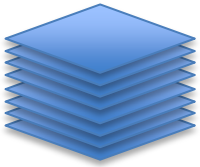
Local Scale: Snapping Results

		Intake Location Snapping Case				
Region	# Intakes	Adjacent	Proximate	Waterbody	Watershed	Unresolved
Florida	10	1	6	7	7	3
New England	595	7	38	64	89	506
Ohio	179	17	35	41	11	134
Total	784	25	79	112	107	643

1. 'Adjacent' case: Nearest 3x3 pixel array within 300 m of reported PWS location
2. 'Proximate' case: Nearest 3x3 pixel array within 900 m of reported PWS location
3. 'Waterbody' case: All pixels of nearest waterbody within 900 m of the reported PWS location containing >= 9 pixels
4. 'Watershed' case: All pixels within the same HUC12 watershed as the reported PWS location if watershed contains >= 9
5. 'Unresolved' case: No candidate pixels satisfy criteria. Pixel nearest to reported PWS location recorded for QA/QC

Image: MODIS Aqua - 10/9/2011 - NASA

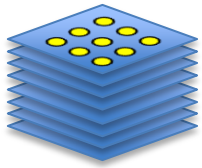
Local Scale: Overview



CI-multi Imagery

PWS Intake

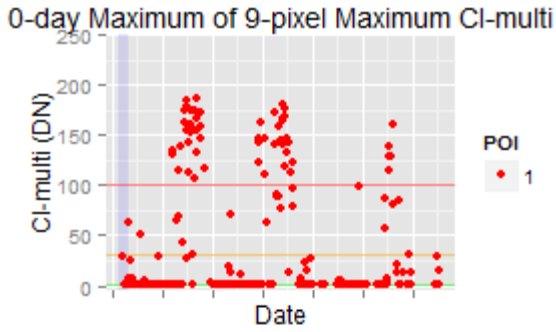
Snap Locations



Time-Series Extract

Scene	x_1y_1	x_1y_2	x_2y_1	x_2y_2
t_1	NA	1	NA	1
t_2	5	10	79	1
t_3	23	13	5	1
t_4	35	1	6	1

Spatial Aggregate



Temporal Aggregate

$$Bloom\ Freq = \frac{n_{obs > threshold}}{n_{obs}}$$

Report

Image: MODIS Aqua - 10/9/2011 - NASA

Simulated data, does not represent true intake locations.

Risk Categories By Quantile

		Risk		
Case	Region	Low	Moderate	High
Adjacent	Florida	0	0	1
	New England	0	7	0
	Ohio	7	5	5
	All	7	12	6
Proximate	Florida	0	0	6
	New England	6	26	6
	Ohio	15	12	8
	All	21	38	20
Waterbody	Florida	1	2	4
	New England	27	37	0
	Ohio	8	33	0
	All	36	72	4
Watershed	Florida	0	1	6
	New England	22	49	18
	Ohio	5	5	1
	All	27	55	25

$$Bloom\ Frequency = \frac{n_{observations > threshold}}{n_{observations}}$$

Low = BF $\leq$ Q1
Mod = Q1 < BF $\leq$ Q3
High = BF > Q3

Image: MODIS Aqua - 10/9/2011 - NASA

Objectives

- National scale
  - Estimate spatially resolvable features
- Regional scale
  - Calculate temporal coverage
  - Calculate bloom frequency
- Local scale
  - Relate features to nearby observations
  - Extract time series
  - Summarize across space and time

Image: MODIS Aqua - 10/9/2011 - NASA

An approach to monitoring cyanobacteria blooms at surface drinking water intakes using satellite imagery

Thank You!

- John Darling – EPA, ORISE mentor
- Blake Schaeffer – EPA, CyAN
- Erin Urquhart – EPA, ORISE
- Amy Davis – EPA, ORISE
- CyAN Team
- EnviroAtlas Team
- …and many more

John M. Clark – Clark.John@epa.gov
ORISE Participant at ORD/USEPA

References:

- Graham, J. L., Loftin, K. A., & Kamman, N. (2009). Monitoring recreational freshwaters. *Lakeline*, *29*, 18-24.
- Lunetta, R. S., Schaeffer, B. A., Stumpf, R. P., Keith, D., Jacobs, S. A., & Murphy, M. S. (2015). Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. *Remote Sensing of Environment*, *157*, 24-34.

R packages:

- raster, ggplot2, plyr, shiny, reshape2, dismo, maptools, xts

This project was supported in part by an appointment to the Research Participation Program at the Office of Research and Development, U.S. Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and EPA.

OAK RIDGE INSTITUTE FOR SCIENCE AND EDUCATION
Managed by ORAU for DOE