Application of GAMs to Explain Trends in Tidal Water Quality

Rebecca Murphy (UMCES at CBPO)
Jeni Keisman (USGS)

March 27, 2019

STAC Quarterly Meeting

University of Maryland CENTER FOR ENVIRONMENTAL SCIENCE

Chesapeake Bay Program
Science. Restoration. Partnership.

Outline

- GAM project overview/development steps
- Key development steps since STAC review
  - Method change analyses
  - Flow or salinity adjustment
  - CRAN package
  - Documentation/manuscript
- How the results are used
  - Bay-wide tidal trends annual production
  - Research collaborations
  - Ongoing work

Motivation

Advancing our understanding of factors affecting trends in water quality requires adoption of analytical approaches that can:

- represent potential nonlinearities in the long-term patterns now being observed with more than 30 years of water quality data;
- explore environmental factors that may affect those trends.

2014 MEOWQT Workshop Finding

- The GAM (General Additive Model) approach is a promising mathematical tool for detecting and describing trends in estuarine water quality parameters.

- Full implementation the GAMs technique will require incorporation of additional temporally-distributed hydrologic processes that influence mixing and flow in increasingly more saline areas of the estuary (i.e., the oligo-, meso-, and poly-haline zones).

2014 MEOWQT Workshop Recommendation

- The CBP should continue to develop and apply GAMs to the appropriate response variables (i.e., nutrients, sediments, DO) in tidal waters.

- Developers of the GAMs method should solicit input from the estuarine research community to guide construction of GAMs for the Chesapeake Bay.

- In order to fully implement GAMs as a standardized method for evaluating water quality trends in the tidal waters, functionality that enables automated analysis must be developed.

- The CBP should submit the GAMs technique to a rigorous peer review process before establishing it as the primary tool for estimating trends in estuarine water quality parameters.

Project Overview

- 2014-present: Collaborative development team implementing Generalized Additive Model (GAM) approach
- 2015-present: Ongoing research collaborations using GAMs
- 2016-17: STAC review of the approach
- 2017: MDDNR & VADEQ adoption of the GAM approach to compute annual tidal trends
- 2017-2018: Incorporated method change approach & flow-adjustment
- 2018: ‘baytrends’ package published on R repository ‘CRAN’
- 2018: MDDNR & VADEQ adoption of updated approach with flow-adjustment and method-change approach for annual trends
- 2019: Peer reviewed manuscript

Basic approach

A GAM represents the constituent of interest as the sum of multiple smooth (possibly non-linear) functions of explanatory variables

Water quality =
linear(date) + s(date) + s(doy) + ti(date, doy)

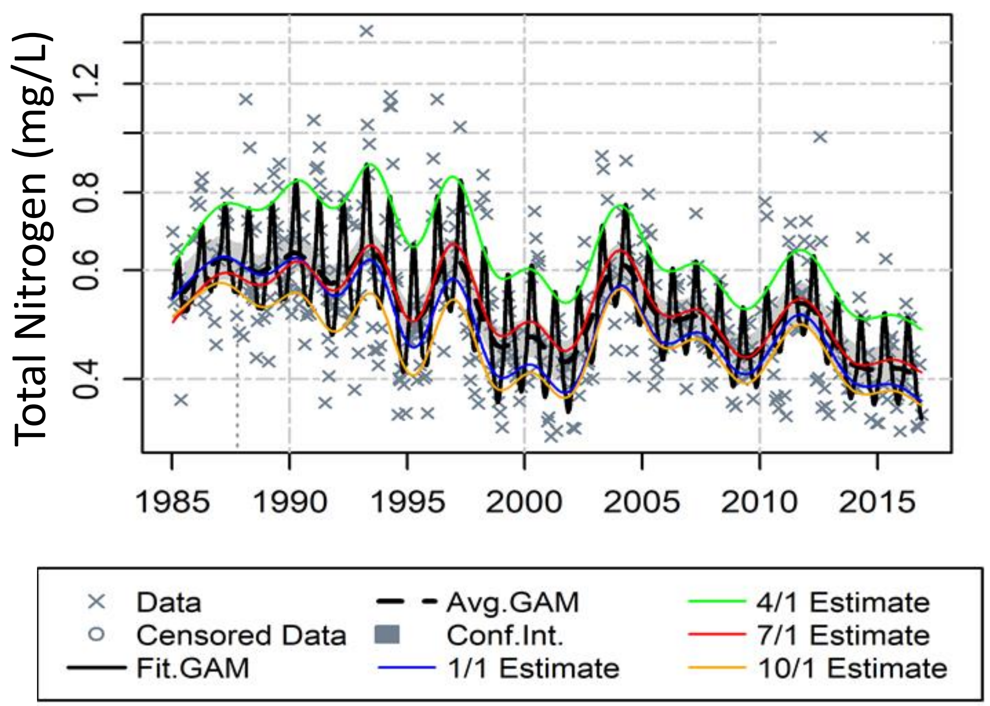
Functions can be linear

Smoothly-varying non-linear “spline” functions

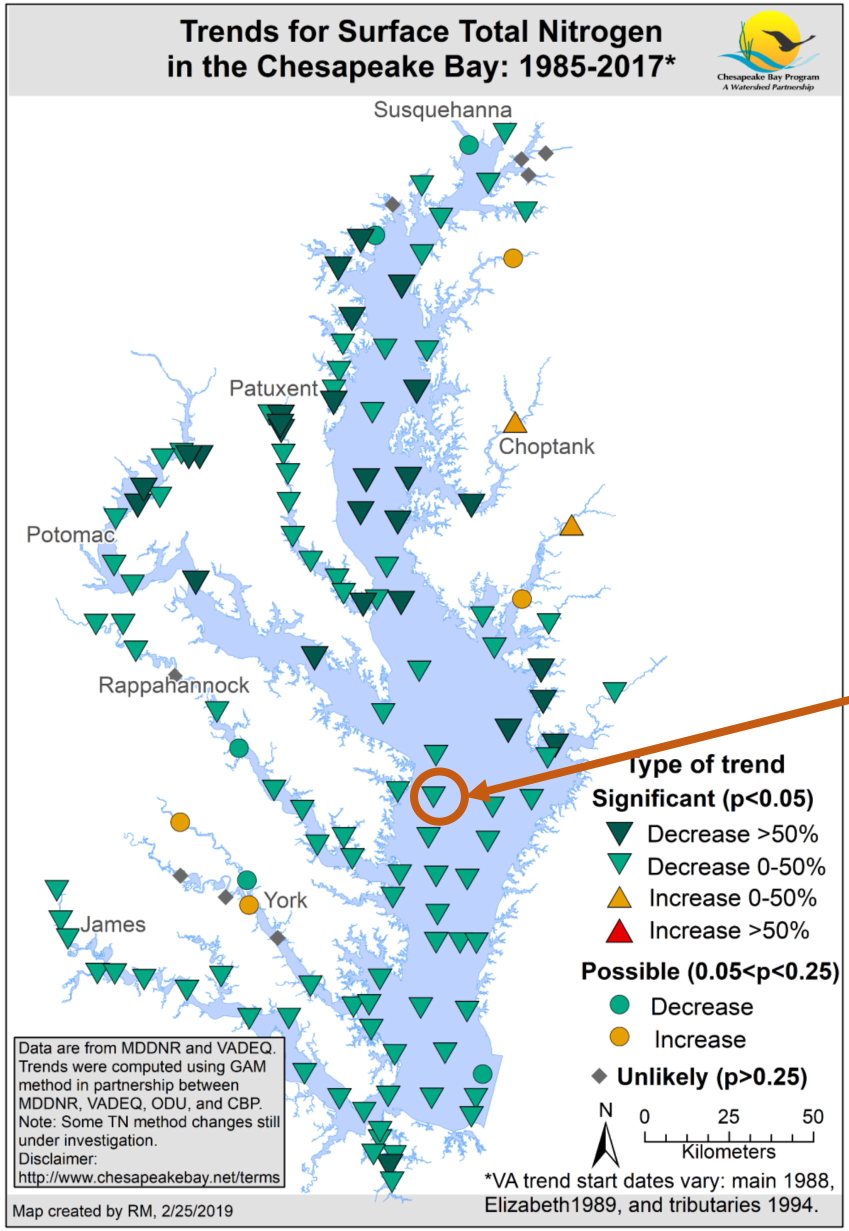
And multi-dimensional smooth functions

s = spline smooth functions
doy = day of year

Example: TN in mid-Bay at CB5.4 Surface

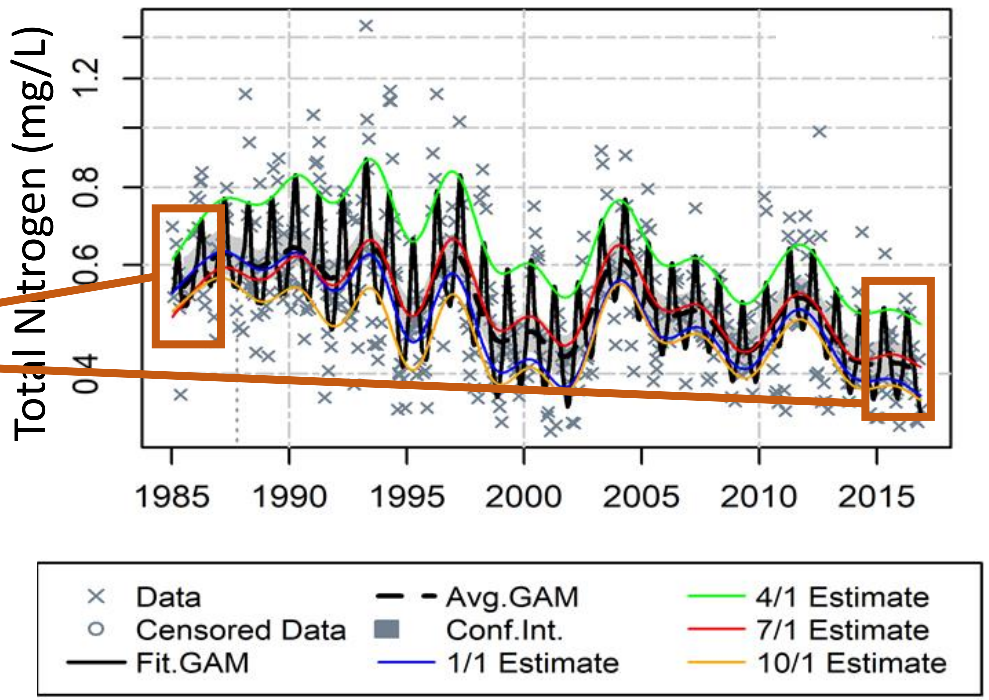


Basic approach: Percent change



change = -35%
$p<0.001$

Example: TN in mid-Bay at CB5.4 Surface



Basic approach

Water quality =
linear(date) + s(date) + s(doy) + ti(date, doy)

- Basic approach tells us:
  - Observed mean pattern over time
  - Seasonal changes
  - Trends over time
- These trends are influenced by wet/dry periods
- Also needed an approach to answer the question: if river flow had been average, what would the trends look like?

Key updates since STAC review

- Flow or salinity adjustment
- Method or lab change analyses (i.e., interventions)
- Publicly-available CRAN package ‘baytrends’
- Manuscript in review

Flow or salinity-adjustment

Water quality =
linear(date) + s(date) + s(date) + ti(date,doy) +
s(flw_sal) + ti(flw_sal,doy) + ti(flw_sal, date) + ti(flw_sal,doy,date)

Same as previous model

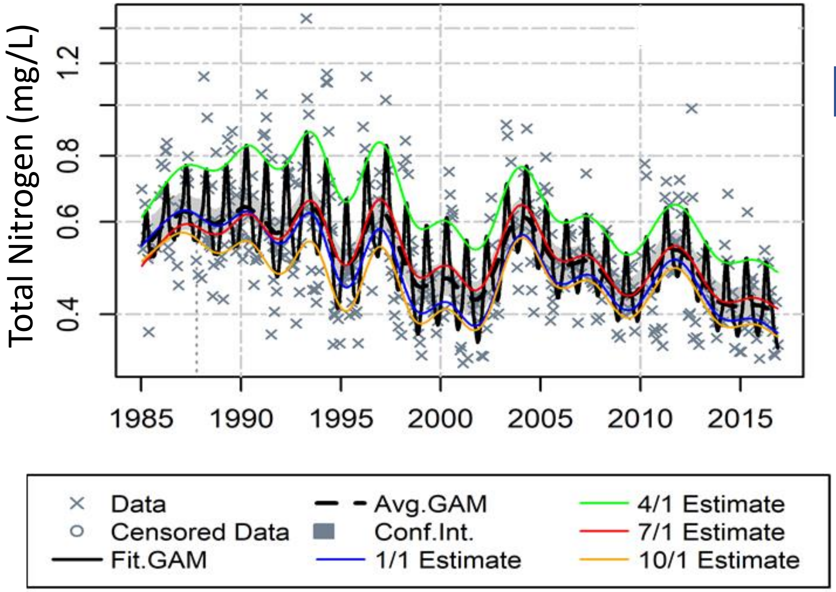
Additional smooth on flow plus all interactions

flw_sal is either:

- Salinity at same place & time
- Pre-processed flow from a RIM station averaged over preceding n days (selected by correlation)

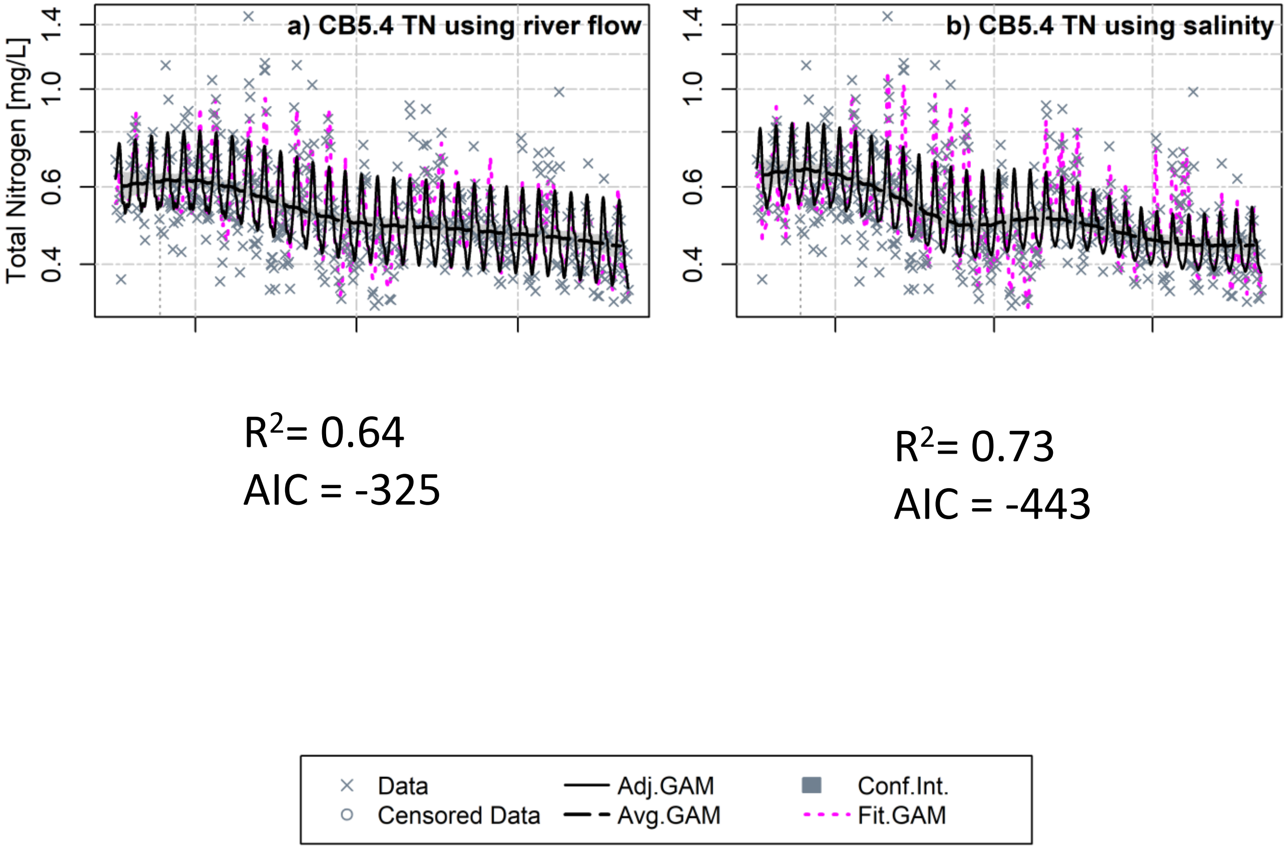
Flow or salinity-adjustment

Basic Model:
TN in mid-Bay at CB5.4 Surface



R^2= 0.52
AIC = -195

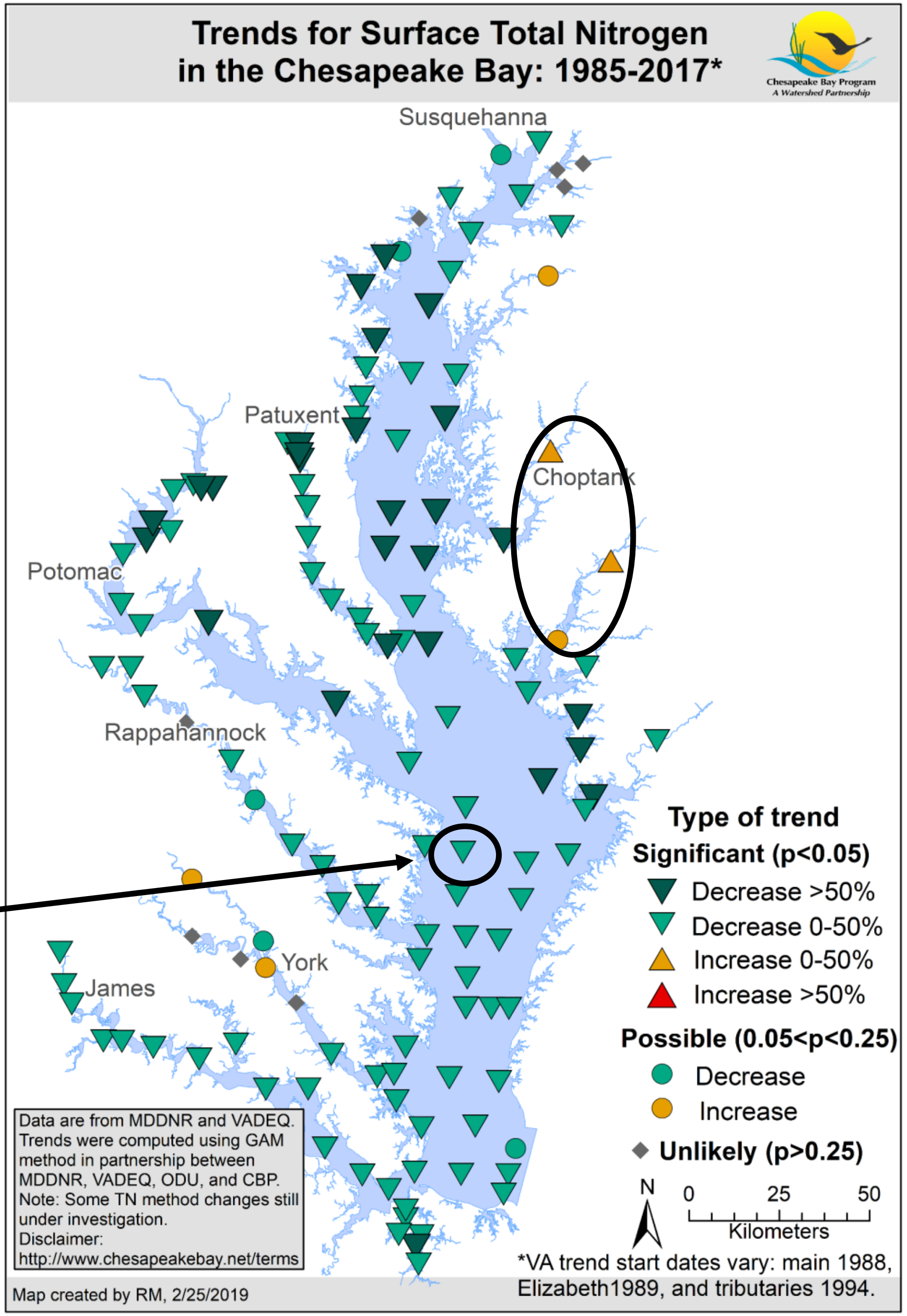
Magenta: Model with 90 days of Susquehanna flow or CB5.4 salinity
Black: Flow-or-salinity adjusted



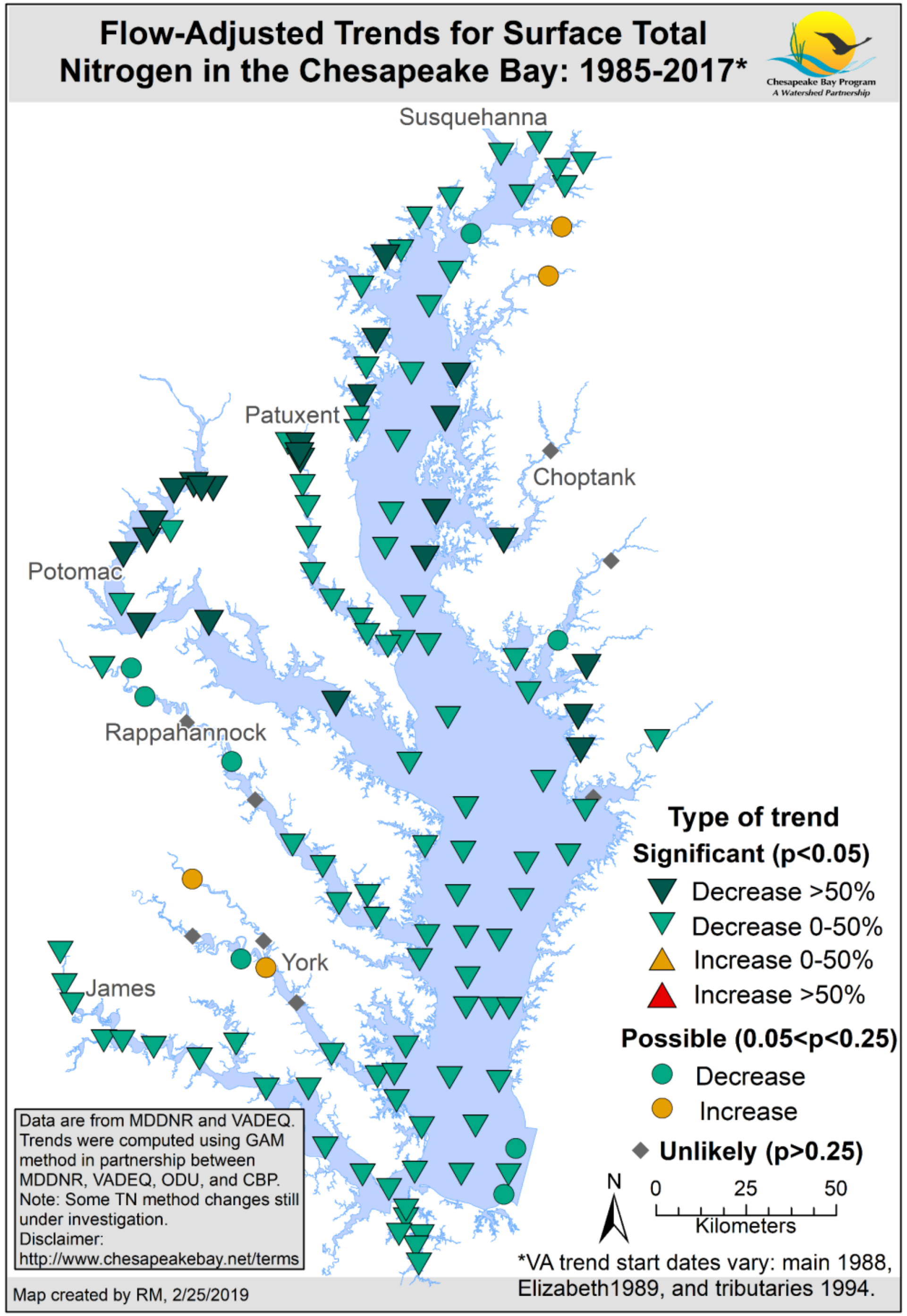
R^2= 0.64
AIC = -325

R^2= 0.73
AIC = -443

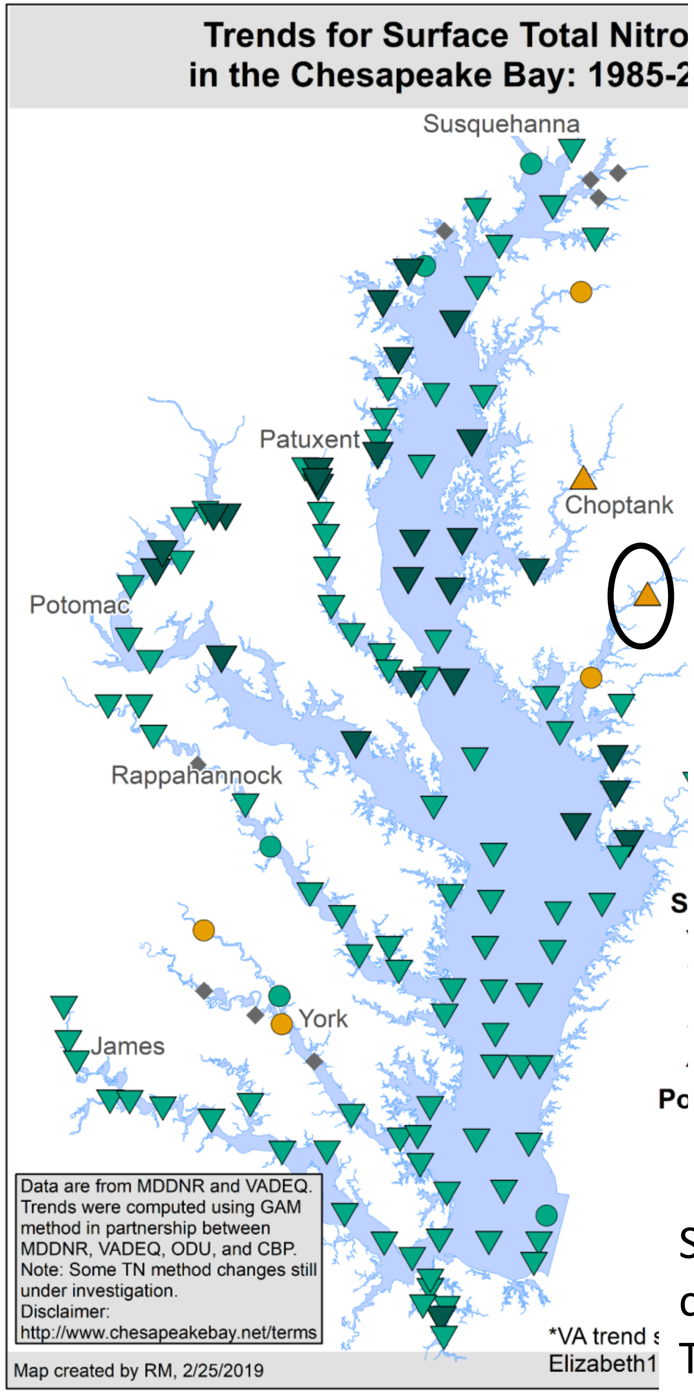
Compare maps with and without flow-adjustment

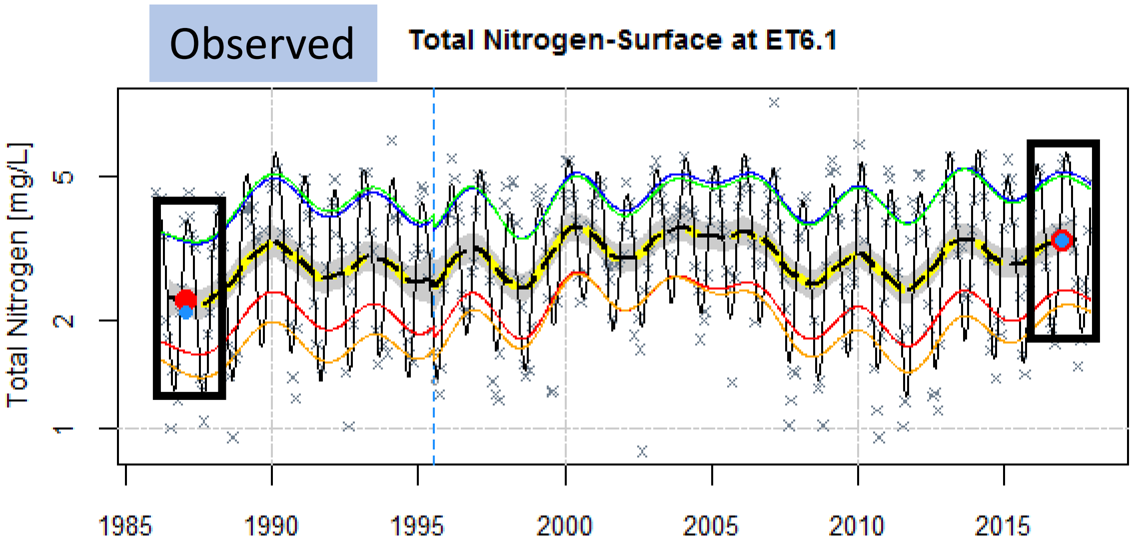


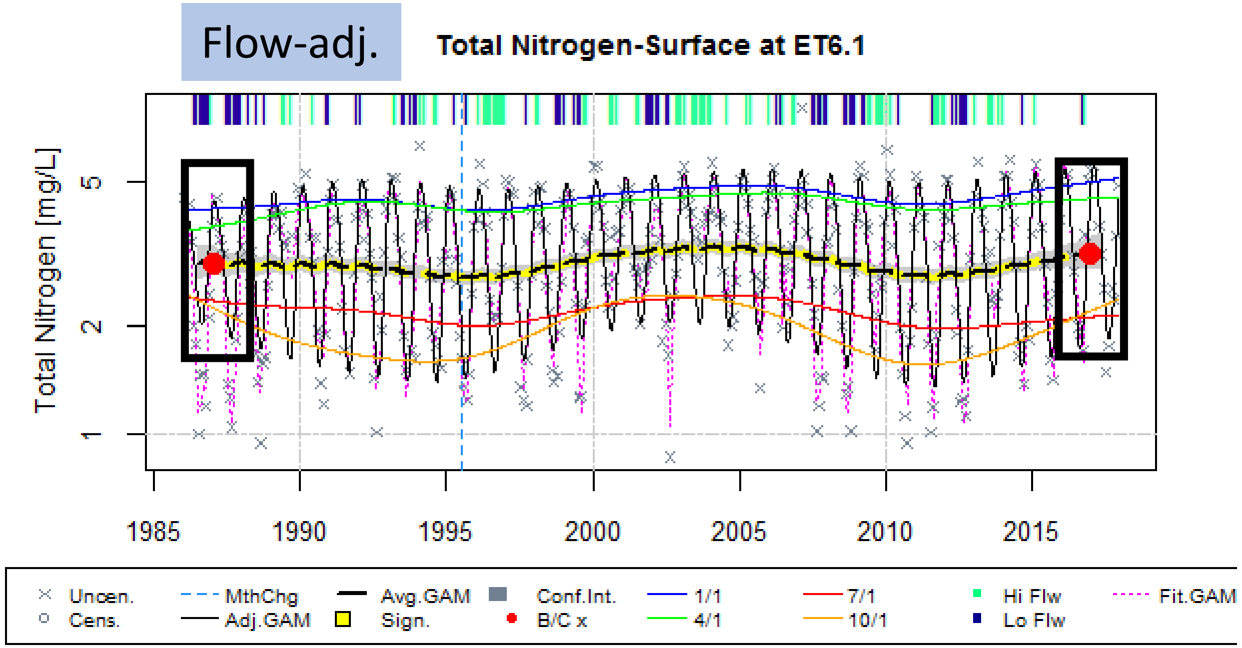
Previous example



Example change with flow-adjustment







Still a slight increase with flow-adjustment, but no longer confident in it because the early low flow years pulled down the TN in the 1980s

Method or lab change analysis: Interventions

Water quality =
linear(date) + s(date) + s(doy) + ti(date, doy) ← Original model
+ intervention ← Factor term (e.g., 0 or 1) before and after change
+ all flow terms ← Model could include or not include flow

- Intervention investigations proved challenging (as discussed by STAC reviewers)
- Work still underway on dissolved nutrients and TSS
- Currently changes included in our annual trends for TN and TP

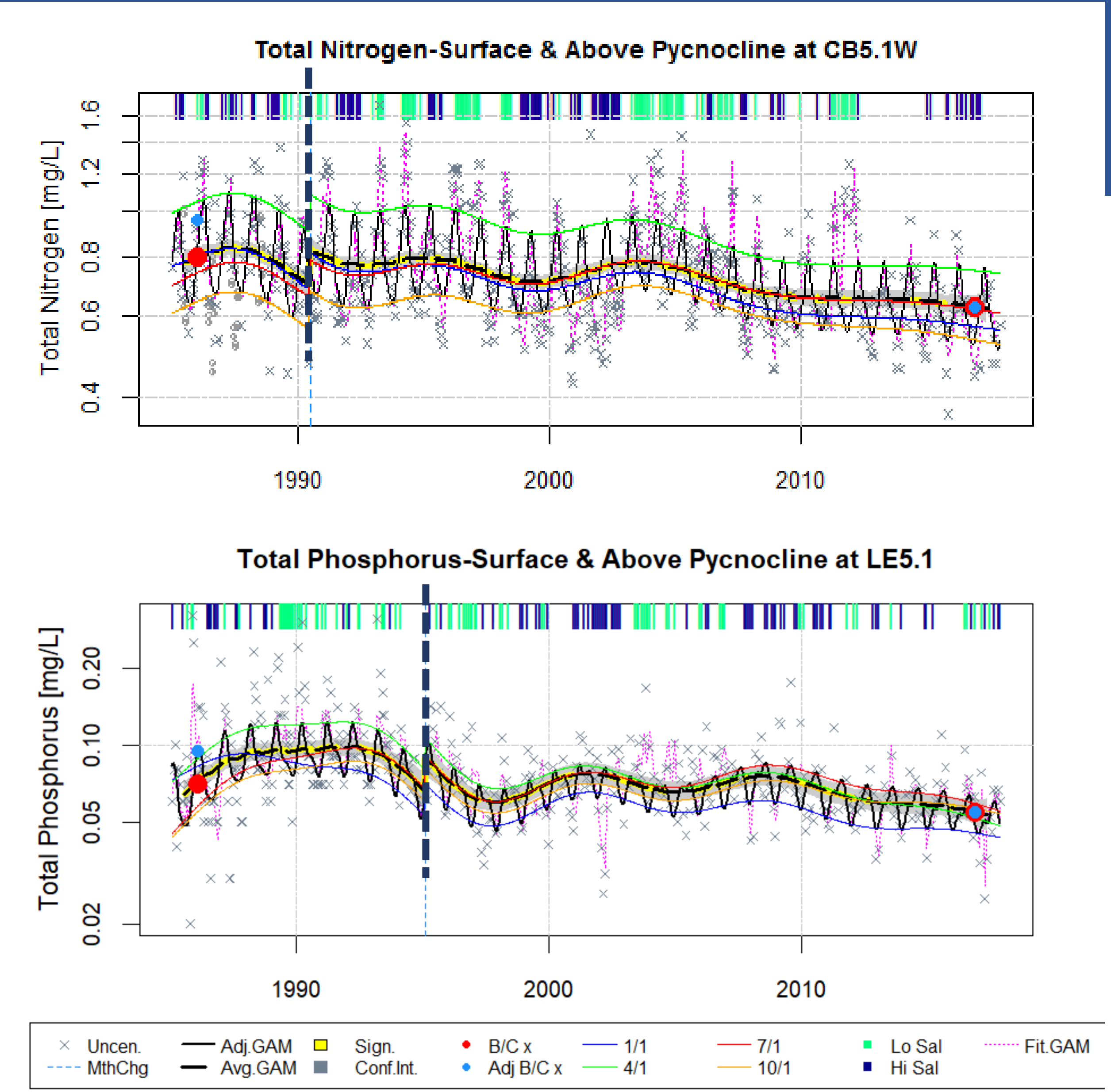
Method or lab change analysis:

TN change from TKNW+NO23 → PN + TDN

Intervention = 0.19, p<0.001

TP change from direct measure → PP + TDP

Intervention = 0.30, p=0.04



Package on CRAN

https://cran.r-project.org/web/packages/baytrends/index.html

baytrends: Long Term Water Quality Trend Analysis

Enable users to evaluate long-term trends using a Generalized Additive Modeling (GAM) approach. The model development includes selecting a GAM structure to describe nonlinear seasonally-varying changes over time, incorporation of hydrologic variability via either a river flow or salinity, the use of an intervention to deal with method or laboratory changes suspected to impact data values, and representation of left- and interval-censored data. The approach has been applied to water quality data in the Chesapeake Bay, a major estuary on the east coast of the United States to provide insights to a range of management- and research-focused questions.

Version:	1.1.0
Depends:	R (≥ 3.2.0), lubridate, mgcv
Imports:	XML, dataRetrieval, digest, gdata, memoise, methods, plyr, survival, zCompositions
Suggests:	devtools, fitdistrplus, imputeTS, knitr, nlme, pander, readxl, rmarkdown, sessioninfo, testthat
Published:	2019-03-14
Author:	Rebecca Murphy, Elgin Perry, Jennifer Keisman, Jon Harcum, Erik W Leppo
Maintainer:	Erik Leppo <Erik.Leppo at tetratech.com>
License:	GPL-3
URL:	https://github.com/tetratech/baytrends
NeedsCompilation:	no
Materials:	README NEWS
CRAN checks:	baytrends results

Downloads:

Reference manual:	baytrends.pdf
Vignettes:	Data Sets QW
Package source:	baytrends_1.1.0.tar.gz
Windows binaries:	r-devel: baytrends_1.1.0.zip, r-release: baytrends_1.1.0.zip, r-oldrel: baytrends_1.1.0.zip
OS X binaries:	r-release: baytrends_1.1.0.tgz, r-oldrel: baytrends_1.1.0.tgz
Old sources:	baytrends archive

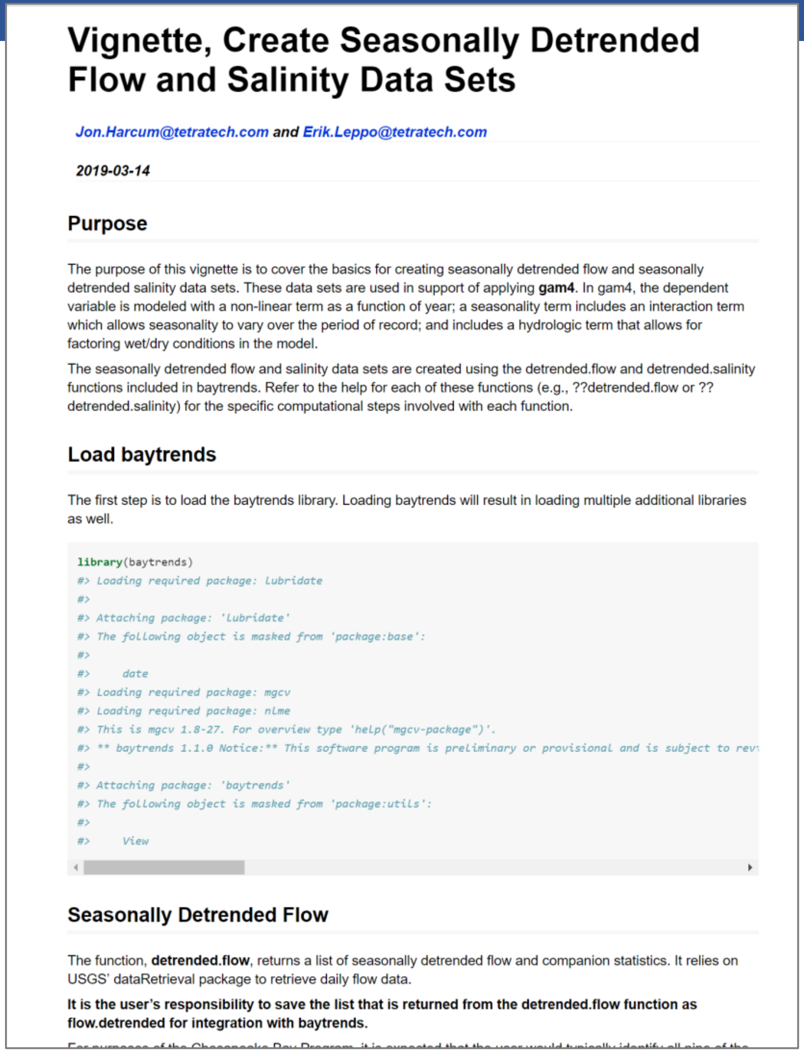
Linking:

Please use the canonical form `https://CRAN.R-project.org/package=baytrends` to link to this page.

Documentation

- Murphy, R.R., E. Perry, J. Harcum, and J. Keisman. 2019. **A Generalized Additive Model approach to evaluating water quality: Chesapeake Bay Case Study.** *Environmental Modelling and Software*. In Review.

- In-package documentation including overview, vignettes, and function instructions



Vignette, Create Seasonally Detrended Flow and Salinity Data Sets

Jon.Harcum@tetratech.com and *Erik.Leppo@tetratech.com*

2019-03-14

Purpose

The purpose of this vignette is to cover the basics for creating seasonally detrended flow and seasonally detrended salinity data sets. These data sets are used in support of applying **gam4**. In gam4, the dependent variable is modeled with a non-linear term as a function of year; a seasonality term includes an interaction term which allows seasonality to vary over the period of record; and includes a hydrologic term that allows for factoring wet/dry conditions in the model.

The seasonally detrended flow and salinity data sets are created using the detrended.flow and detrended.salinity functions included in baytrends. Refer to the help for each of these functions (e.g., ??detrended.flow or ??detrended.salinity) for the specific computational steps involved with each function.

Load baytrends

The first step is to load the baytrends library. Loading baytrends will result in loading multiple additional libraries as well.

```
library(baytrends)
#> Loading required package: lubridate
#>
#> Attaching package: 'lubridate'
#> The following object is masked from 'package:base':
#>
#>     date
#> Loading required package: mgcv
#> Loading required package: nlme
#> This is mgcv 1.8-27. For overview type 'help("mgcv-package")'.
#> ** baytrends 1.1.0 Notice:** This software program is preliminary or provisional and is subject to rev
#>
#> Attaching package: 'baytrends'
#> The following object is masked from 'package:utils':
#>
#>     View
```

Seasonally Detrended Flow

The function, **detrended.flow**, returns a list of seasonally detrended flow and companion statistics. It relies on USGS' dataRetrieval package to retrieve daily flow data.

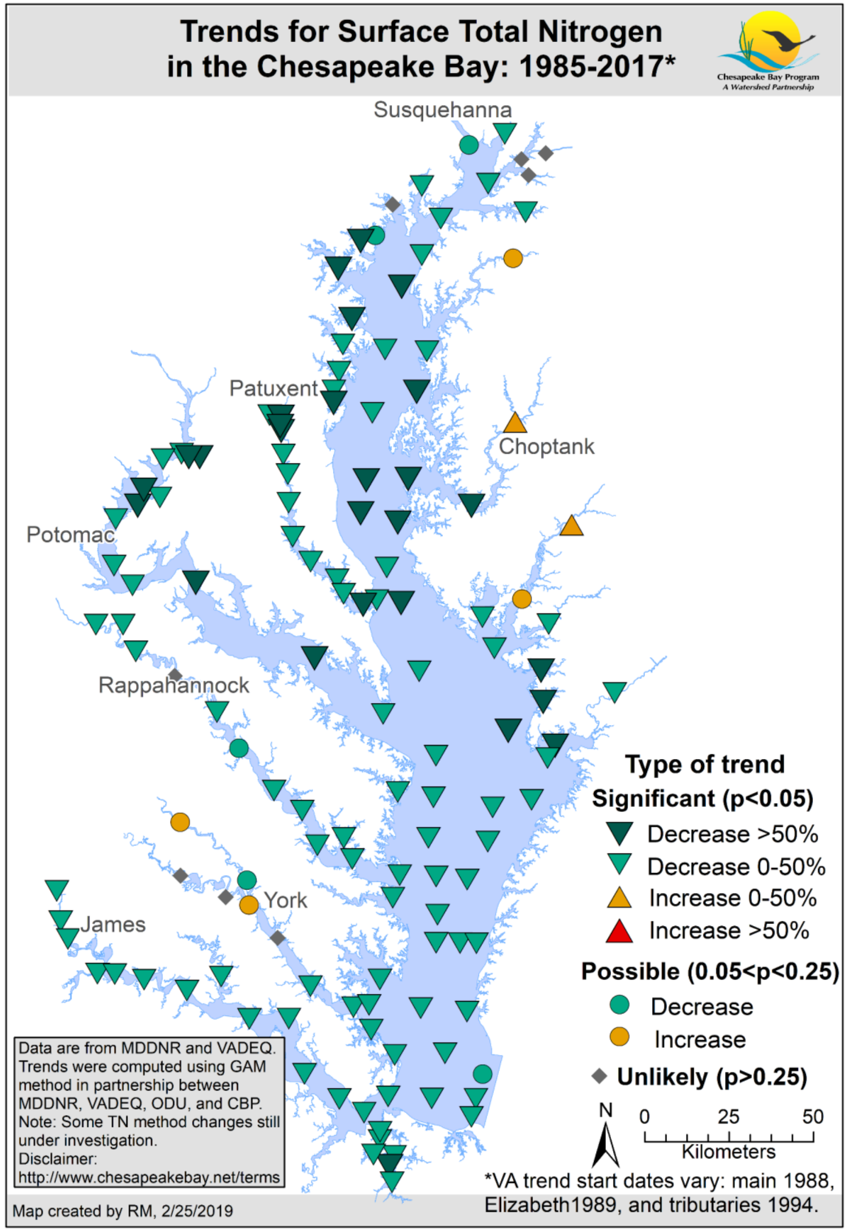
It is the user's responsibility to save the list that is returned from the detrended.flow function as flow.detrended for integration with baytrends.

Chesapeake GAMs Applications

- Annual trends in tidal water quality
- Academic research collaborations
- GAMs team ongoing work: hypothesis testing

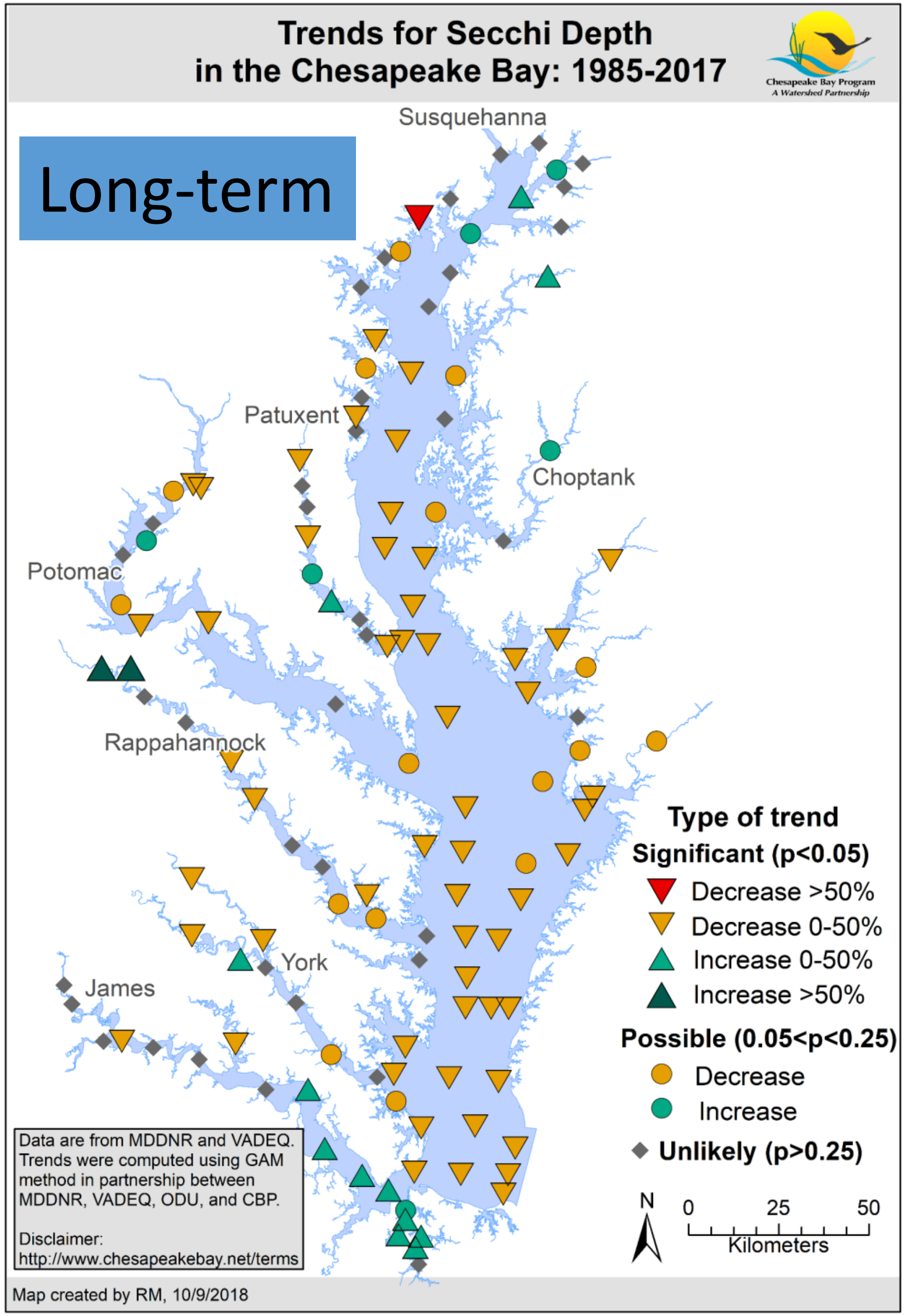
Tidal trends annual production

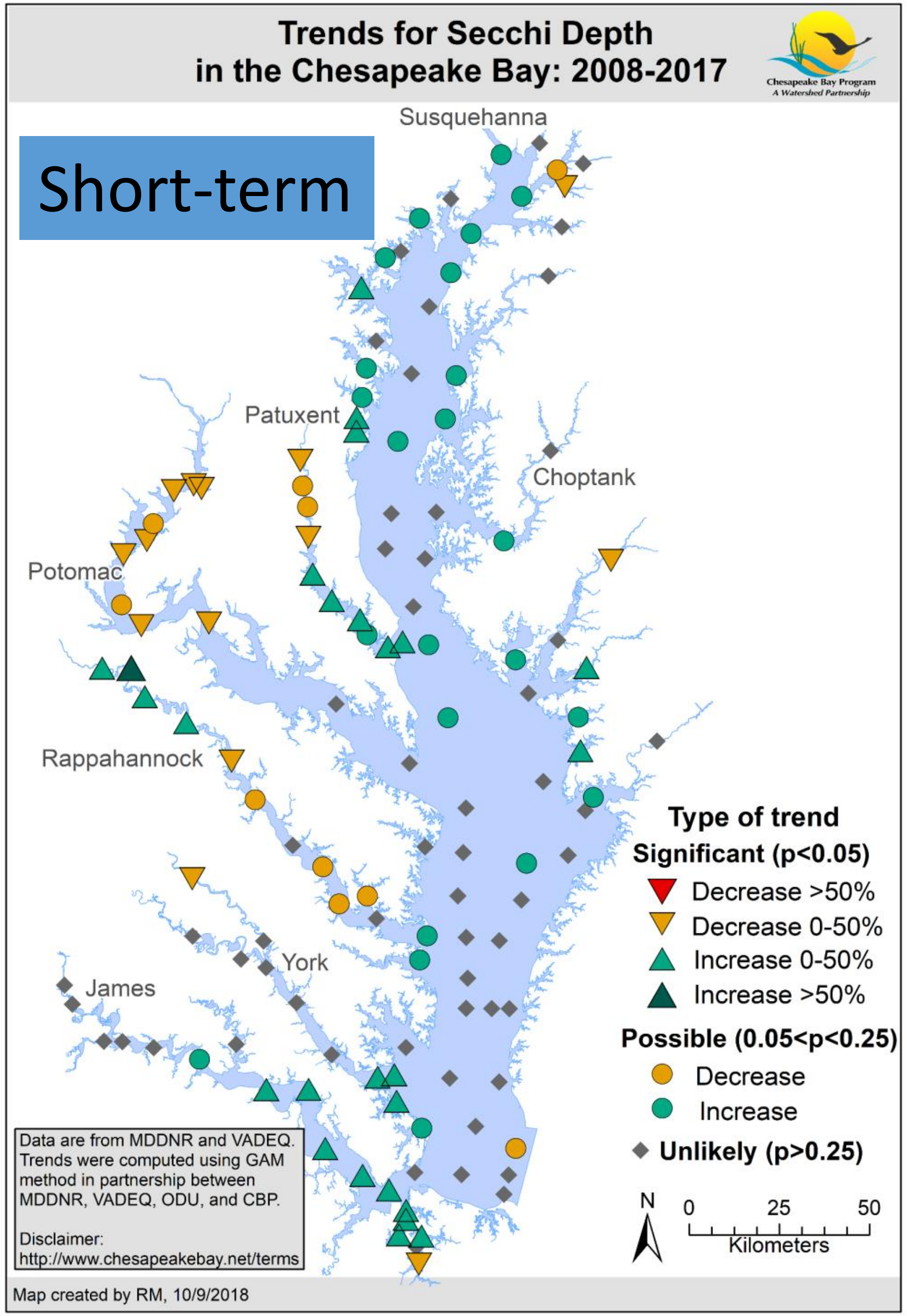
- MDDNR and VADEQ compute trends annually (with CBP, ODU and USGS)
- GAM-based approach through R package
- Scope:
  - ~150 stations
  - TN, TP, DIN, PO4, secchi, DO, TSS, chlorophyll-a, water temp, salinity
  - Surface & bottom
  - Annual and/or seasonal; short-term & long-term
  - Observed & flow-or-salinity-adjusted trends



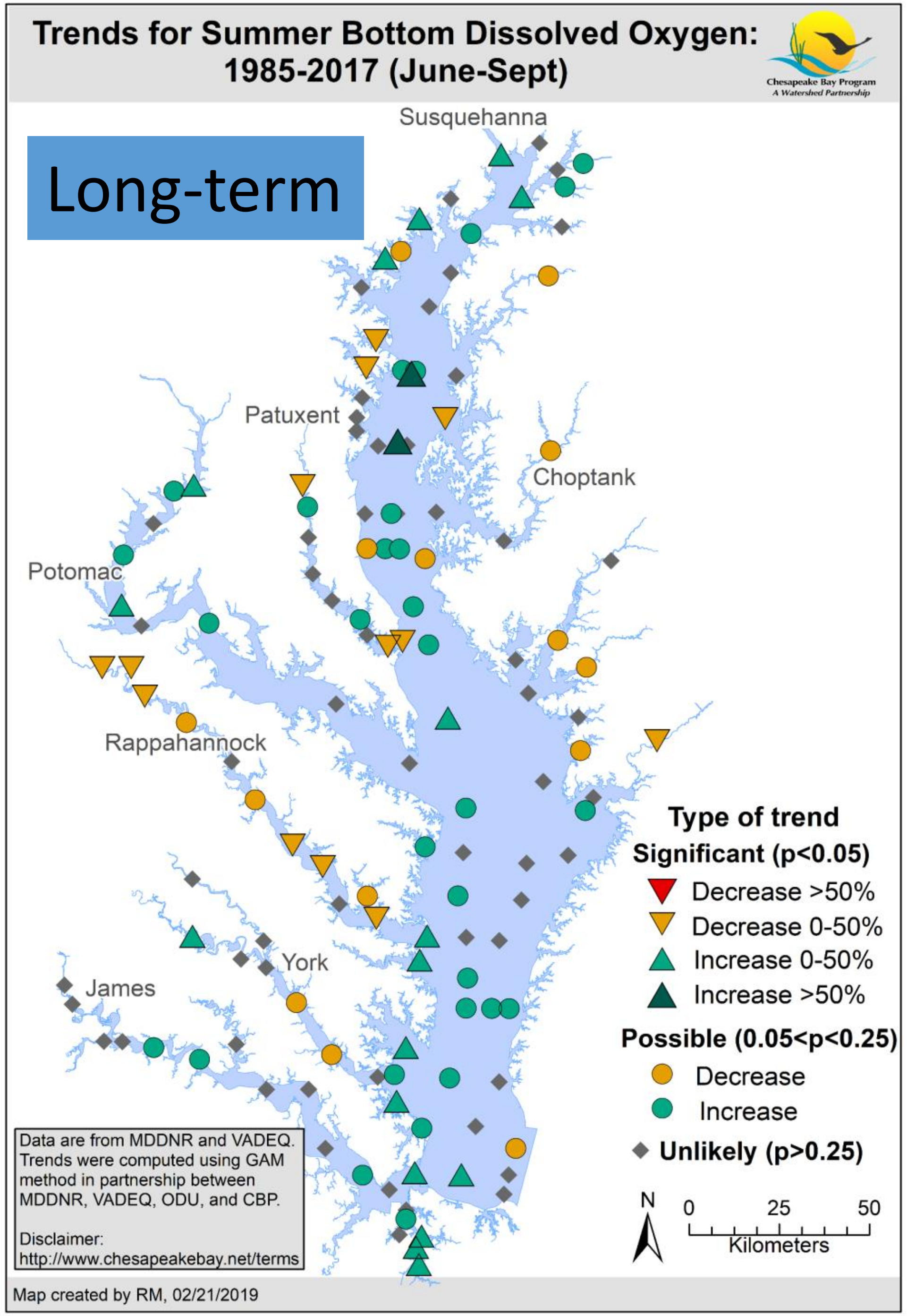
Annually:

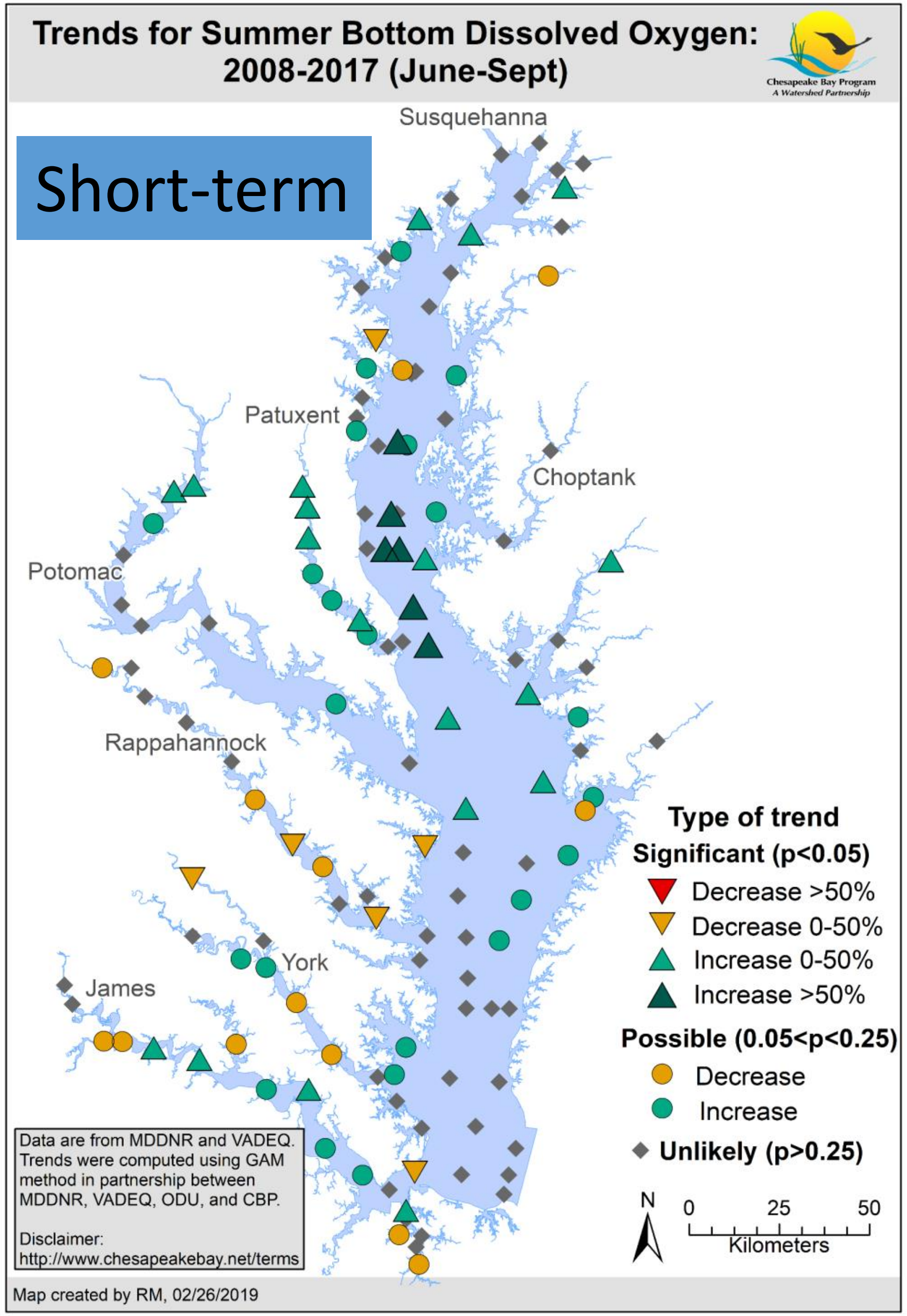
Bay-wide picture of trends





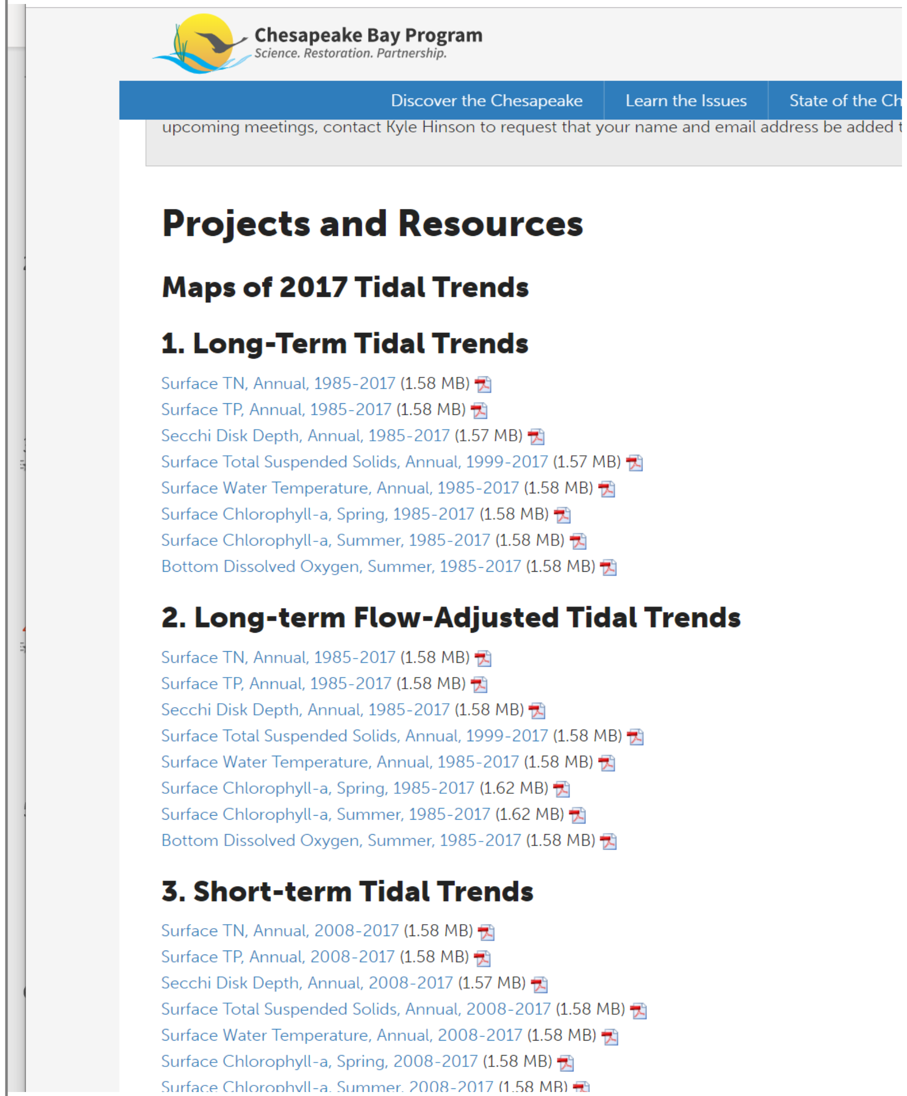
Annually:
Bay-wide picture of trends





Annually: Updated online

Integrated Trends Analysis Team Website


Chesapeake Bay Program
Science. Restoration. Partnership.

Discover the Chesapeake | Learn the Issues | State of the Ch

upcoming meetings, contact Kyle Hinson to request that your name and email address be added t

Projects and Resources

Maps of 2017 Tidal Trends

1. Long-Term Tidal Trends

- Surface TN, Annual, 1985-2017 (1.58 MB)
- Surface TP, Annual, 1985-2017 (1.58 MB)
- Secchi Disk Depth, Annual, 1985-2017 (1.57 MB)
- Surface Total Suspended Solids, Annual, 1999-2017 (1.57 MB)
- Surface Water Temperature, Annual, 1985-2017 (1.58 MB)
- Surface Chlorophyll-a, Spring, 1985-2017 (1.58 MB)
- Surface Chlorophyll-a, Summer, 1985-2017 (1.58 MB)
- Bottom Dissolved Oxygen, Summer, 1985-2017 (1.58 MB)

2. Long-term Flow-Adjusted Tidal Trends

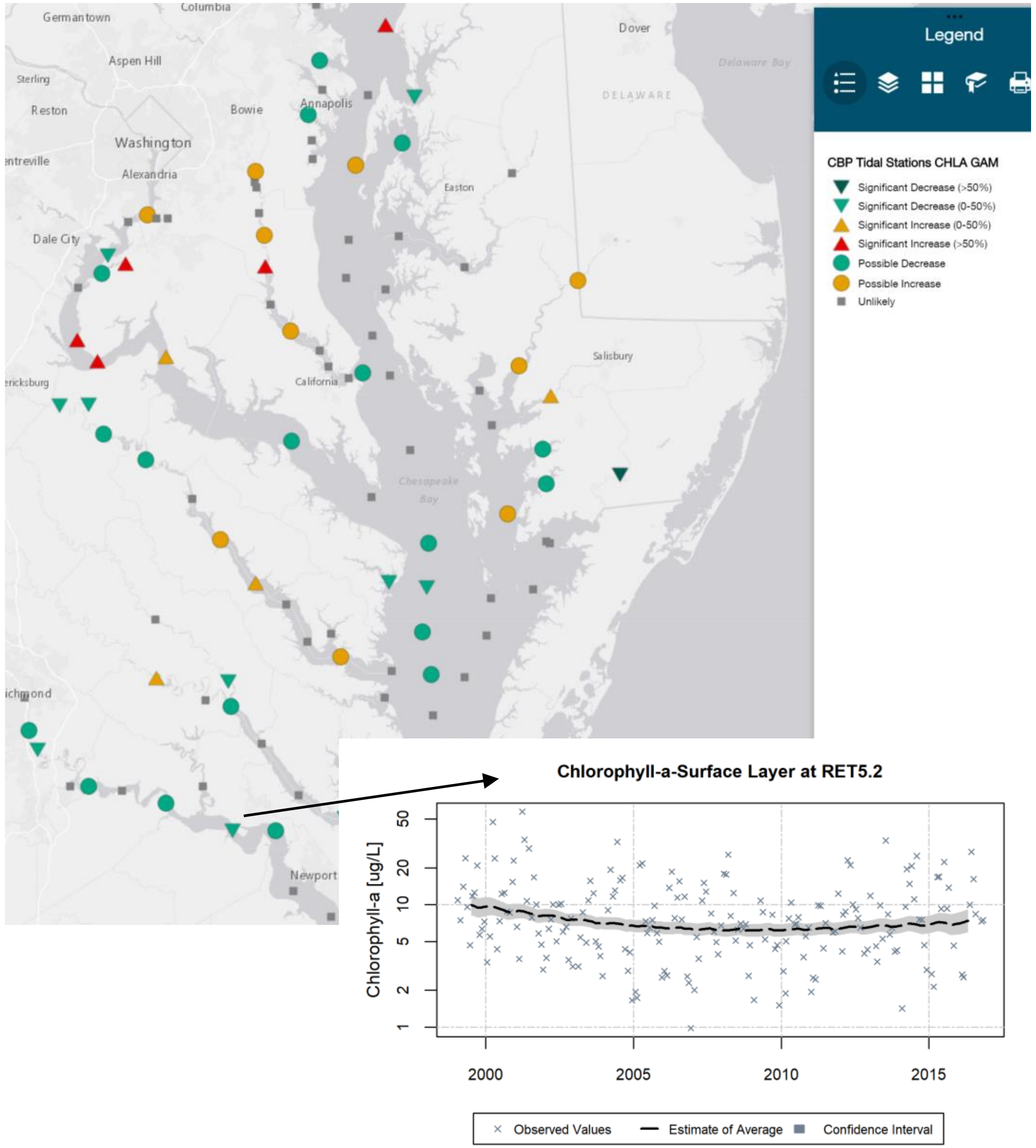
- Surface TN, Annual, 1985-2017 (1.58 MB)
- Surface TP, Annual, 1985-2017 (1.58 MB)
- Secchi Disk Depth, Annual, 1985-2017 (1.58 MB)
- Surface Total Suspended Solids, Annual, 1999-2017 (1.58 MB)
- Surface Water Temperature, Annual, 1985-2017 (1.58 MB)
- Surface Chlorophyll-a, Spring, 1985-2017 (1.62 MB)
- Surface Chlorophyll-a, Summer, 1985-2017 (1.62 MB)
- Bottom Dissolved Oxygen, Summer, 1985-2017 (1.58 MB)

3. Short-term Tidal Trends

- Surface TN, Annual, 2008-2017 (1.58 MB)
- Surface TP, Annual, 2008-2017 (1.58 MB)
- Secchi Disk Depth, Annual, 2008-2017 (1.57 MB)
- Surface Total Suspended Solids, Annual, 2008-2017 (1.58 MB)
- Surface Water Temperature, Annual, 2008-2017 (1.58 MB)
- Surface Chlorophyll-a, Spring, 2008-2017 (1.58 MB)
- Surface Chlorophyll-a, Summer, 2008-2017 (1.58 MB)

https://www.chesapeakebay.net/who/group/integrated_trends_analysis_team

WIP Dashboard

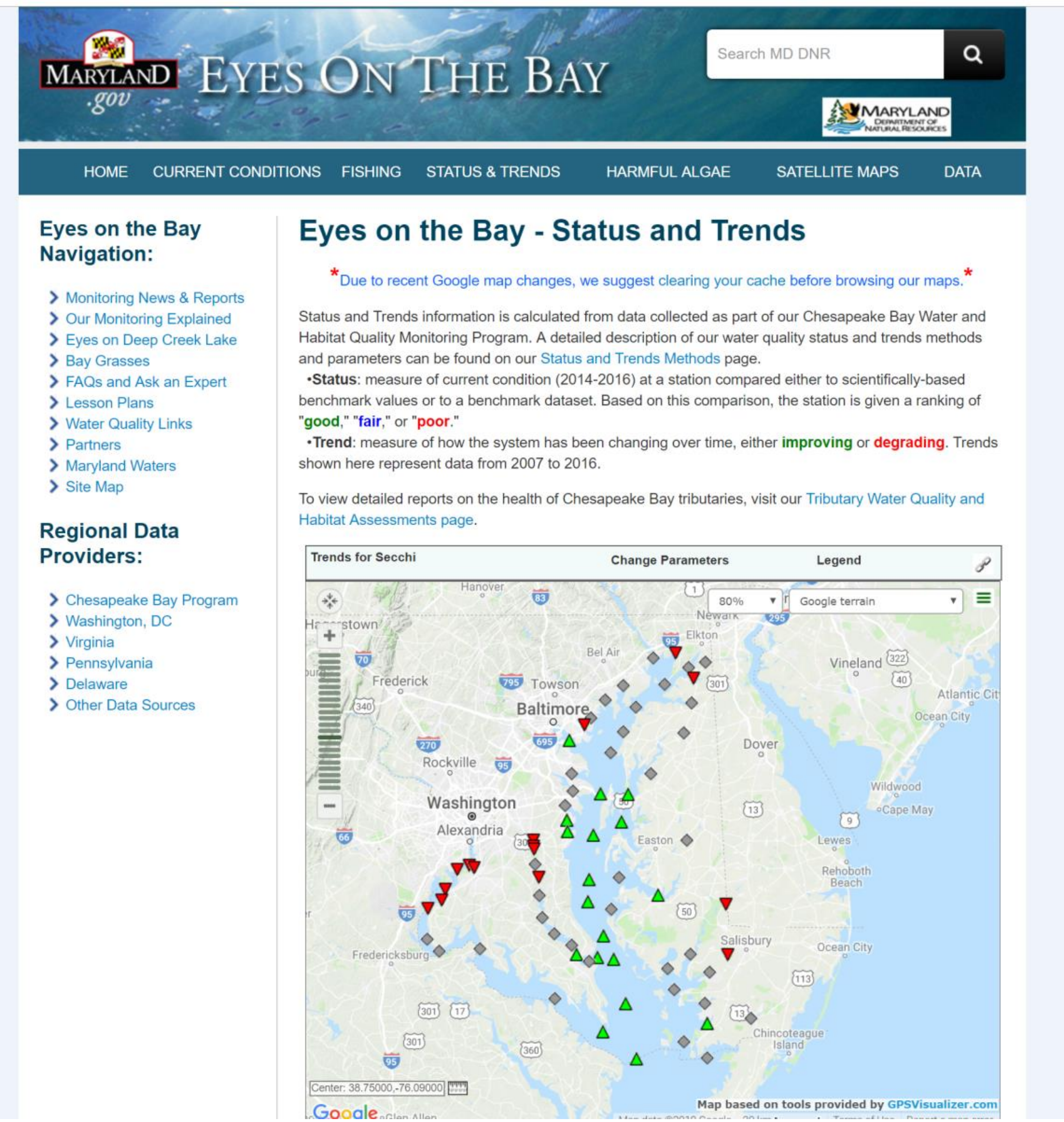


https://gis.chesapeakebay.net/wip/wqmonitoring/

How the results are used

Informing public, stakeholders, and management communities:

- CBP websites (earlier slide)
- MD and VA reports and websites (i.e., Eyes on the Bay)
- CBP and partner presentations and products (e.g., WQGIT, MWCOG)
- Stakeholder presentations (e.g., Potomac riverkeepers)



http://eyesonthebay.dnr.maryland.gov/eyesonthebay/statustrends.cfm

How the results are used: Research collaborations

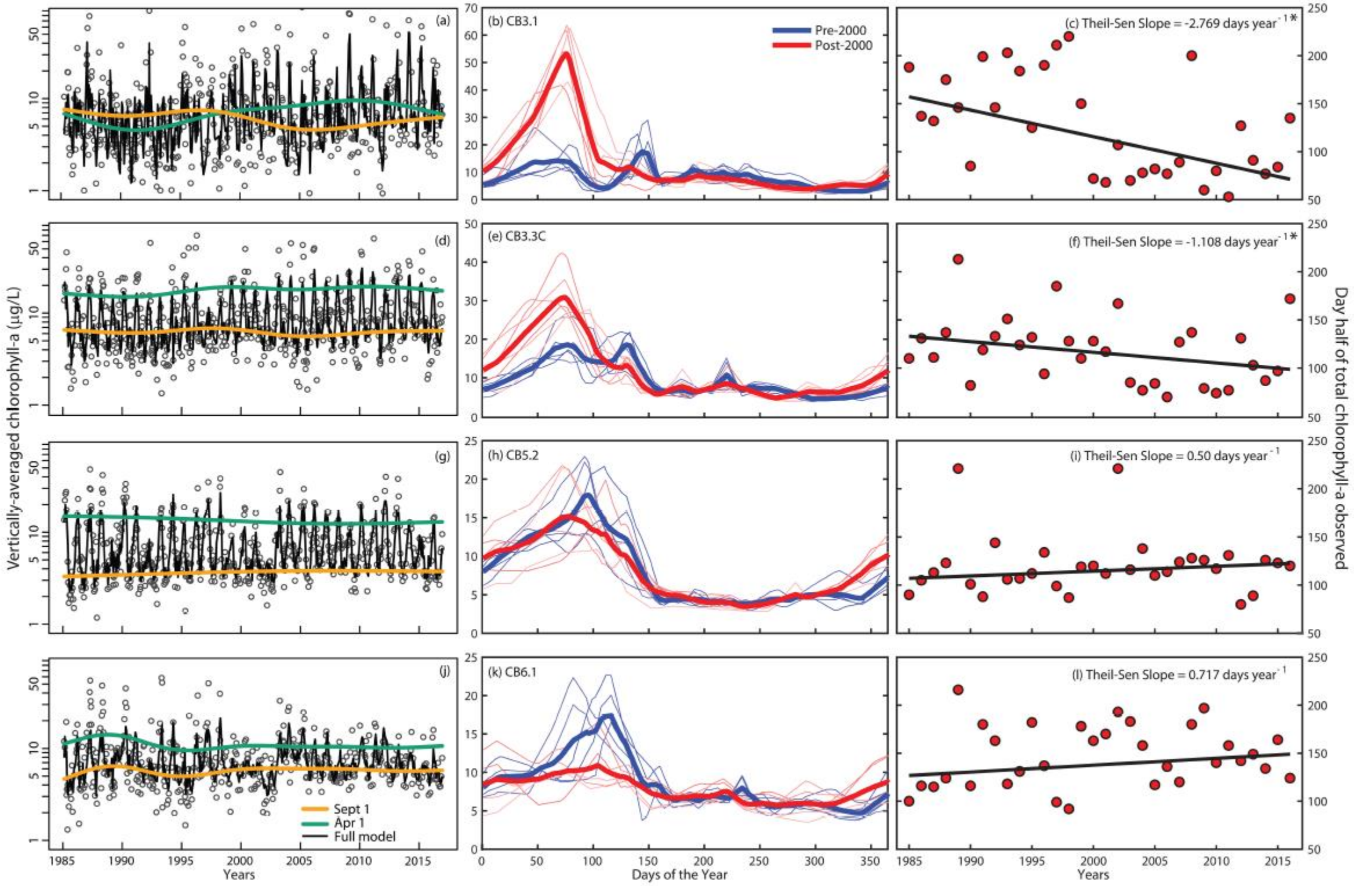
Chl-a and nitrogen seasonality: Testa, Murphy, Brady, and Kemp. 2018. **Nutrient- and Climate-Induced Shifts in the Phenology of Linked Biogeochemical** Cycles in a Temperate Estuary. *Frontiers in Marine Science* 5.

Water Clarity: Keisman, Friedrichs, Buchanan, Cornwell, Lane, Porter, Testa, Trice, Zhang, Zimmerman, Batiuk, Blomquist, Lyubchich, Moore, Murphy, Noe, Orth, Sanford, 2018. **Understanding and explaining over 30 years of water-clarity trends in Chesapeake Bay: Previous work and new insights.** Edgewater, MD. STAC Publication Number 18-XXX.

SAV: Lefcheck, Orth, Dennison, Wilcox, Murphy, Keisman, Gurbisz, Hannam, B. Landry, Moore, Patrick, Testa, Weller, and Batiuk. 2018. **Long-term nutrient reductions lead to the unprecedented recovery of a temperate coastal region.** *Proceedings of the National Academy of Sciences* 115(14) 3658-3662.

Statistical methods: Beck and Murphy. 2017. **Numerical and Qualitative Contrasts of Two Statistical Models for Water Quality Change in Tidal Waters.** *Journal of the American Water Resources Association (JAWRA)* 53(1):197–219.

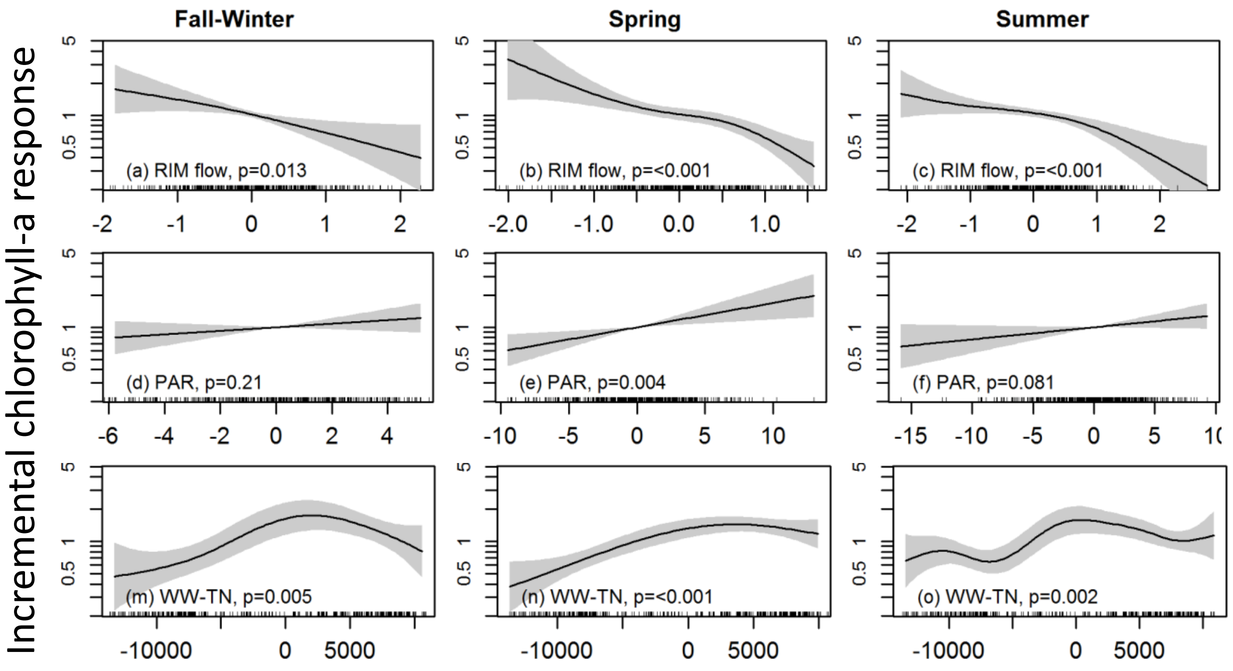
Mainstem chlorophyll-a seasonal changes analyzed multiple ways



From Testa et al. 2018

How the results are used: On-going investigations

Hypothesis testing with GAMs:
chla~ s(flow) + s(PAR) + s(WW-TN) *[All by season]*



Excerpt from Harris et al. study in-progress

- Patuxent and Rappahannock; Potomac exploratory analyses (Perry, Keisman, Murphy; Harris et al.)
- Maximum chlorophyll-a response (Murphy, Keisman)
- Phosphorus trends and changes in mainstem (Testa, Brady, Zhang and Murphy)

GAM implementation team

- Statistical Development: Elgin Perry
- baytrends package: Jon Harcum, Erik Leppo (Tetra Tech)
- Key user input and feedback: Renee Karrh (MDDNR), Mike Lane (ODU)

For more info, contact:

Jeni Keisman, USGS jkeisman@usgs.gov
Rebecca Murphy, UMCES at CBP rmurphy@chesapeakebay.net