Progress Report: Update on STAC Reviews for the Midpoint Assessment

STAC Review Panel for the Generalized Additive Model (GAM) Approach for Water Quality Trends in Tidal Waters

CBP Group requesting review: Scientific and Technical Analysis and Reporting (STAR) Team's Integrated Trends Analysis Team (ITAT)

CBP contacts: Rebecca Murphy, Jeni Keisman (with support from Elgin Perry, Jon Harcum)

Review panel:

Hugh Ellis (JHU) Chair – Environmental systems analysis, air pollution
Carl Friedrichs (VIMS) – Estuarine dynamics and water quality
Slava Lyubchich (UMCES) – Environmental statistics, trend clustering
Pang Du (VT) – Statistician, very familiar with GAM methods

Motivation:

In March 2014, the "Estimating Land Management Effects on Water Quality Status and Trends" (MEOWQT) STAC workshop recommended that GAM methods be adopted and applied trend detection and analysis. The previous CBP tool (Seasonal Kendall technique) could only be applied to monotonic variations.

What is a Generalized Additive Model (GAM) Approach for detecting and analyzing trends?

A GAM is a statistical model in which the expected value of a response variable E(y) can be modeled as a constant offset plus the sum of functions of explanatory variables:

$E(y) = a + f_1(x_1) + f_2(x_2) + f_3(x_3) + \ldots$

(General Linear Models, including multiple linear regression, are a subset of GAMs. In multiple linear regression, f_1, f_2, f_3 … are constant multipliers.)

In GAMs, f_1, f_2, f_3 … are usually "scatterplot smoothing functions" (typically types of splines), which (i) maximize the match to y while (ii) constraining their degree of smoothness (i.e., how many degrees of freedom are "used up" by the complexity of the splines).

f_1, f_2, f_3 … are often nonparametric functions meaning that their shapes are be entirely determined by the data rather than by set of parameters (as in a polynomial equation).

GAMs allow for diverse model shapes from linear to nonlinear – including patterns that change direction over time.

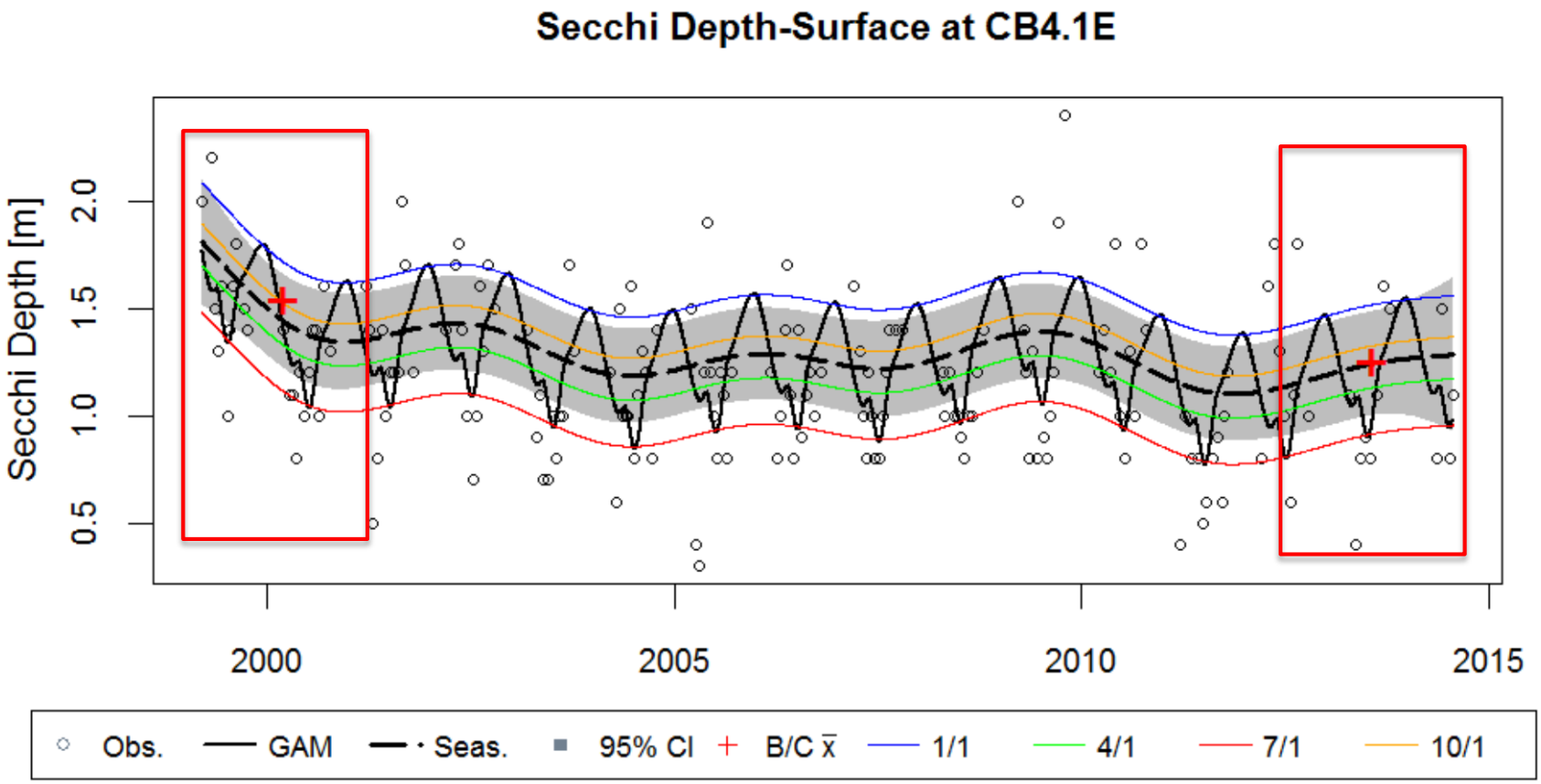

Table: GAM Analysis of Variance.

Type	Source	edf	F-stat	p-value
parametric terms	cyear	1.00	4.2950	0.0398
smoothed terms	s(cyear)	7.29	2.5963	0.0113
	s(doy)	5.05	3.8682	<0.0001
	ti(cyear,doy)	0.00	0.0000	0.7712

Table: GAM Parameter Coefficients.

Parameter	Estimate	Std. Err.	t value	p-value
(Intercept)	1.072385	0.053882	19.9025	<0.0001
cyear	-0.207781	0.100259	-2.0724	0.0398

Table: Estimates of Change from 1999-2014.

Calculation	Estimate
Baseline mean	1.5399
Current mean	1.2546
Estimated difference	-0.2853
Std. Err. difference	0.1168
95% Confidence interval for difference	(-0.5141 , -0.0564)
Difference p-value	0.0156
Period of Record Percent Change Estimate (%)	-18.53%

GAM Approach for Tidal Trends Review Questions:

1. Were resource and references materials adequate for conducting this review?
2. Does ‘mgcv’ R address the STAC workshop recommendation to apply and automate GAM?
3. Comment on the three model options (gam0, gam1, gam2).
4. Is the GAM % Change calculation sufficient for concluding whether trends are up or down?
5. Is it appropriate to base conclusions on log-transformed results?
6. Is the Maximum Likelihood method to account for censored data reasonable?
7. Comment on suggested approaches to account for changes in lab and sampling methods.
8. Comment on continuing research toward a comprehensive flow-adjustment procedure.
9. Recommend any additional issues the CBP investigate regarding application of GAM.
10. Are there other technical approaches that can supplement the application of GAM?

Large majority of review questions were very easily addressed:

1. Were resource and references materials adequate for conducting this review?

- Yes, impressively so. Also, our minor comments were well addressed in the CBP response

2. Does 'mgcv' R address the STAC workshop recommendation to apply and automate GAM?

- Yes, this R package is an excellent choice. And our comments were well addressed.

3. Comment on the three model options (gam0, gam1, gam2).

- All three options (linear, non-linear, non-linear + interactions) can easily be retained.

5. Is it appropriate to base conclusions on log-transformed results?

- Yes, if applied carefully. And our minor comments were well addressed in the CBP response

6. Is the Maximum Likelihood method to account for censored data reasonable?

- Yes, if applied carefully. And our minor comments were well addressed in the CBP response.

7. Comment on suggested approaches to account for changes in lab and sampling methods.

- Continue following your proposed approaches. And our comments were well addressed.

8. Comment on continuing research toward a comprehensive flow-adjustment procedure.

- The draft flow-adjustment procedure is a promising extension of the modeling approach.

4. Is the GAM % Change calculation sufficient for concluding whether trends are up or down?

From GAM background material: “The percent change approach is to average the GAM-generated estimates over the first two years, or baseline period, and last two years, or current period. The standard error of this estimate can be computed as well (Appendix 2.4A). With the estimate and standard error, we can obtain tests of significance and confidence intervals.”

Concern: “aren’t the neighboring monthly output from mgcv less independent of each other than were the observed data? It’s not clear that they can be pooled in the simple way described in the report in order to calculate standard errors on the two-year averages at each end of the multi-year prediction.”

Response: “The estimate of percent change is a linear function of the parameter vector, which is in turn a linear function of the data. Thus its variance is readily available through a quadratic form defined by the covariance matrix of the GAM-estimates. Because our estimate of percent change is a linear function of the data, our estimate of the SE of percent change also propagates from the raw data and not from the twelve monthly values which are actually an intermediate computation.”

4. Is the GAM % Change calculation sufficient for concluding whether trends are up or down?

Concern: What makes this approach superior to using more traditional methods for comparing means (based on raw observations)?"

Response: A t-test or Wilcoxon rank sum test… assumes that all the variability within the two periods is random… Failing to account for [the] seasonal pattern will result in overestimation of stochastic variability and hence inflate the false negative rate of the test… One might consider extending traditional methods to include stratifying the data by month to compare the first and last two years. [This] conceptually equates to what we are doing… If the model is expanded to include terms for deterministic effects of flow, salinity, or wind, it will be easy to adjust the time comparison for these effects as well.

9. Recommend any additional issues the CBP investigate regarding application of GAM.

Suggestion: Allow seasonality to be turned on and off, and allow input to be annual averages so that the results of Harding can be reproduced as a test case.
Response: These are both relatively simple modifications which we will work on.

Suggestion: Allow seasonally-adjusted T, S, stratification, wind, TSS, nutrients, etc. to be utilized as additional explanatory variables.
Response: For now, these would be research-oriented efforts. We would gladly share the package with researchers who would like to incorporate climatic variables beyond river flow.

Concern: How to best assign uncertainties to the predictions of the individual model components of GAM appears unresolved.
Response: The 'mgcv' GAM documentation states "In general the p-values behave well, but neglecting smoothing parameter uncertainty means that they may be somewhat too low [with] highly correlated covariates." As long as we do not precisely think about a cutoff of something like $p<0.05$, but instead use use the p-values on the individual components to understand the relative variability of the components, we should be fine.

10. Are there other technical approaches that can supplement the application of GAM?.

Suggestion:

Recent STAC reports emphasize the importance of using of multiple models for management in the Chesapeake Bay. The use of multiple statistical analyses is especially cost-effective. In a larger sense across the entire CBP, "multiple model" comparisons between statistical models and process-based numerical models should also be promoted, funded, and otherwise facilitated.

The WRTDS approach and Seasonal Kendall trend technique should also be included going forward, as well as simple generalized linear models (GLMs) (multiple regression/ANCOVA). The degree of consistency among simple GLMs and various nonparametric models will always be enlightening. Another possibly relevant approach is Multivariate Adaptive Regression Splines (MARS) approach, especially for fitting multiple response variables.

Response:

We agree that a multiple-models approach is important, hence our comparison with WRTDS and working comparisons with Seasonal Kendall. We will continue to conduct these on a case-by-case basis, and will encourage and work with any researchers who want to compare their methods to this one. We will investigate the use of MARS for cases where we may want to be using two multiple response variables.

Detailed Response to: Scientific and Technical Advisory Committee Review of the Generalized Additive Model (GAM) Approach for Water Quality Trends in Tidal Waters April 2017

Summary

We are pleased with the reviewers' positive response to our GAM application to tidal water quality analysis in Chesapeake Bay and helpful suggestions. Below is a summary of the suggestions provided by the reviewers and our timeline for implementing them.

1. Recommendations we are currently working on (approximate 6 month time-frame):

- Develop a comprehensive list of method and lab changes by location and parameter for each sampling program.
- Test the method/lab-changes as interventions and begin to deal with related challenges such as multiple method changes in a short time period.
- Continue flow-adjustment method development.

2. **Recommendations that will be incorporated in the near-term (approximate one-year time-frame):**

- **Documentation additions: Our near-term plans are to compile documentation on the currently in-development version of the package which includes the intervention method and accounts for censored data. This documentation will include much of what the STAC review team evaluated, and their suggestions including:**
  - **More details on the Seasonal Kendall technique being replaced,**
  - **Documenting how the GAM approach built into 'mgcv' compares to the general category of GAM analyses,**
  - **Time cost experiments,**
  - **Documentation of defaults and other options,**
  - **Description of abbreviations in the legend,**
  - **More thorough description of uncertainty bounds and interpretation of component p-values, and**
  - More details explaining the percent change computation and computation of standard errors.
- Build flow-adjustment functionality into the R package.
- Publically available documentation of method and lab changes that are relevant to long-term trend analysis.
- Functionality to turn seasonality off and analyze once-per-year data.
- Continue sharing the package with research teams, and using GAMs as part of scientific research projects aimed towards hypothesis testing.

3. Recommendations that will be considered in future years:

- Putting the package on CRAN.
- Based on further research on the EM approach applied to GAMs, incorporate modifications to the algorithm that more accurately account for the uncertainty in the model.
- Address concurvity between smooth model components, as the statistical research develops on this topic.
- Additional functions built into the package that allow for more explanatory variables and a hypothesis-testing approach for research-purposes.
- Consider spatial correlations between stations in grouped-station analyses.
- Consider mixed model approach and residual autocorrelation.
- Encourage and participate in multiple-models approaches to statistical tidal trends analyses.