

Comparison of WRTDS and GAMs for evaluating long-term trends in chlorophyll

Marcus W. Beck¹ Rebecca Murphy²

¹ORISE, USEPA, Gulf Ecology Division, beck.marcus@epa.gov

²UMCES at Chesapeake Bay Program, rmurphy@chesapeakebay.net

October 16, 2015

Since the last call...

- Application of GAMs and WRTDS to 30 year time series of monthly chlorophyll at LE1.2 and TF1.6
- Development of similar methods for model fitting
- Model comparisons
- Development of simulated datasets to evaluate flow-normalization
- Simulation comparisons
- Conclusions

Spatial and temporal observations

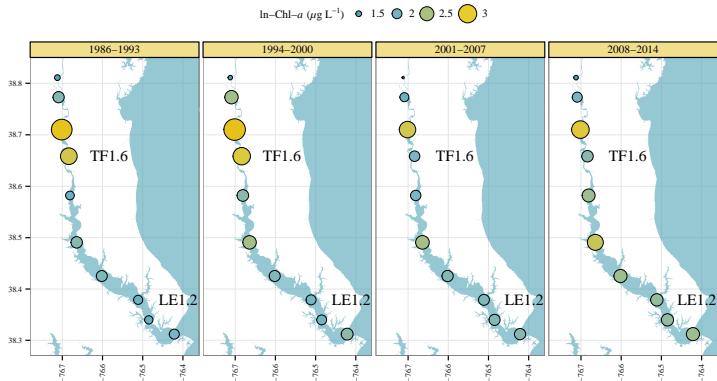


Figure : Observed chlorophyll at each monitoring station in the Patuxent, grouped by annual, seasonal, and flow periods.

Spatial and temporal observations

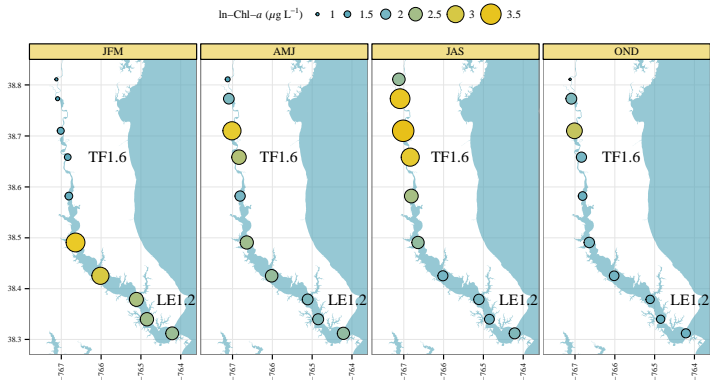


Figure : Observed chlorophyll at each monitoring station in the Patuxent, grouped by annual, seasonal, and flow periods.

Spatial and temporal observations

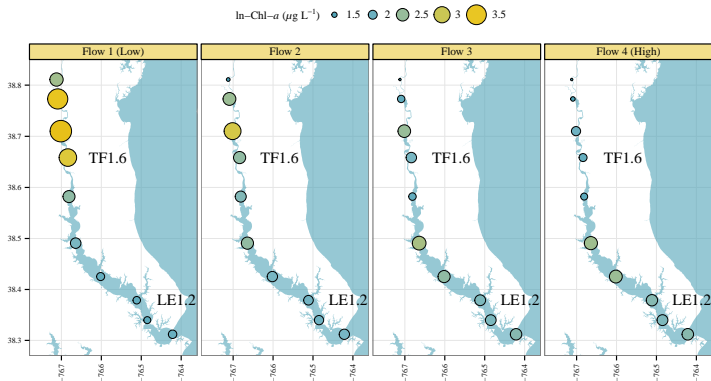


Figure : Observed chlorophyll at each monitoring station in the Patuxent, grouped by annual, seasonal, and flow periods.

Model applications

Both models used vertically-integrated chlorophyll, monthly timestep

LE1.2: $\ln chl a \sim \text{time} + \textit{salinity}$

TF1.6: $\ln chl a \sim \text{time} + \textit{flow}$

Fits evaluated for whole time series and annual/seasonal/flow aggregations:

- predicted to observed, GAM predicted to WRTDS predicted
- Trends in flow-normalized results (average and % change overall, by time period)

Model comparisons of flow-normalized results with simulated datasets

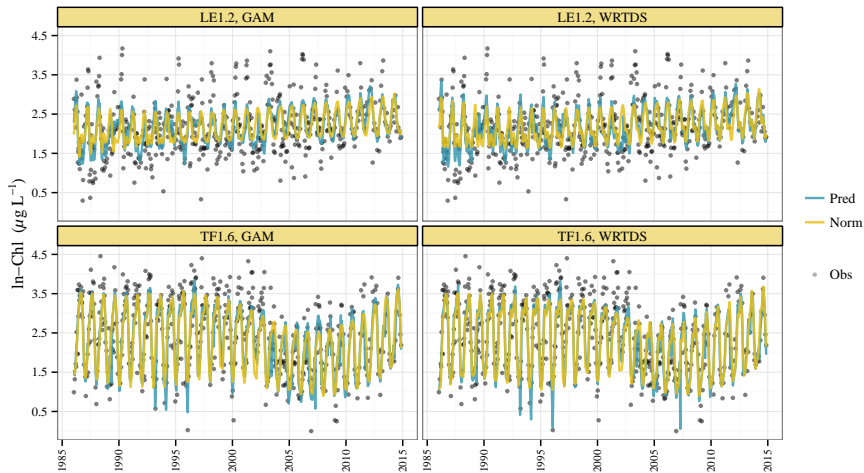
Development of comparable methods

Objective: Compare models on the same playing field

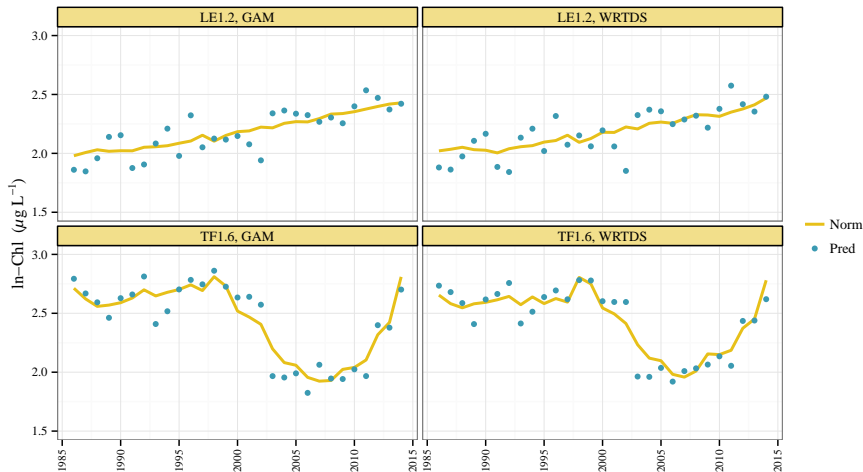
Problem: Need methods to prevent over-fitting for comparability

- GAMs: identify optimal smoothing parameter to balance fit and ‘wiggleness’
 - ▶ Iterative generalized cross-validation with penalized likelihood maximization to solve for model coefficients
- WRTDS: identify optimal window widths for time, discharge (salinity or flow), and season
 - ▶ k-fold cross-validation with search algorithm (‘limited memory BFGS quasi-Newton method’) that chooses window widths

Model comparisons



Model comparisons



Model comparisons

Table : RMSE of observed to predicted ln-chlorophyll.

Period	LE1.2		TF1.6	
	GAM	WRTDS	GAM	WRTDS
All				
	0.54	0.51	0.54	0.52
Annual				
1986-1993	0.54	0.50	0.53	0.49
1994-2000	0.52	0.50	0.58	0.58
2001-2007	0.63	0.60	0.54	0.53
2008-2014	0.39	0.36	0.49	0.44
Seasonal				
JFM	0.61	0.58	0.53	0.49
AMJ	0.69	0.64	0.60	0.58
JAS	0.38	0.35	0.48	0.46
OND	0.41	0.38	0.55	0.54
Flow				
1 (Low)	0.40	0.36	0.48	0.46
2	0.47	0.42	0.56	0.54
3	0.61	0.57	0.56	0.52
4 (High)	0.64	0.63	0.56	0.54

Model comparisons

Table : Comparison of predicted results between models.

Period	LE1.2		TF1.6	
	Ave. diff.	RMSE	Ave. diff.	RMSE
All				
	-0.11	0.15	0.01	0.17
Annual				
1986-1993	0.18	0.16	-0.78	0.17
1994-2000	0.53	0.15	-1.09	0.19
2001-2007	-0.95	0.14	0.48	0.14
2008-2014	-0.18	0.14	3.12	0.18
Seasonal				
JFM	2.91	0.14	-5.02	0.22
AMJ	-3.42	0.17	0.93	0.14
JAS	5.03	0.14	-0.10	0.17
OND	-5.25	0.14	2.08	0.17
Flow				
Flow 1 (Low)	0.19	0.16	-0.09	0.12
Flow 2	-0.83	0.16	0.73	0.15
Flow 3	0.19	0.15	0.84	0.20
Flow 4 (High)	0.03	0.13	-1.62	0.20

Model comparisons

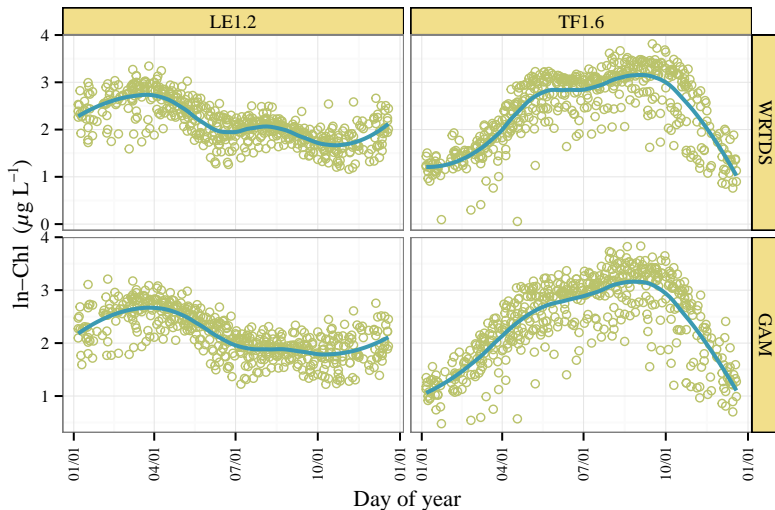


Figure : Seasonal variation from model predictions.

Model comparisons

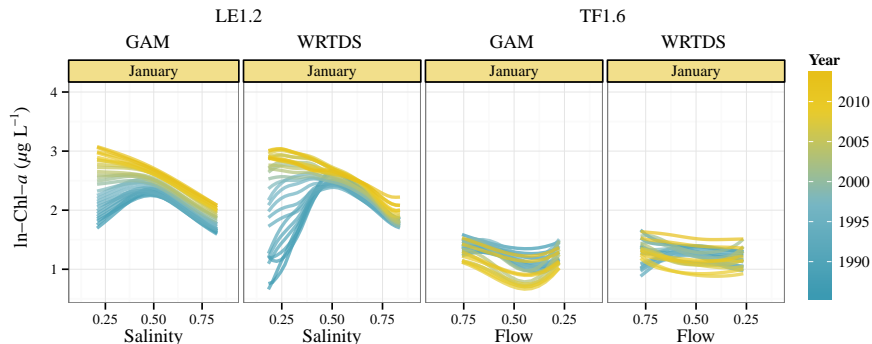


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

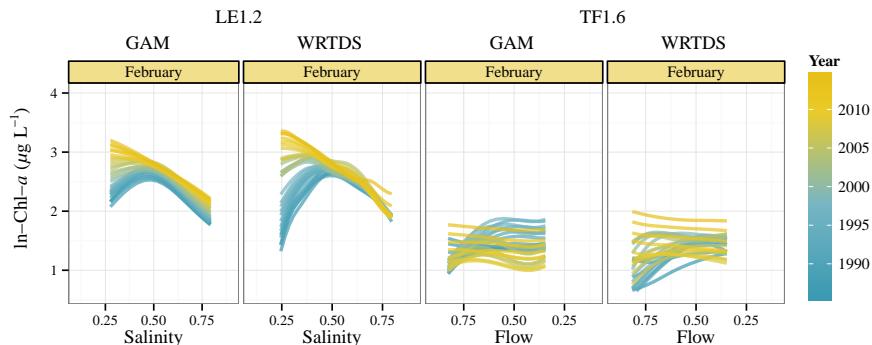


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

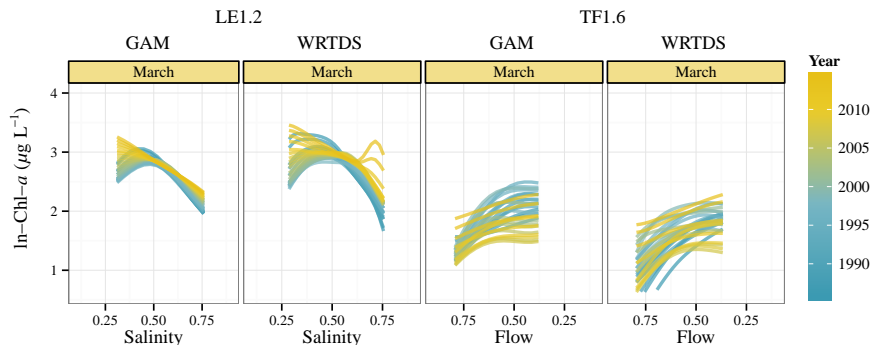


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

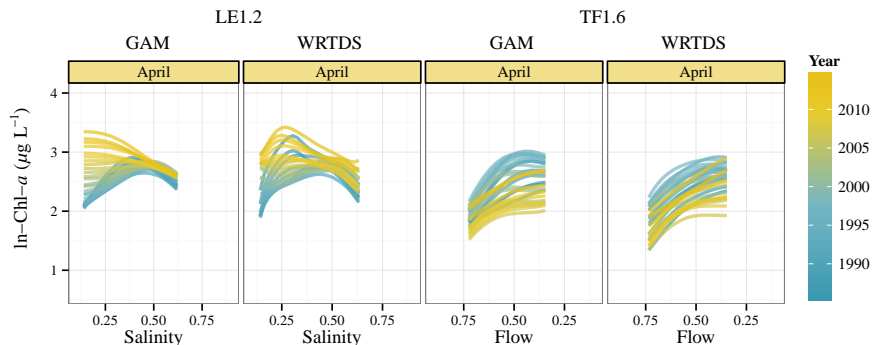


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

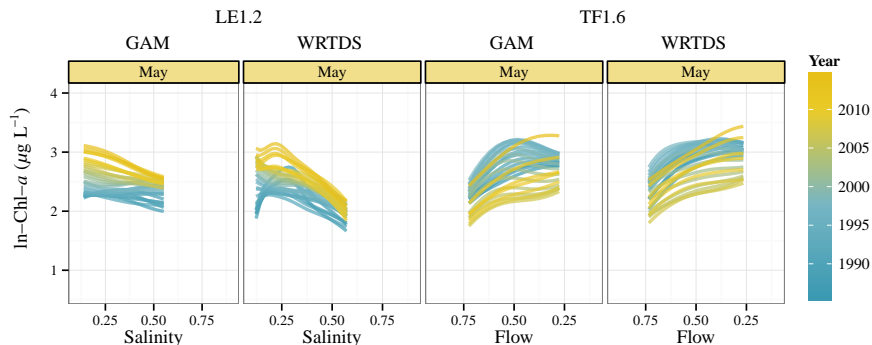


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

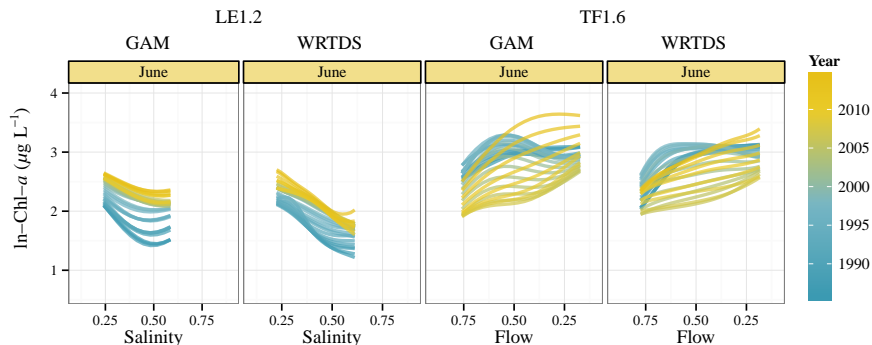


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

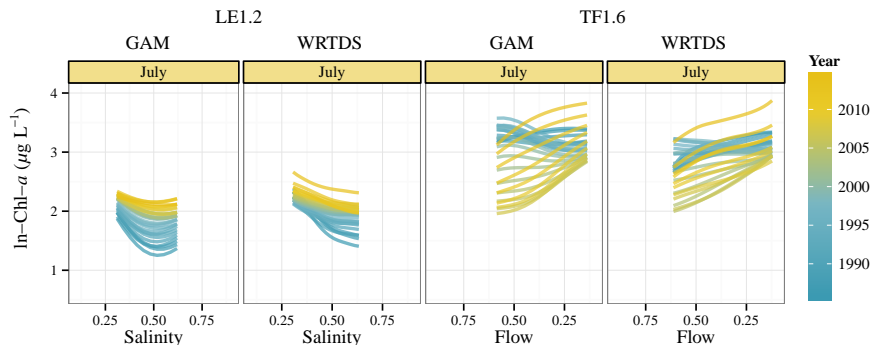


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

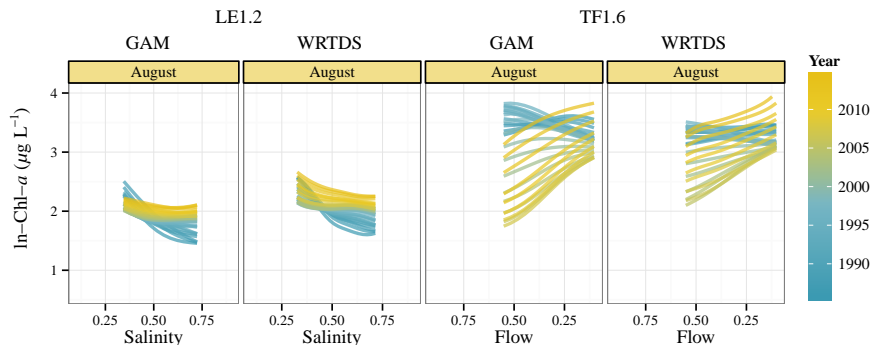


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

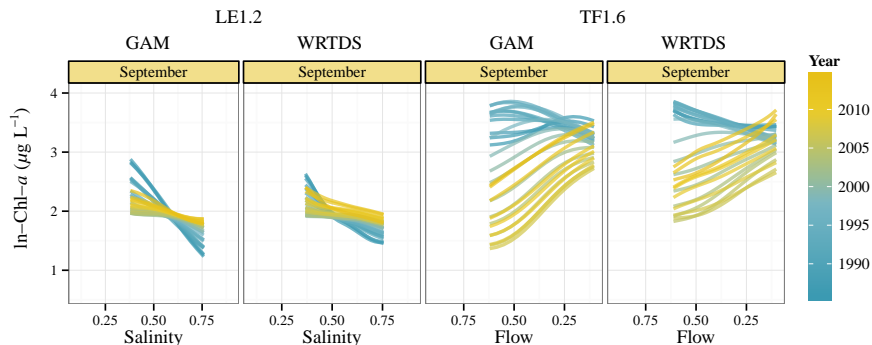


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

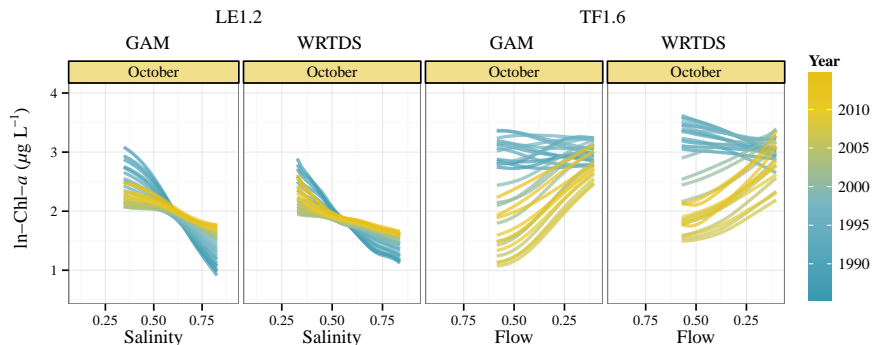


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

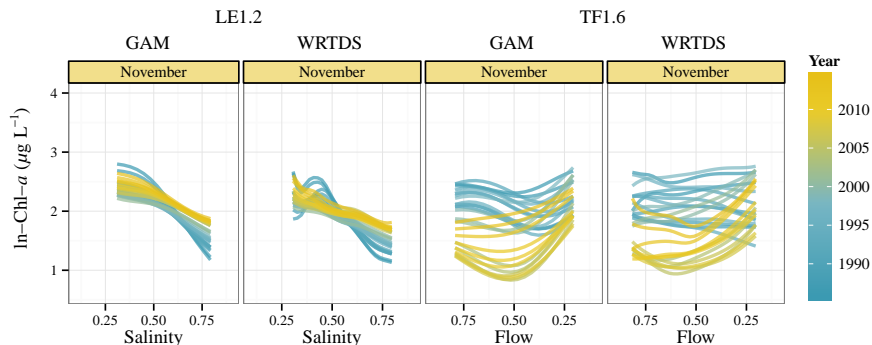


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Model comparisons

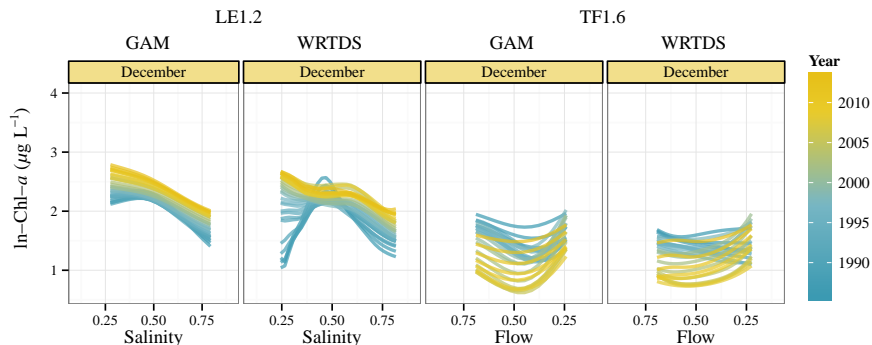


Figure : Changes in the relationship between chlorophyll and flow across the time series with separate plots by month, model, and station. The scales of salinity and flow are reversed for trend comparison with units in proportion of the total range for each month.

Development of simulated datasets

Objective: Evaluate ability of each model to reproduce flow-normalized trends

Problem: The true flow-normalized trends are not known and can only be empirically estimated

We simulated monthly datasets following techniques in Beck et al. 2015 and Hirsch et al. 2015

- Daily time series: Bowie gage discharge, Jug Bay fluorescence
- Overall: $Chl_{obs} = Chl_{flo} + Chl_{bio}$
- Three monthly datasets with different flow components: none, constant, increasing

Development of simulated datasets

Simulated data behaved as expected: three datasets with different flow contributions

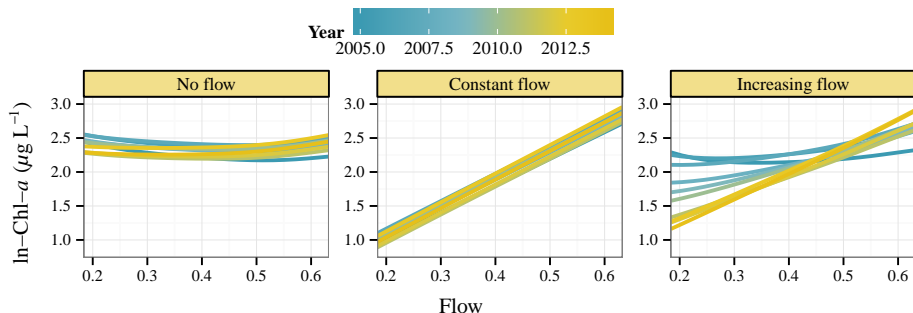


Figure : WRTDS predictions from August for three simulated datasets with different flow contributions.

Simulation comparisons

Objective: How well do flow-normalized predictions for each model reproduce Chl_{flo}

Simulated time series: $Chl_{obs} = Chl_{flo} + Chl_{bio}$

Table : Performance summaries (RMSE) of model predictions for the three simulated time series (no flow, constant flow, and increasing flow effect).

Simulations	$Chl_{obs} \sim \widehat{Chl}_{obs}$	$Chl_{bio} \sim \widehat{Chl}_{bio}$
No flow		
GAM	0.51	0.53
WRTDS	0.50	0.52
Constant flow		
GAM	0.51	0.58
WRTDS	0.53	0.57
Increasing flow		
GAM	0.51	0.54
WRTDS	0.50	0.52

Conclusions

- WRTDS prediction errors always less than GAMs
- Seasonal patterns were slightly different between models
- GAM estimates were more ‘stable’ during periods with little data
- No clear differences in flow-normalization abilities
- Interesting trends in Patuxent
 - ▶ Chlorophyll increasing lower estuary (LE1.2), mainstem influences
 - ▶ Multi-year signal of Isabel at TF1.6, flushing and low chlorophyll
 - ▶ Distinct changing relationships of chlorophyll with flow by station

Conclusions

Journal venue? Modelling vs ecosystem dynamics? Two papers?

Extra

For comparing each model's *predictions to observed*, at both sites:

$$RMSE_{fit} = \sqrt{\frac{\sum_{i=1}^n (Chl_i - \widehat{Chl}_i)^2}{n}}$$

For comparing *predictions between models*, at both sites:

$$RMSE_{btw} = \sqrt{\frac{\sum_{i=1}^n (\widehat{Chl}_{WRTDS,i} - \widehat{Chl}_{GAM,i})^2}{n}}$$

$$\text{Average difference} = \left(\frac{\sum_{i=1}^n \widehat{Chl}_{WRTDS,i} - \sum_{i=1}^n \widehat{Chl}_{GAM,i}}{\sum_{i=1}^n \widehat{Chl}_{GAM,i}} \right) * 100$$

Simulated datasets:

- Daily time series: Bowie gage discharge, Jug Bay fluorescence
- Overall: $Chl_{obs} = Chl_{flo} + Chl_{bio}$
- From discharge: $Chl_{flo} = I \left(\widehat{Q}_{seas} + \sigma \cdot \varepsilon_{Q, sim} \right)$
- From fluorescence: $Chl_{bio} = \widehat{Chl}_{seas} + \sigma \cdot \varepsilon_{Chl, sim}$
- Indicator I changes to simulate changing flow component

