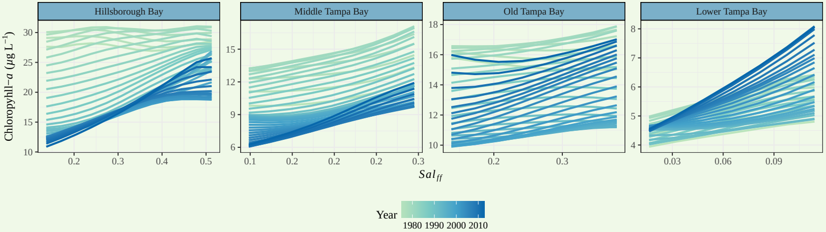
Quantitative approaches for biological assessment: Examples from Minnesota Lakes to Florida Estuaries

Marcus W. Beck, Ph.D.

USEPA National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, beck.marcus@epa.gov, Phone: 8509342480

May 22, 2017



Assessing environmental condition

How do we collect and use data?

The foundation of environmental management is a strong monitoring network [National Research Council, 1990]

Monitoring provides information for decision-making based on apparent trends...

What are the changes in environmental condition over time?

Are these changes 'good' or 'bad' based on our management objectives?

What may have caused these changes?

Assessing environmental condition

How do we collect and use data?

The good news: We are getting better at monitoring - standardized, automated, increased coverage, real-time/continuous

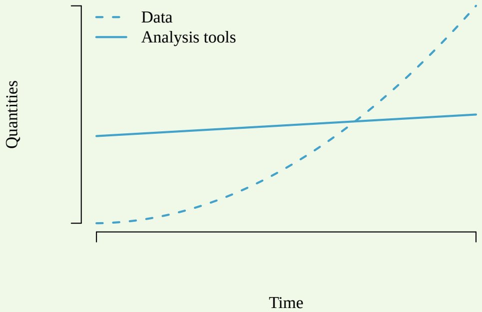
The bad news: Our ability to use these data for decision-making has not kept pace with availability!

Assessing environmental condition

How do we collect and use data?

The good news: We are getting better at monitoring - standardized, automated, increased coverage, real-time/continuous

The bad news: Our ability to use these data for decision-making has not kept pace with availability!



Assessing environmental condition

How do we collect and use data?

Today's talk: My experience using environmental data to understand causes and dynamics of water quality change

- ***Case 1***: Developing and quantifying response of a biotic index for Minnesota lakes
- ***Case 2***: Describing historical changes in eutrophication from long-term datasets in Tampa Bay

Assessing environmental condition

How do we collect and use data?

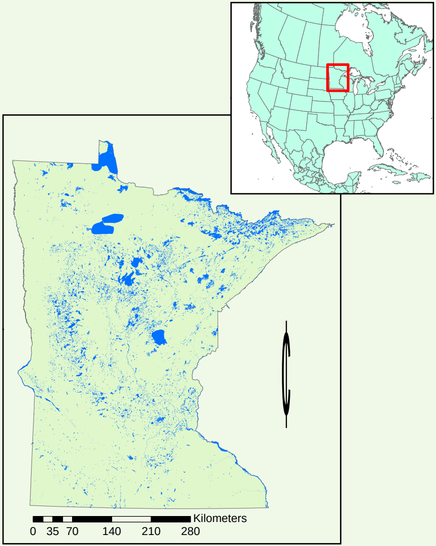
Today's talk: My experience using environmental data to understand causes and dynamics of water quality change

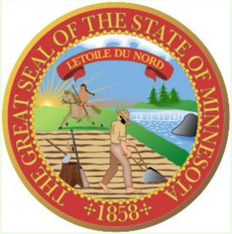
- ***Case 1***: Developing and quantifying response of a biotic index for Minnesota lakes
- ***Case 2***: Describing historical changes in eutrophication from long-term datasets in Tampa Bay

Each case addresses the challenges of ***evaluating change*** with biological endpoints and ***developing quantitative tools*** to improve efficiency and understanding

Case 1: Minnesota lakes

Evaluating biological response





Case 1: Minnesota lakes

Evaluating biological response

The macrophyte IBI can be used to evaluate relative lake condition by monitoring and evaluating aquatic plant metrics [Beck et al., 2010, Beck et al., 2013]

Case 1: Minnesota lakes

Evaluating biological response

The macrophyte IBI can be used to evaluate relative lake condition by monitoring and evaluating aquatic plant metrics [Beck et al., 2010, Beck et al., 2013]

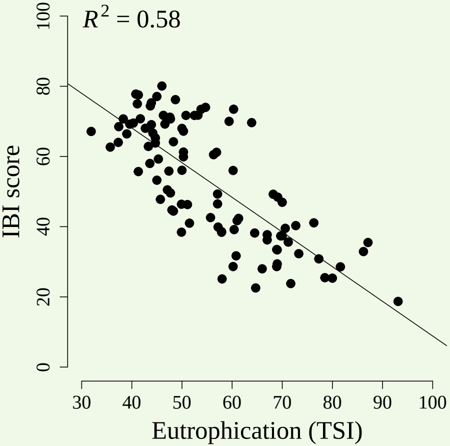
Includes eight metrics, summed to get one IBI score per lake

- MAXD: Maximum depth of plant growth
- LITT: Percentage of littoral zone vegetated
- OVER: Number of species with frequency occurrence >10%
- EMFL: Relative frequency of emergent-floating species
- SUBM: Relative frequency of submersed species
- SENS: Relative frequency of sensitive species
- TOLR: Relative frequency of tolerant species
- TAXA: Number of native taxa

Case 1: Minnesota lakes

Evaluating biological response

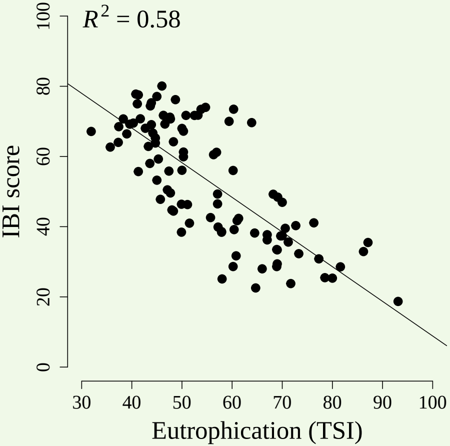
Related to changes in water quality



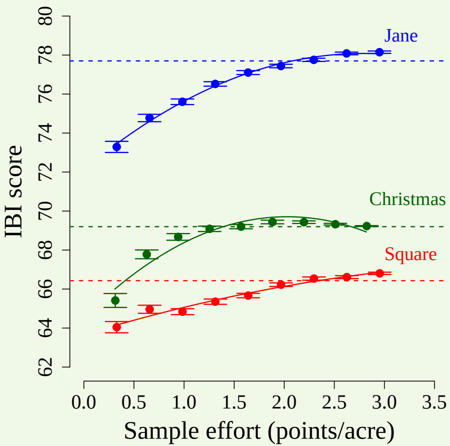
Case 1: Minnesota lakes

Evaluating biological response

Related to changes in water quality



High precision given sampling effort



Case 1: Minnesota lakes

Evaluating biological response

How appropriate is the IBI for characterizing effects of multiple stressors? Will it work within an assessment/impairment framework?

Case 1: Minnesota lakes

Evaluating biological response

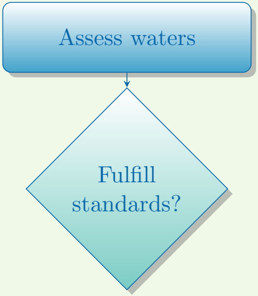
How appropriate is the IBI for characterizing effects of multiple stressors? Will it work within an assessment/impairment framework?

Assess waters

Case 1: Minnesota lakes

Evaluating biological response

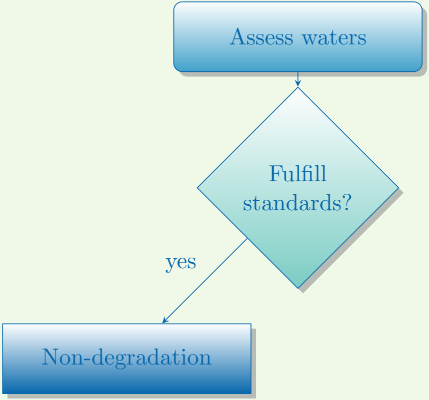
How appropriate is the IBI for characterizing effects of multiple stressors? Will it work within an assessment/impairment framework?



Case 1: Minnesota lakes

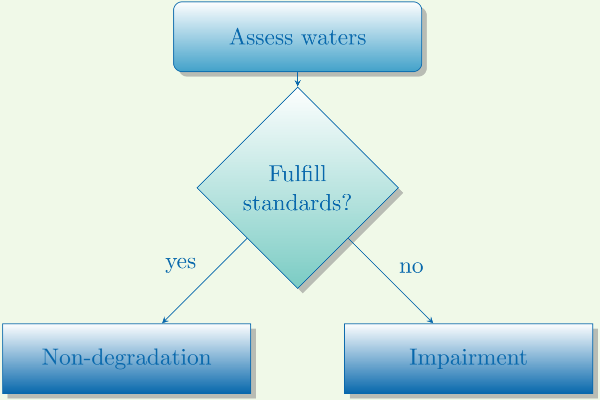
Evaluating biological response

How appropriate is the IBI for characterizing effects of multiple stressors? Will it work within an assessment/impairment framework?



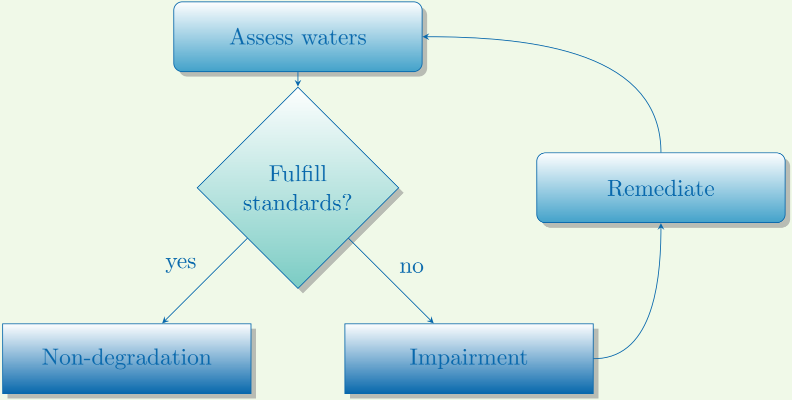
Case 1: Minnesota lakes
Evaluating biological response

How appropriate is the IBI for characterizing effects of multiple stressors? Will it work within an assessment/impairment framework?



Case 1: Minnesota lakes
Evaluating biological response

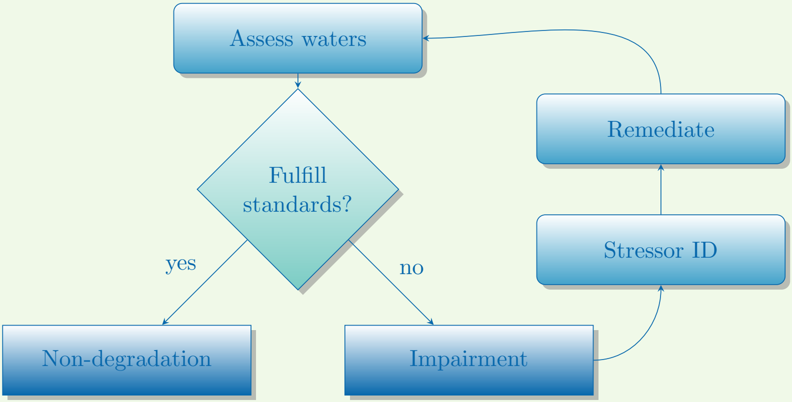
How appropriate is the IBI for characterizing effects of multiple stressors? Will it work within an assessment/impairment framework?



Case 1: Minnesota lakes

Evaluating biological response

How appropriate is the IBI for characterizing effects of multiple stressors? Will it work within an assessment/impairment framework?



Case 1: Minnesota lakes

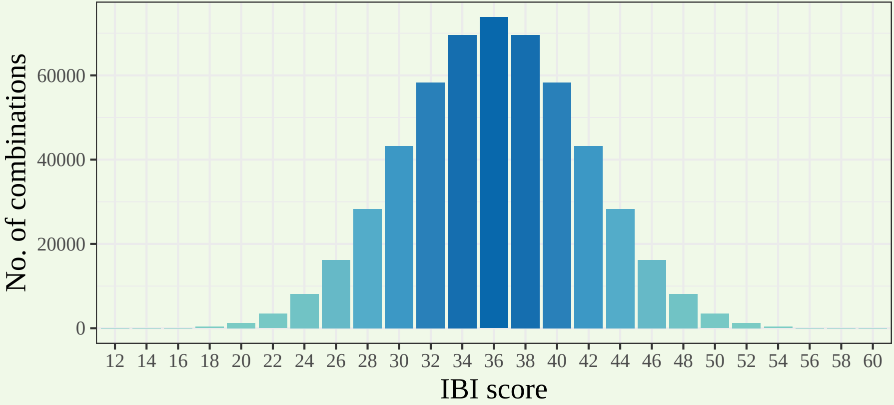
Evaluating biological response

Consider an IBI with 12 metrics, each scored 1, 3, or 5

How many different combinations of metrics lead to the same score?

Consider an IBI with 12 metrics, each scored 1, 3, or 5

How many different combinations of metrics lead to the same score?



Case 1: Minnesota lakes

Evaluating biological response

Develop and implement a framework for evaluating the macrophyte IBI to inform its use in biological monitoring:

Case 1: Minnesota lakes

Evaluating biological response

Develop and implement a framework for evaluating the macrophyte IBI to inform its use in biological monitoring:

1. How well does the index distinguish between signal and noise?

Case 1: Minnesota lakes

Evaluating biological response

Develop and implement a framework for evaluating the macrophyte IBI to inform its use in biological monitoring:

1. How well does the index distinguish between signal and noise?
2. Can information on stressors and their effects be quantified with certainty?

Case 1: Minnesota lakes

Evaluating biological response

Develop and implement a framework for evaluating the macrophyte IBI to inform its use in biological monitoring:

1. How well does the index distinguish between signal and noise?
2. Can information on stressors and their effects be quantified with certainty?
3. What stressors primarily influence index response?

Case 1: Minnesota lakes

Evaluating biological response

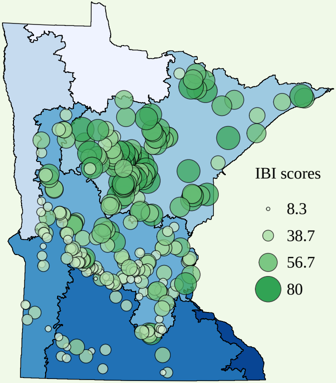
Develop and implement a framework for evaluating the macrophyte IBI to inform its use in biological monitoring:

1. How well does the index distinguish between signal and noise?
2. Can information on stressors and their effects be quantified with certainty?
3. What stressors primarily influence index response?
4. How appropriate is a multimetric index for characterizing effects of multiple stressors?

Case 1: Minnesota lakes

Evaluating biological response

- Dataset of 332 vegetation surveys, courtesy of MNDNR [Beck et al., 2014]
- Numerous covariates describing lake characteristics and anthropogenic stressors



- lake surface area
- maximum lake depth
- trophic state index
- growing degree days
- percent agriculture in wshed
- percent impervious surfaces in wshed
- density of groundwater wells in wshed
- wshed area to lake area
- crop productivity index of wshed
- dock density
- ...

Case 1: Minnesota lakes
Evaluating biological response

Ecological and numerical complexity warrants the use of creative solutions

Neural networks to model IBI response

Case 1: Minnesota lakes

Evaluating biological response

Ecological and numerical complexity warrants the use of creative solutions

Neural networks to model IBI response

- Essentially a large, non-linear regression model free of assumptions that can handle multivariate response

Case 1: Minnesota lakes

Evaluating biological response

Ecological and numerical complexity warrants the use of creative solutions

Neural networks to model IBI response

- Essentially a large, non-linear regression model free of assumptions that can handle multivariate response
- Models relationships among variables using a network that mimics neuronal structure of the human brain

Case 1: Minnesota lakes

Evaluating biological response

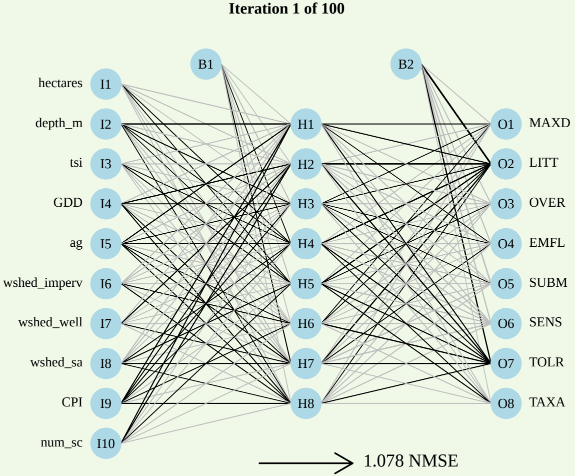
Ecological and numerical complexity warrants the use of creative solutions

Neural networks to model IBI response

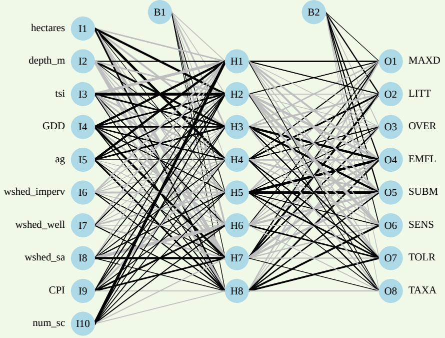
- Essentially a large, non-linear regression model free of assumptions that can handle multivariate response
- Models relationships among variables using a network that mimics neuronal structure of the human brain
- 'Supervised' neural networks are meant for prediction but network information can be used to infer causation

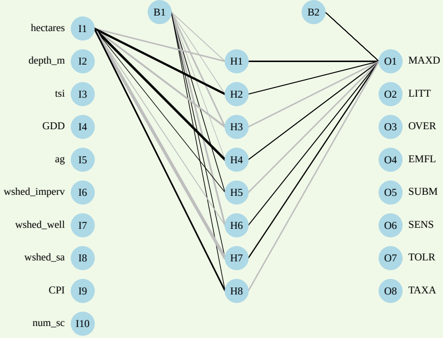
Case 1: Minnesota lakes

Evaluating biological response

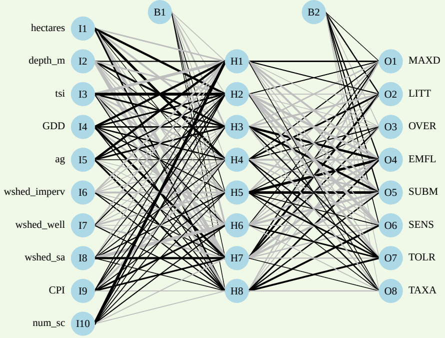


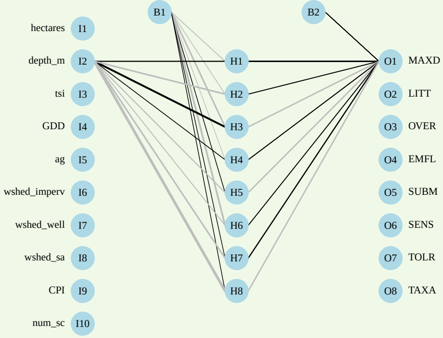
Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]





Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]

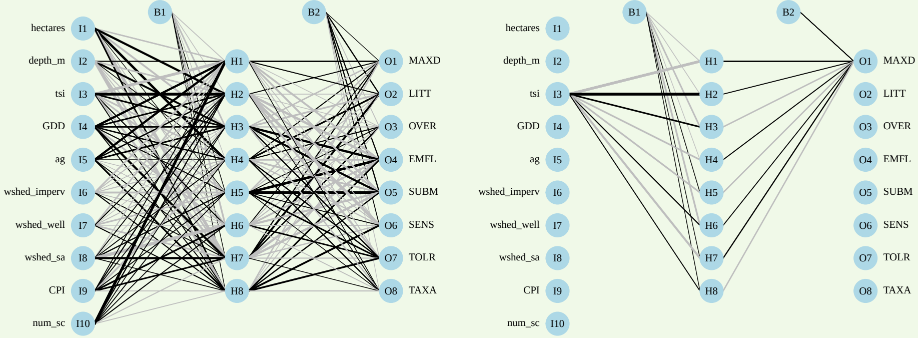




Case 1: Minnesota lakes

Evaluating biological response

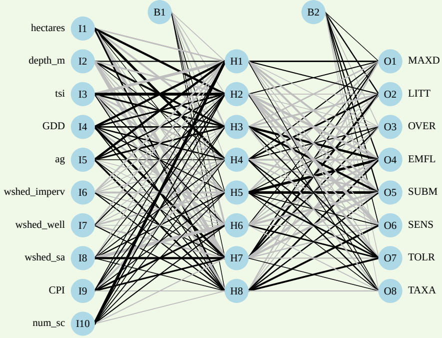
Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]

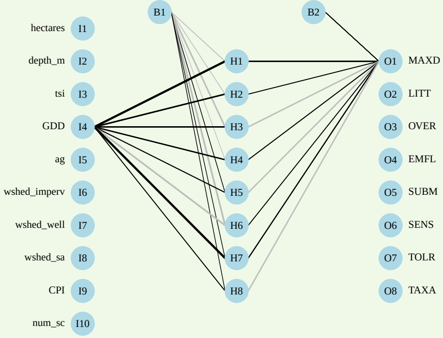


Case 1: Minnesota lakes

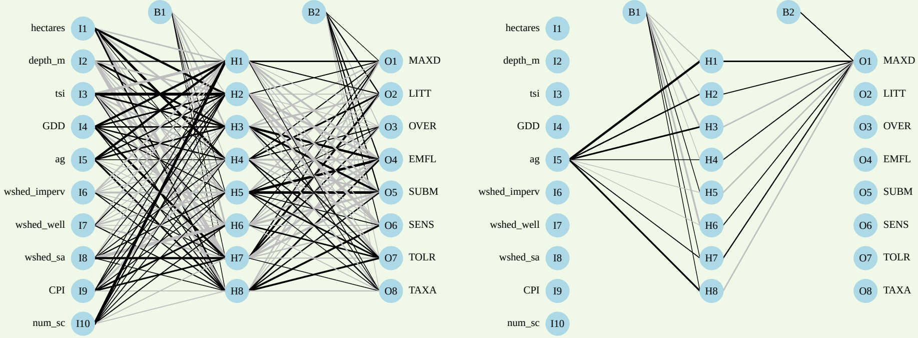
Evaluating biological response

Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]





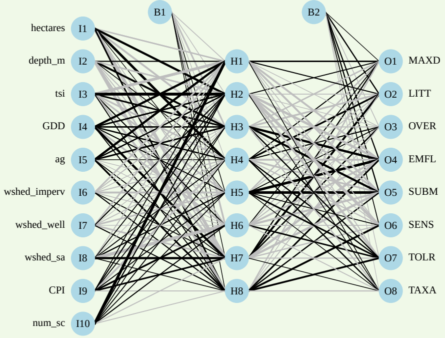
Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]

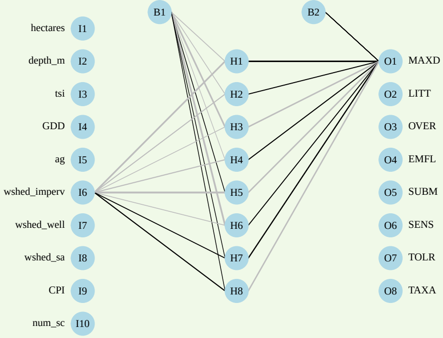


Case 1: Minnesota lakes

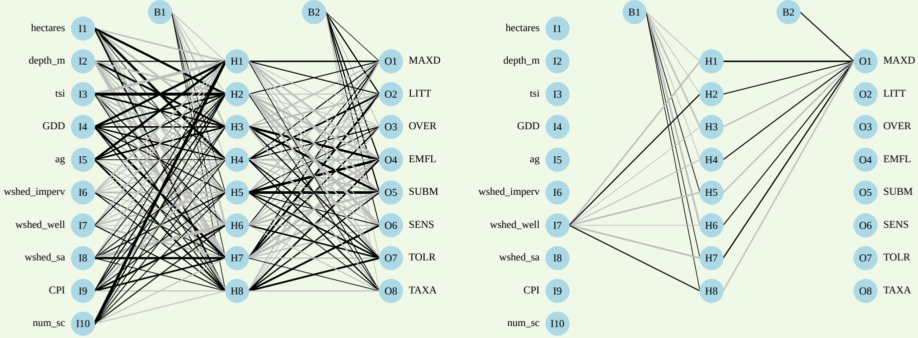
Evaluating biological response

Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]

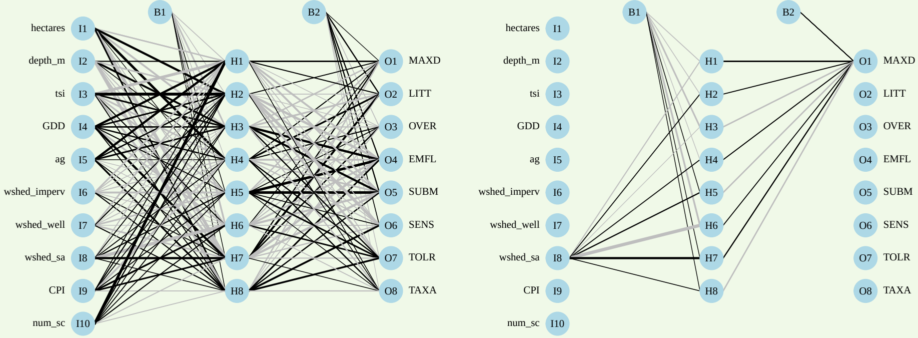




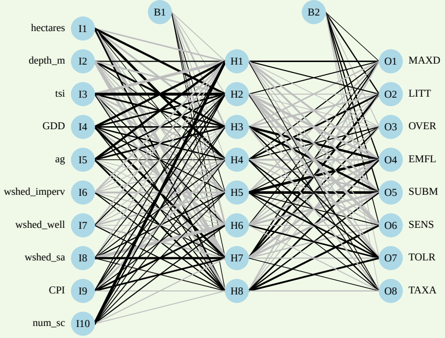
Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]

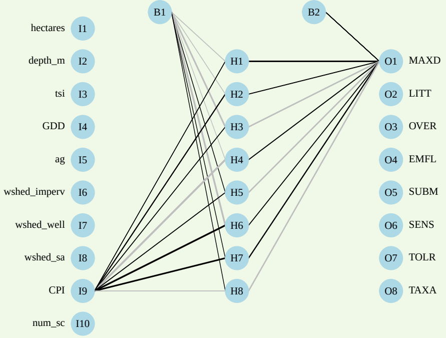


Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]



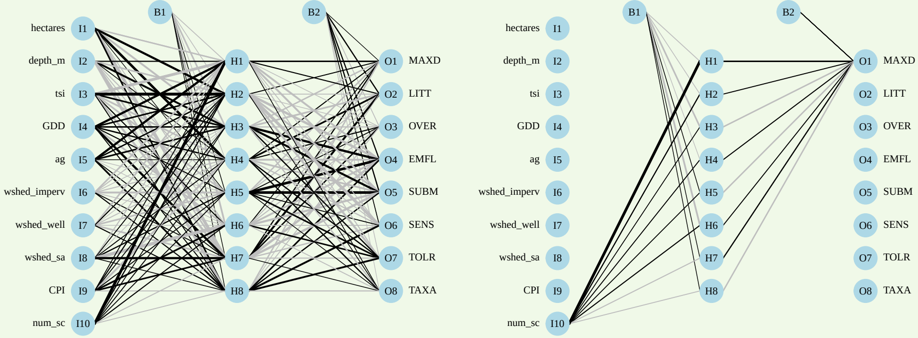
Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]



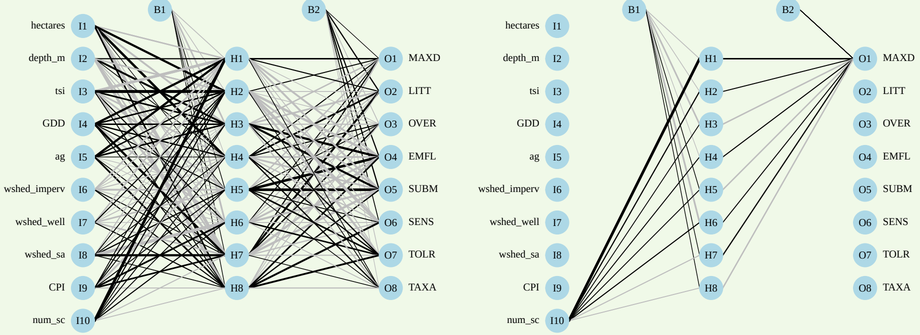


Case 1: Minnesota lakes
Evaluating biological response

Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]



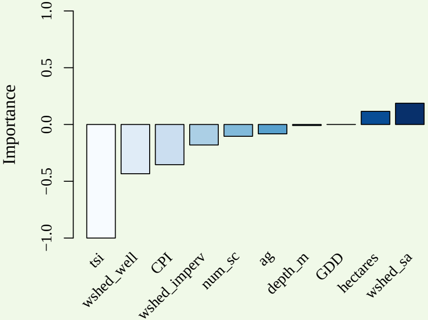
Modification of Garson's algorithm to determine relative importance of variables [Beck, in press]



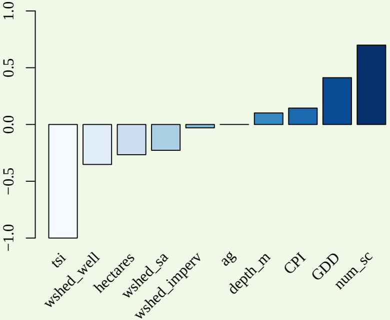
Relative importance is summation of product of weights between layers

Case 1: Minnesota lakes

Evaluating biological response



(a) IBI scores



(b) MAXD metric

Figure: Examples of relative importance of explanatory variables based on weights between layers in optimal neural networks.

Case 1: Minnesota lakes

Evaluating biological response

Neural networks are powerful enough to model noise in the data

The model may be specific to peculiarities the training dataset

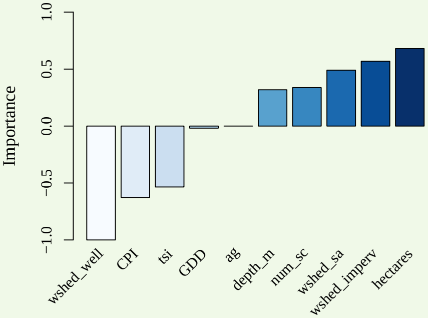
Uncertainty of variable importance must be quantified - bootstrap!

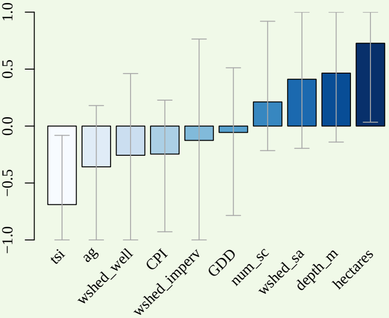
Case 1: Minnesota lakes
Evaluating biological response

Neural networks are powerful enough to model noise in the data

The model may be specific to peculiarities the training dataset

Uncertainty of variable importance must be quantified - bootstrap!





Case 1: Minnesota lakes

Evaluating biological response

Lots of uncertainty associated with input contributions...

- One of ten relationships for IBI scores with explanatory variables
- Four of 80 relationships for metrics with explanatory variables

Case 1: Minnesota lakes

Evaluating biological response

Lots of uncertainty associated with input contributions...

- One of ten relationships for IBI scores with explanatory variables
- Four of 80 relationships for metrics with explanatory variables
- IBI negatively related to lake trophic state
- MAXD, OVER, and TAXA negatively related to lake trophic state
- TAXA positively related to lake size

Case 1: Minnesota lakes

Evaluating biological response

Lots of uncertainty associated with input contributions...

- One of ten relationships for IBI scores with explanatory variables
- Four of 80 relationships for metrics with explanatory variables
- IBI negatively related to lake trophic state
- MAXD, OVER, and TAXA negatively related to lake trophic state
- TAXA positively related to lake size

What is the source of this uncertainty? Sample sizes, method, neural network?

Potentially competing objectives of an IBI: certainty vs simplicity

Case 1: Minnesota lakes

Evaluating biological response

Potentially competing objectives of an IBI: certainty vs simplicity

> *IBI relies on multiparameters, a requirement when the system to be evaluated is complex. [Karr et al., 1986]*

> *The resulting index allows people without specialized expertise to describe overall condition and to make informed decisions that will affect the health of those resources. [Karr and Chu, 1999]*

Case 1: Minnesota lakes
Evaluating biological response

Potentially competing objectives of an IBI: certainty vs simplicity

IBI relies on multiparameters, a requirement when the system to be evaluated is complex. [Karr et al., 1986]

The resulting index allows people without specialized expertise to describe overall condition and to make informed decisions that will affect the health of those resources. [Karr and Chu, 1999]

Combining responses into an index hides the component responses, thereby obscuring causation. [Suter, 1993]

Case 2: Florida estuaries

Evaluating long-term water quality datasets

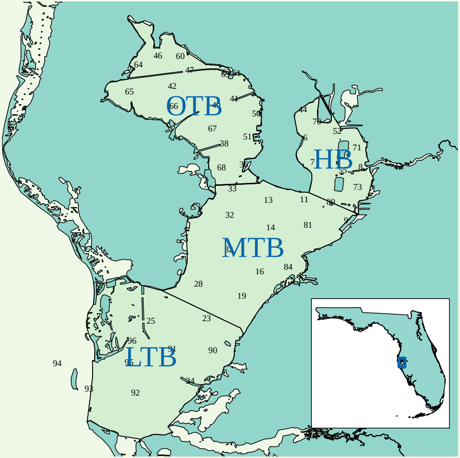
USEPA Gulf Ecology Division - guidance to Florida DEP and others on criteria development for coastal areas

Case 2: Florida estuaries

Evaluating long-term water quality datasets

- Four bay segments
- Monthly wq data at 50 stations from 1974 to present
- Longitudinal profile of nutrient load and salinity

Data from [TBEP (Tampa Bay Estuary Program), 2011]



Case 2: Florida estuaries

Evaluating long-term water quality datasets

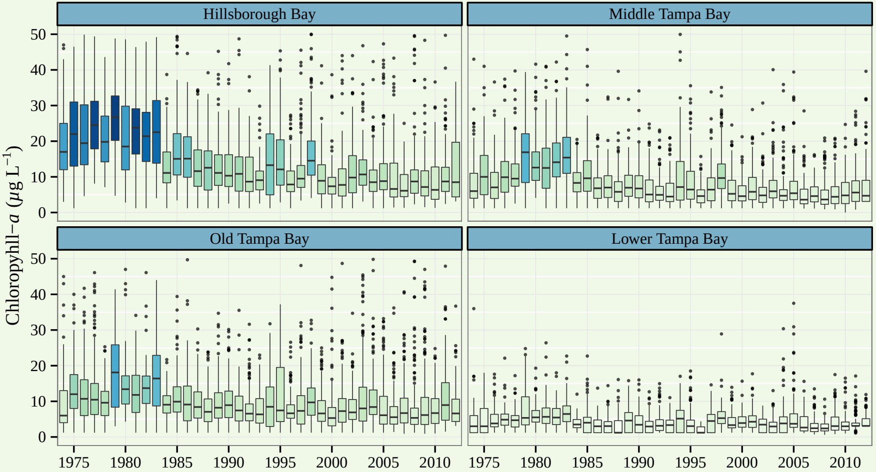


Figure: Annual trends in chlorophyll for each bay segment.

Case 2: Florida estuaries

Evaluating long-term water quality datasets

What affects our interpretation of chlorophyll response to nutrients?

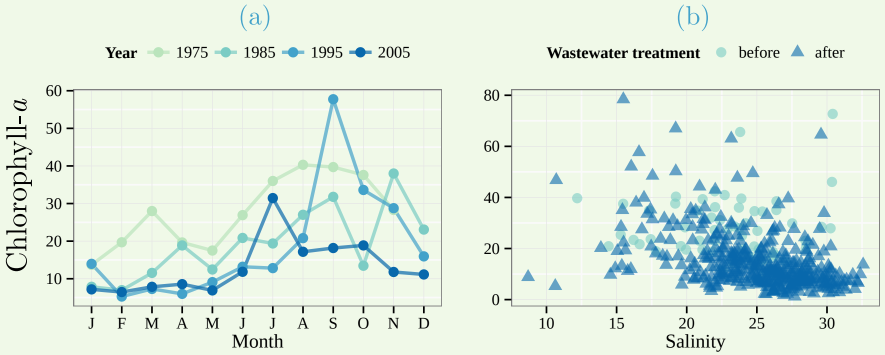


Figure: Variation in chlorophyll by (a) time and (b) salinity and management in Hillsborough Bay. Panel (a) is colored before and after wastewater treatment in 1979.

Case 2: Florida estuaries

Evaluating long-term water quality datasets

Study objective

Adapt and apply a nutrient response model for estuaries that leverages the descriptive capabilities of large datasets [Beck and Hagy III, 2015]

Case 2: Florida estuaries

Evaluating long-term water quality datasets

Study objective

Adapt and apply a nutrient response model for estuaries that leverages the descriptive capabilities of large datasets [Beck and Hagy III, 2015]

Questions of concern – Can we...

- ...provide a natural history of water quality that is temporally consistent with drivers of change?

Case 2: Florida estuaries

Evaluating long-term water quality datasets

Study objective

Adapt and apply a nutrient response model for estuaries that leverages the descriptive capabilities of large datasets [Beck and Hagy III, 2015]

Questions of concern – Can we...

- ...provide a natural history of water quality that is temporally consistent with drivers of change?
- ...improve our understanding of the nutrient-response paradigm in estuaries?

Case 2: Florida estuaries

Evaluating long-term water quality datasets

Weighted **R**egression on **T**ime, **D**ischarge, and **S**eason

- Describes a time series in the context of these parameters, locally fitted
- Useful to describe long-term trends, ie., multi-decadal time series
- Evaluation of flow-normalized trends, hypothesis generation

Case 2: Florida estuaries

Evaluating long-term water quality datasets

Weighted **R**egression on **T**ime, **D**ischarge, and **S**eason

- Describes a time series in the context of these parameters, locally fitted
- Useful to describe long-term trends, ie., multi-decadal time series
- Evaluation of flow-normalized trends, hypothesis generation

Developed by [Hirsch et al., 2010] for pollutants in stream/rivers

Adapted for tidal waters by [Beck and Hagy III, 2015]

Case 2: Florida estuaries

Evaluating long-term water quality datasets

How does it work?

$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 Sal + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

N: nitrogen (or other response endpoint)
t: time
Sal: Salinity (or other flow-related variable)

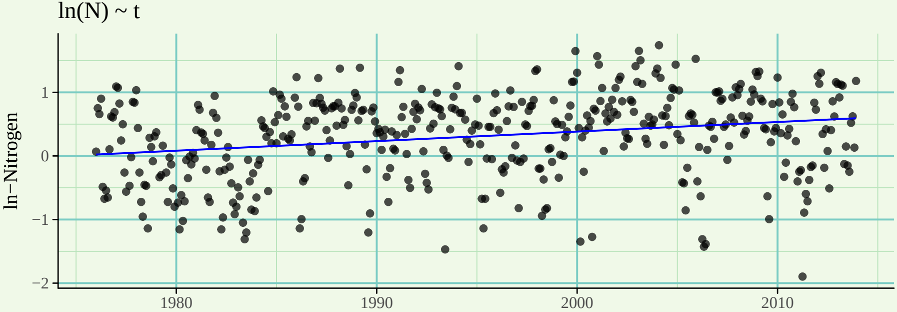
Case 2: Florida estuaries

Evaluating long-term water quality datasets

How does it work?

$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 Sal + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

N: nitrogen (or other response endpoint)
t: time
Sal: Salinity (or other flow-related variable)



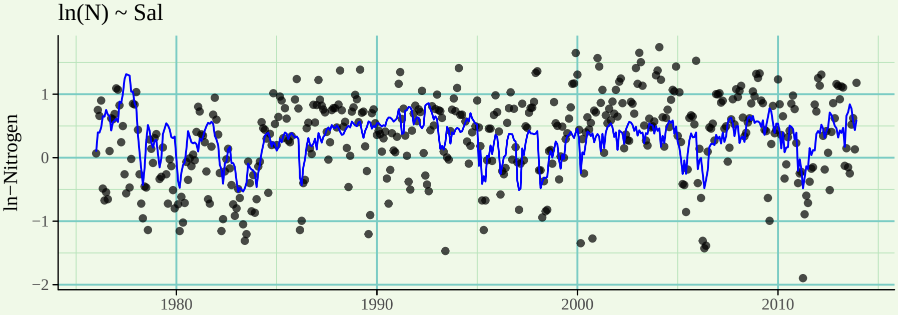
Case 2: Florida estuaries

Evaluating long-term water quality datasets

How does it work?

$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 Sal + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

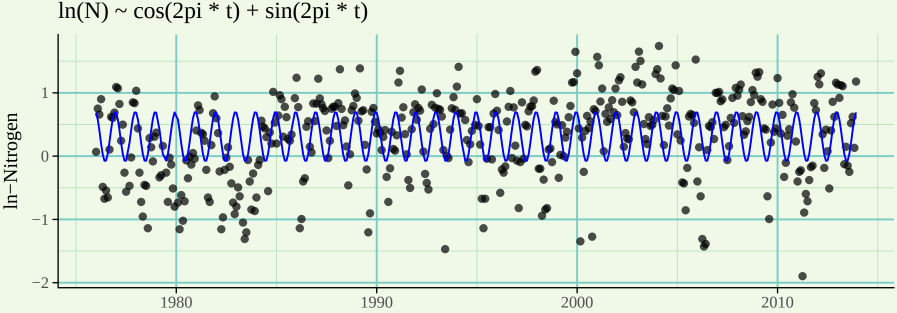
N: nitrogen (or other response endpoint)
t: time
Sal: Salinity (or other flow-related variable)



How does it work?

$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 Sal + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

N: nitrogen (or other response endpoint)
t: time
Sal: Salinity (or other flow-related variable)



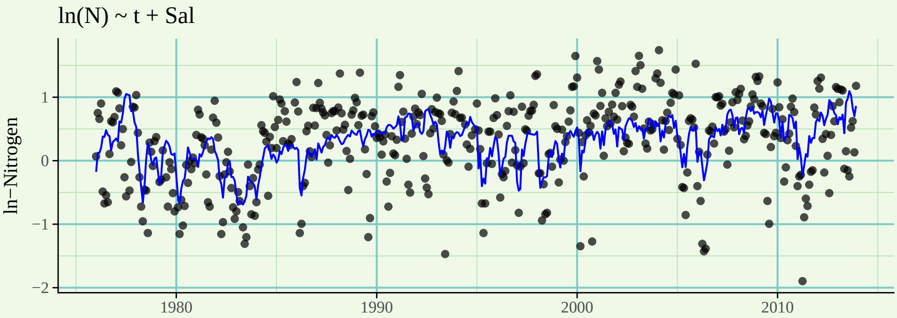
Case 2: Florida estuaries

Evaluating long-term water quality datasets

How does it work?

$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 Sal + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

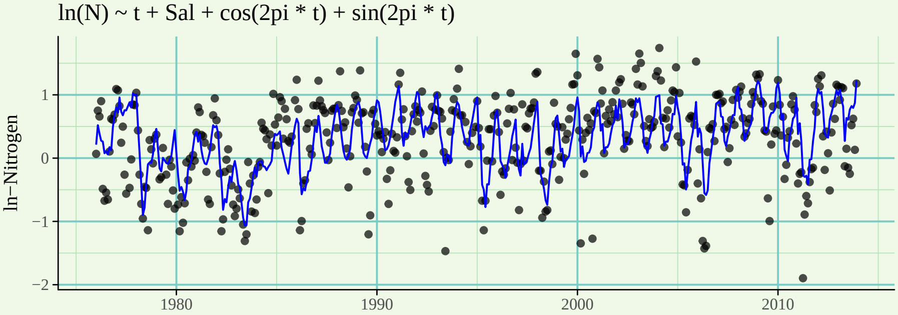
N: nitrogen (or other response endpoint)
t: time
Sal: Salinity (or other flow-related variable)



How does it work?

$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 Sal + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

N: nitrogen (or other response endpoint)
t: time
Sal: Salinity (or other flow-related variable)



Case 2: Florida estuaries

Evaluating long-term water quality datasets

This is not the whole story...

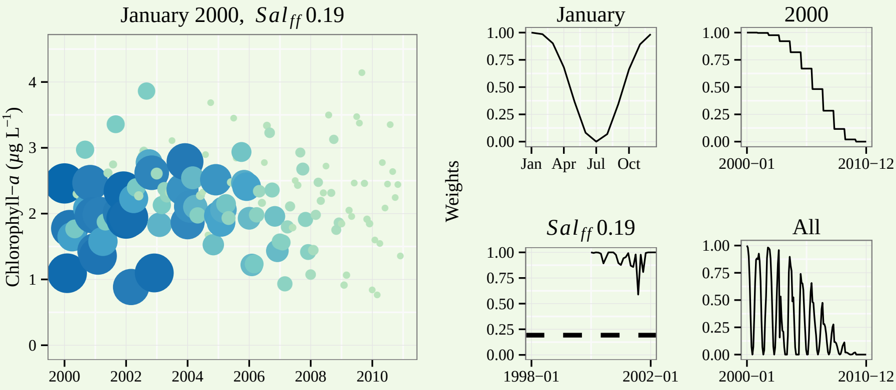
$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 Sal + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

One parameter set to many parameter sets - a moving window regression

Within each window, a unique regression is fit, weighted by the local salinity, time, and season

Similar to a loess/spline smooth but specific to the effects of these three variables on the response

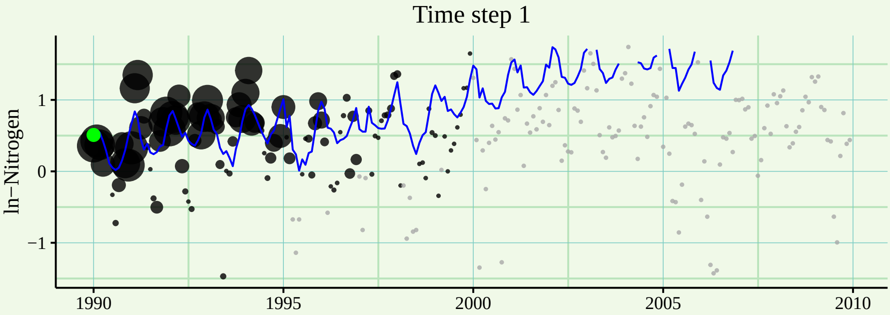
How does weighted regression work?



Case 2: Florida estuaries

Evaluating long-term water quality datasets

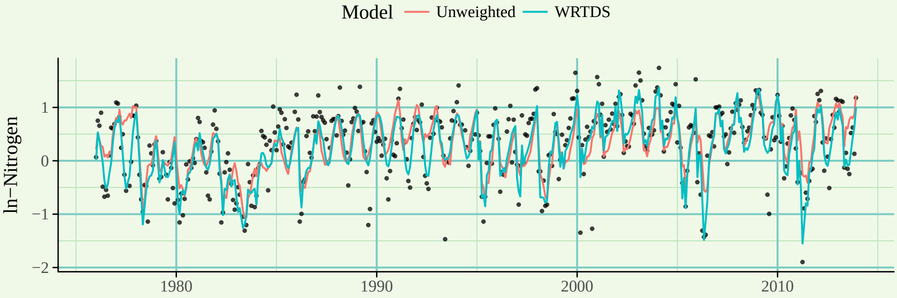
Points: observed time series (black are weighted, grey is zero weight)
Green point: observation at the center of the regression
Blue line: Global model with weights specific to the window
Red line: Accumulated WRTDS model



Case 2: Florida estuaries

Evaluating long-term water quality datasets

RMSE fit for unweighted = 0.58, WRTDS = 0.36



This gives us improved predictions of chlorophyll dynamics...

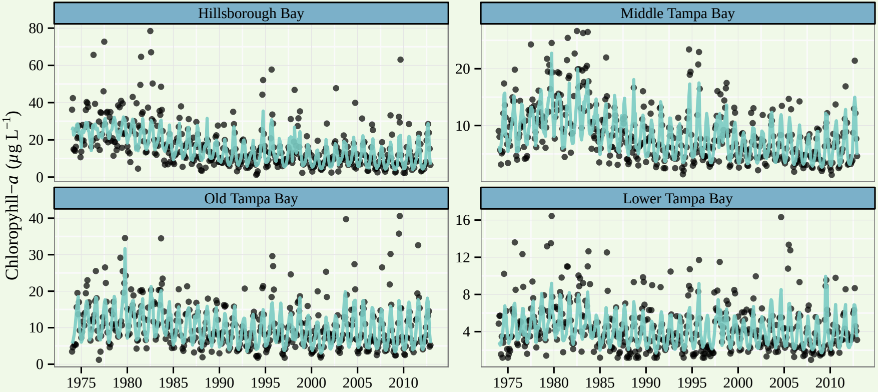
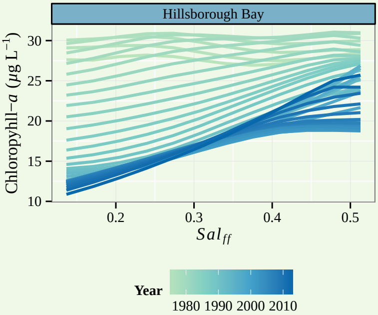


Figure: Predicted and observed monthly chlorophyll by segment.

Case 2: Florida estuaries

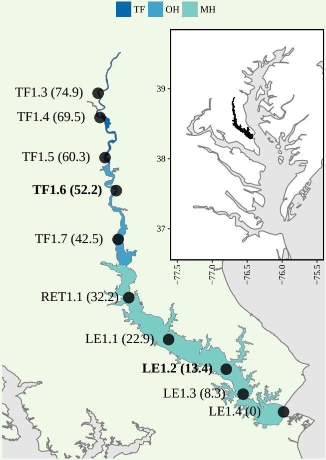
Evaluating long-term water quality datasets

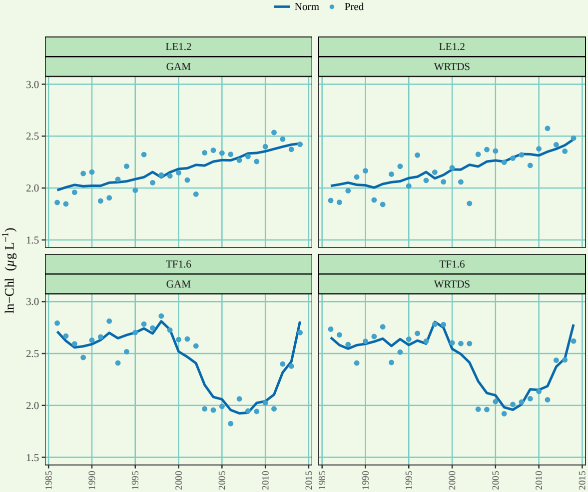
Because the model is dynamic, we have parameters describing the relationship of chlorophyll with other factors specific to different time periods



- Early period (light blue) - point-sources
- Late period (dark blue) - non-point sources
- Chlorophyll shows increasing response to freshwater input in recent years

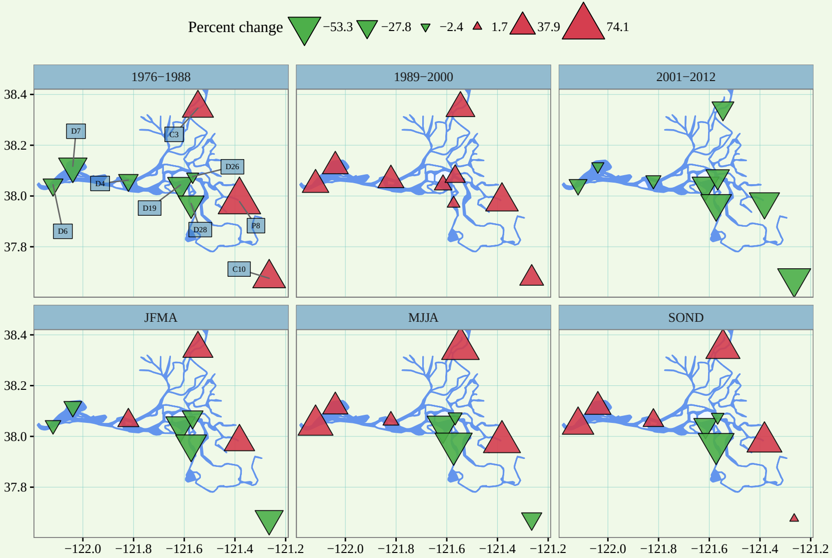
Comparing models for trend evaluation [Beck and Murphy, 2017]





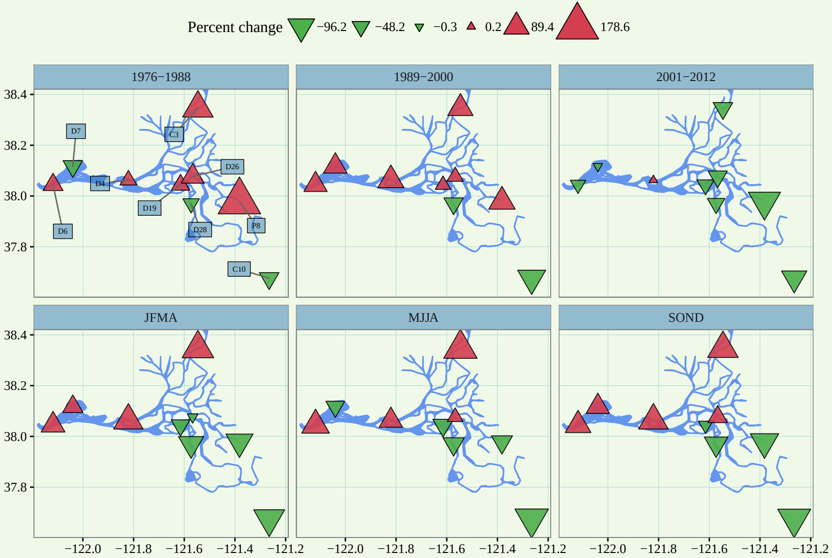
Application to other systems

Nitrogen dynamics in the Delta - DIN



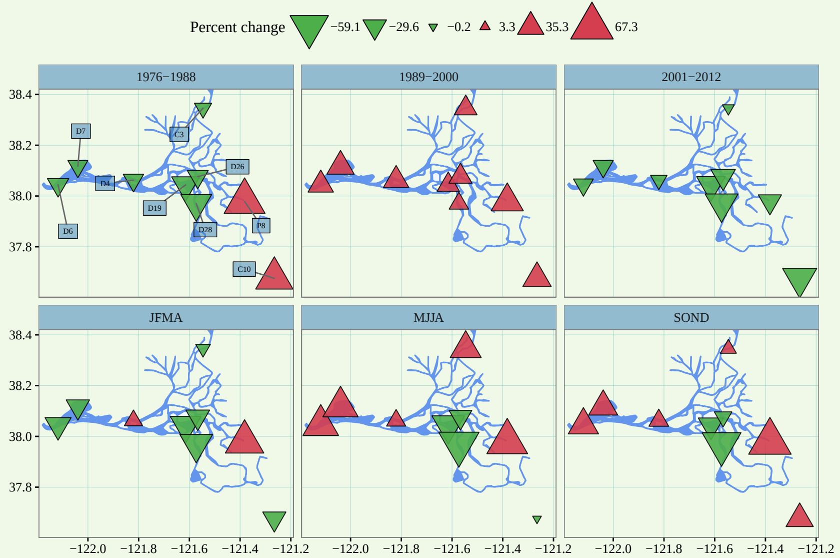
Application to other systems

Nitrogen dynamics in the Delta - ammonium



Application to other systems

Nitrogen dynamics in the Delta - nitrite/nitate



Conclusions

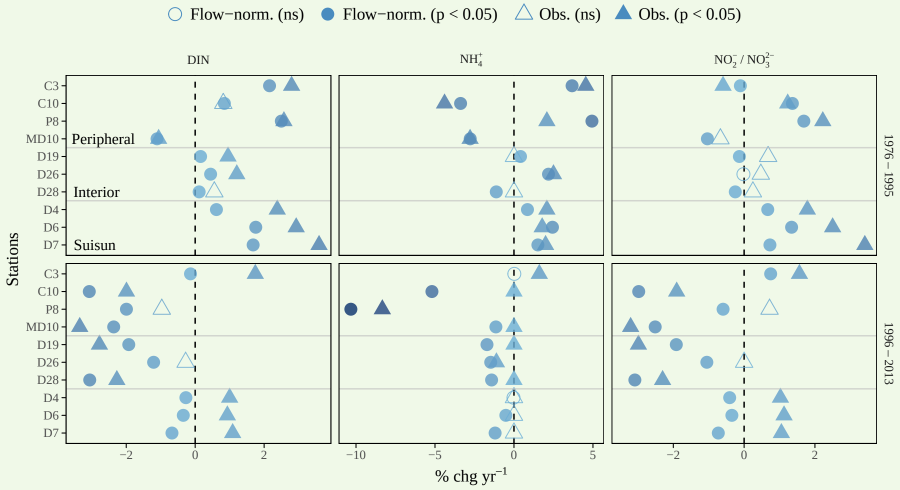
What does this mean for Tampa Bay and elsewhere?

- Better description of biological endpoints can change conclusions

Conclusions

What does this mean for Tampa Bay and elsewhere?

- Better description of biological endpoints can change conclusions



Conclusions

Observational data are not particularly telling...
...so we use ***indicators*** and ***quantitative methods*** to ***decompose*** the observations

Conclusions

Observational data are not particularly telling...
...so we use ***indicators*** and ***quantitative methods*** to ***decompose*** the observations

Chosen method depends on the question

Conclusions

Observational data are not particularly telling...
...so we use ***indicators*** and ***quantitative methods*** to ***decompose*** the observations

Chosen method depends on the question

- More complete description of trends
- Better link to causal events
- More comprehensive evaluation of site-specific issues
- Deconstruct the past to predict the future

Acknowledgments:

Dissertation committee and advisors: B. Vondracek, L. Hatch, D. Pereira, S. Weisberg

Research staff and employees at USEPA Gulf Ecology Division - especially J. Hagy, M. Murrell

Field staff and data managers at Hillsborough County Environmental Protection Commission

Research staff and employees at the San Francisco Estuary Institute

Funding sources and contact:

beck.marcus@epa.gov, 8509342480

Blog: https://beckmw.wordpress.com/

GitHub: https://github.com/fawda123

Twitter: @fawda123

References I

Beck MW, Hagy III JD. 2015.
Adaptation of a weighted regression approach to evaluate water quality trends in an estuary.
Environmental Modelling and Assessment, 20(6):637–655.

Beck MW, Hatch LK, Vondracek B, Valley RD. 2010.
Development of a macrophyte-based index of biotic integrity for Minnesota lakes.
Ecological Indicators, 10(5):968–979.

Beck MW, Murphy RR. 2017.
Numerical and qualitative contrasts of two statistical models for water quality change in tidal waters.
Journal of the American Water Resources Association, 53(1):197–219.

Beck MW, Vondracek B, Hatch LK. 2013.
Environmental clustering of lakes to evaluate performance of a macrophyte index of biotic integrity.
Aquatic Botany, 108:16–25.

Beck MW, Wilson BN, Vondracek B, Hatch LK. 2014.
Application of neural networks to quantify the utility of indices of biotic integrity for biological monitoring.
Ecological Indicators, 45:195–208.

Hirsch RM, Moyer DL, Archfield SA. 2010.
Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs.
Journal of the American Water Resources Association, 46(5):857–880.

Karr JR, Chu EW. 1999.
Restoring Life in Running Waters: Better Biological Monitoring.
Island Press, Washington, DC.

References II

Karr JR, Fausch KD, Angermeier PL, Yant PR, Schlosser IJ. 1986.
Assessing biological integrity in running waters: a method and its rationale.
Technical Report Special Publication 5, Illinois Natural History Survey, Champaign, Illinois.

National Research Council. 1990.
Managing Troubled Waters: The Role of Environmental Monitoring.
National Academy Press, Washington, DC.

Suter GW. 1993.
A critique of ecosystem health concepts and indexes.
Environmental Toxicology and Chemistry, 12(9):1533–1539.

TBEP (Tampa Bay Estuary Program). 2011.
Tampa Bay Water Atlas.
http://www.tampabay.wateratlas.usf.edu/. (Accessed October, 2013).