Time series topic 1: Weighted regression

Marcus W. Beck[1]

[1]USEPA NHEERL Gulf Ecology Division
Email: beck.marcus@epa.gov

Objectives for the session (2:00 - 3:00)

- What is weighted regression
- The WRTDStidal package
- Application to NERRS data
  - Fitting a model
  - Evaluating a model
  - Viewing a model

Interactive portion

Follow along as we go:

- flash drive
- online: swmprats.net 2016 workshop tab

Interactive portion

Follow along as we go:

- flash drive
- online: swmprats.net 2016 workshop tab

You will run examples whenever you see this guy:

Is everything installed?

We will use the WRTDStidal package

Option 1, from the R Console prompt:

```
install.packages('devtools')
library(devtools)
install_github('fawda123/WRTDStidal')
library(WRTDStidal)
```

Is everything installed?

We will use the WRTDStidal package

Option 1, from the R Console prompt:

```
install.packages('devtools')
library(devtools)
install_github('fawda123/WRTDStidal')
library(WRTDStidal)
```

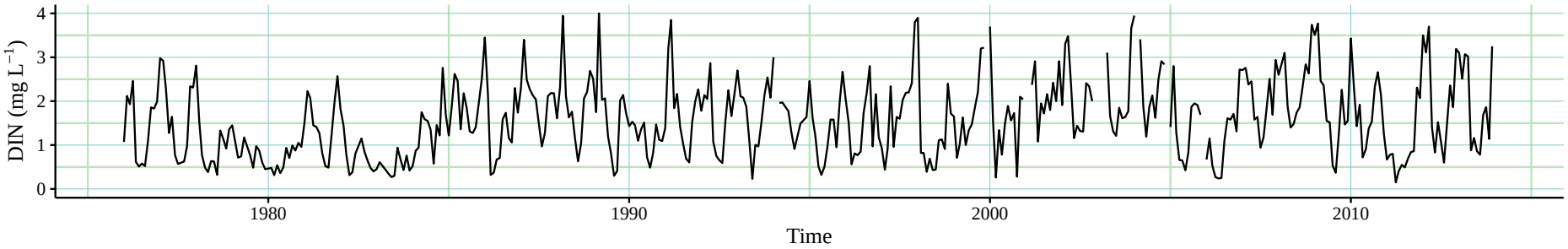
Option 2, install the source file from the flash drive:

```
# change as needed
path_to_file <- 'C:/Users/mbeck/Desktop/WRTDStidal-0.0.49.9000.tar.gz'

# install, load
install.packages(path_to_file, repos = NULL, type="source")
library(WRTDStidal)
```

Model theory and background - what is WRTDS

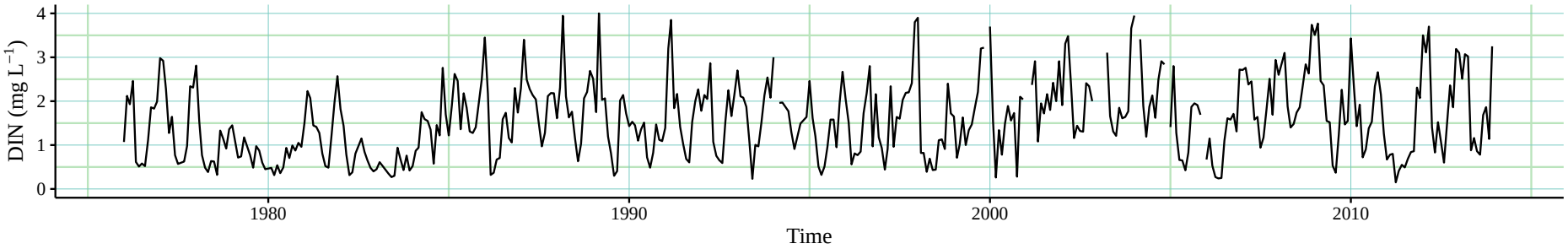
Observed data represents effects of many processes



Climate	***Local***	***Regional/historical***
precipitation	light/turbidity	watershed inputs
temperature	residence time	point sources
wind events	invasive species	management actions
ENSO effects	trophic effects	flow changes

Model theory and background - what is WRTDS

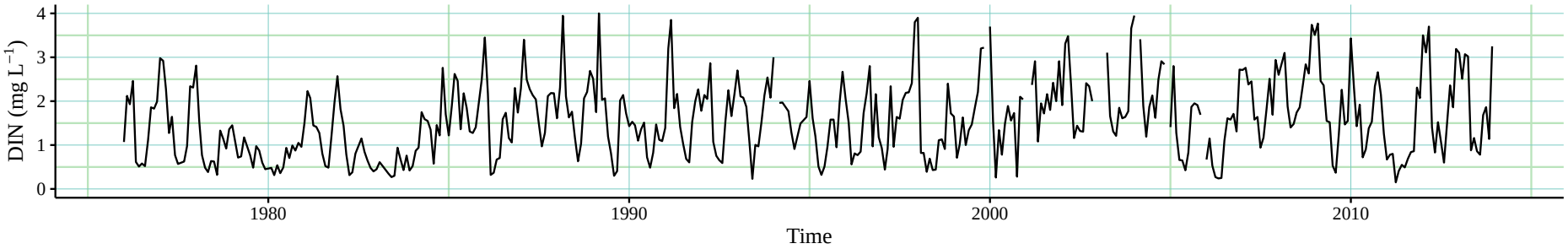
Observed data represents effects of many processes



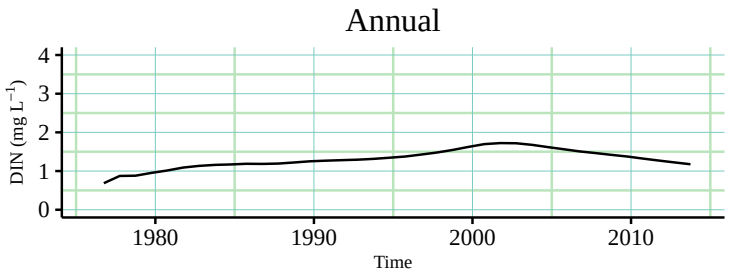
Models should describe components to evaluate effects

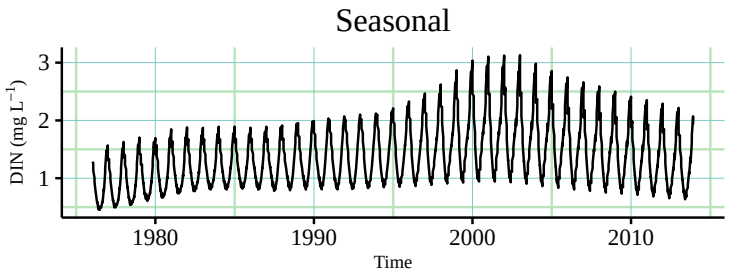
Model theory and background - what is WRTDS

Observed data represents effects of many processes



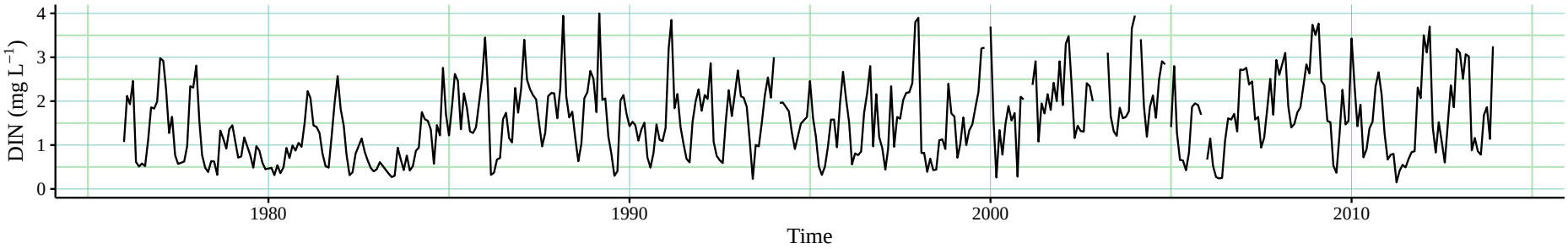
Models should describe components to evaluate effects



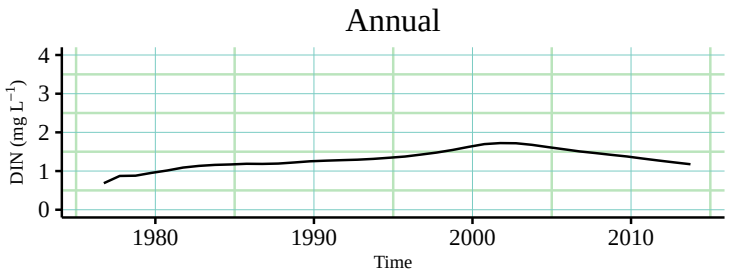


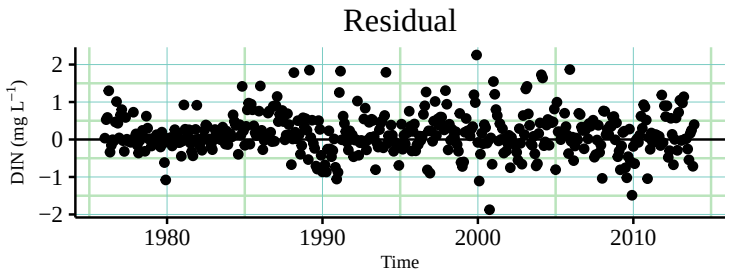
Model theory and background - what is WRTDS

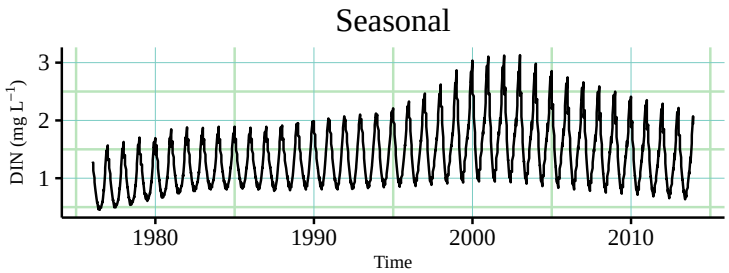
Observed data represents effects of many processes

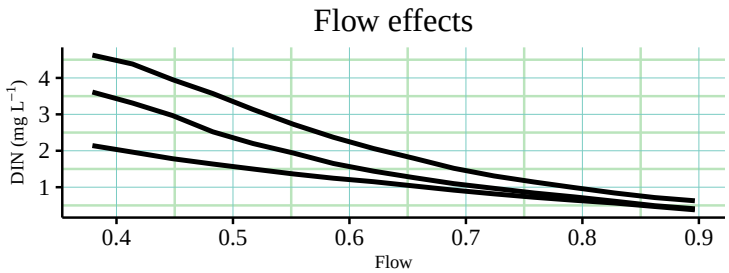


Models should describe components to evaluate effects









Model theory and background - what is WRTDS

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

Model theory and background - what is WRTDS

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]

Model theory and background - what is WRTDS

How does it work?

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

N: nitrogen (or other response endpoint)

t: time

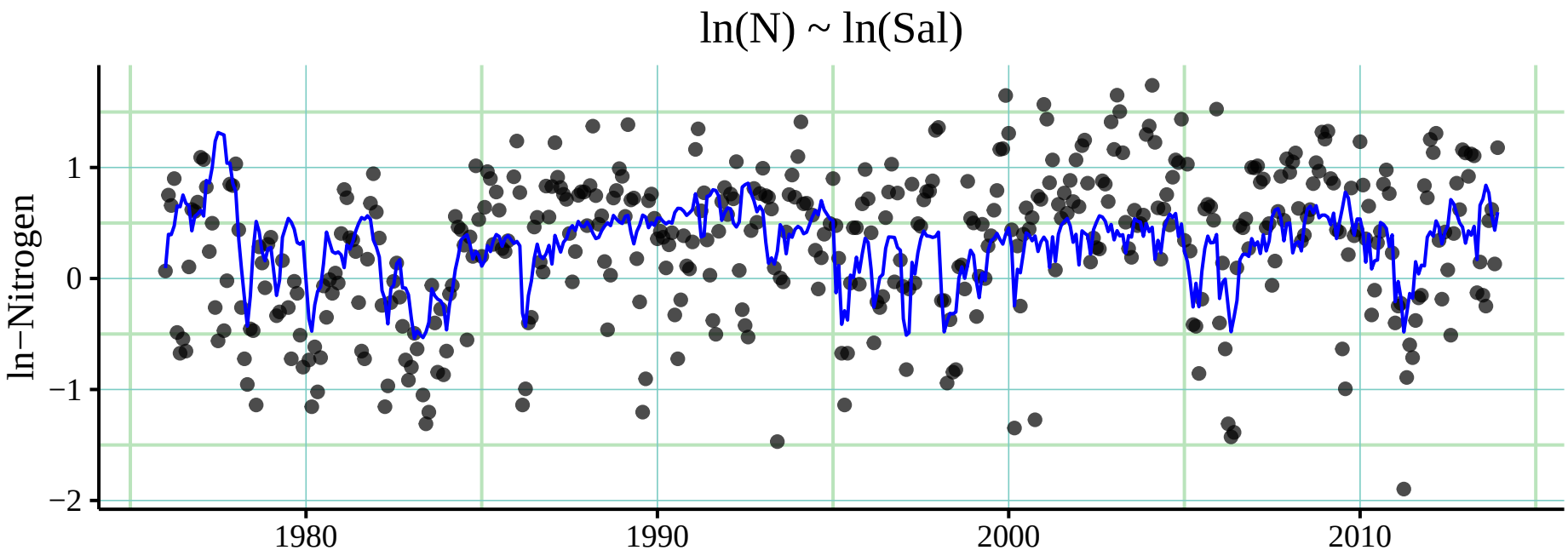
Sal: Salinity (or other flow-related variable)

Model theory and background - what is WRTDS

How does it work?

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

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

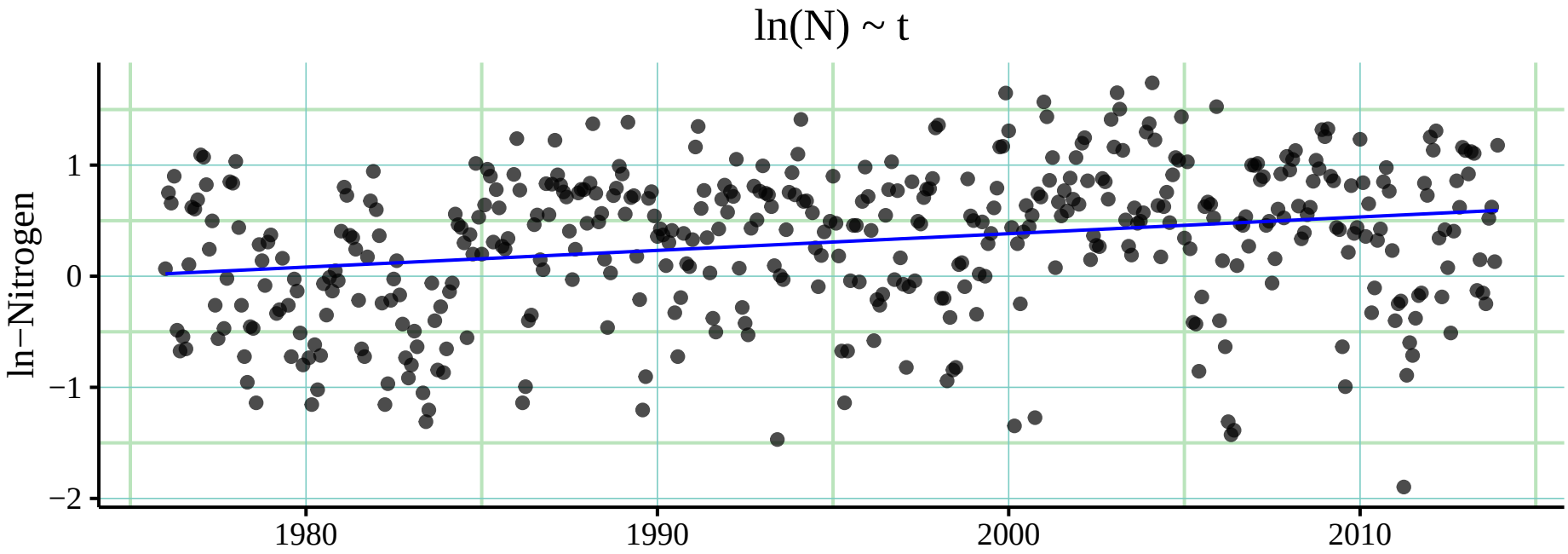


Model theory and background - what is WRTDS

How does it work?

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

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

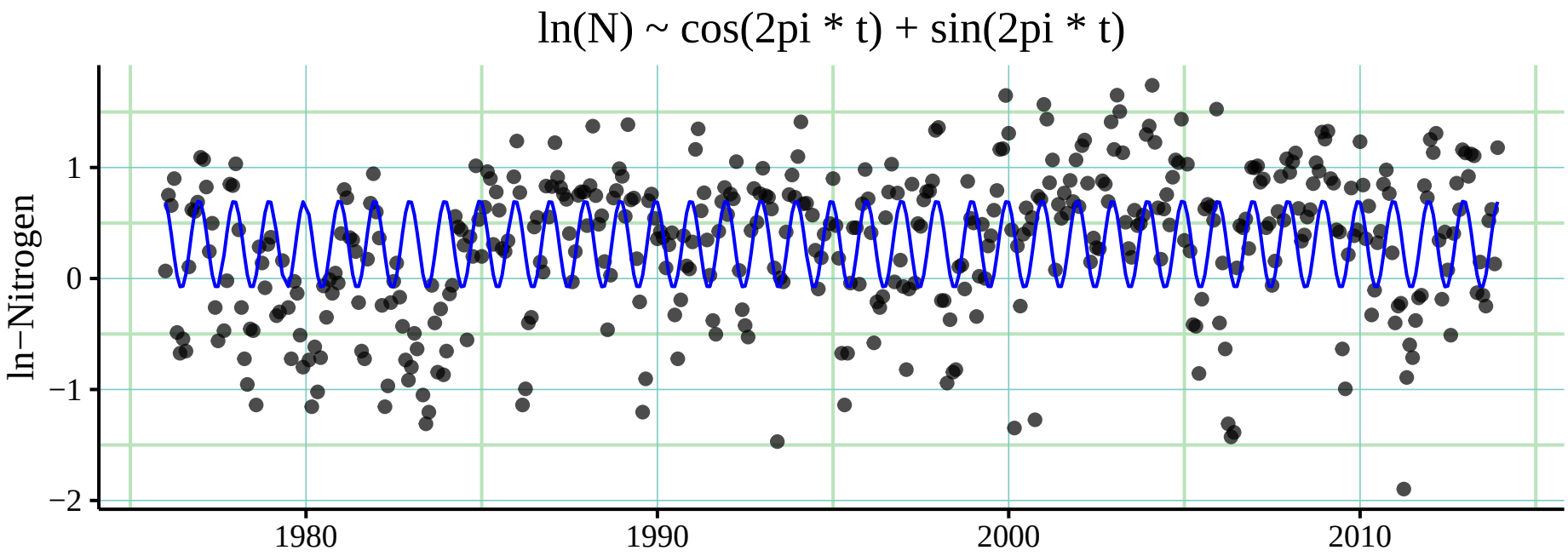


Model theory and background - what is WRTDS

How does it work?

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

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

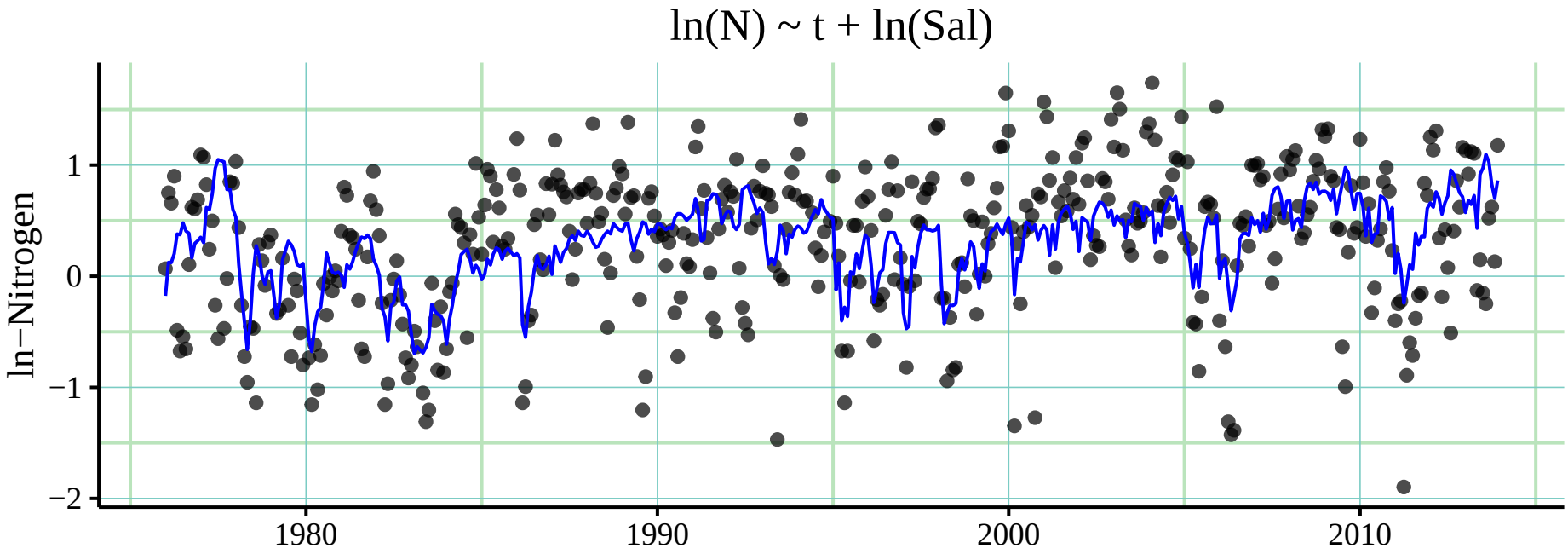


Model theory and background - what is WRTDS

How does it work?

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

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

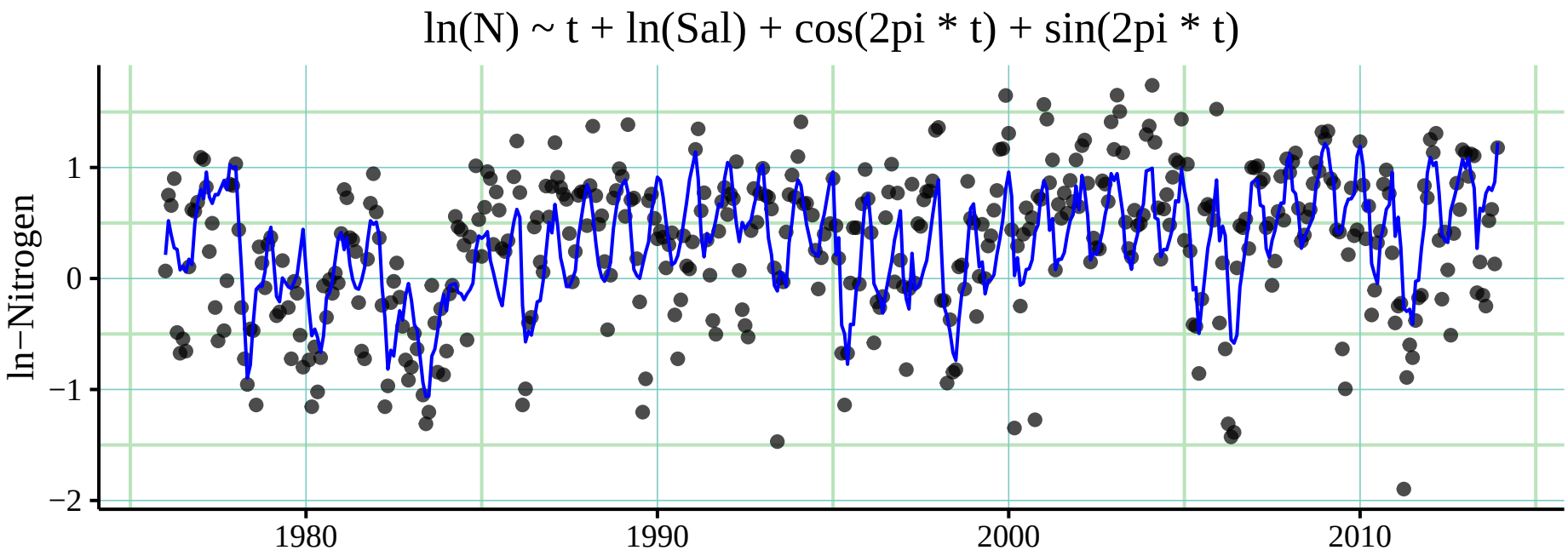


Model theory and background - what is WRTDS

How does it work?

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

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



Model theory and background - what is WRTDS

This is not the whole story...

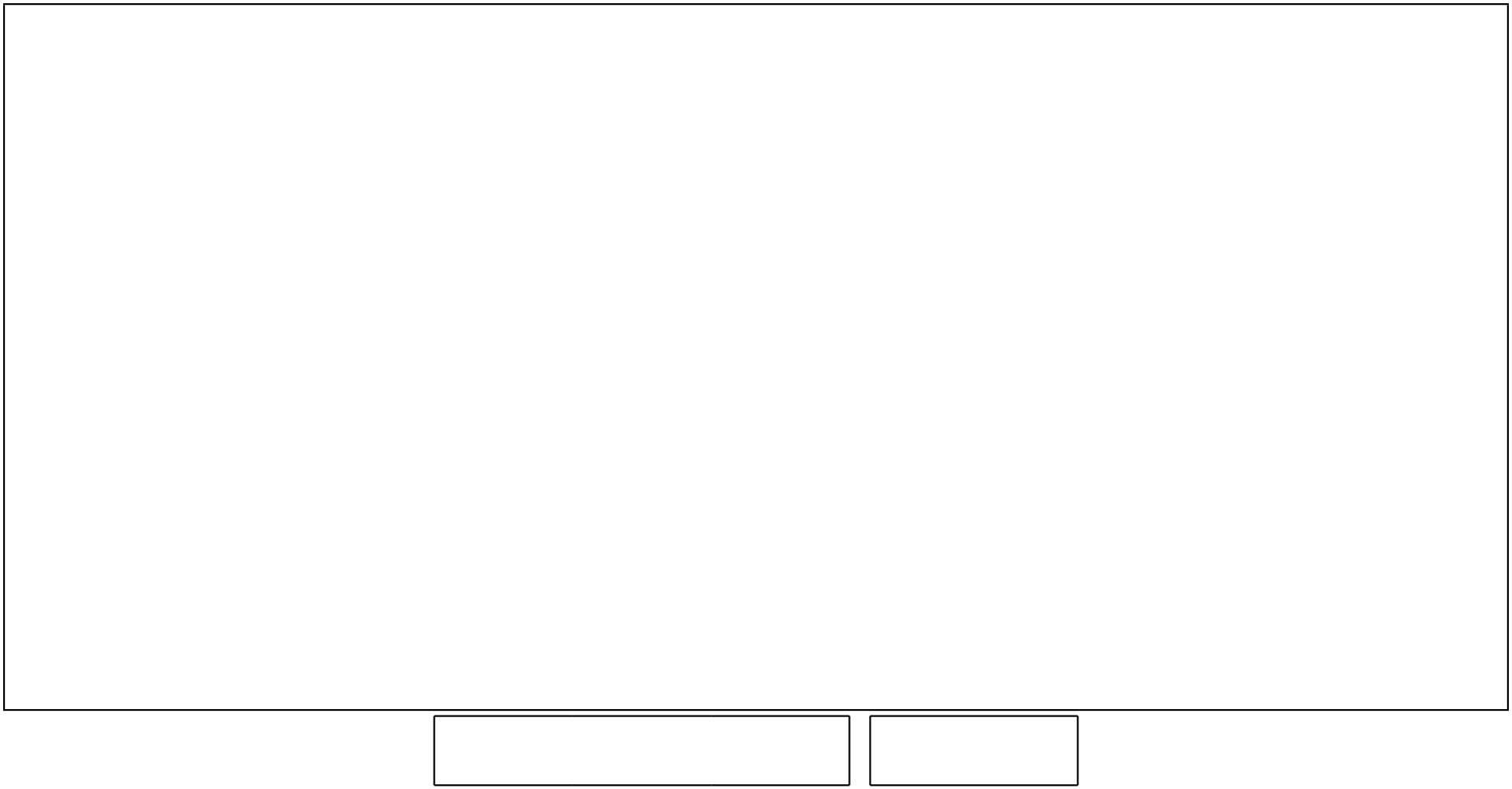
$$\ln(N) = \beta_0 + \beta_1 t + \beta_2 \ln(Sal) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$

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

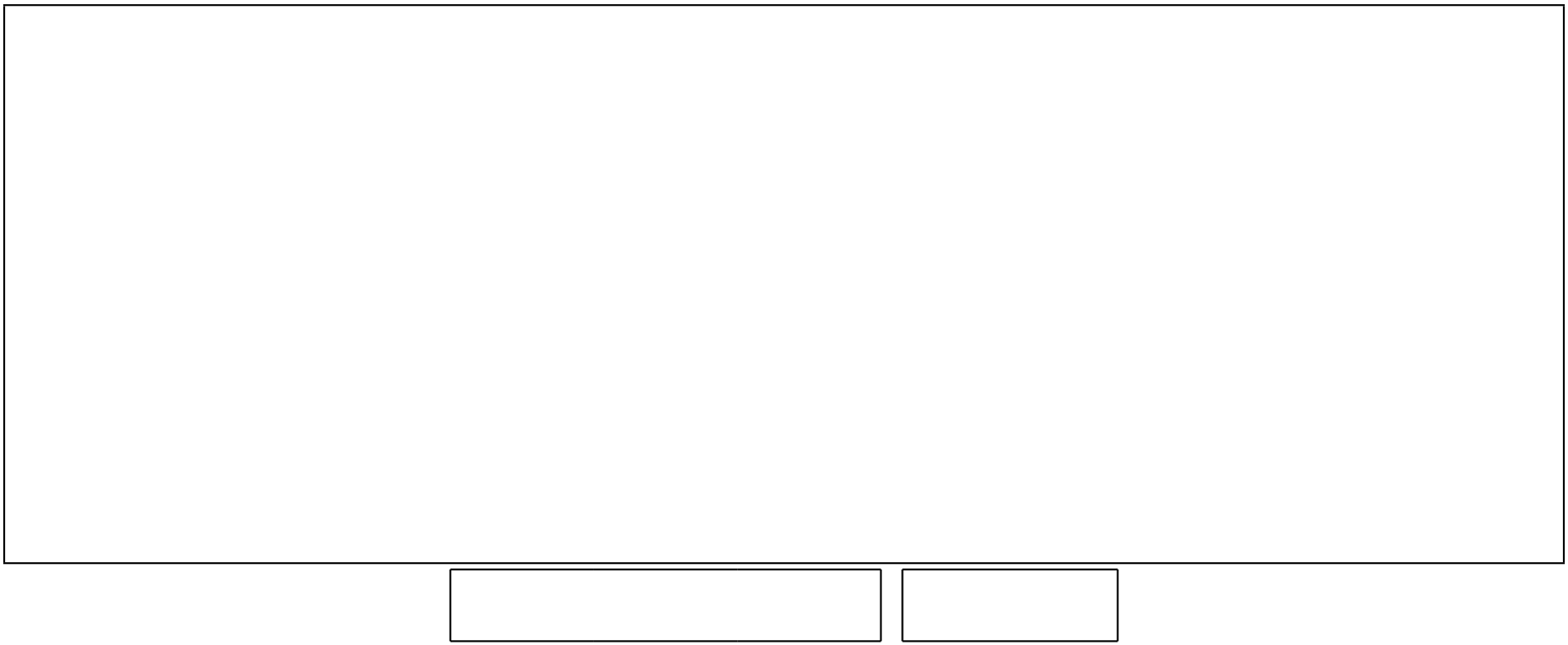
Similar to a loess/spline smooth but specific to the effects of these three variables on the response

Model theory and background - what is WRTDS



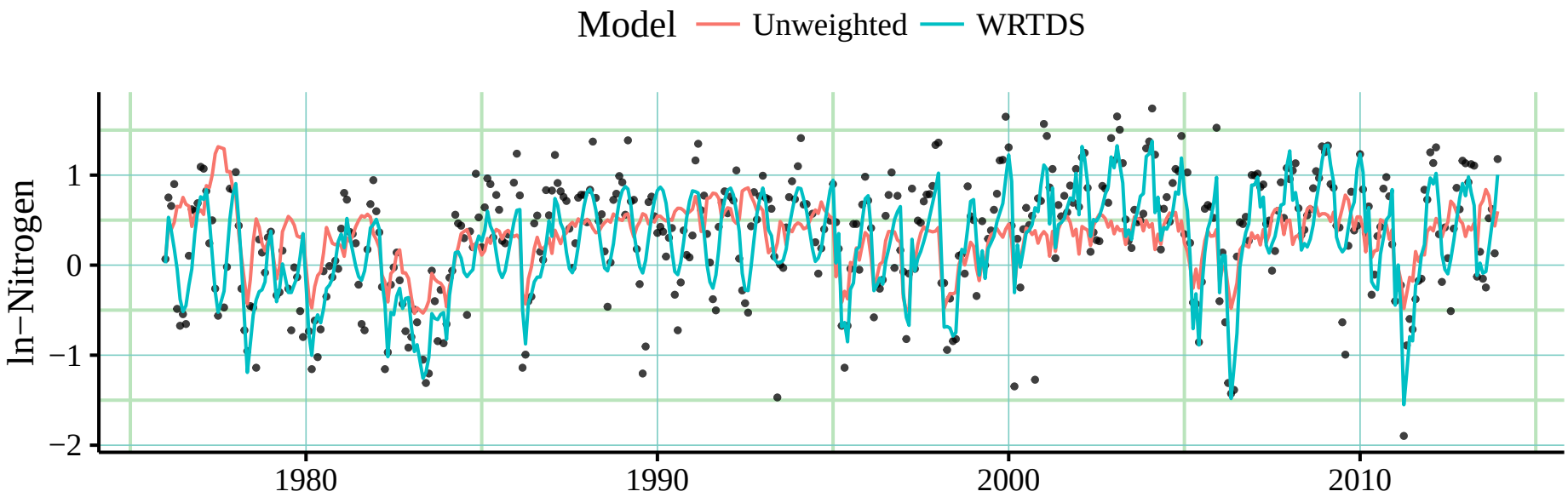
Model theory and background - what is WRTDS

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



Model theory and background - what is WRTDS

RMSE fit for unweighted = 0.58, WRTDS = 0.36



Model theory and background - what is WRTDS

All you need to know:

- Describe a response variable in relation to time, salinity (discharge), and season

Model theory and background - what is WRTDS

All you need to know:

- Describe a response variable in relation to time, salinity (discharge), and season
- A 'moving window, weighted regression', like a loess/spline smooth

Model theory and background - what is WRTDS

All you need to know:

- Describe a response variable in relation to time, salinity (discharge), and season
- A 'moving window, weighted regression', like a loess/spline smooth
- WRTDStidal package does all of it for you - including viz

Model theory and background - what is WRTDS

All you need to know:

- Describe a response variable in relation to time, salinity (discharge), and season
- A 'moving window, weighted regression', like a loess/spline smooth
- WRTDStidal package does all of it for you - including viz

...let's not forget about flow-normalization, more about this later

Using WRTDS with NERRS data

Up next... Time Series Topic 2: Decomposition

Questions??

References

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.

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.