SWMPrats: A community of practice for NERRS data analysis

Marcus W. Beck[1] Todd D. O'Brien[2]

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

[2]NOAA/NMFS COPEPOD Project
Email: todd.obrien@noaa.gov

May 4, 2015

Overview

- The genesis of SWMPrats
- Features of SWMPrats.net
  - SWMPr
  - widgets
  - forum
- Continuing work and engaging the larger community

Genesis of SWMPrats

As of April 29th, $>$ 58 million SWMP data records available on CDMO

An invaluable data source but...

- No recent comparative analyses between systems
- No simple tools for trend analysis at individual sites

These needs were identified in 2013 annual meeting, led to a workshop at the 2014 meeting to focus on time series analysis

Genesis of SWMPrats

One-day training workshop at 2014 annual meeting

- Attended by over 70 NERRS staff, representing 19 of 28 reserves
- General focus on time series analysis, simple applications with SWMP data
- Pre/post workshop materials, including an R package for SWMP

Genesis of SWMPrats

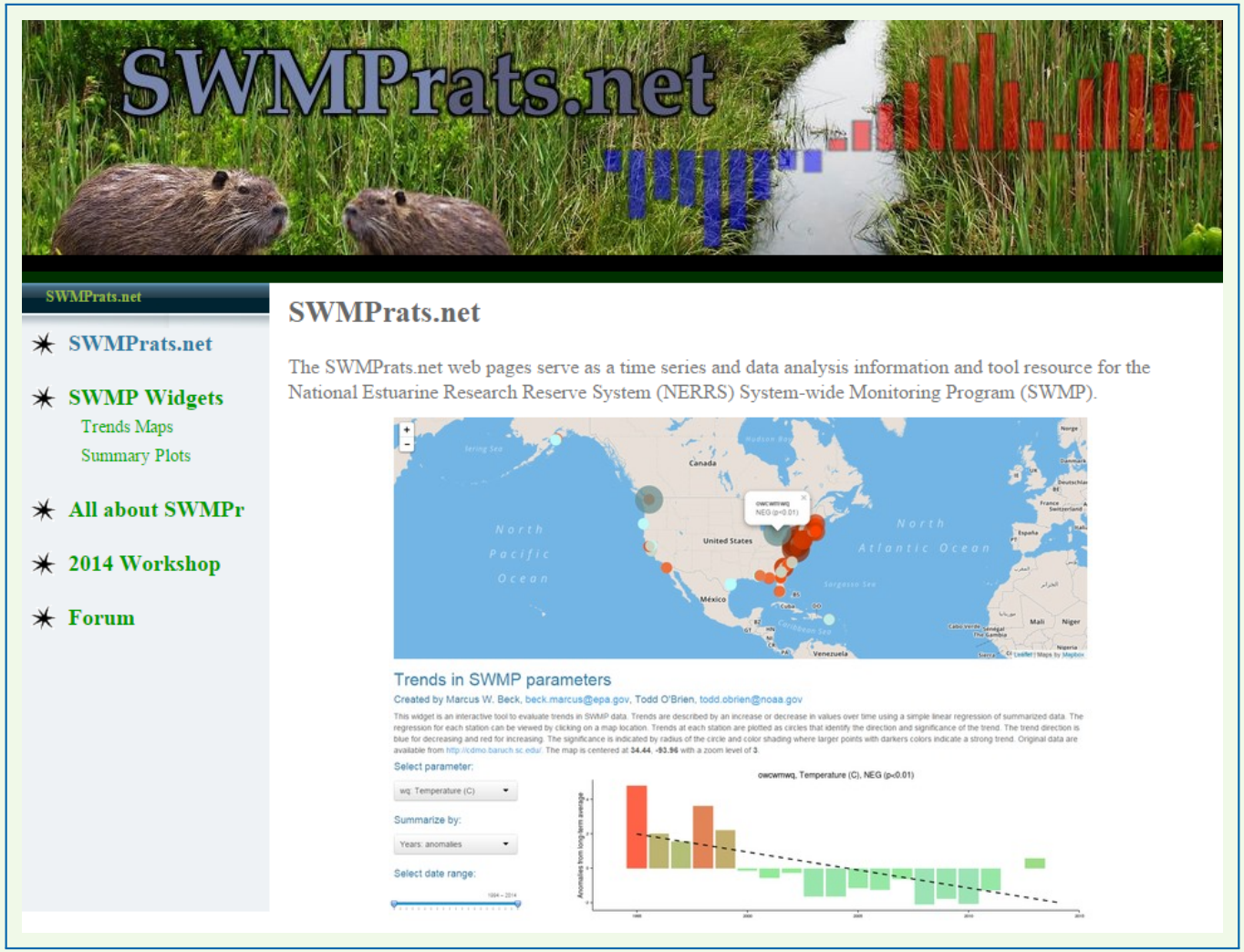
A working group was formed from this meeting

System-***W***ide ***M***onitoring ***P***rogram ***R*** ***A***nd ***T***ime ***S***eries

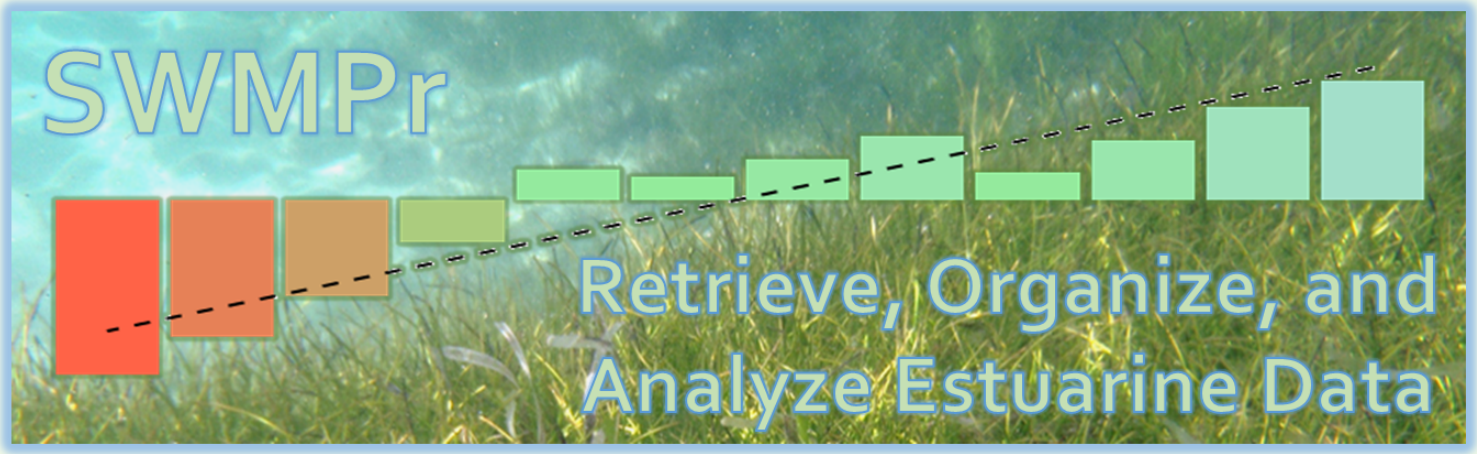
SWMPrats.net is our base of operations...

SWMPrats.net

A time series and data analysis information and tool resource



SWMPrats.net: The SWMPr package



SWMPr is an open-source R package, v2.0.0 is now available

```
> # install/load from R
> install.packages('SWMPr')
> library(SWMPR)
```

Currently working on a manuscript to describe the package in detail

SWMPrats.net: The SWMPr package

SWMPr functions are grouped into three categories that describe their use in the 'data workflow'

Retrieve	*Organize*	*Analyze*
all_params	comb	aggreswmp
all_params_dtrng	qaqc	aggremetab
import_local	qaqcchk	ecometab
import_remote	rem_reps	decomp
single_param	setstep	decomp_cj
site_codes	subset	hist
site_codes_ind		lines
		na.approx
		plot
		plot_metab
		plot_summary
		smoother

SWMPrats.net: The SWMPr package

The software addresses the tedious but necessary challenges of analyzing time series, specific to SWMP

What are some challenges?

- Dealing with 'bad' data
- Subsetting by date ranges, parameters
- Combining data from different sites
- Standardizing time steps
- ...and analysis

SWMPrats.net: The SWMPr package

Proof of concept, import and combine wq and weather data from Apalachicola Bay

```
> # import data
> met <- import_remote('apaebmet')
> wq <- import_remote('apacpwq')
>
> # combine, two hours time step
> # only overlapping date ranges
> dat <- comb(met, wq, timestep = 120,
+   method = 'intersect')
```

Try this with excel...

SWMPrats.net: The SWMPr package

```
> head(dat, 4)
```

```
##             datetimestamp temp spcond sal do_pct do_mgl depth
## 1 2001-12-31 23:00:00   NA     NA  NA     NA     NA    NA
## 2 2002-01-01 01:00:00   12     37  24    104     10     2
## 3 2002-01-01 03:00:00   12     40  26     99      9     2
## 4 2002-01-01 05:00:00   11     42  26     98      9     2
##   cdepth level clevel ph turb chlfluor atemp rh   bp wspd
## 1     NA    NA     NA NA   NA       NA     4 69 1017    4
## 2     NA    NA     NA NA    3       NA     3 75 1017    3
## 3     NA    NA     NA NA    4       NA     2 77 1018    3
## 4     NA    NA     NA NA    5       NA     1 82 1019    4
##   maxwspd wdir sdwdir totpar totprcp totsorad
## 1      NA  347     NA      0      NA        NA
## 2      NA    9     NA      0      NA        NA
## 3      NA  331     NA      0      NA        NA
## 4      NA    0     NA      0      NA        NA
```

SWMPrats.net: The SWMPr package

Time series analysis can range from very general to very specific

SWMPr functions include...

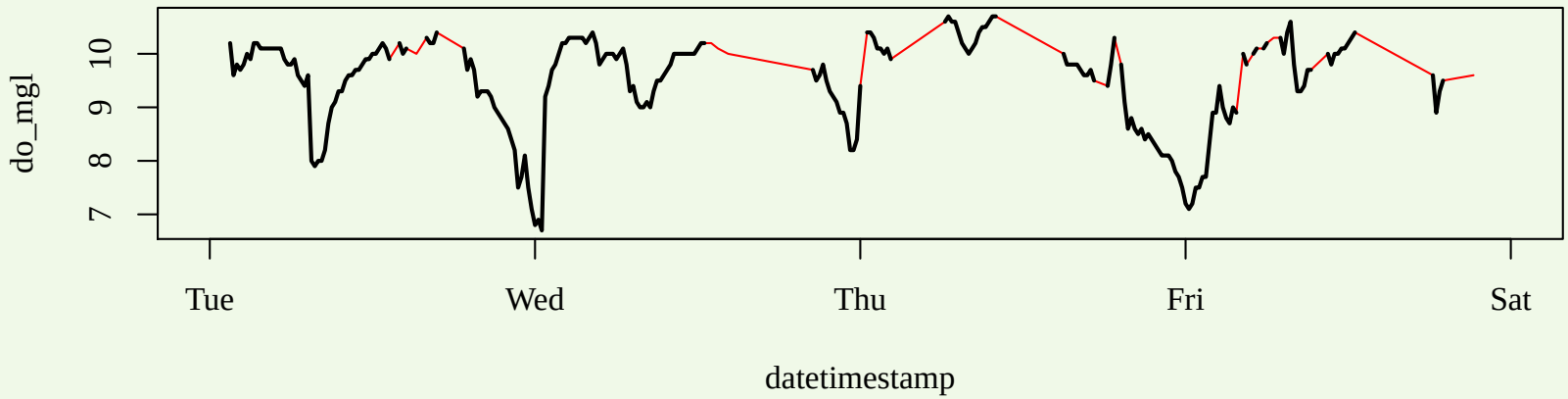
General

- `na.approx`
- `smoother`
- `aggreswmp`
- `plot`
- `hist`
- `lines`

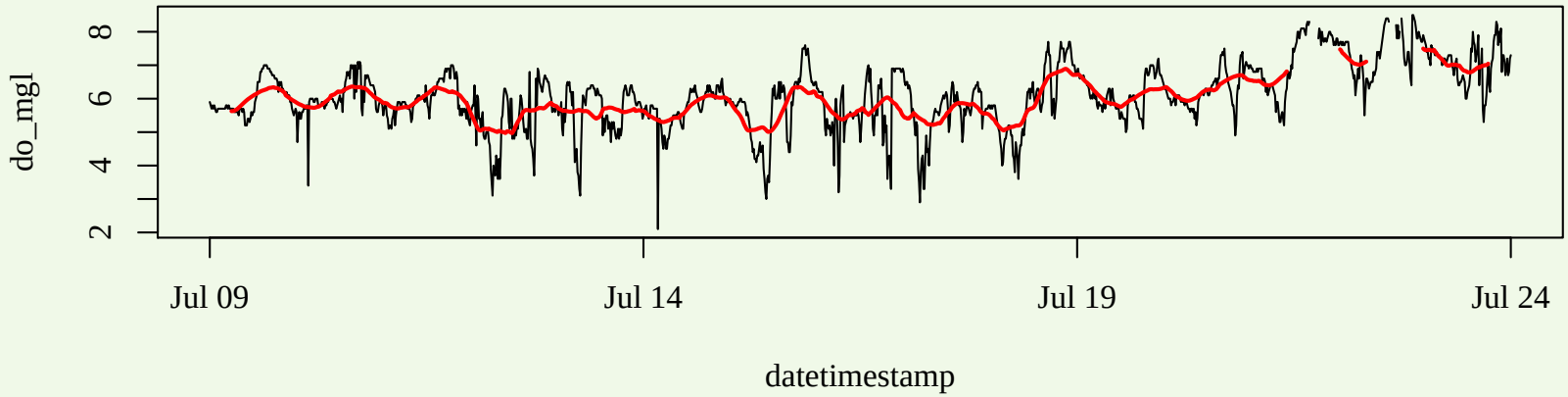
Specific

- `decomp`
- `decomp_cj`
- `ecometab`
- `aggremetab`
- `plot_metab`
- `plot_summary`

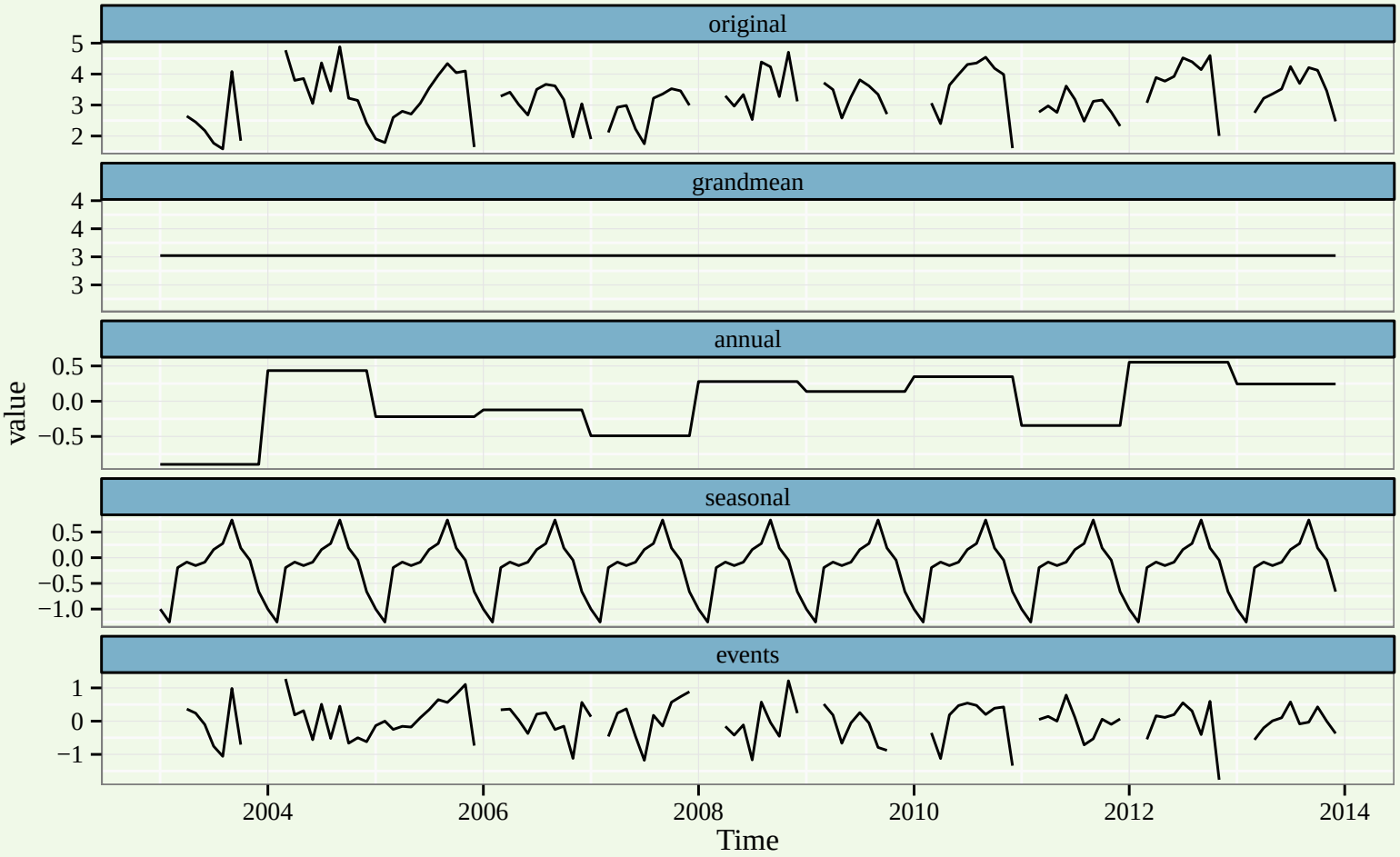
Fill missing data with `na.approx`



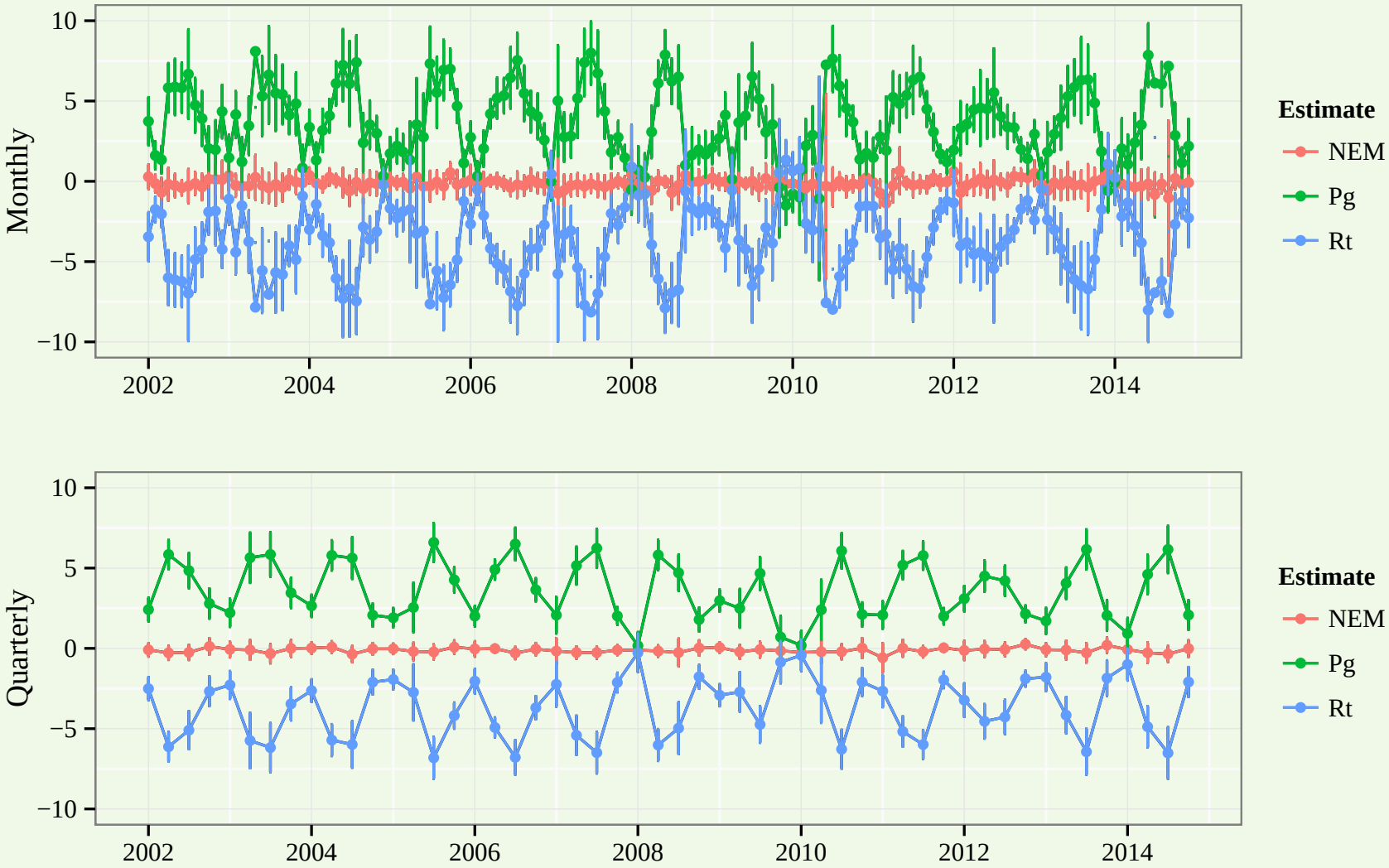
Smooth data with `smoother`



Time series decomposition with `decomp_cj` (chl-a at cbmocnut)



Estimate ecosystem metabolism with `ecometab` (apadbwq)



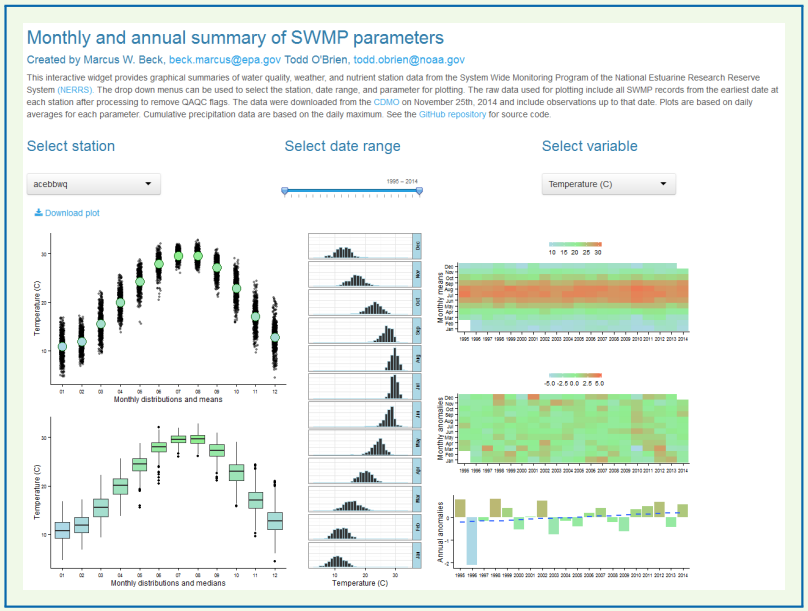
SWMPrats.net: Widgets

The most common question - what is the change over time at my site?

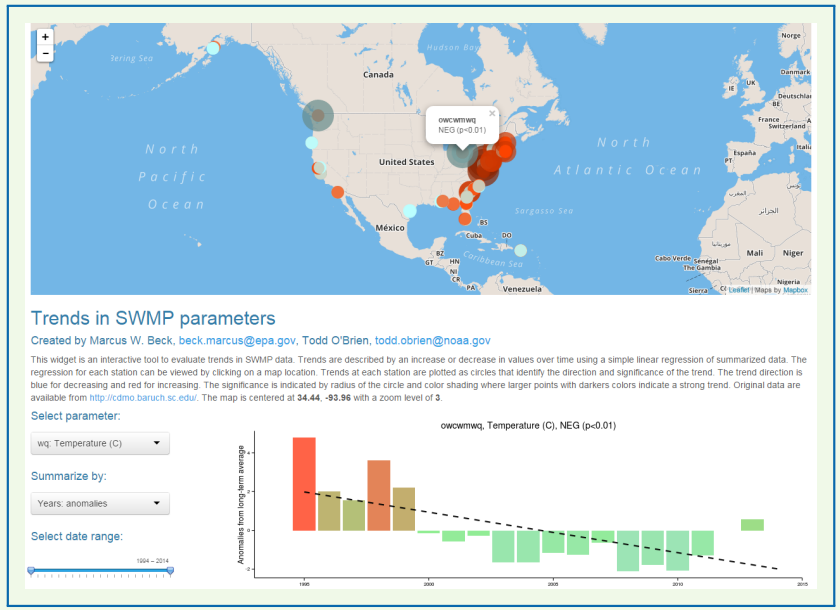
The functions in SWMPr can help, but it's easier to interact!

Two apps on SWMPrats.net can help visualize trends

Summary plots

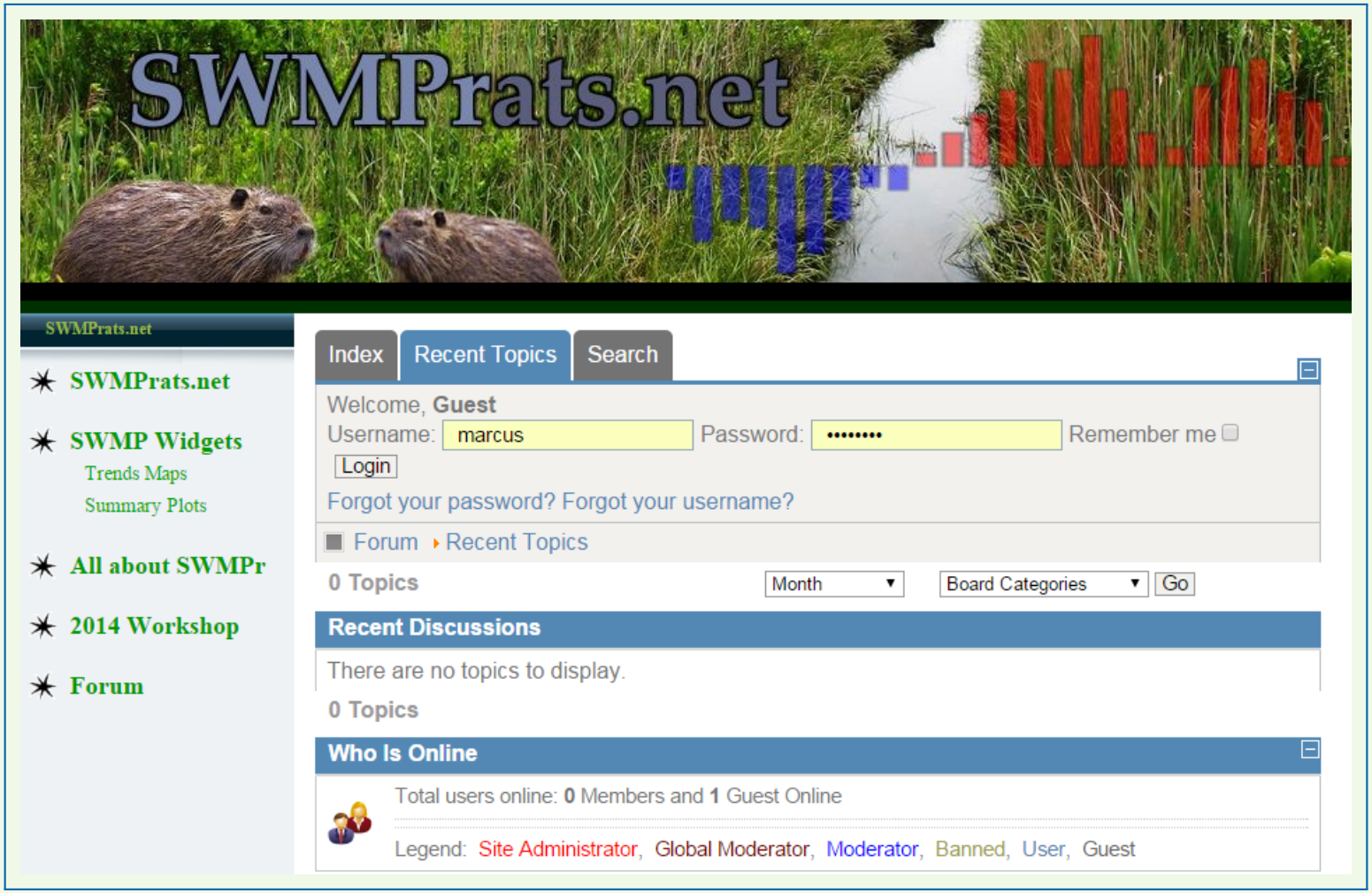


Trends map



SWMPrats.net: Forum

Last but not least, a discussion forum for all things analytical



Continuing work and engagement

SWMPrats.net is in its infancy but already seeing heavy use

- SWMPr downloaded 289 times from R network (as of April 29)
- Apps have been used 345 hours (as of April 29)

Continuing development of packages/apps - submit suggestions/bug reports via email or on GitHub (preferred)

Plan for greater engagement with the forum - soliciting moderators, suggested topics

Additional training workshops??

Continuing work and engagement

Contacts: beck.marcus@epa.gov, todd.obrien@noaa.gov

To get this presentation: https://github.com/fawda123/swmprats_pres

Summary app: http://swmprats.net/swmp-widgets/summary-plots

Trends map app: http://swmprats.net/swmp-widgets/trends-maps

Visit the development site for the most recent version of SWMPr: https://github.com/fawda123/SWMPr