NERRS / SWMP

Training Workshop: *R Intro & SWMPr*

October 25, 2015

SWMPr overview, retrieve, and organize

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

Objectives for the session

- Why and what is SWMPr?

- How can data get from CDMO into R using SWMPr?

- What is the basic structure of a `swmpr` data object?

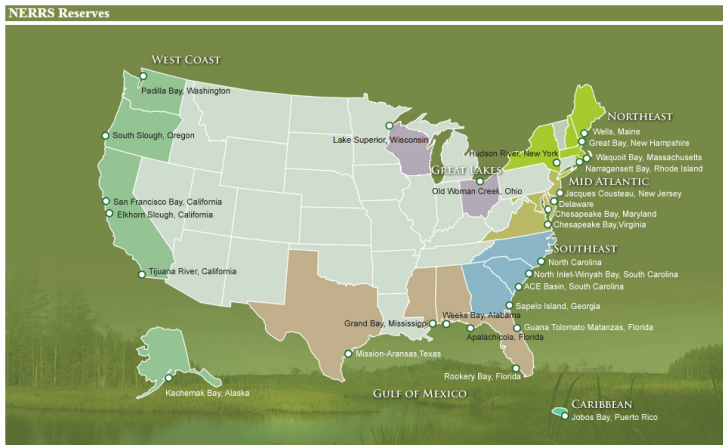- What is data organization and how can SWMPr help?

Interactive portion

We will use the swmpr1.Rproj project for this session

- location on flash drive

- location online

'double-click' the file to open in RStudio...

Why and what is SWMPr?

SWMP - System Wide Monitoring Program, initiated in 1995 to provide continuous monitoring data at over 300 stations in 28 US estuaries

Why and what is SWMPr?

CDMO (link) is your one-stop shop for retrieving SWMP data

Why and what is SWMPr?

The raw data will look like this...

	A	B	C	D	E	F	G	H	I	J	K	L
1	StationCo	isSWMP	DateTimeStamp	Historical	Provision:	CollMeth	REP	F_Record	PO4F	F_PO4F	NH4F	F_NH4F
2	apacpnut	P	1/10/2012 10:20	0	1	1	1		0.003	<-4> [SBL]	0.03	<0>
3	apacpnut	P	2/7/2012 11:41	0	1	1	1		0.005	<0>	0.019	<0>
4	apacpnut	P	3/5/2012 11:51	0	1	1	1		0.003	<-4> [SBL]	0.041	<0>
5	apacpnut	P	4/4/2012 10:30	0	1	1	1		0.003	<-4> [SBL]	0.043	<0>
6	apacpnut	P	5/9/2012 10:12	0	1	1	1		0.003	<0>	0.053	<0>
7	apacpnut	P	5/9/2012 10:15	0	1	1	2		0.003	<-4> [SBL]	0.022	<0>
8	apacpnut	P	5/9/2012 10:20	0	1	1	3		0.003	<0>	0.016	<0>
9	apacpnut	P	6/5/2012 8:30	0	1	1	1		0.003	<-4> [SBL]	0.04	<0>
10	apacpnut	P	7/3/2012 9:58	0	1	1	1 {CSM}		0.004	<0>	0.094	<0>
11	apacpnut	P	7/3/2012 9:59	0	1	1	2 {CSM}		0.004	<0>	0.066	<0>
12	apacpnut	P	7/3/2012 10:01	0	1	1	3 {CSM}		0.005	<0>	0.069	<0>
13	apacpnut	P	8/7/2012 9:53	0	1	1	1 {CSM}		0.003	<-4> [SBL]	0.05	<0>
14	apacpnut	P	9/5/2012 10:56	0	1	1	1		0.003	<-4> [SBL]	0.026	<0>
15	apacpnut	P	10/2/2012 9:22	0	1	1	1		0.003	<-4> [SBL]	0.042	<0>
16	apacpnut	P	10/2/2012 9:27	0	1	1	2		0.003	<-4> [SBL]	0.024	<0>
17	apacpnut	P	10/2/2012 9:32	0	1	1	3		0.003	<0>	0.042	<0>
18	apacpnut	P	11/6/2012 10:30	0	1	1	1		0.003	<-4> [SBL]	0.07	<0>
19	apacpnut	P	11/26/2012 11:39	0	1	1	1		0.003	<-4> [SBL]	0.041	<0>

Why and what is SWMPr?

What are the challenges for evaluating SWMP data?

Why and what is SWMPr?

What are the challenges for evaluating SWMP data?

- Knowing what we want

- Dealing with QAQC columns and removing 'bad' observations

- Data we don't want... extra columns or irrelevant parameters

- Combining data for comparison

- Issues inherent with time series, e.g., missing data

- Others?

Why and what is SWMPr?

What: An R package to **augment** existing CDMO services and to provide a **bridge** to analysis

Why and what is SWMPr?

What: An R package to **augment** existing CDMO services and to provide a **bridge** to analysis

Why: There are many challenges working with SWMP data... a toolkit for addressing these challenges will be useful

Why and what is SWMPr?

What: An R package to **augment** existing CDMO services and to provide a **bridge** to analysis

Why: There are many challenges working with SWMP data... a toolkit for addressing these challenges will be useful

How: Use the SWMPr functions to **retrieve**, **organize**, and **analyze** SWMP data

Why and what is SWMPr?

Some housekeeping...

```r
# install from CRAN (only do once)
install.packages('SWMPr')

# load for your current session
library(SWMPr)
```

https://cran.r-project.org/web/packages/SWMPr/index.html

Why and what is SWMPr?

Uses an **object-oriented** structure... data are imported into R as a swmpr data object, with functions built to use this object

What are the **retrieve**, **organize**, and **analyze** functions?

Run this code one line at a time... What comes up?

```
help.search('retrieve', package = 'SWMPr')
help.search('organize', package = 'SWMPr')
help.search('analyze', package = 'SWMPr')
```

Why and what is SWMPr?

Uses an **object-oriented** structure... data are imported into R as a swmpr data object, with functions built to use this object

What are the **retrieve**, **organize**, and **analyze** functions?

Run this code one line at a time... What comes up?

```
help.search('retrieve', package = 'SWMPr')
help.search('organize', package = 'SWMPr')
help.search('analyze', package = 'SWMPr')
```

What about this?

```
?import_local
```

Any useful information?

Getting SWMP data into R

Let's get some data into R!

The **retrieval** functions do two things:

Import data directly from the CDMO:

```
all_params
all_params_dtrng
single_param
site_codes
site_codes_ind
```

These functions require registering your IP address with CDMO

Import data from a local path:

```
import_local
```

Imports data obtained from (and only from) the zip downloads feature

Getting SWMP data into R

The 'zip_ex' folder in the project is a sample dataset that looks exactly like a folder you get from CDMO

Let's import some data from that folder...

```
# get data for apacpwq, all years

# location of data
mypath <- 'zip_ex'

# import and assign to 'dat'
dat <- import_local(mypath, 'apacpwq', trace = T)
```

Getting SWMP data into R

The 'zip_ex' folder in the project is a sample dataset that looks exactly like a folder you get from CDMO

Let's import some data from that folder...

```
# get data for apacpwq, all years

# location of data
mypath <- 'zip_ex'

# import and assign to 'dat'
dat <- import_local(mypath, 'apacpwq', trace = T)
```

What about this?

```
dat2 <- import_local(mypath, 'apacp2012', trace = T)
dat3 <- import_local(mypath, 'apadbnut', trace = T)
```

Structure of the `swmpr` data object

Now we have data in our 'workspace' that we can organize/analyze

Try running the following...

```
head(dat)
tail(dat)
View(dat)
str(dat)
attributes(dat)
```

Structure of the `swmpr` data object

Now we have data in our 'workspace' that we can organize/analyze

Try running the following...

```
head(dat)
tail(dat)
View(dat)
str(dat)
attributes(dat)
```

How are the data organized?

What are the column names?

What are the attributes?

Structure of the `swmpr` data object
The `swmpr` object is a `data.frame` and a list of attributes

```
head(dat, 3)

##          datetimestamp temp f_temp spcond f_spcond sal f_sal do_pct f_do_pct
## 1 2011-01-01 00:00:00   11    <0>    44      <0>   28   <0>    68      <0>
## 2 2011-01-01 00:15:00   11    <0>    44      <0>   28   <0>    68      <0>
## 3 2011-01-01 00:30:00   11    <0>    44      <0>   28   <0>    68      <0>
##   do_mgl f_do_mgl depth f_depth cdepth f_cdepth level f_level clevel f_clevel
## 1      6      <0>     2     <0>      2      <3>    NA     <-1>     NA       NA
## 2      6      <0>     2     <0>      2      <3>    NA     <-1>     NA       NA
## 3      6      <0>     2     <0>      2      <3>    NA     <-1>     NA       NA
##   ph f_ph turb f_turb chlfluor f_chlfluor
## 1  8 <0>    3    <0>       NA       <-1>
## 2  8 <0>    3    <0>       NA       <-1>
## 3  8 <0>    2    <0>       NA       <-1>


names(attributes(dat))

## [1] "names"      "row.names"  "class"      "station"    "parameters"
## [6] "qaqc_cols"  "date_rng"   "timezone"   "stamp_class"


attr(dat, 'parameters')

## [1] "temp"    "spcond"  "sal"     "do_pct"  "do_mgl"  "depth"
## [7] "cdepth"  "level"   "clevel"  "ph"      "turb"    "chlfluor"
```

Data organization with SWMPr

First problem is solved... we know how to get SWMP data from CDMO into R:

- Download a dataset from zip downloads

- Find where the data have downloaded

- Import using `import_local`

- Have a look at the data (`head`, `View`, `attributes`)

- Lost? Check the help files: `?import_local`

Now we can think about preprocessing or organizing prior to analysis

Data organization with SWMPr

Problem 1 solved... we know how to get SWMP data from CDMO into R:

- Download a dataset from zip downloads

- Find where the data have downloaded

- Import using import_local

- Have a look at the data (head, View, attributes)

- Lost? Check the help files: ?import_local

Now we can think about preprocessing or organizing prior to analysis

NERRS / SWMP

Training Workshop: *R Intro & SWMPr*

October 25, 2015

Questions??