

SWMP data and retrieval

Marcus W. Beck¹ Todd D. O'Brien²

¹ORISE, USEPA NHEERL Gulf Ecology Division

Email: beck.marcus@epa.gov

²NOAA/NMFS COPEPOD Project

Email: todd.obrien@noaa.gov

Objectives and agenda

- Objectives

- ▶ What are the various ways data are obtained from SWMP?
- ▶ What are some issues that need to be addressed before importing into a statistical program for time series analysis?

Objectives and agenda

- Objectives

- ▶ What are the various ways data are obtained from SWMP?
- ▶ What are some issues that need to be addressed before importing into a statistical program for time series analysis?

- Agenda

- ▶ Brief overview of SWMP network and available data
- ▶ Format and potential issues with output data
- ▶ Retrieving and importing the data

Interactive portion

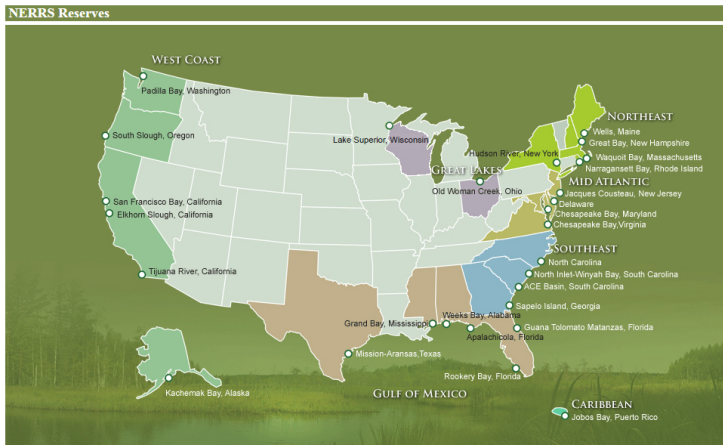
You can follow along later in this module:

- dataset1
- script1

Interactive!

Overview of SWMP and available data

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



<http://nerrs.noaa.gov/ReservesMap.aspx>



Overview of SWMP and available data

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

Home	About CDMO	About Data	Get Data	Web Services	Contact CDMO
					
View / Download Data		Real Time Monitoring Data		CDMO News	
 Requested Citation Format		<div>Choose Reserve... ▾</div> <div>GTMPCMET 10/08/14 09:45 AM GTMPCVQ 10/08/14 09:45 AM</div>  <div>Air Temperature: 27.8 °C (82 °F) Wind Speed: 1.1 m/Sec (02 mph) Water Temperature: 22.7 °C (73 °F) Salinity: 7.1 PPT Dissolved Oxygen: 4.7 mg/L</div>		<p>The CDMO is excited to announce the launch of our new SWMP Mobile application. Near real-time SWMP data is now available on your smartphone or tablet at: www.nerrsdata.org/mobile</p> <hr/> <p>Our Data Export System has been updated and now has enhanced graphing capabilities! Want to easily export or graph data? If so, check out our Data Export System!</p>	

Overview of SWMP and available data

A wide range of data can be requested... a few records for one site to all records for multiple sites

Requests can return a lot of data so make sure you have clear objectives

Check the [available data](#) before making a request!

- station names
- data types
- date ranges
- parameters

Format and potential issues with output data

To orient yourself, understand the NERRS/SWMP naming convention

Site (reserve), **station**, and **parameter type** are identified by a 7 or 8 character name

Format and potential issues with output data

To orient yourself, understand the NERRS/SWMP naming convention

Site (reserve), ***station***, and ***parameter type*** are identified by a 7 or 8 character name

E.g., elkcwmet

Format and potential issues with output data

To orient yourself, understand the NERRS/SWMP naming convention

Site (reserve), **station**, and **parameter type** are identified by a 7 or 8 character name

E.g., elkcwmet

- elk: site, Elkhorn Slough
- cw: station, Caspian Weather Station
- met: parameter type (weather)

Format and potential issues with output data

The fundamental unit of data is the 'station' defined by a parameter type

The parameters for a station are specific to the parameter type

Format and potential issues with output data

The fundamental unit of data is the 'station' defined by a parameter type

The parameters for a station are specific to the parameter type

Nutrients

po4f, chla_n, no3f,
no2f, nh4f, no23f,
ke_n, urea

Water quality

temp, spcond, sal,
do_pct, do_mgl,
depth, cdepth, level,
clevel, ph, turb,
chlfluor

Meteorology

atemp, rh, bp, wspd,
maxwspd, wdir,
sdwdir, totpar,
totprcp, cumprcp,
totsorad

Format and potential issues with output data

The raw data will look like this...

	A	B	C	D	E	F	G	H	I	J	K	L
1	StationCo	isSWMP	DateTimeStamp	Historical	Provisional	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>

Format and potential issues with output data

What are the challenges for evaluating SWMP data??

Format and potential issues with output data

What are the challenges for evaluating SWMP data??

- Knowing what we want (I can't help with this)

Format and potential issues with output data

What are the challenges for evaluating SWMP data??

- Knowing what we want (I can't help with this)
- Dealing with QAQC columns and removing 'bad' observations

Format and potential issues with output data

What are the challenges for evaluating SWMP data??

- Knowing what we want (I can't help with this)
- Dealing with QAQC columns and removing 'bad' observations
- Data we don't want... extra columns or irrelevant parameters

Format and potential issues with output data

What are the challenges for evaluating SWMP data??

- Knowing what we want (I can't help with this)
- Dealing with QAQC columns and removing 'bad' observations
- Data we don't want... extra columns or irrelevant parameters
- Combining data for comparison

Format and potential issues with output data

What are the challenges for evaluating SWMP data??

- Knowing what we want (I can't help with this)
- 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

Format and potential issues with output data

What are the challenges for evaluating SWMP data??

- Knowing what we want (I can't help with this)
- 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

We will learn how to handle most of these challenges!

Overview of the SWMP_r package

What: An R package for retrieving, organizing and analyzing SWMP data

Overview of the SWMP_r package

What: An R package for retrieving, organizing and analyzing SWMP data

Why: There are many challenges for working with SWMP data... a toolkit for addressing these challenges will be useful (I hope!)

Overview of the SWMP_r package

What: An R package for retrieving, organizing and analyzing SWMP data

Why: There are many challenges for working with SWMP data... a toolkit for addressing these challenges will be useful (I hope!)

How:

- Install R/RStudio on your computer (done already!)
- Install the SWMP_r package (dont already!)
- Use the SWMP_r functions to **retrieve**, **organize**, and **analyze** SWMP data

Overview of the SWMP_r package

This is where SWMP_r lives - [https://github.com/fawda123/SWMP_r](https://github.com/fawda123/SWMPr)

The screenshot shows the GitHub repository page for **fawda123 / SWMP_r**. The repository is described as "r package for accessing, processing, and evaluating data from SWMP of NERRS". It has 21 commits, 1 branch, 0 releases, and 1 contributor. The current branch is **master**. The commit history shows a recent commit by **fawda123** 23 minutes ago, with the message "added method to na.approx for swmpr objects". The commit details show a list of files added, including **R**, **README_files**, **data/zip_ex**, **man**, **.Rbuildignore**, **.Rprofile**, **.gitignore**, and **DESCRIPTION**. The right sidebar shows the repository's code, issues, pull requests, wiki, and settings. The HTTPS clone URL is <https://github.com/fawda123/SWMPr>.

Overview of the SWMP_r package

What is provided in the SWMP_r package?

Retrieve

```
all_params  
all_params_dtrng  
single_param  
import_local
```

Organize

```
qaqc.swmpr  
qaqcchk.swmpr  
subset.swmpr  
setstep.swmpr  
comb.swmpr
```

Analyze

```
aggregate.swmpr  
smoother.swmpr  
na.approx.swmpr  
plot.swmpr  
hist.swmpr  
lines.swmpr  
decomp.swmpr  
map_reserve
```

Built around the concept of ***object-oriented programming*** - retrieval functions return a data type with specific methods to organize and analyze

Overview of the SWMP_r package

To view the help file for any function (including examples for most):

```
?all_params
```

all_params {SWMP_r}

R Documentation

Import current station records from the CDMO

Description

Import current station records from the CDMO starting with the most current date, CDMO equivalent of `exportAllParamsXMLNew`

Usage

```
all_params(station_code, Max = 100)
```

Arguments

`station_code` chr string of station, 7 or 8 characters

`Max` numeric value for number of records to obtain from the current date, maximum of 100

Value

Returns a `swmpr` object, all available parameters including QAQC columns

Overview of the SWMP_r package

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
```

These functions require [registering your IP address](#) with CDMO
















Import data from a local path:

```
import_local
```

Allows import of data obtained from (and only from) the [zip downloads](#) feature

Overview of the SWMP_r package

After unzipping, data from [zip downloads](#) will have separate .csv files for each station and year

Name	Date modified	Type	Size
 apacpnut2011.csv	9/19/2014 7:04 AM	Microsoft Excel C...	3 KB
 apacpnut2012.csv	9/19/2014 7:04 AM	Microsoft Excel C...	3 KB
 apacpnut2013.csv	9/19/2014 7:04 AM	Microsoft Excel C...	3 KB
 apacpwq2011.csv	9/19/2014 7:06 AM	Microsoft Excel C...	5,481 KB
 apacpwq2012.csv	9/19/2014 7:06 AM	Microsoft Excel C...	5,472 KB
 apacpwq2013.csv	9/19/2014 7:06 AM	Microsoft Excel C...	5,567 KB
 apadbnut2011.csv	9/19/2014 7:06 AM	Microsoft Excel C...	3 KB
 apadbnut2012.csv	9/19/2014 7:06 AM	Microsoft Excel C...	3 KB
 apadbnut2013.csv	9/19/2014 7:06 AM	Microsoft Excel C...	3 KB
 apadbwq2011.csv	9/19/2014 7:08 AM	Microsoft Excel C...	5,407 KB
 apadbwq2012.csv	9/19/2014 7:08 AM	Microsoft Excel C...	5,483 KB
 apadbwq2013.csv	9/19/2014 7:08 AM	Microsoft Excel C...	5,337 KB
 apaebmet2011.csv	9/19/2014 7:10 AM	Microsoft Excel C...	5,453 KB
 apaebmet2012.csv	9/19/2014 7:10 AM	Microsoft Excel C...	5,401 KB
 apaebmet2013.csv	9/19/2014 7:11 AM	Microsoft Excel C...	5,669 KB

Overview of the SWMP_r package

Use the following to import some data into R...

Open script1.R, change the path to where you have the folder 'dataset1'

```
# get data for apacpwq, all years  
  
# location of data  
mypath <- 'C:/data/dataset1'  
  
# import and assign to 'dat'  
dat <- import_local(mypath, 'apacpwq', trace = T)
```

The console will return some informative text...

Overview of the SWMP_r package

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

```
head(dat)
```

```
##           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>
## 4 2011-01-01 00:45:00   11  <0>    44    <0>   28  <0>    68    <0>
## 5 2011-01-01 01:00:00   11  <0>    44    <0>   29  <0>    68    <0>
## 6 2011-01-01 01:15:00   11  <0>    44    <0>   29  <0>    67    <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
## 4      6    <0>     2  <0>      2    <3>    NA  <-1>    NA    NA
## 5      6    <0>     2  <0>      2    <3>    NA  <-1>    NA    NA
## 6      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>
## 4  8 <0>    1  <0>    NA    <-1>
## 5  8 <0>    2  <0>    NA    <-1>
## 6  8 <0>    1  <0>    NA    <-1>
```



NERRS / SWMP

Data Analysis Workshop: *Time Series*

November 17, 2014

Questions??