

dbhydroR: An R package to access the DBHYDRO Environmental Database

Joseph Stachelek

June 5, 2015

1 Introduction

This document introduces the `dbhydroR` package and its associated functions. These functions are aimed at improving programmatic workflows that query the DBHYDRO Environmental Database. HTTP requests are facilitated by the `httr` (Wickham 2015) and `RCurl` (Lang 2015) packages.

2 Package installation

The R package `dbhydroR` is distributed via a `.tar.gz` (analagous to `.zip`) package archive file. This package contains the source code for package functions. In RStudio, it can be installed by navigating to `Tools -> Install Packages... -> Install from: -> Package Archive File`. Computers running the Windows operating system can only install binary `.zip` package archive files unless they have additional compiler software installed. The `dbhydroR` binary package can be installed by running the following command from the R console:

```
> install.packages("B:\\restoration_sciences\\projects\\joe_stachelek  
+                 \\R\\dbhydro_0.1-1.zip", type="win.binary",  
+                 repos=NULL, dependencies=TRUE)
```

Once installed, the package can be loaded using the following command:

```
> library(dbhydroR)
```

3 Composing database queries

The workhorse `dbhydroR` function is `dbhydro_get`. This function takes four required arguments. The user must specify a station ID, a test name, and a date range.

Station IDs can be located on the [SFWMD Google Earth portal](#). A list of available test names can be found in the [Appendix](#) to this document. Dates must be specified in YYYY-MM-DD format (e.g. 2015-02-26). The following set of examples retrieve measurements between March 2011 and May 2012. They can be run from the R console by issuing the command:

```
> example(dbhydro_get)
```

- One variable at one station

```
> dbhydro_get(station_id="FLAB08", date_min="2011-03-01",  
+             date_max="2012-05-01", test_name="CHLOROPHYLLA-SALINE")
```

- One variable at multiple stations

```
> dbhydro_get(station_id=c("FLAB08", "FLAB09"), date_min="2011-03-01",  
+             date_max="2012-05-01", test_name="CHLOROPHYLLA-SALINE")
```

- One variable at a wildcard station

```
> dbhydro_get(station_id=c("FLAB0%"), date_min="2011-03-01",  
+             date_max="2012-05-01", test_name="CHLOROPHYLLA-SALINE")
```

- Multiple variables at multiple stations

```
> dbhydro_get(station_id=c("FLAB08", "FLAB09"), date_min="2011-03-01",  
+             date_max="2012-05-01", test_name=c("CHLOROPHYLLA-SALINE",  
+                                               "SALINITY"))
```

By default, `dbhydro_get` returns a "cleaned output". First, the cleaning function `dbhydro_clean` converts the raw output from native DBHYDRO "long" format (each piece of data on its own row) to "wide" format (each site x variable combination in its own column) using the `reshape2` package ([Wickham 2007](#)). Next, the QA blanks/flags and DBHydro database identifiers are stripped. Setting the `raw` flag to `TRUE` will force `dbhydro_get` to retain this information. An example query that retains this information is shown below.

```
> dbhydro_get(station_id="FLAB08", date_min="2011-03-01",  
+             date_max="2011-05-01", test_name="CHLOROPHYLLA-SALINE",  
+             raw=TRUE)
```

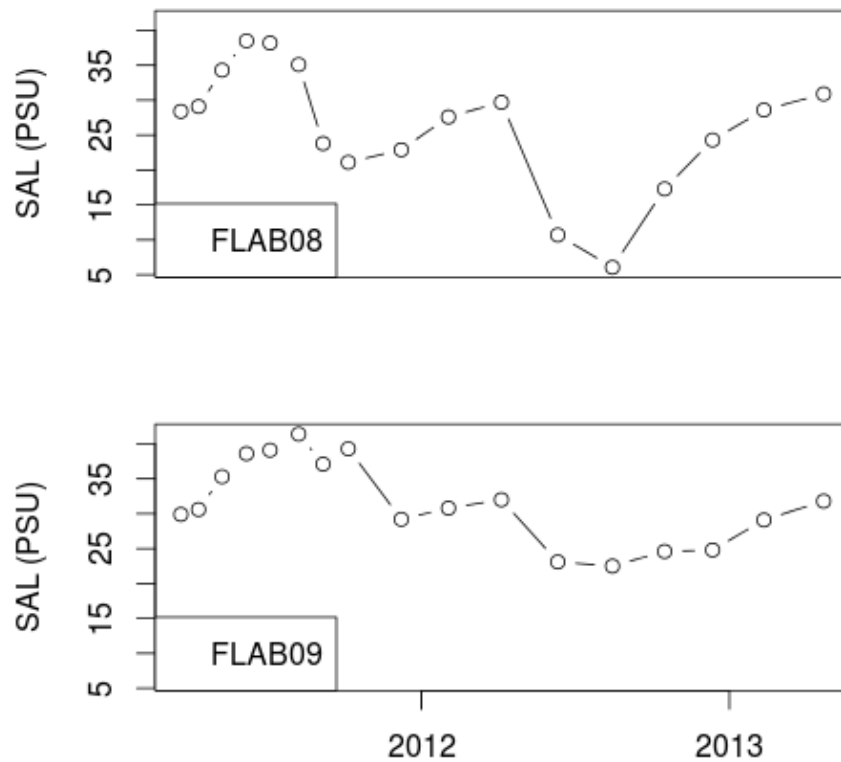
4 Plotting

Once data has been retrieved with `dbhydro_get` in non-raw form (or cleaned with `dbhydro_clean`), and assigned to an R object, plots can be produced using the `dbhydro_plot` function. This function will create a dedicated panel of figures for each of the available variables. The following example retrieves and plots measurements between March 2011 and May 2014. It can be run from the R console by issuing the command:

```
> example(dbhydro_plot)
```

or

```
> dt<-dbhydro_get(station_id=c("FLAB08","FLAB09"),  
+                 date_min="2011-03-01",date_max="2014-05-01",  
+                 test_name=c("CHLOROPHYLLA-SALINE","SALINITY"))  
> dbhydro_plot(dt,label="bottomleft",abb=3)
```



5 Appendix

5.1 Test names

There are many test names available in DBHYDRO. These are detailed in the following table.

Code
ALKALINE PHOSPHATASE
AMMONIA-N
CARBON, TOTAL ORGANIC
CAROTENOIDS
CHLOROPHYLL-A
CHLOROPHYLL-A(LC)
CHLOROPHYLL-A, CORRECTED
CHLOROPHYLL-B
CHLOROPHYLL-B(LC)
CHLOROPHYLL-C
CHLOROPHYLLA-SALINE
DEPTH,TOTAL
DISSOLVED OXYGEN
KJELDAHL NITROGEN,TOTAL
NITRATE+NITRITE-N
NITRITE-N
ORP
PAR-LIGHT ATTEN COEF
PH,FIELD
PHEOPHYTIN
PHEOPHYTIN-A(LC)
PHOSPHATE,ORTHO AS P
PHOSPHATE,TOTAL AS P
SALINITY
SECCHI DISK DEPTH
SILICA
SP CONDUCTIVITY,FIELD
SP CONDUCTIVITY,LAB
TEMP
TOTAL NITROGEN
TURBIDITY

5.2 Further reading

See section on URL-based data access in the [DBHYDRO Browser User's Guide](#)

References

Duncan Temple Lang. *RCurl: General network (HTTP/FTP/...) client interface for R*, 2015. URL <http://CRAN.R-project.org/package=RCurl>. R package version 1.95-4.6.

Hadley Wickham. Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12):1–20, 2007. URL <http://www.jstatsoft.org/v21/i12/>.

Hadley Wickham. *httr: Tools for Working with URLs and HTTP*, 2015. URL <http://CRAN.R-project.org/package=httr>. R package version 0.6.1.