HBEF Dashboard Documentation

The HBEF dashboard was created using R, shiny, and a MySQL database. It is hosted on a Digital Ocean server.

# For Users

## Accessing Dashboard

The HBEF dashboard can be accessed from the HBEF Team Data Portal link at the top of the current HBEF site (<http://hbef.streampulse.org/>).

## Important Things to Note When Uploading/Downloading data

* Whether you upload ‘current’ data, make sure the names of your data columns match those of the data template exactly. (I’d suggest using the “DataTemplate\_current.xlsx” file.)
* You may only upload .csv files, i.e. comma-separated values.
* If at all possible, do not use commas anywhere in the data. In places where commas are used, they will be replaced with a semicolon. (This ensures that the .csv files download properly.)
* The “date” and “datetime” columns are temperamental, and only seem to work if the entire column is formatted very specifically. For “date”, it needs to be formatted as ‘Date’ in the “3/14/12” format. (To do this, highlight the entire “date” column, right-click on your highlighted column, select “Format Cells…”, select “Date” from the Category menu, and choose the ‘3/14/12’ format from the Type menu–this is important, even choosing the ‘3/14/2012’ format will cause problems, stick to the mm/dd/yy format.) For “datetime”, it needs to be formatted in the “3/14/12 13:30” format, also found in the “Date” Category menu.
* Ideally, blank spaces should be filled with the letters “NA”.
  + If “date” or “timeEST” has no data, must be NA (not -9999 as done previously)
* uniqueID’s:
  + The “uniqueID” column can be created with the following function when cell B2 contains the site, cell C2 contains the date, and cell D2 contains the time: =CONCATENATE(B2, "\_", TEXT(C2, "yyyymmdd"), "\_", TEXT(D2,"[hh]mm"))
  + The “uniqueID” column is easiest to create when the “date” column is formatted as a date (the default date format type in Microsoft Excel works, i.e. 3/14/12), and the “timeEST” column is formatted as time (in the format of military time, e.g. 13:30). \***Note:\*** **In fact, it appears the data only uploads properly when they columns have been formatted as such.**
  + If you see errors when you upload the data, check that the “uniqueID” column has no duplicates (and no spaces)
* If a brand new site is included in the data (i.e. not on the current sites lists within the dashboard), tell the developer before uploading the data. The developer needs to add this site to the list of sites in the code.
* When you download data, you can ignore the “refNo” (reference numbers) column. This is a column that is more useful within the MySQL database. Also, do not assume that data exists for every reference number in the range you see; i.e. there will always be gaps in the sequence of numbers you see in “refNo.”

## Important Things to Note with Data & Edits Table

If a cell turns red after you make a change, you may be entering the wrong type of data in the column, and this change will not be saved. For example, if you enter text (e.g. “testing”) in the *pH* column, it will turn red because you can only enter numbers (e.g. “4.53”) in this cell.

Like with data upload, the same guideline for the *notes* and *sampleType* columns apply here - if you can avoid using commas in the notes, that'd be great!

# For Developers

## Accessing Dashboard Files

The files pertaining to the HBEF dashboard can be found in the restricted\_QAQC folder of <https://github.com/HBEF/hbef>.

## Dashboard Purpose

The long term ecological research project at Hubbard Brook Experimental Forest in New Hampshire has been collecting weekly stream and precipitation samples since the 1960’s.

The team carrying out this work needed a centralized place to load, view, and edit the incoming streams of data.

There was an explicit need to upload, view, and edit the data that is currently under review (what we call ‘current’ data), in addition to viewing the data that has already been QA/QC’d and finalized (what we call ‘historical’ data).

This HBEF dashboard was created to fill these needs.

## Developer Notes

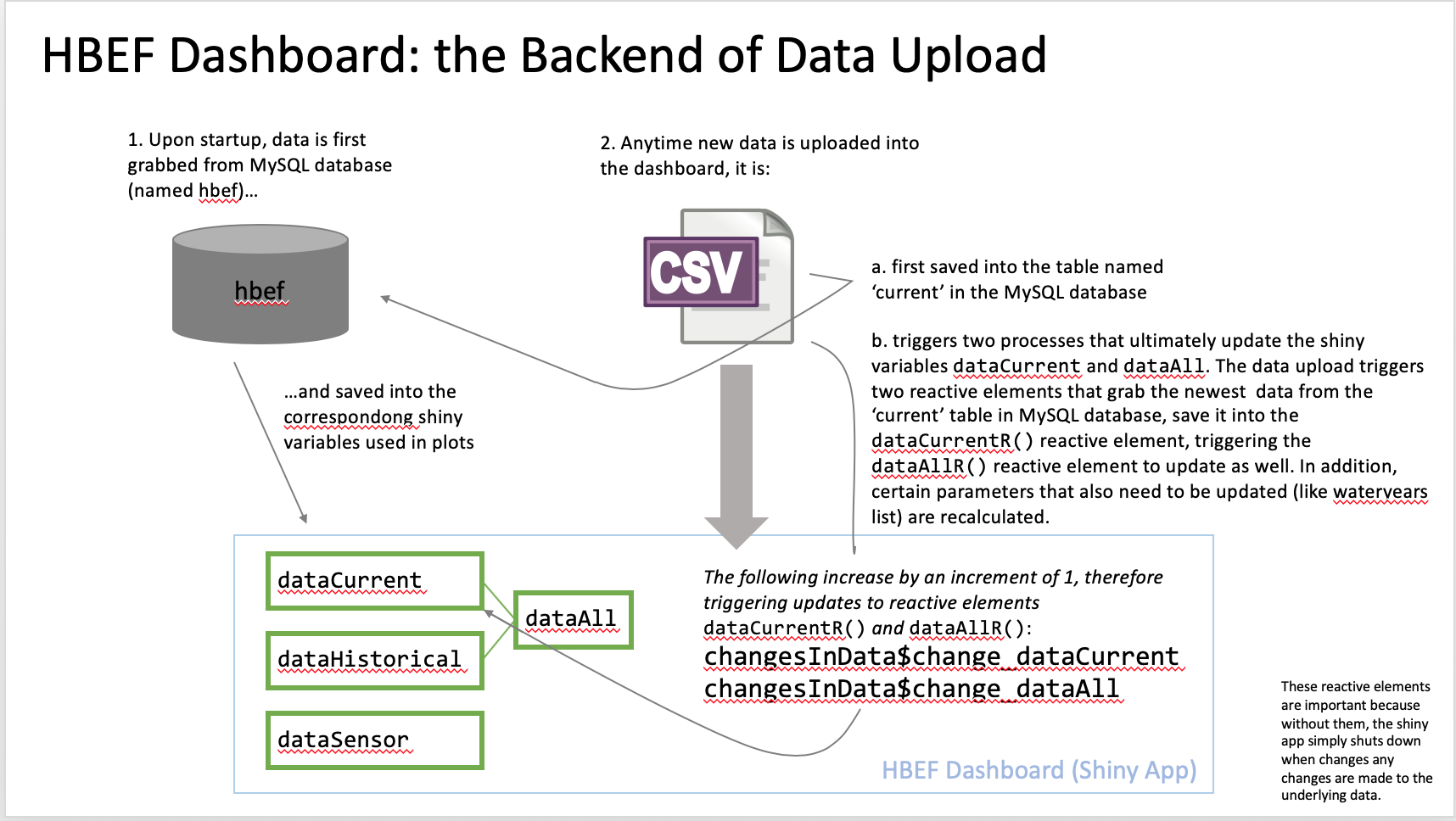
See VersionControlWorkflow\_GitHubMap.pdf document for ideal workflow.

Generally, R & shiny code written into local hbef/restricted\_QAQC folder, which is merged into GitHub user HBEF’s hbef repository (user is HBEF, repo is named hbef).

Code from GitHub user HBEF then pushed to Digital Ocean server. The content in this folder is what’s shown on the hbef.streampulse.org website.

Data is appended to or extracted from MySQL database, which is also set up on Digital Ocean server.

See DataUpload\_Backend\_Explanation.pptx for guidelines on how data uploads were constructed on the backend. The main shiny variables used are dataCurrent, data Historical, dataAll, and dataSensor. Summary slide here:



## Data Structure Notes

### Variable Naming conventions

Variables begin with lowercase letters.

All input variables are written in all caps, for example DATE4.

A number at the end of a variable indicates what panel it’s a part of. So DATE4 is the date input from Panel 4 in the app.

The main identifier of the data is the “uniqueID”, which contains the site\_date\_time information. If there is more than one entry with a certain site\_date\_time, the additional entries have a “\_Dup” (for duplicate), and further “\_Dup2,” “\_Dup3,” etc., added to the end of the uniqueID.

### R Packages Used

Panel 1 – graphs made with dygraphs package

Panel 2 – graphs made with dygraphs package

Panel 3 – graphs made with dygraphs package

Panel 4 – graphs made with ggplot2 package

Panel 5 – editable table made with rhandsontable package

And additional helpful package to troubleshoot reactivity is reactlog: <https://cran.r-project.org/web/packages/reactlog/vignettes/reactlog.html>

### Descriptions of Tables within *hbef* MySQL Database & the Variables They’re Assigned to:

*current* This table contains the most recent lab data generated after samples

have been collected from the field. The initial lab analyses are run at Hubbard Brook, before being sent to Durham for solute analysis. At this point in time (Jul 2019), this data is collected by Brenda Minicucci and Tammy Wooster, and then processed further by Jeffrey Merriam.

**This data is assigned to the** dataCurrent **variable in the shiny app.**

*historical* All the data from the ‘current’ table that has gone through QA/QC, and

has been approved for archiving by all members of the team.

Some notes on data columns here: both stream and precipitation data are

included here. They can be distinguished from the site data column. Precipitation sites before 2013 have N or S as a site, while values after 2013 have the actual rain gage site names.

**This data is assigned to the** dataHistorical **variable in the shiny app.**

**The** dataAll **variable in the shiny app is simply the** dataCurrent **and** dataHistorical **tables joined together.**

*sensor2* Streamflow and precipitation data that comes from the Forest Service’s

ETI sensors.

**This data is assigned to the** dataSensor **variable in the shiny app.**

*sensor* This table is not used.

For historical data, several solutes or data were measured with different instruments throughout time. See Don Buso *et al.* 2000 for more details.

## Some history of pH and Al measurements

**pH**

YearsNotes/Instrument Used for Measurement

1964-1972 Beckman analog meter (Zeromatic or model N)

1972-1996 Orion analog meters (models 401, 407A)

1996-2011 Orion digital meters (models 710A and 940) [710A used in long

term database.]

2011-2015 3 Star and Orion digital meters

2015-present Metrohm meter begins to be used

*2015-2017* entries in database from Orion 940 and Metrohm

*Sept. 19, 2017* Orion 904 fails.

*Oct. 11, 2017* 3 Star meter replaces Orion, preferred over Metrohm

**Aluminum (Al)**

Years Notes/Instrument Used for Measurement

1964-1970 Spectrophotometer (Beckman model B)

1976-2008 Spectrophotometer, dual-beam (Coleman PE model 55, or

Shimadzu model 60)

*In sites*: W-6, RG-1, RG-11

*1995-present* *In sites*: W-7, W-8, W-9

Technicon AutoAnalyzer II?

2008-present Aluminum measurements stop at The Cary Institute, and are

transferred to Forest Service, who use Total Monomeric Aluminum

measurements [this is used in long term database].

*In sites*: W-6, RG-1, RG-11

*2011-present* *In sites*: W-1, W-2, W-3, W-4, W-5, RG-23, and Hubbard Brook

## Authors

Carolina Simao Roe-Raymond

*Additional Contributors*

Mike Vlah\*

Aaron Berdanier

original shiny code: <https://github.com/berdaniera/StreamPULSE>

Developers of hbef.streampulse.org:

Richard Marinos (led effort)

Camila Vargas (worked on coding)

Annie Lott (worked on coding)

The HBEF team members heavily involved in the creation of this dashboard were:

Dr. Emily Bernhardt, Dr. Emma Rosi, Tammy Wooster, Brenda Minicucci, and Jeffrey Merriam.

\*Current primary contact

## Miscellaneous Notes from Dashboard Development

### MySQL Database Table Construction

For data with decimal numbers, chose to use DECIMAL data type instead of FLOAT/DOUBLE because it seems FLOAT/DOUBLE can be less precise and cause errors. See: <https://dev.mysql.com/doc/refman/5.7/en/problems-with-float.html>

For data with DECIMAL numbers, most values were given 4 decimal places, but didn’t need as many digits. However, MySQL requires number of digits (M) to be greater than or equal to number of decimals places (D), therefore many number of digits had to be made bigger than they needed to be.

Database name in MySQL: hbef

### Restricted Site Info

Info given by Richard Marinos:

“It seems that the restricted directory got overwritten due to Annie cloning an updated version of the GitHub repo to the directory. I thought I had taken out the restricted folder from the local repo index, but I guess I hadn't. Or maybe she had deleted everything in the folder to start fresh. I dunno.

Anyway, I rebuilt the directory structure so you can put stuff in the restricted zone. The folder is /home/hbef/shiny/restricted and any HTML you want to serve restricted goes in /home/hbef/HTML/restricted.

The site config file that establishes this configuration is /etc/nginx/sites-enabled/hbefshiny . I had to set up authentication through the http server because Shiny only offers native authentication for the pro version. It works just fine though. If you need to change the http server configuration, be very careful because the streampulse sites also rely on nginx. The fact that we have individual site config files should sandbox our stuff somewhat so you should be OK to edit the file above, but definitely be careful if you end up playing with /etc/nginx/nginx.config”

Update:

The restricted directory folder used for this app is actually now restricted\_QAQC.

# References

Buso, D.C., Likens, G.E., Eaton, J.S. 2000. Chemistry of Precipitation, Streamwater, and Lakewater from the Hubbard Brook Ecosystem Study: A Record of Sampling Protocols and Analytical Procedures. United States Department of Agriculture, Forest Service, Northeastern Research Station. General Technical Report NE-275.