Introduction to USGSHydroOpt

Samuel Christel[1] and Steve Corsi[1]

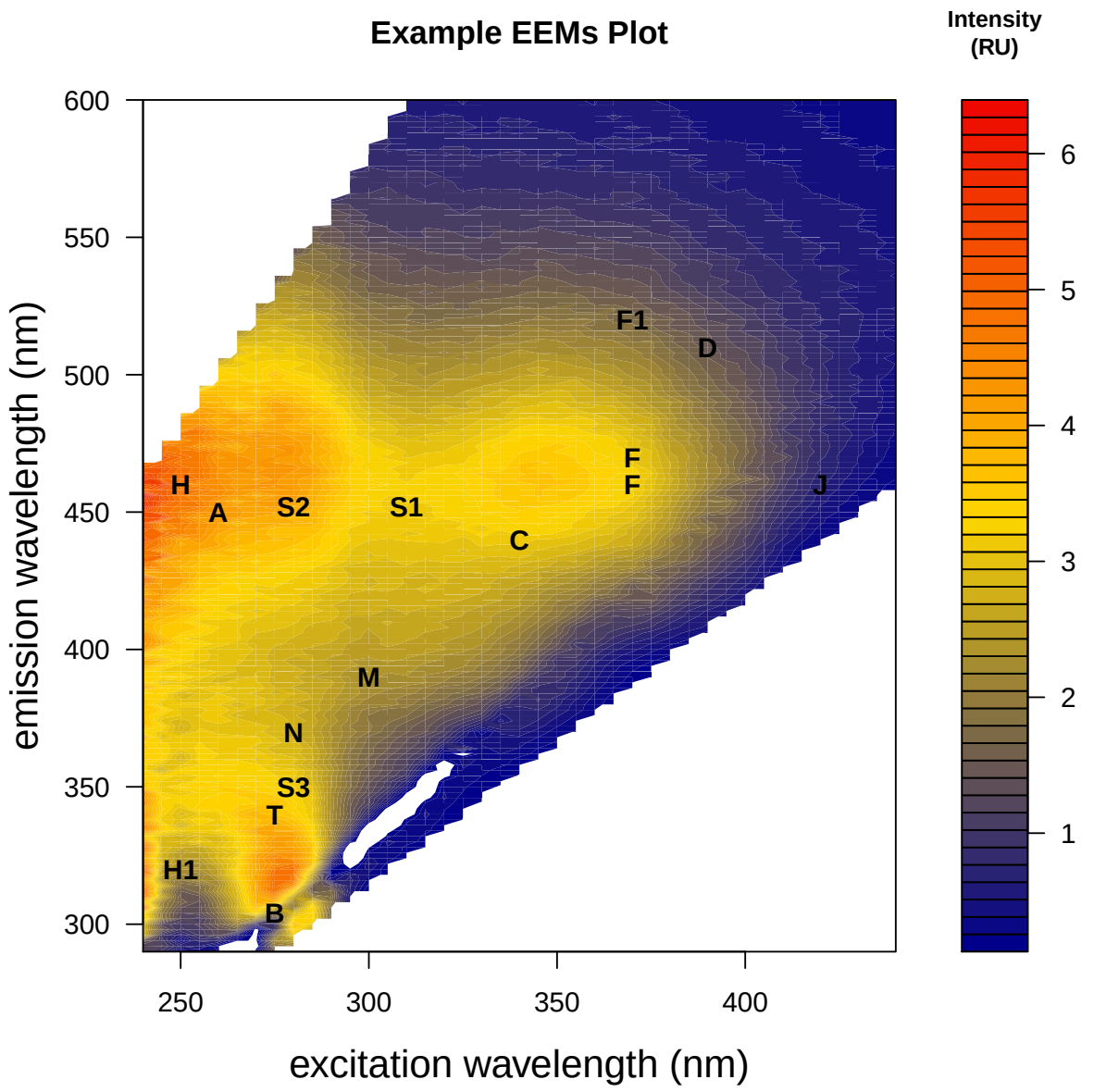
[1] *United States Geological Survey*

June 24, 2014

Contents

1 Introduction to USGSHydroOpt

The USGSHydroOpt package was created to streamline the process of creating optical summary variables and excitation-emission (EEMs) plots for absorbance and fluoresence data collected from various freshwater sources. Examples of optical summary variables that can be produced with this package include various absorbance peaks, The functions in this package were designed to operate on dataframes with a standardized structures. This package is not amenable to dataframes that do not fit the prescribed format. The example dataframes in this package illustrate exactly how dataframes should be formatted, and the examples illustrate how the functions operate on the dataframes. Depicted below is an example of an EEMs plot produced with this package.




```
NULL
```

2 Dataframe Formatting for USGS HydroOpt

The functions contained in USGSHydroOpt operate on dataframes with defined structures. Users interested in using USGSHydroOpt should format dataframes according to the structures defined in this section.

2.1 Absorbance Data

Absorbance data used by functions in USGSHydroOpt should be formatted such that each sample occupies a column, and one column contains the wavelength (nm) for which the absorbance measurement was mea-

sured (*See example below*). The column with the wavelengths does not need to be called "wavelengths," as it is named in the example dataframe below. Since this package was developed primarily for USGS activities, the default for naming samples is "gr" then the sample number. This convention was started by the USGS California Water Sciences Center (CA WSC) and the USGS Wisconsin Water Science Center (WI WSC) follows the same naming convention to ensure standardization.

```
     gr13307    gr13351    gr13353    gr13357 wavelengths
1  0.0001296 -0.0003505 -0.0003480 -0.0002695         750
2 -0.0002367  0.0000305 -0.0000915 -0.0001407         749
3 -0.0001582 -0.0004900 -0.0006325 -0.0001534         748
4 -0.0004642 -0.0000105 -0.0000932 -0.0000245         747
5 -0.0002551 -0.0000653  0.0000841 -0.0001615         746
6 -0.0001842 -0.0002135  0.0002082  0.0001429         745
```

2.2 Fluoresence Data

Fluoresence data used by functions in USGSHydroOpt should also be formatted such that each sample occupies a column, and one column contains the excitation emission wavelength pairs (nm) for which the fluoresence measurment was measured (*See example below*). The column with the excitation emission pairs does not need to be called "Wavelength.Pairs," as it is in the example dataframe below. Again since this package was developed for USGS activities, the default sample naming convention is "gr" followed by the sample number.

```
Error:  object 'dfFluor' not found
```

2.3 Spectral Slopes Data

Information on the upper and lower wavelength (nm) for which a spectral slope should be calculated needs to be stored in a dataframe if USGSHydroOpt is used. The dataframe should contain exactly three columns. The first column should contain the upper wavelength, the second column should contain the lower wavelength, and the third column should contain the name of the spectral slope being calculated (*See example below*). The columns need to be in this exact order, although the names of the columns may be different. The data types for each column are integer, integer, and character, respectively. More spectral slopes can be added to the table than specified in the example dataframe below.

```
Error:  object 'dfsags' not found
```

2.4 Optical Summary Data

This is the dataframe that contains many of the summary optical variables that can be produced using functions in USGSHydroOpt (*See example below*). The functions in USGSHydroOpt calculate summary optical variables and add to a dataframe formatted according to the example below. The example dataframe below is how the WI WSC stores optical summary variables. Note that this dataframe can contain other columns with metadata, for example, the sample data and time, the sample ID, or whether or not the sample went through QA/QC.

```
  GRnumber          B      T      A       J FI_2005    A254
1  gr13307  0.050997 0.1826 0.4553 0.03705   1.639 0.05228
2  gr13351 -0.245602 0.4085 3.0783 0.19110   1.456 0.43531
3  gr13353 -0.175220 0.7794 6.8624 0.56309   1.523 0.68274
4  gr13357  0.111561 0.2691 0.9451 0.06969   1.572 0.08698
5  gr13360 -0.001569 0.4593 2.8254 0.37231   1.563 0.28605
6  gr13363  0.052137 0.4892 1.9518 0.19098   1.583 0.19283
```

However, also note that summary optical variable names in the dataframe must be identical to those specified in the table below.

```
 [1] "OB1"         "OB2"         "OB3"         "S1.50"
 [5] "S2.50"       "S3.50"       "S1.25"       "S2.25"
 [9] "S3.25"       "Mrange.25"   "Mrange.50"   "B"
[13] "T"           "M"           "A"           "C"
[17] "N"           "D"           "F"           "J"
[21] "S1"          "S2"          "S3"          "H1"
[25] "H2"          "F1"          "F2"          "W"
[29] "LT1"         "LT2"         "LT3"         "LA"
[33] "HIX_2002"    "FI_2005"     "FI_2001"     "FreshI"
[37] "A254"        "A275"        "A280"        "A290"
[41] "A295"        "A350"        "A370"        "A400"
[45] "A412"        "A440"        "A488"        "A510"
[49] "A532"        "A555"        "A650"        "A676"
[53] "A715"        "Sag275_290"  "Sag290_350"  "Sag350_400"
[57] "Sag412_676"  "Aresids"
```

2.5 Excitation-Emission (EEMs) Peak Data

The EEMs peak data contains three columns listing the name of a characterized EEM peak along with the corresponding wavelengths (nm) (*See example below*). The first column, "Peak," contains the name of the characterized EEM peak. The next two columns contain the excitation and emission wavelengths (nm) at which a given peak occurs. The column names must be identical to those displayed in the example below, although the order of the columns can be different.

```
   Peak ExCA EmCA
1     B  275  304
2     T  275  340
3     M  300  390
4     A  260  450
5     C  340  440
6     N  280  370
7     D  390  510
8     F  370  460
9     J  420  460
10   S1  310  452
11   S2  280  452
12   S3  280  350
```

```
13    H  250  460
14   H1  250  320
15    F  370  470
16   F1  370  520
```