

Import Onset HOBOLogger Data into R

Matthew Van Scoyoc, National Park Service

April 22, 2022

Contents

1	Introduction	1
2	Installation	1
3	Import Data into R	2
3.1	Using <i>import_hobo()</i>	2
4	Summarize Data	4
4.1	Precipitation (Event) Data with <i>raindance()</i>	4
4.2	Temperature and Relative Humidity Data with <i>sundance()</i>	4
5	Processing HOBOLogger files	5
6	Reporting Errors and Issues	6
7	Next Steps	6

1 Introduction

This R package imports data from Onset temperature, relative humidity, and precipitation (event) data loggers into R. Data collected in the field using Onset loggers are exported to comma delimited (csv) files using the Onset HOBOLogger application. This package imports the csv files into R and summarizes the data.

The structure of the csv files generated from HOBOLogger vary wildly. They can have anywhere from 4 to 10 columns and the logger details are usually in two “hidden” columns following the data. This makes data from the temperature, relative humidity, and precipitation loggers difficult to read into R. This package imports data collected by Onset loggers used in the Southeast Utah Group (SEUG) national parks long-term vegetation monitoring program (LTVMP) and then summarizes these data. This package was written to be used with the *dataprocessR* package, <https://github.com/scoyoc/dataprocessR>, that exports the raw and summarized data to the SEUG LTVMP database (a Microsoft Access database).

2 Installation

This package is available on GitHub at <https://github.com/scoyoc/raindancer>. Dependent packages include *dplyr*, *glue*, *lubridate*, *RODBC*, *stringr*, *tibble*, *tidyr*, and *utils*. Suggested packages include *janitor*, *knitr*, *rmarkdown*, and *readr*.

```
if (!"devtools" %in% installed.packages()[, "Package"]) {  
  install.packages("devtools")  
}
```

```
devtools::install_github("scoyoc/raindancer")
library("raindancer")
```

3 Import Data into R

There are two functions that import the csv files generated from HOBOWare into R, *import_hobo_2008()* and *import_hobo()*. Below is a table of Onset loggers that these functions can import data from.

Table 1. Onset loggers used in the SEUG LTVMP.

Product	Element	Year Used
H07 Logger	PRCP	2008-2019
HOB0 UA-003-64 Pendant Temp/Event	PRCP & TEMP	2019-present
H08 Logger	TEMP	2008-2019
HOB0 UA-001-64 Pendant Temp	TEMP	2019-present
HOB0 U23-001 Temp/RH	TEMP & RH	2019-present

3.1 Using *import_hobo()*

The functions *import_hobo_2008()* and *import_hobo()* return a list with three components:

1. **file_info** is a one row data frame that contains information about the file, the logger, and data.
2. **details** is a data frame of logger and sampling event information.
3. **data_raw** is a data frame of raw data.

Lets start by usgin *list.files()* to bring a list of csv files into R.

```
file_list <- list.files(path = "C:/path/to/data", pattern = ".csv",
                        full.names = TRUE, recursive = FALSE)
```

There are some files included in this package for examples, so we'll use these for the vignette.

```
file_list <- list.files(path = system.file("extdata", package = "raindancer"),
                        pattern = ".csv", full.names = TRUE, recursive = FALSE)
print(file_list)
#> [1] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/2010_h01st_prpc.csv"
#> [2] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/2010_h01st_temp.csv"
#> [3] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/2012_i01gp_prpc.csv"
#> [4] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/2012_i01gp_temp.csv"
#> [5] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/20200416_A03_PRCP.csv"
#> [6] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/20200416_A03_TEMP_RH.csv"
#> [7] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/20200416_A08_PRCP.csv"
#> [8] "C:/Users/mvanscoyoc/Documents/R/library/raindancer/extdata/20200416_A08_TEMP_RH.csv"
```

The first four csv files are from 2010 and 2012, so lets start by using *import_hobo_2008()*.

```
dat.1 <- import_hobo_2008(file_list[1])
```

Let's examine the components of the list returned by the *import_hobo_2008()* function. The first component returns information about the file, the logger, and the data.

```
dplyr::glimpse(dat.1$file_info)
#> Rows: 1
#> Columns: 10
#> $ FileName      <chr> "2010_h01st_prpc.csv"
#> $ PlotID        <chr> "H01"
```

```
#> $ Element      <chr> "PRCP"
#> $ Product      <chr> "H07 Logger"
#> $ SerialNumber  <chr> "707910"
#> $ LaunchName    <chr> "H1 Square Tower"
#> $ DeploymentNumber <chr> "10"
#> $ LaunchTime    <chr> "05/06/09 11:04:53 AM GMT-06:00"
#> $ FirstSampleTime <chr> "05/06/09 11:04:53 AM GMT-06:00"
#> $ LastSampleTime <chr> "04/08/10 06:20:45 PM GMT-06:00"
```

The second component returns a three column data frame with information about the logger and sampling event.

```
dat.1$details
#> # A tibble: 21 x 3
#>   FileName      Details      Value
#>   <chr>         <chr>      <chr>
#> 1 2010_h01st_prcp.csv Battery at Launch Unknown
#> 2 2010_h01st_prcp.csv Deployment Number 10
#> 3 2010_h01st_prcp.csv Device Memory 32768
#> 4 2010_h01st_prcp.csv First Sample Time 05/06/09 11:04:53 AM GMT-06:00
#> 5 2010_h01st_prcp.csv Full Series Name Events
#> 6 2010_h01st_prcp.csv Last Sample Time 04/08/10 06:20:45 PM GMT-06:00
#> 7 2010_h01st_prcp.csv Launch Name H1 Square Tower
#> 8 2010_h01st_prcp.csv Launch Time 05/06/09 11:04:53 AM GMT-06:00
#> 9 2010_h01st_prcp.csv Product H07 Logger
#> 10 2010_h01st_prcp.csv Samples 1
#> # ... with 11 more rows
```

And the third component returns a data frame of raw data.

```
dplyr::glimpse(dat.1$data_raw)
#> Rows: 1,458
#> Columns: 6
#> $ FileName <chr> "2010_h01st_prcp.csv", "2010_h01st_prcp.csv", "2010_h01st_prc~
#> $ PlotID <chr> "H01", "H01", "H01", "H01", "H01", "H01", "H01", "H01", "H01"~
#> $ DateTime <dtm> 2009-05-06 11:04:53, 2009-05-06 11:04:54, 2009-05-06 11:04:5~
#> $ Element <chr> "PRCP", "PRCP", "PRCP", "PRCP", "PRCP", "PRCP", "PRCP", "PRCP", "PRCP~
#> $ Value <dbl> 79569, 79570, 79571, 79572, 79573, 79574, 79575, 79576, 79577~
#> $ Units <chr> "Events", "Events", "Events", "Events", "Events", "Events", "Events", "~
```

In 2019 SEUG resource staff upgraded the loggers in hopes of preventing logger failure and maintaining the weather data set for the LTVMP. The new loggers had a different file structure that the older loggers, requiring a new function to import data into R. Files five through eight are examples from the new loggers. Let's use `import_hobo()` to bring data from one of these files into R and examine the information and data.

```
dat.6 <- import_hobo(file_list[6])
str(dat.6)
#> List of 3
#> $ file_info:'data.frame': 2 obs. of 10 variables:
#> ..$ FileName : chr [1:2] "20200416_A03_TEMP_RH.csv" "20200416_A03_TEMP_RH.csv"
#> ..$ PlotID : chr [1:2] "A03" "A03"
#> ..$ Element : chr [1:2] "TEMP" "RH"
#> ..$ Product : chr [1:2] "HOBO U23-001 Temp/RH" "HOBO U23-001 Temp/RH"
#> ..$ SerialNumber : chr [1:2] "20547370" "20547370"
#> ..$ LaunchName : chr [1:2] "A03_RH" "A03_RH"
#> ..$ DeploymentNumber: chr [1:2] "1" "1"
```

```
#> ..$ LaunchTime      : chr [1:2] "03/26/19 08:15:04 AM GMT-06:00" "03/26/19 08:15:04 AM GMT-06:00"
#> ..$ FirstSampleTime : chr [1:2] "03/26/19 05:00:00 PM GMT-06:00" "03/26/19 05:00:00 PM GMT-06:00"
#> ..$ LastSampleTime  : chr [1:2] "04/16/20 03:00:00 PM GMT-06:00" "04/16/20 03:00:00 PM GMT-06:00"
#> $ details : tibble [46 x 3] (S3: tbl_df/tbl/data.frame)
#> ..$ FileName: chr [1:46] "20200416_A03_TEMP_RH.csv" "20200416_A03_TEMP_RH.csv" "20200416_A03_TEMP_RH.csv" ...
#> ..$ Details : chr [1:46] "Product" "Serial Number" "Version Number" "Manufacturer" ...
#> ..$ Value : chr [1:46] "HOB0 U23-001 Temp/RH" "20547370" "1.10" "Onset Computer Corporation" ...
#> $ data_raw : tibble [37,146 x 6] (S3: tbl_df/tbl/data.frame)
#> ..$ FileName: chr [1:37146] "20200416_A03_TEMP_RH.csv" "20200416_A03_TEMP_RH.csv" "20200416_A03_TEMP_RH.csv" ...
#> ..$ PlotID : chr [1:37146] "A03" "A03" "A03" "A03" ...
#> ..$ DateTime: POSIXct[1:37146], format: "2001-01-20 00:00:00" "2001-01-20 00:00:00" ...
#> ..$ Element : chr [1:37146] "RH" "TEMP" "RH" "TEMP" ...
#> ..$ Value : num [1:37146] 88.1 -13.4 90.5 -14.1 90.8 ...
#> ..$ Units : chr [1:37146] "%RH" "F" "%RH" "F" ...
```

4 Summarize Data

Let's summarize these data up to daily values now that we have data read into R. The data are either event data from precipitation gauges or recorded at set intervals through out the day for temperature and relative humidity data.

4.1 Precipitation (Event) Data with *raindance()*

The precipitation loggers record an event every time the bucket inside the rain gauge fills with 0.254 mm of water and tips to trigger an event. The data recorded is simply an event (e.g., event 1, event 2, event 3, and so on). The *raindance()* function calculates hourly precipitation, provides the number of tips per hour, and provides an estimate of intensity with maximum tips per minute in a given hour. This function requires the *data_raw* component from *import_hobo()* or *import_hobo_2008()* functions and returns a data frame of hourly data.

```
raindance(dat.1$data_raw)
#> # A tibble: 8,100 x 6
#> # Groups:   PlotID, DateTime [8,100]
#>   PlotID DateTime      Element PRCP_mm Tips MaxTips_min
#>   <chr>   <dtm>         <chr>    <dbl> <dbl>      <dbl>
#> 1 H01    2009-05-06 01:00:00 PRCP         0     0         0
#> 2 H01    2009-05-06 02:00:00 PRCP         0     0         0
#> 3 H01    2009-05-06 03:00:00 PRCP         0     0         0
#> 4 H01    2009-05-06 04:00:00 PRCP         0     0         0
#> 5 H01    2009-05-06 05:00:00 PRCP         0     0         0
#> 6 H01    2009-05-06 06:00:00 PRCP         0     0         0
#> 7 H01    2009-05-06 07:00:00 PRCP         0     0         0
#> 8 H01    2009-05-06 08:00:00 PRCP         0     0         0
#> 9 H01    2009-05-06 09:00:00 PRCP         0     0         0
#> 10 H01   2009-05-06 10:00:00 PRCP         0     0         0
#> # ... with 8,090 more rows
```

4.2 Temperature and Relative Humidity Data with *sundance()*

Temperature and relative humidity loggers record data at set intervals through out the day. The *sundance()* function summarizes these data to daily values, providing mean, minimum, maximum, and the number of measurements (n) for a given day. This function also returns the time that the minimum and maximum were recorded. This function requires the *data_raw* component from *import_hobo()* or *import_hobo_2008()*

functions and returns a data frame of daily data.

```
sundance(dat.6$data_raw)
#> # A tibble: 776 x 10
#> # Groups:   PlotID, Date, Element [776]
#>   PlotID Date      Element Mean    Min    Max      n MinTime MaxTime Units
#>   <chr> <date>    <chr> <dbl> <dbl> <dbl> <int> <chr> <chr> <chr>
#> 1 A03 2001-01-20 RH      90.6  80.3  96.8    48 15:00 07:00 %RH
#> 2 A03 2001-01-20 TEMP    -8.18 -15.3  -3.01    48 05:00 15:30 F
#> 3 A03 2001-02-20 RH      94.2  86.2  98.2    48 14:30 00:00 %RH
#> 4 A03 2001-02-20 TEMP    -4.78  -6.20  -2.33    48 00:00 16:30 F
#> 5 A03 2001-03-20 RH      92.3  76.3  97.5    48 15:00 04:00 %RH
#> 6 A03 2001-03-20 TEMP    -5.98 -12.4  -0.591    48 23:30 15:30 F
#> 7 A03 2001-04-20 RH      96.2  84.7  99.7    48 14:30 23:30 %RH
#> 8 A03 2001-04-20 TEMP    -5.75 -12.6  -2.04    48 00:30 14:00 F
#> 9 A03 2001-05-20 RH      98.3  95.2  100    48 15:30 01:00 %RH
#> 10 A03 2001-05-20 TEMP    -5.46  -6.80  -3.81    48 08:30 12:30 F
#> # ... with 766 more rows
```

5 Processing HOBO files

The `process_hobo()` function was developed to summarize all the csv files in a directory or folder. It evaluates the elements of the data from the `file_info` component and uses `raindance()` or `sundance()` to summarize the data. This function requires an object returned from `import_hobo()` or `import_hobo_2008()` and adds the summarized data to the original list, returning a four component list. An effective way to use this function is with `lapply()`.

```
lapply(file_list[2:4], function(this_file){
  dat <- import_hobo_2008(this_file) |> process_hobo()
  print(basename(this_file))
  print(data.class(dat)); print(names(dat))
  dat$data
})
#> [1] "2010_h01st_temp.csv"
#> [1] "list"
#> [1] "file_info" "details"    "data_raw"  "data"
#> [1] "2012_i01gp_prcp.csv"
#> [1] "list"
#> [1] "file_info" "details"    "data_raw"  "data"
#> [1] "2012_i01gp_temp.csv"
#> [1] "list"
#> [1] "file_info" "details"    "data_raw"  "data"
#> [[1]]
#> # A tibble: 338 x 10
#> # Groups:   PlotID, Date, Element [338]
#>   PlotID Date      Element Mean    Min    Max      n MinTime MaxTime Units
#>   <chr> <date>    <chr> <dbl> <dbl> <dbl> <int> <chr> <chr> <chr>
#> 1 H01 2009-05-06 TEMP      76.4  53.9  86.6    11 11:11 03:59 F
#> 2 H01 2009-05-07 TEMP      67.8  42.5  88.0    20 06:23 03:59 F
#> 3 H01 2009-05-08 TEMP      67.1  45.4  83.0    20 06:23 03:59 F
#> 4 H01 2009-05-09 TEMP      67.3  46.1  83.7    20 06:23 03:59 F
#> 5 H01 2009-05-10 TEMP      67.1  46.1  85.8    20 06:23 03:59 F
#> 6 H01 2009-05-11 TEMP      67.9  45.4  88.0    20 06:23 03:59 F
#> 7 H01 2009-05-12 TEMP      70.0  46.8  88.0    20 06:23 03:59 F
```

```

#> 8 H01 2009-05-13 TEMP 67.0 46.1 85.1 20 06:23 03:59 F
#> 9 H01 2009-05-14 TEMP 65.6 41.0 85.1 20 06:23 05:11 F
#> 10 H01 2009-05-15 TEMP 68.3 48.2 85.1 20 05:11 05:11 F
#> # ... with 328 more rows
#>
#> [[2]]
#> # A tibble: 8,988 x 6
#> # Groups: PlotID, DateTime [8,988]
#> PlotID DateTime Element PRCP_mm Tips MaxTips_min
#> <chr> <dtm> <chr> <dbl> <dbl> <dbl>
#> 1 I01 2011-04-18 01:00:00 PRCP 0 0 0
#> 2 I01 2011-04-18 02:00:00 PRCP 0 0 0
#> 3 I01 2011-04-18 03:00:00 PRCP 0 0 0
#> 4 I01 2011-04-18 04:00:00 PRCP 0 0 0
#> 5 I01 2011-04-18 05:00:00 PRCP 0 0 0
#> 6 I01 2011-04-18 06:00:00 PRCP 2.29 8 2
#> 7 I01 2011-04-18 07:00:00 PRCP 4.32 17 1
#> 8 I01 2011-04-18 08:00:00 PRCP 2.79 11 1
#> 9 I01 2011-04-18 09:00:00 PRCP 0.508 2 1
#> 10 I01 2011-04-18 10:00:00 PRCP 0 0 0
#> # ... with 8,978 more rows
#>
#> [[3]]
#> # A tibble: 398 x 10
#> # Groups: PlotID, Date, Element [398]
#> PlotID Date Element Mean Min Max n MinTime MaxTime Units
#> <chr> <date> <chr> <dbl> <dbl> <dbl> <int> <chr> <chr> <chr>
#> 1 I01 2011-04-06 TEMP 48.7 41.0 71.8 11 10:22 11:34 F
#> 2 I01 2011-04-07 TEMP 47.9 38.0 59.4 20 05:34 03:10 F
#> 3 I01 2011-04-08 TEMP 45.0 30.9 53.2 20 10:22 11:34 F
#> 4 I01 2011-04-09 TEMP 40.0 31.7 47.5 20 10:22 03:10 F
#> 5 I01 2011-04-10 TEMP 37.5 13.0 54.6 20 05:34 03:10 F
#> 6 I01 2011-04-11 TEMP 40.1 16.9 63.5 20 05:34 03:10 F
#> 7 I01 2011-04-12 TEMP 46.8 23.3 67.7 20 12:46 03:10 F
#> 8 I01 2011-04-13 TEMP 50.3 25.1 67.7 20 05:34 03:10 F
#> 9 I01 2011-04-14 TEMP 43.8 34.9 56.7 20 05:34 03:10 F
#> 10 I01 2011-04-15 TEMP 40.2 14.0 62.2 20 05:34 04:22 F
#> # ... with 388 more rows

```

6 Reporting Errors and Issues

The data from 2021 were mostly formatted the same as data from 2020, with a few exceptions. There is a likely that some future csv file will have anomalies that will require improvements to the *import_hobo()* function. Please submit any problems on the Issues page of this GitHub repository, <https://github.com/scoyoc/raindancer/issues>, or contact the author of the package if this happens.

7 Next Steps

This package was designed to work with the dataprocessR package. See the dataprocessR vignette (currently in development) for how to export these data to the SEUG LTVMP database.