

HOBO Paired Sensor Cleanup Walkthrough

Kara Dobson

July 31, 2020

This doc will describe the steps taken to clean the L0 HOBO data in the `HOBO_pairedsensor_cleanup_L0.R` script. The cleanup script is located in the warmXtrophic Github repository.

First, the existing data is cleared from the environment and the needed packages are loaded in.

```
# Clear all existing data
rm(list=ls())

#Load packages
for (package in c("tidyverse")) {
  if (!require(package, character.only=T, quietly=T)) {
    install.packages("package")
    library(package, character.only=T)
  }
}
```

Then, functions are sourced in and the working directory is set to the warmXtrophic shared drive.

```
# Source functions
source("~/warmXtrophic/kara/HOBO_scripts/HOBO_functions.R")

# Set working directory to Google Drive
# **** Update with the path to your Google drive on your computer
setwd("/Volumes/GoogleDrive/Shared drives/SpaCE_Lab_warmXtrophic/data/")
```

The data is then read in as csv files, separately for both U and H sensors. The U and H sensor data is then merged based on the date_time column. The 1516 dataframes do not have separate U and H files.

```
##### KBS Pair 1
#Read in H
KBS_1_1516 <- read.csv("L0/KBS/sensor_data/2015_2016/KBS_1.csv")
KBS_1H_2017 <- read.csv("L0/KBS/sensor_data/2017/09_01_2017/KBS_1H_09012017.csv")
KBS_1H_2018 <- read.csv("L0/KBS/sensor_data/2018/07_12_2018 (stations)/KBS_1H_07122018.csv", skip=1)
KBS_1H_2019 <- read.csv("L0/KBS/sensor_data/2019/10_07_2019/KBS_1H_10072019.csv", skip=1)
KBS_1H_2020 <- read.csv("L0/KBS/sensor_data/2020/04_05_2020/KBS_1H_04052020.csv", skip=1)

#Read in U
KBS_1U_2017 <- read.csv("L0/KBS/sensor_data/2017/09_01_2017/KBS_1U_09012017.csv")
KBS_1U_2018 <- read.csv("L0/KBS/sensor_data/2018/07_12_2018 (stations)/KBS_1U_07122018.csv", skip=1)[,1:6]
KBS_1U_2019 <- read.csv("L0/KBS/sensor_data/2019/10_07_2019/KBS_1U_10072019.csv", skip=1)[,1:6]
KBS_1U_2020 <- read.csv("L0/KBS/sensor_data/2020/04_05_2020/KBS_1U_04052020.csv", skip=1)[,1:6]

# Merge H and U data
```

```
KBS_1_2017 <- merge(KBS_1H_2017, KBS_1U_2017, by="Date_Time", all.x=T, all.y=T)
KBS_1_2018 <- merge(KBS_1H_2018, KBS_1U_2018, by="Date_Time..GMT.04.00", all.x=T, all.y=T)
KBS_1_2019 <- merge(KBS_1H_2019, KBS_1U_2019, by="Date_Time..GMT.04.00", all.x=T, all.y=T)
KBS_1_2020 <- merge(KBS_1H_2020, KBS_1U_2020, by="Date_Time..GMT.04.00", all.x=T, all.y=T)
```

The merged dataframes are then added to a list so functions can be applied over every dataframe. lapply is used to apply the necessary functions to the list. The functions can be found in the HOB0_functions.R script.

```
#Apply functions
list_pairk1 <- list(KBS_1_1516=KBS_1_1516, KBS_1_2017=KBS_1_2017, KBS_1_2018=KBS_1_2018, KBS_1_2019=KBS_1_2019)
list_pairk1 <- lapply(list_pairk1, change_pair_names)
list_pairk1 <- lapply(list_pairk1, change_POSIX)
list_pairk1 <- lapply(list_pairk1, remove_col, name=c('X', 'X..x', 'X..y'))
list_pairk1[2:4] <- lapply(list_pairk1[2:4], f_to_c)
list_pairk1 <- lapply(list_pairk1, remove_outliers)
```

These steps are then repeated for every paired sensor, at both KBS and UMBS. Occasionally, the column names don't match the names given to the function, so they are manually renamed as shown below.

```
#Manually rename columns with different names
names(list_pairk2$KBS_2_2017)[names(list_pairk2$KBS_2_2017)=="X2H_ambient_soil_moist_5cm"] <- "XH_ambient_soil_moist_5cm"
names(list_pairk2$KBS_2_2018)[names(list_pairk2$KBS_2_2018)=="Water.Content..m..m...LGR.S.N..10736967.."] <- "Water.Content..m..m...LGR.S.N..10736967.."]
names(list_pairk2$KBS_2_2019)[names(list_pairk2$KBS_2_2019)=="Water.Content..m..m...LGR.S.N..10736967.."] <- "Water.Content..m..m...LGR.S.N..10736967.."]
names(list_pairk2$KBS_2_2020)[names(list_pairk2$KBS_2_2020)=="Water.Content..m..m...LGR.S.N..10736967.."] <- "Water.Content..m..m...LGR.S.N..10736967.."]
```

The data import for the U sensors at UMBS is different, due to sensors being reset. Seen below, two separate files for 2018 are imported (2018a and 2018b) and then bound, due to the sensors being reset on the first given date. For 2020, two separate files are also imported (2020a and 2020b), but they could not be bound due to differing column names, so the names are manually changed in order to bind the files

```
UMBS_1U_2017 <- read.csv("L0/UMBS/sensor_data/2017/08_15_2017/UMBS_1U_08152017.csv")
UMBS_1U_2018a <- read.csv("L0/UMBS/sensor_data/2018/06_25_2018/UMBS_1U_06252018.csv", skip=1)
UMBS_1U_2018b <- read.csv("L0/UMBS/sensor_data/2018/09_19_2018/UMBS_1U_09192018.csv", skip=1)
UMBS_1U_2018 <- rbind(UMBS_1U_2018a, UMBS_1U_2018b)
UMBS_1U_2019 <- read.csv("L0/UMBS/sensor_data/2019/09_10_2019/UMBS_1U_09102019.csv", skip=1)[,1:6]
UMBS_1U_2020a <- read.csv("L0/UMBS/sensor_data/2020/05_13_2020/UMBS_1U_05132020.csv", skip=1)[,1:6]
UMBS_1U_2020b <- read.csv("L0/UMBS/sensor_data/2020/06_12_2020/UMBS_1U_06122020.csv", skip=1)[,1:6]
names(UMBS_1U_2020a)[names(UMBS_1U_2020a)=="Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"] <- "Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"
names(UMBS_1U_2020a)[names(UMBS_1U_2020a)=="Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"] <- "Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"
names(UMBS_1U_2020a)[names(UMBS_1U_2020a)=="Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"] <- "Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"
names(UMBS_1U_2020a)[names(UMBS_1U_2020a)=="Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"] <- "Temp...F..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"
names(UMBS_1U_2020b)[names(UMBS_1U_2020b)=="Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"] <- "Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"
names(UMBS_1U_2020b)[names(UMBS_1U_2020b)=="Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"] <- "Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_amb"
names(UMBS_1U_2020b)[names(UMBS_1U_2020b)=="Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"] <- "Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"
names(UMBS_1U_2020b)[names(UMBS_1U_2020b)=="Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"] <- "Temp...C..LGR.S.N..10737620..SEN.S.N..10737620..LBL..1U_warn"
UMBS_1U_2020 <- rbind(UMBS_1U_2020a, UMBS_1U_2020b)
```

For both KBS and UMBS, the final step is to create a .RData file with the list output for each station, which can be used in the HOB0_pairedsensor_merge_L1.R script to merge the data for each year and each station.

```
#Create .RData file
save(list_pairk1, list_pairk2, list_pairk3, file="L1/HOB0_data/HOB0_paired_sensor_data/KBS/KBS_pairedsensor.RData")
```