

DATA-SPECIFIC README

Before reading this document please read the BIFoR data readme that applies to all of the BIFoR data and check the accompanying contributors table which contains a data accessibility statement and other important details.

If this data is tabular check the accompanying variable definitions table for information on the variables.

data_ID: [weather](#)

Last updated: [2022-12-05](#)

DATA FORMAT

.csv (and .dat for the very raw data in data loggers)

R SCRIPTS (if applicable)

MH_combine_loggers.R

- imports and combines tables from data loggers at BIFoR FACE and exports them in an organised form (e.g. one file per year)

bifor_weather_additional.R

- downloads additional weather data from nearby weather stations at Gnosall from the weather underground network (<https://www.wunderground.com/wundermap>) and converts units

bifor_weather_additional_combine.R

- combines multiple datasets from wunderground from multiple weather stations
- performs some cleaning (thresholds, inner fences to discard outliers, see data cleaning below)
- summarises to every 15 min to get a consistent dataset

MH_weather.R

- imports and combines weather data from BIFoR FACE loggers and wunderground
- performs cleaning (thresholds, outer fences, faults log)
- consolidate highfreq dataset to create “whole site” parameters from multiple sensors at different locations (e.g. top/bottom of met towers), including filling gaps using the available sources
- calculate additional parameters (e.g. air vapour pressure deficit, VPD, see below)
- consolidate highfreq dataset to have 15min resolution and create summaries (30min, 1hr, 1day, 1month, 1year)
- exports consolidated and organised datasets (e.g. one file per year)

Weather

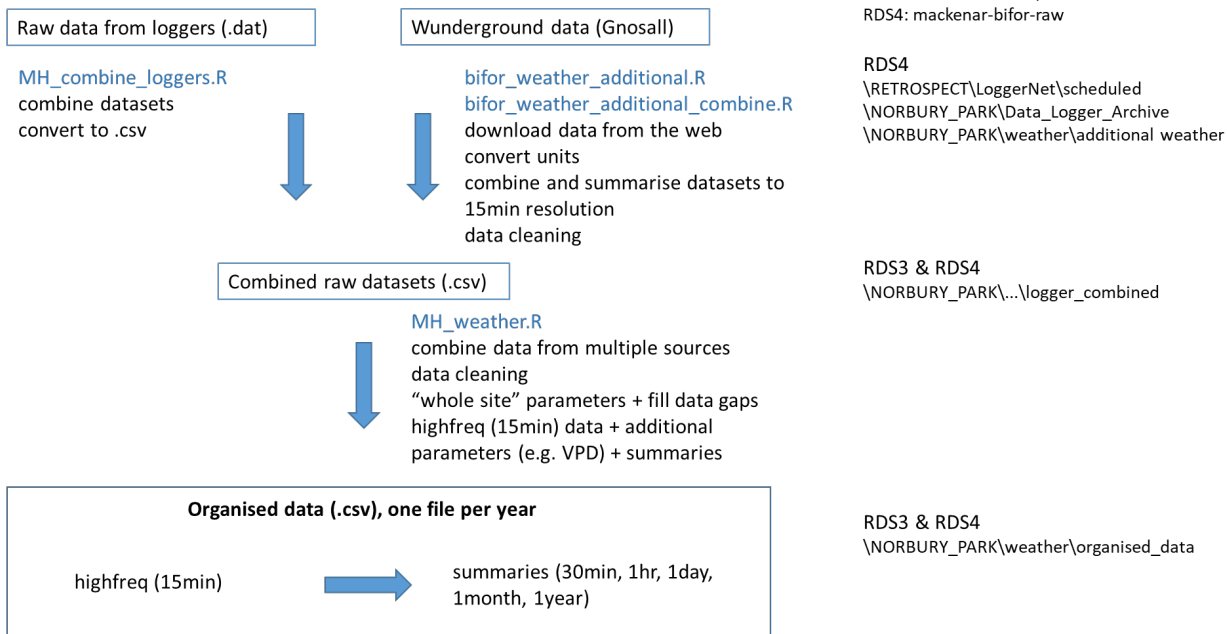


Figure 1: Data processing diagram.

RAW DATA

Raw data are from sensors located in four met towers surrounding (outside) the BIFoR FACE forest. They are originally saved from dataloggers in .dat format and later converted to .csv. See header below for more details. The raw data comes with variable sampling frequency, e.g. every 6 min, 10 min or 15 min. To make a consistent dataset comparable with other sources (e.g. weather underground network) the highfreq dataset has been summarised to every 15 min. The highfreq dataset (.csv) has also been through some quality control procedures (see data cleaning below).

HEADER raw data (.dat from data loggers)

The headers have changed over time, these are the possibilities (updated in April 2021):

```
header_met0 <- c("TIMESTAMP", "RECORD", "AirTC_A", "RH_A", "AirTC_A_Min",  
"AirTC_D", "RH_D", "AirTC_D_Min", "PAR_LI190_Top", "NetRAD_2", "NetRAD_2_Tot",  
"AmbP_Avg", "Rain_mm_1_Tot", "Rain_mm_2_Tot")  
header_met1 <- c("TIMESTAMP", "RECORD", "AirTC_A", "RH_A", "AirTC_A_Min",  
"AirTC_D", "RH_D", "AirTC_D_Min", "PAR_LI190_Top", "NetRAD_2", "NetRAD_2_Tot",  
"AmbP_Avg", "Rain_mm_1_Tot", "Rain_mm_2_Tot", "WindSpeed_25M_Avg", "WindDir_25M")  
header_met2 <- c("TIMESTAMP", "RECORD", "AirTC_A", "RH_A", "AirTC_A_Min",  
"AirTC_D", "RH_D", "AirTC_D_Min", "PAR_LI190_Top", "NetRAD_2", "NetRAD_2_Tot",  
"AmbP_Avg", "WindSpeed_25M_Avg", "WindDir_25M", "Rain_mm_1_Tot", "Rain_mm_2_Tot")
```

A is top of the tower (approx 25m), D is bottom near the ground level.
Rain columns, 1 is top, 2 is bottom.

Note: the organised datasets have different variable names, see below.

ORGANISED DATA

High frequency data consolidated to 15 min resolution were used to calculate summaries (every 30 min, 1 hr, 1 day, 1 month, 1 year). One file per year.

HEADER ORGANISED DATA

See var_def table.

Suffix:

_met: parameter measured from the BIFoR FACE met towers

_top: parameter measured on the top of the BIFoR FACE met towers (~25m)

_bot: parameter measured on the bottom of the BIFoR FACE met towers (~2m)

_gnosall: parameter derived as the mean of a number of sensors from the weather underground network (<https://www.wunderground.com/wundermap>) located at approximately 3-4 km from the BIFoR FACE site. These values are available for comparison and were used to fill gaps in the BIFoR FACE dataset.

_reservoir: data from rain gauge located just North of the BIFoR FACE forest next to a reservoir. Only used to fill data gaps if both the FACE and Gnosall data were missing (only before March 2019).

_min: minimum for the selected time period

_max: maximum for the selected time period

DATA CLEANING

The raw data (.dat format, resolution ≤ 15 min) have been cleaned of evident errors using physical thresholds and assigned NA (i.e. values physically impossible or safely considered unrealistic given the site conditions, see below). Additional clear outliers have been removed (i.e. assigned NA) after visual inspection. In addition, the 15min datasets from the Gnosall weather stations and from the BIFoR FACE met towers were combined after removing outliers falling outside inner or outer fences thresholds (i.e. $< q_{25} - A * \text{interquartile range}$ or $> q_{75} + A * \text{interquartile range}$, where $A = 1.5$ for inner fences and 3 for outer fences). Extensive checks on the data have not been carried out.

The parameters satisfying the following conditions were assigned NA [Note: | means OR]:

- $\text{air_temp} < -30$ | $\text{air_temp} > 50$
- $\text{RH} \leq 1$ | $\text{RH} > 100$
- $\text{amb_P} \leq 910$ | $\text{amb_P} \geq 1080$ (https://en.wikipedia.org/wiki/List_of_atmospheric_pressure_records_in_Europe)
- $\text{WS} < 0$ | $\text{WS} > 50$
- $\text{pcpn} < 0$ | $\text{pcpn} > 50$
- $\text{net_rad} < -100$

To minimise data gaps the following datasets have been used, in order of priority:

1. BIFoR FACE met tower dataset
2. Gnosall dataset (weather underground network: <https://www.wunderground.com/wundermap>)
3. Reservoir rain gauge dataset (for precipitation prior to March 2019. Note: the reservoir is immediately North of the BIFoR FACE forest, within Norbury Park Estate)

ADDITIONAL INFORMATION

The parameters SVP and VPD are calculated based on the consolidated air temperature and relative humidity (air_temp and RH, see definitions for more details) after cleaning.

Prior to the installation of the BIFoR FACE met towers (8 March 2019) the data from wunderground has been used as default. Although wunderground has some quality checks in place it is not possible to verify all the sensors used, their location and the maintenance procedures. The average of multiple weather stations (the number varied over the years) has been used to improve reliability and outliers have been assigned NA (see data cleaning). This data might be slightly less reliable, especially in 2016.

In addition to this dataset additional historical weather datasets with daily or hourly resolution are available from the CEDA archive (Centre for Environmental Data Analysis, <http://archive.ceda.ac.uk/>). Only individual applications to access the CEDA datasets are accepted therefore these are not available on the RDS. Nearby locations are Shawbury (RAF base) and Newport-salop. These datasets should be of high quality but they are usually 1 or 2 years behind and taken from greater distance from the site. However, they are very useful for historical analysis. [Note that in the hourly rainfall dataset from CEDA every 12 rows there is an extra row with the summary (sum) of the previous 12 hours. These extra rows most likely need removing]