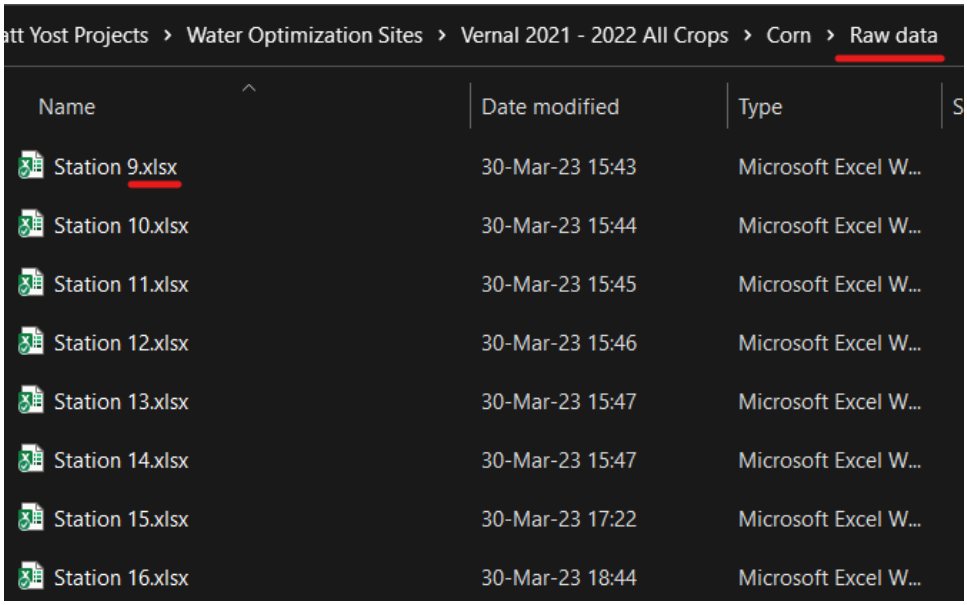
ET estimation protocol for SMET model

1. Prepare the working directory.

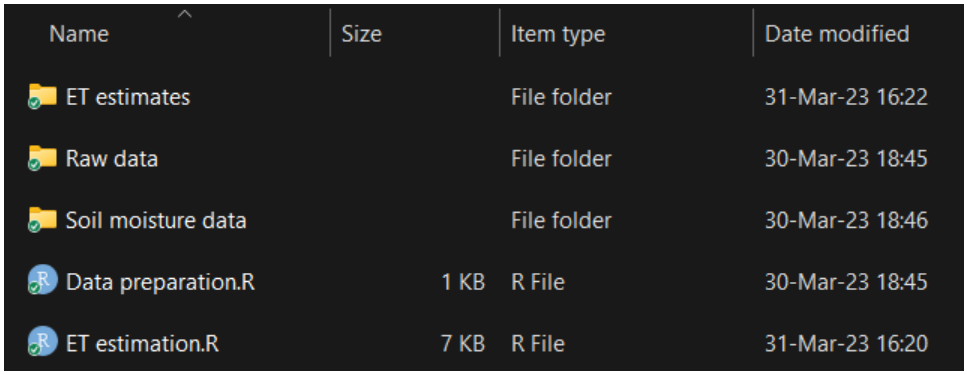
i. Create a folder for each study site, year, and crop.

ii. Create a "Raw data", a "Soil moisture data", and a "ET estimate" folder within the study site folder.

iii. Copy the raw soil moisture data into the "Raw data" folder. Name each data sheet "Station [number]" and save as an .xlsx file.



iv. Copy the "Data preparation.R" and "ET estimation.R" files into the folder (https://github.com/OliverHargreaves/SMET-demo)



2. Data preparation - The SMET model only needs one soil moisture measurement per day (preferably around midnight) and daily ETr values. The input data needs to have a date column, an ETr column and the volumetric water content.

i. Extract the midnight soil moisture value from the raw dataset. Can be done in excel or using the "Data preparation.R" code.

Notes on "Data preparation.R": If using the code make sure that the raw data spreadsheet only has one line of header, otherwise R studio will not interpret the values correctly. The station ID number

needs to be inputted manually on line 8. If the raw data has multiple years the code can extract the desired year but it has to be selected manually on line 20, otherwise skip or delete this part.

```
7  # Upload data ----
8  StationID=16 # Station ID number: needs to be inputted manually
9  data.path=file.path('Raw data', paste0('Station ', StationID, '.xlsx'))
10 data=read_xlsx(data.path)
```

```
18 # Select the year of interest if the data is from multiple years, otherwise delete/skip
19 data$Year=format(data$Date, "%y")       # Create a column with the year
20 data=data[data$Year=='21',]             # Keep only the 2021 values
```

ii. Check whether the data has gaps. Small gaps (1-5 days) can be interpolated if the soil moisture trend is monotonous (steadily decreasing). If there are large gaps in the data, the model cannot provide an estimate.

iii. Download the ETr data (https://climate.usu.edu/mchd/index.php) and check for gaps if there are any interpolate them. Copy the ETr values into the soil moisture data sheet, making sure the dates align.

iv. Save as an .xlxs file in the 'Soil moisture' folder. The .xlxs extension ensures that the code will recognize the date column properly.

I7

	A	B	C	D	E	F	G	H
1	Date	ETr	s1	s2	s3	s4	s5	s6
2	2021-05-13 23:45:00 UTC	7	30.6	29.4	30.0	27.1	27.3	28.7
3	2021-05-14 23:45:00 UTC	7	31.2	29.4	30.0	27.1	27.3	28.7
4	2021-05-15 23:45:00 UTC	8	35.0	29.5	30.1	27.3	27.3	28.7
5	2021-05-16 23:45:00 UTC	8	34.6	29.6	30.1	27.3	27.2	28.6
6	2021-05-17 23:45:00 UTC	7	33.8	29.7	30.2	27.3	27.2	28.6
7	2021-05-18 23:45:00 UTC	7	33.8	29.8	30.2	27.3	27.2	28.5

3. Running the model
   i. Open the "ET estimation.R" code.
   ii. There are 2 manual inputs into the code: the station ID number at line 8 and sensor depths (in inches) at line 30, 31, and 32.

```
7  # Data - Soil moisture data needs to be in a folder named "Soil moisture data" in the
   working directory of the code----
8  StationID=16 # Station ID number: needs to be inputted manually
9  data.path=file.path('Soil moisture data', paste0('Station ', StationID, '.xlsx'))
10 data=read_xlsx(data.path)
```

```
29 # Soil moisture sensor depths (inches) - Other than the station ID this is the only
   manual input into the code. If more than 3 sensors are used it needs to be modified.
30 d1=6   # 1st sensor is at 6"
31 d2=18  # 2nd sensor is at 18"
32 d3=30  # 3rd sensor is at 30"
```

While running the code R studio will create plots that facilitate QAQC and make sure the results make sense. The comments in the code should help with plot interpretation.

Note: Be sure to check that the volumetric soil water content is expressed in percentage (1-100) and not in fraction (0-1). If the soil moisture is in fraction either transform it into percentage in excel before running the code OR modify the code by deleting the 100 division in line 43 and 50.

```
41 for (i in 1:n) {data$SM.N[i]=c(z1*data$s1[i]+
42                                 z2*data$s2[i]+
43                                 z3*data$s3[i])/100}
48 for (i in 1:n) {data$SM.S[i]=c(z1*data$s4[i]+
49                                 z2*data$s5[i]+
50                                 z3*data$s6[i])/100}
```

4. Interpreting the results

The code is set up to save the results in the "ET estimates" folder as an .xlxs file with all the soil moisture data, the intermediate steps, and the cumulative ET. Since each station has two substations the output file will have two cumulative ET estimates, for the 1:3 (N) and 4:6 (S) ports respectively. These values have an accuracy of ±7% according to the calibrations performed with EC towers in Vernal and Modena.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Date	s1	s2	s3	s4	s5	s6	ETr	SM.N	SM.S	delta.N	delta.S	ET.N	ET.S	ET_cum.N	ET_cum.S
2	2021-05-13 23:45:00 UTC	30.6	29.4	30	27.1	27.3	28.7	7	228.6	211.8			0.0	0.0	0.0	0.0
3	2021-05-14 23:45:00 UTC	31.2	29.4	30	27.1	27.3	28.7	7	230.0	211.8	1.4	0.0	6.0	6.0	6.0	6.0
4	2021-05-15 23:45:00 UTC	35	29.5	30.1	27.3	27.3	28.7	8	239.2	212.3	9.2	0.5	6.8	6.8	12.8	12.8
5	2021-05-16 23:45:00 UTC	34.6	29.6	30.1	27.3	27.2	28.6	8	238.5	211.8	-0.7	-0.5	3.7	3.7	16.6	16.5
6	2021-05-17 23:45:00 UTC	33.8	29.7	30.2	27.3	27.2	28.6	7	237.2	211.8	-1.3	0.0	3.5	6.0	20.1	22.5
7	2021-05-18 23:45:00 UTC	33.8	29.8	30.2	27.3	27.2	28.5	7	237.4	211.5	0.2	-0.3	6.0	3.1	26.1	25.6
8	2021-05-19 23:45:00 UTC	33.5	29.8	30.3	27.3	27.1	28.5	7	237.1	211.2	-0.4	-0.2	3.2	3.1	29.2	28.7
9	2021-05-20 23:45:00 UTC	34.4	33.3	30.4	28.4	27.2	28.5	11	247.4	214.0	10.4	2.7	9.4	9.4	38.7	38.1
10	2021-05-21 23:45:00 UTC	34.1	32.7	30.4	28	27.2	28.7	5	245.4	213.7	-2.1	-0.3	3.0	2.3	41.7	40.4
11	2021-05-22 23:45:00 UTC	33.9	32.3	30.5	27.8	27.2	28.7	4	244.3	213.2	-1.1	-0.5	2.2	1.9	43.8	42.3
12	2021-05-23 23:45:00 UTC	33.7	32.1	30.4	27.8	27.2	28.6	7	243.1	212.9	-1.2	-0.3	3.5	3.1	47.4	45.4

Alfalfa: the model has proven to be resilient to changes in biomass (harvests) in alfalfa, but the cumulative ET value needs to be calculated for each cut individually, this can either be done by running the model for each cut separately and obtaining an output sheet for each one or by running the model for the entire season and calculating the cumulative estimate from the daily ET values.