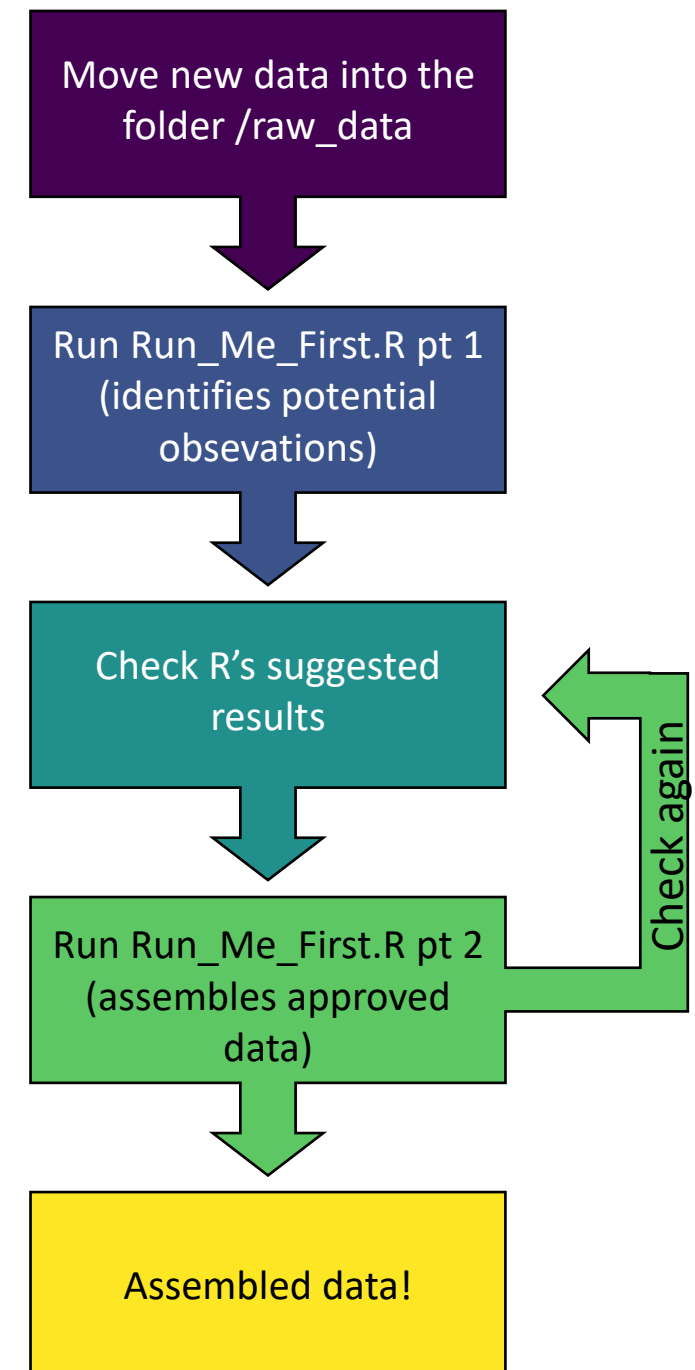# In this ReadMe:

- Overview
  - Concept
  - Terminology

- Getting started (QAQC)
  - File organization
  - Running the first part of Run_Me_First.R

- Script outputs part 1: QAQC files
  - File output locations
  - File descriptions and examples
  - How do the filters work?
  - Validating observations

- Data Assembly
  - Running the rest of Run_Me_First.R

- Script outputs part 2: assembled files
  - File output locations
  - File descriptions and examples

Overview

- This ReadMe describes a series of R scripts and files which are designed to:

  - Provide automated assignment of potential observations

  - Provide tools to assess the quality of potential observations

  - Run one to a large number of input files at the same time

  - Conditionally overwrite (with user permission) previous runs while protecting any edits you have made to previous outputs: in plain English, protects you from overwriting your work

  - Do all of this with a helpful amount of documentation and paper trail

Concept

- The workflow is designed to be as automated as possible

- R will identify potential observations, but you will need to verify them

- R uses a set of filters which look at:
  - The total number of tag pings
  - The rate of noise pings
  - The tag:noise ping ratio
  - The difference (in dB) in tag – noise signal strength

- The outputs include lots of files to help you verify the results and to keep a record of what you have done

- The scripts are designed to be re-run, and allow you to only look at new files OR reanalyze old files

Move new data into the folder /raw_data

Run Run_Me_First.R pt 1 (identifies potential obsevations)

Check R's suggested results

Run Run_Me_First.R pt 2 (assembles approved data)

Check again

Assembled data!

Terminology in this manual

- Code: The ping interval recorded by the receiver (so tag 2100 has code 100)

- Tag: A recognized tag code known to be at large in the wild

- Pseudotag: A code which superficially looks like a real tag number, but which is not known to be at large in the wild

- Ping: A single instance of data being recorded by the receiver. It may be noise, timer tag, a pseudotag, or a tag.

- Incident: One or more successive pings at the same code occurring with less than a 5 minute (300 second) gap

- Observation: An incident which is verified (by R **and a human**) to probably be a real fish

All of this terminology is negotiable! I just needed terms to use here.

Getting started: file organization

The R scripts are written to expect a particular file structure. We can rewrite it, but right now it's anticipating a structure like this:
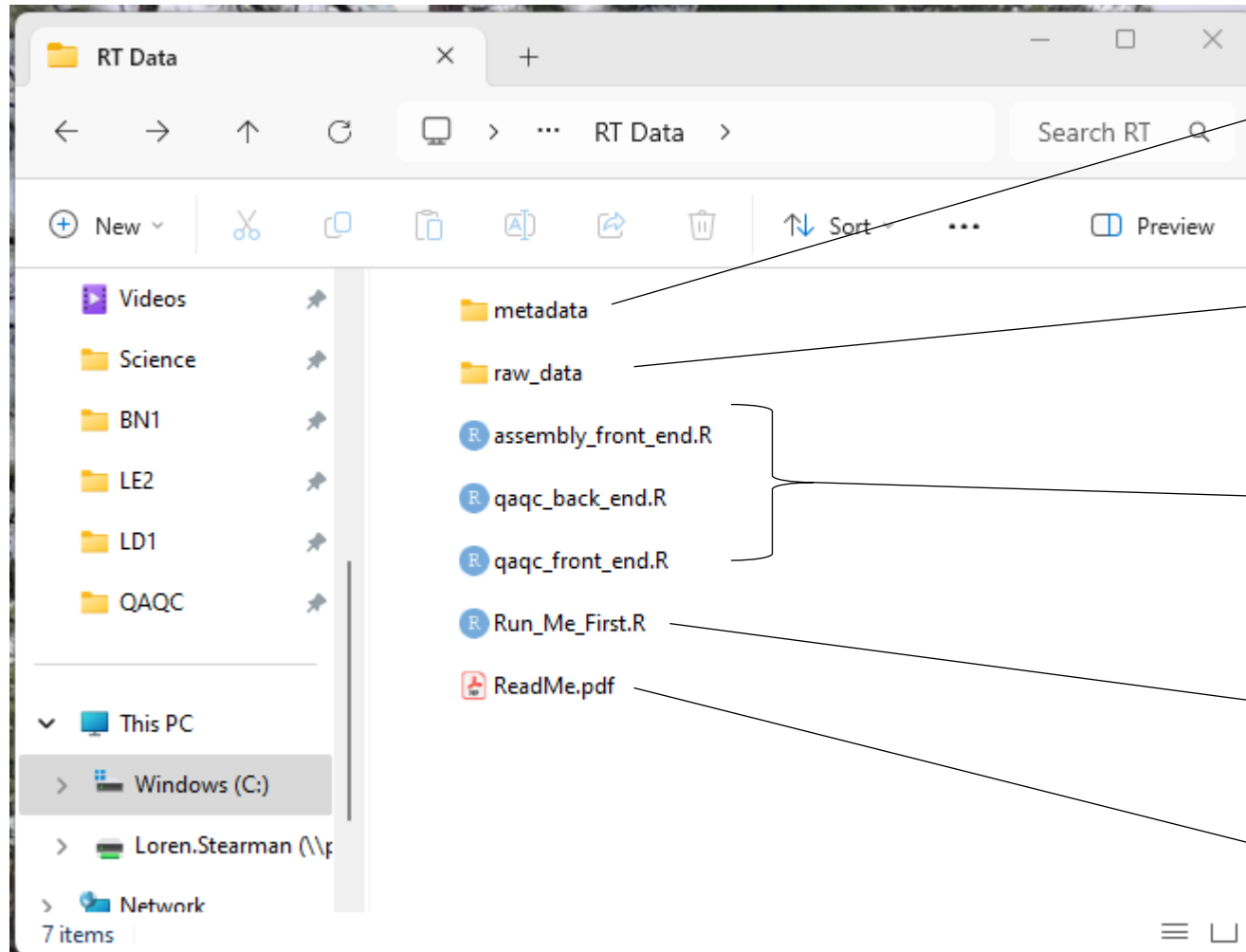
This folder contains your tag ID file and your filter values file

This folder contains the raw data in .hex, .txt, and .csv format. Only .txt files are analyzed here.

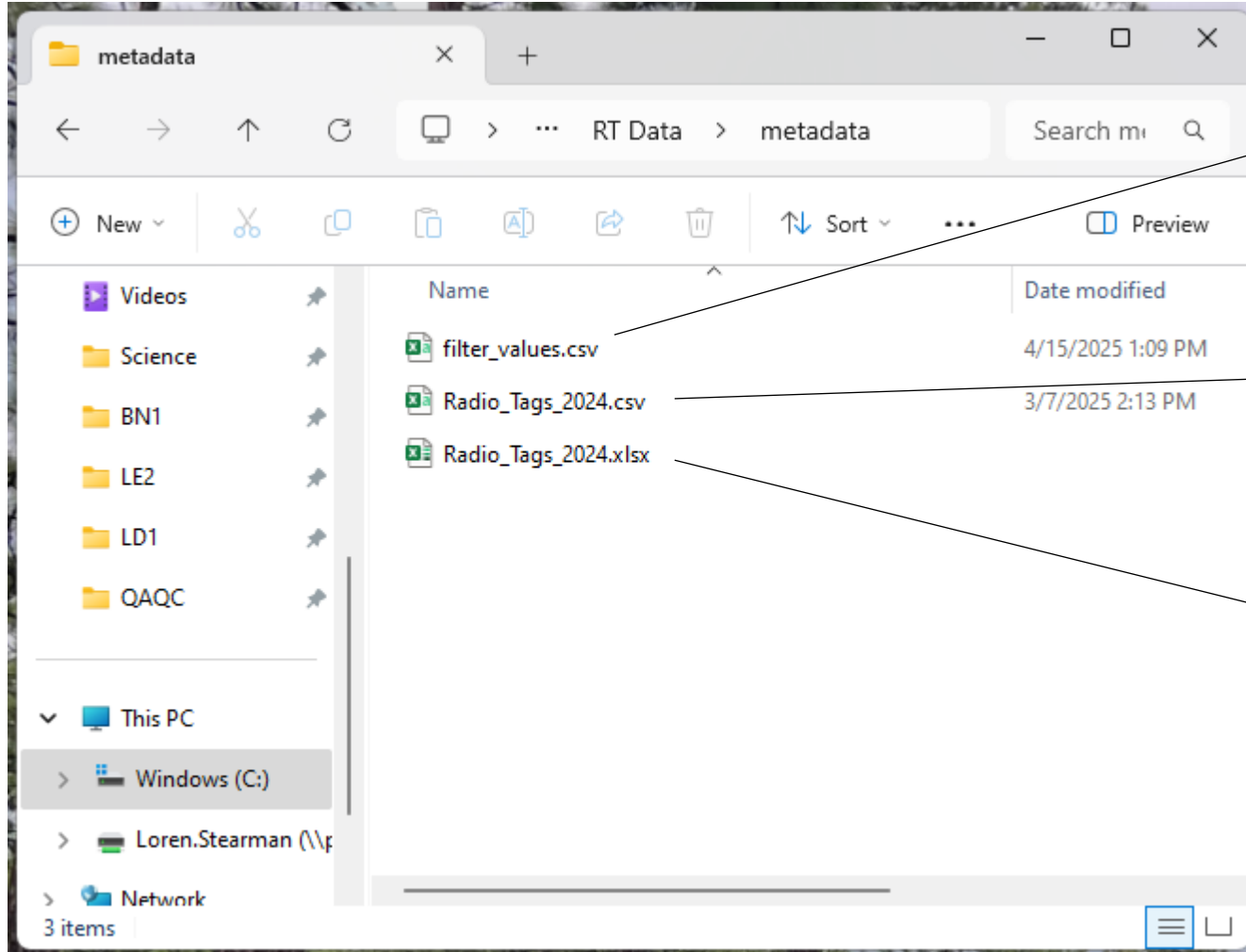These R files make the magic happen. They are called automatically.

This R file is the script that you'll run.

This is the tutorial that you are reading right now

Getting started: file organization

This folder contains the radio tag ID file and the filter values file.

This file contains the filter values used by R to suggest which observations are legit and which are probably not.

This file contains the tag ID information, as well as the metadata for each of the tagged fish.

This file is just the Excel version of the tag ID csv file. It isn't necessary.

Structure of the file "filter_values.csv"

- Filter (character): the type of filter
- Value (numeric): A value for the filter
- Description (character): A text description of the filter
- Automatic Fail (Logical): Does filtering automatically fail the incident if this specific filter criteria is not met?

	A	B	C	D
1	Filter	Value	Description	Automatic Fail?
2	Minimum detections	5	The minimum number of detections in an incident	Yes
3	Noise rate	0.1	The maximum rate of noise pings (per second)	No
4	Tag:noise ping count ratio	0.5	The minimum ratio of tag and noise ping counts	No
5	Tag - noise strength difference	3	The minimum difference in mean tag strength - mean noise strength (dB)	No

Structure of the file "Radio_Tags_2024.csv"

- Tagging Date (date): The date of tagging
- Tag Turn On Date (date): The estimated date of tag turning on
- Fish ID (alphanumeric): A unique ID for each fish
- Total Length (mm) (numeric): The total length in mm
- Weight (g) (numeric): The mass in grams
- Release Location (character): The release location

- PIT Tag Code (alphanumeric): The unique PIT tag code for the fish
- Radio Tag Number (numeric): The channel and code of the radio tag
- Radio Tag Start Time (numeric): The radio tag's start time
- No. of Detections (numeric): The number of detections
- Fish Movement Summary (character): comments on movement of the individual

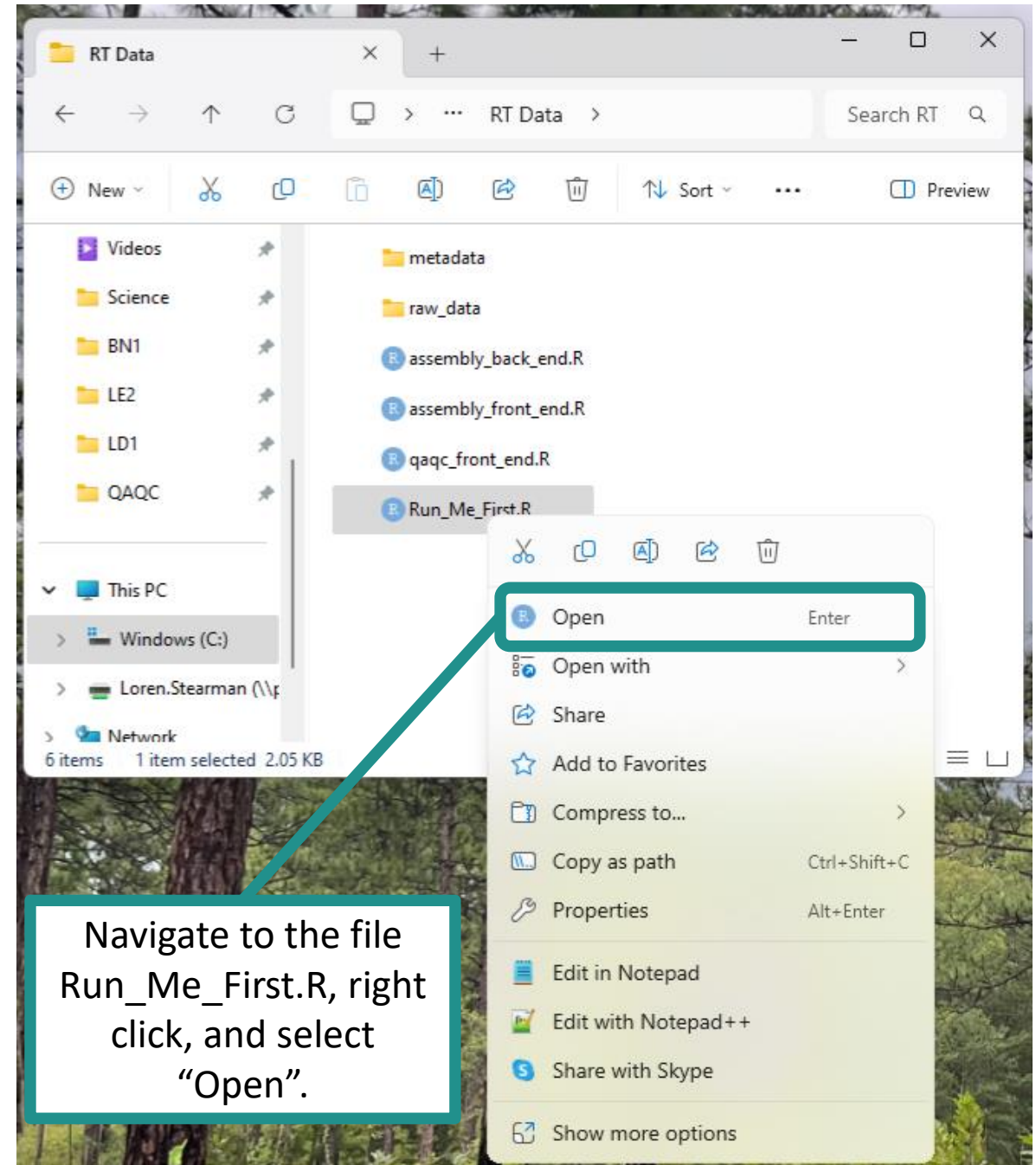|	A	B	C	D	E	F	G	H	I	J	
1	Tagging Date	Tag Turn On Date	Fish ID	Total Length (mm)	Weight (g)	Release Location	PIT Tag Code	Radio Tag Number	Radio Tag Start Time	No. of Detections	Fish Movement Summary
2	9/10/2024	1/8/2025	LGR-Wall-24-0128	465	929	DOWN	3DD.003BD68E86	2060	1110	0	
3	9/16/2024	1/14/2025	LGR-Wall-24-0132	437	787	DOWN	3DD.003BD68E8C	2160	1209	0	
4	9/16/2024	1/14/2025	LGR-Wall-24-0133	468	902	DOWN	3DD.003BD68ED2	2540	1222	29	Detected in ladder entranc
5	9/19/2024	1/17/2025	LGR-Wall-24-0137	549	1318	DOWN	3DD.003BD68EA1	2100	1253	0	
6	9/19/2024	1/17/2025	LGR-Wall-24-0134	533	1422	DOWN	3DD.003BD68EB4	2120	1224	0	
7	9/19/2024	1/17/2025	LGR-Wall-24-0135	464	791	DOWN	3DD.003BD68E9D	2140	1235	0	
8	9/19/2024	1/17/2025	LGR-Wall-24-0138	442	714	DOWN	3DD.003BD68E95	2440	1415	0	
9	9/19/2024	1/17/2025	LGR-Wall-24-0136	465	948	DOWN	3DD.003BD68EAD	2480	1245	0	
10	9/22/2024	1/20/2025	LGR-Wall-24-0141	433	674	DOWN	3DD.003BD68E8E	2460	1223	0	
11	9/22/2024	1/20/2025	LGR-Wall-24-0140	431	752	DOWN	3DD.003BD68E97	2520	1201	0	
12	9/24/2024	1/22/2025	RECAP LGR-Wall-24-0108	443	720	DOWN	3DD.003BD68EE5	2400	1206	-	RECAP 2024

Getting started:
file organization
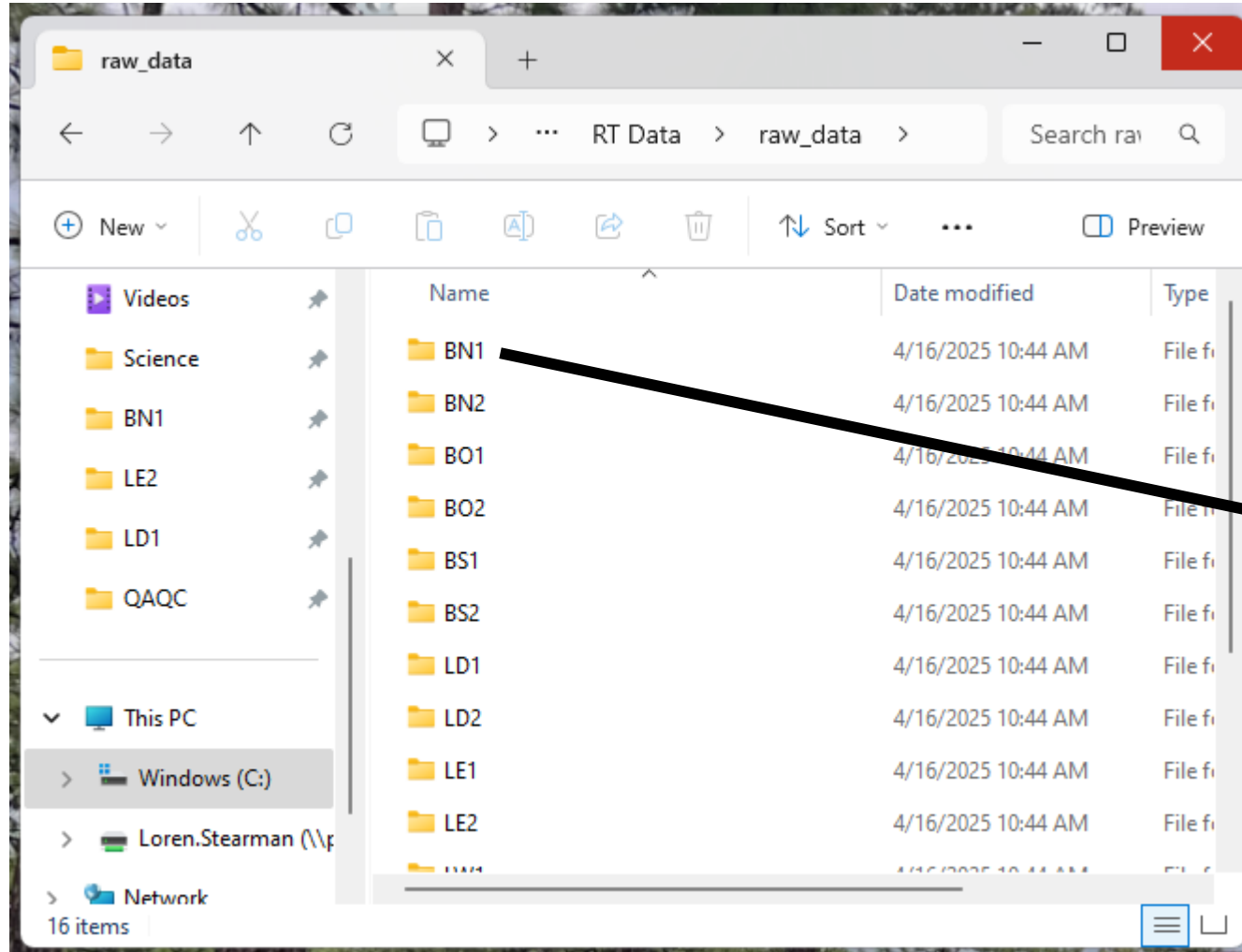
This folder contains the raw data from the receivers. How you organize the data inside of this folder is *entirely up to you*. R runs a recursive search for .txt files within this folder and can handle any structure. Here the data are organized by receiver. The only key is that all raw data should be in the folder /raw_data.
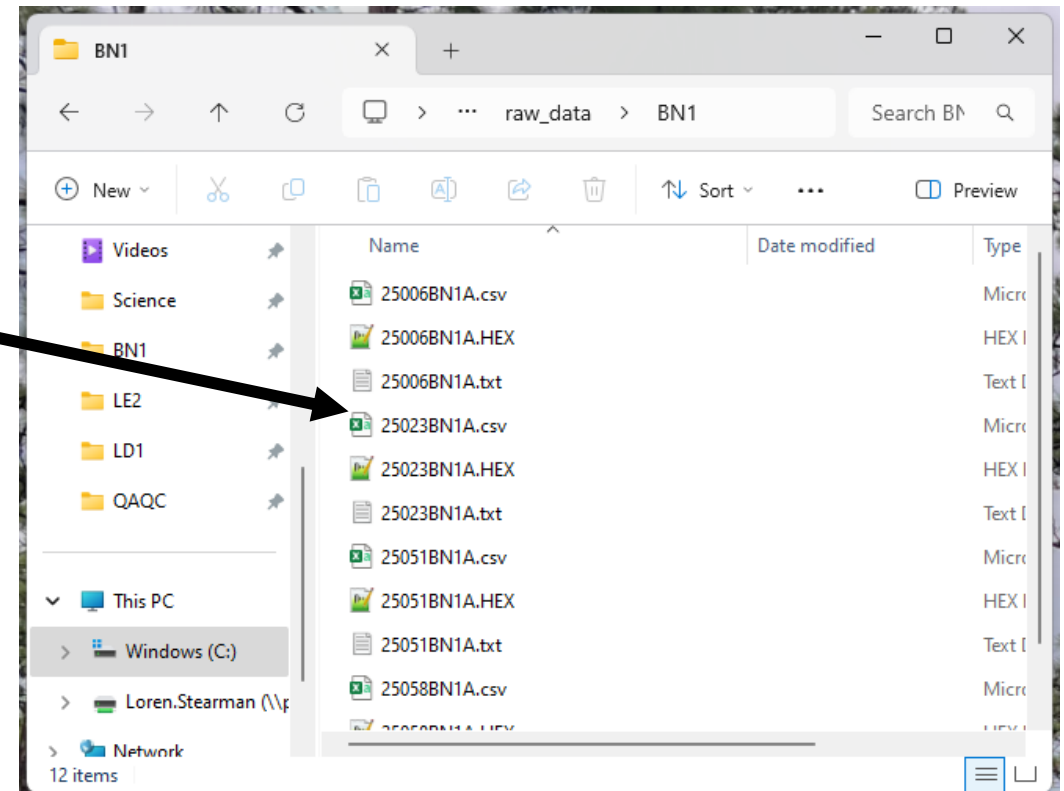
Getting started: Run_Me_First.R

- Navigate to the folder with Run_Me_First.R

- Right click and select "Open"
  - If your registry opens ".R" files with another program (like NotePad+), click "Open With" and select Rstudio

- Opening the file this way sets R's working directory to the correct directory

Navigate to the file Run_Me_First.R, right click, and select "Open".

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function   Addins

Run_Me_First.R

Source on Save

```
13
14  # http://www.apache.org/licenses/LICENSE-2.0
15
16  # Unless required by applicable law or agreed to in writing, software is
17  # distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND,
18  # either express or implied. See the License for the specific language governing
19  # permissions and limitations under the License.
20
21
22
23  #=================================================================#
24  #    START HERE                                                   #
25  #=================================================================#
26
27  # Run this set of lines the first time you run this script, and any time that
28  # you update R. This set of lines checks for and installs necessary packages.
29  # It doesn't hurt to re-run it, but you don't need to once they are installed.
30  # THE CODE WILL NOT RUN WITHOUT THESE PACKAGES
31      all_packages <- installed.packages()
32      if (!"khroma" %in% all_packages) {
33          install.packages("khroma")
34      }
35      if (!"viridis" %in% all_packages) {
36          install.packages("viridis")
37      }
38
39
40  # Run the line below to conduct data management and QAQC.
41  # Highlight the line below and select "Run" from above ----------------------^
42      source("qaqc_front_end.R")
43
```

38:1   (Top Level)   R Script

Console   Terminal   Background Jobs

R 4.4.1 · ~/

Loading required package: permute
Loading required package: lattice
This is vegan 2.6-8
Loading required package: MASS
Loading required package: grDevices
Loading required package: graphics
Loading required package: stats
Loading required package: utils

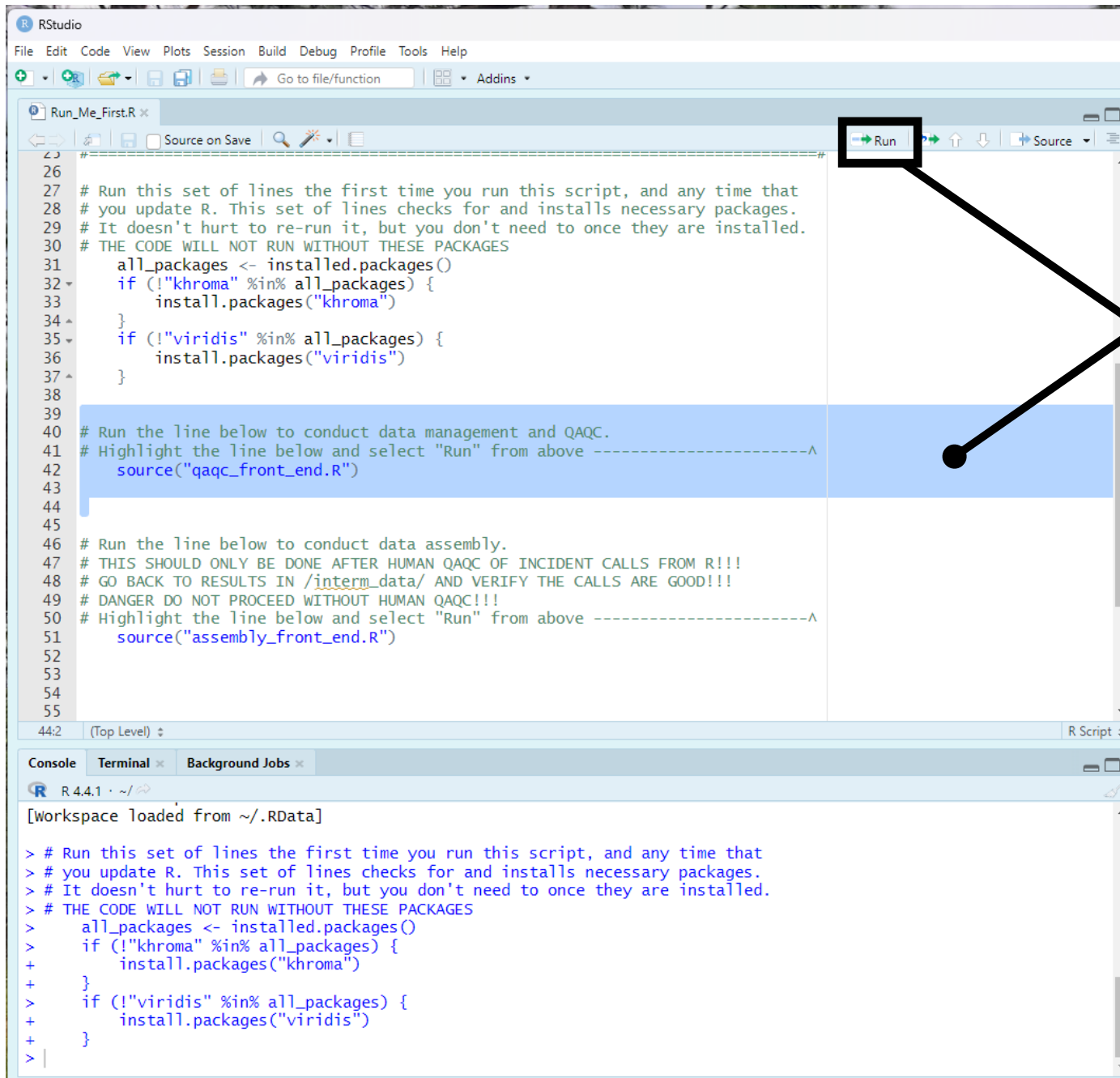Welcome at Wed Apr 16 13:21:00 2025
[Workspace loaded from ~/.RData]

>

This first code chunk installs two key packages. You only need to run it the first time, and then every time you update R. It doesn't hurt to re-run it if you forget…

To run the chunk, highlight it with your mouse, and then select -> Run above.

Note that any line that starts with # is instructions to you. Take the time to read them, they might answer questions that you have.

New Folder   New Blank File   Delete   Rename   More

Home

Name   Size   Modified
.RData   89 KB   Feb 19, 2025, 1:38 PM
.Rhistory   18.2 KB   Apr 11, 2025, 5:34 PM
.Rprofile   290 B   Oct 17, 2024, 8:09 AM
Admin
ArcGIS
Check-Ins
Custom Office Templates
desktop.ini   402 B   Jan 17, 2025, 8:28 AM
Hypothesis.png   66.4 KB   Mar 31, 2025, 2:16 PM
Idaho Meeting
My Music
My Pictures
My Videos
NMDS Outputs.RData   1.4 MB   Dec 4, 2024, 12:34 PM

Next, run this code chunk. This starts the data QAQC process. The process will have popup windows to help guide you through how you want to analyze the data.
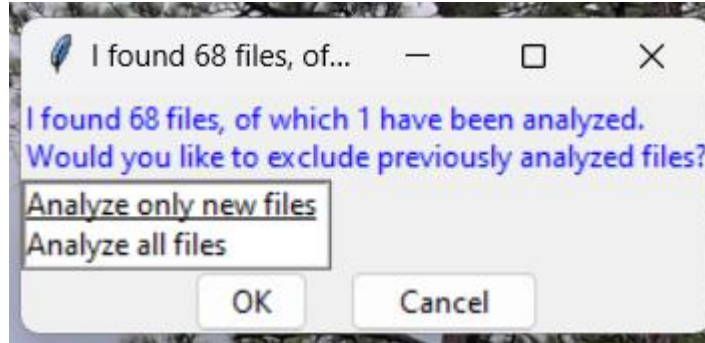
You can analyze all data, a subset, or just one file. If you have analyzed any data before, the script will ask you if you want to reanalyze those files. That can be helpful if you decide to change filter settings.

This could take 1-2 hours to run the first time. Regularly running this every time you bring in new files, and analyzing only the new files, will go much quicker.

```
# Run this set of lines the first time you run this script, and any time that
# you update R. This set of lines checks for and installs necessary packages.
# It doesn't hurt to re-run it, but you don't need to once they are installed.
# THE CODE WILL NOT RUN WITHOUT THESE PACKAGES
    all_packages <- installed.packages()
    if (!"khroma" %in% all_packages) {
        install.packages("khroma")
    }
    if (!"viridis" %in% all_packages) {
        install.packages("viridis")
    }

# Run the line below to conduct data management and QAQC.
# Highlight the line below and select "Run" from above ---------------------^
    source("qaqc_front_end.R")


# Run the line below to conduct data assembly.
# THIS SHOULD ONLY BE DONE AFTER HUMAN QAQC OF INCIDENT CALLS FROM R!!!
# GO BACK TO RESULTS IN /interm_data/ AND VERIFY THE CALLS ARE GOOD!!!
# DANGER DO NOT PROCEED WITHOUT HUMAN QAQC!!!
# Highlight the line below and select "Run" from above ---------------------^
    source("assembly_front_end.R")
```
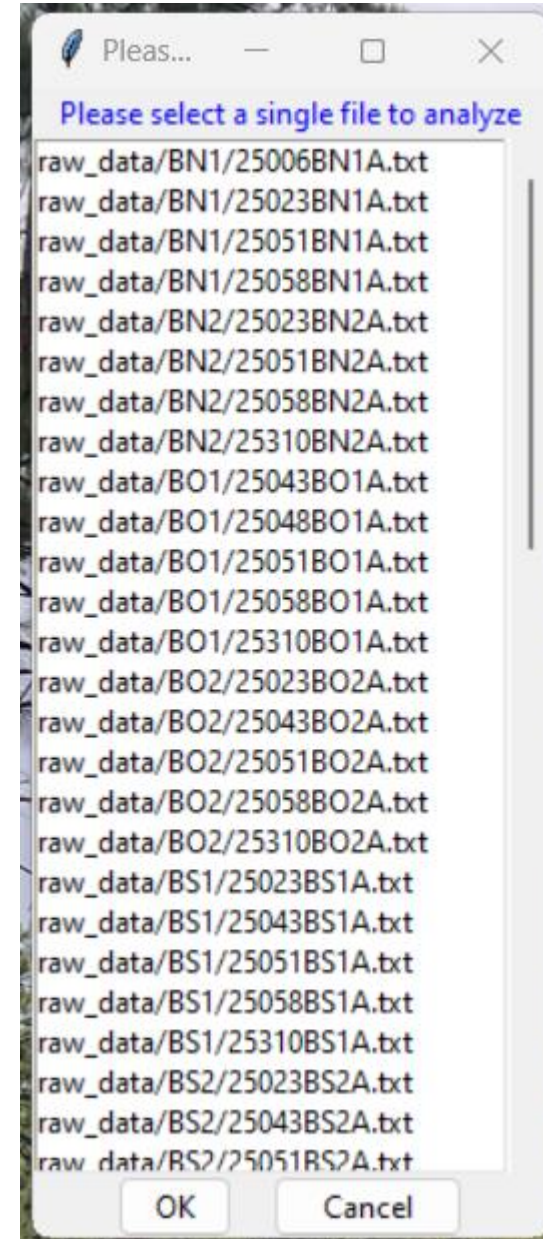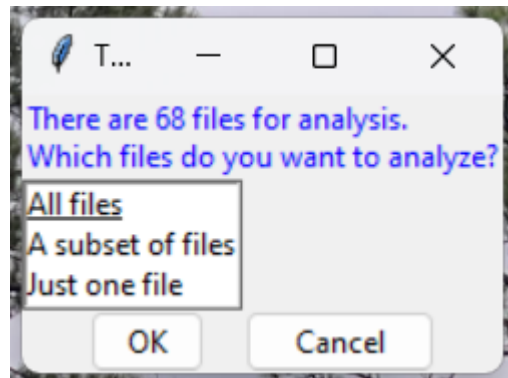
User file selection dialogues

These open when you run the script "qaqc_front_end.R"

1. If R detects that you have analyzed files before, it will open this dialogue asking if you want to analyze only new files, or to reanalyze all files. If it doesn't detect previously analyzed files, this dialogue does not appear.

I found 68 files, of... — ☐ ✕

I found 68 files, of which 1 have been analyzed.
Would you like to exclude previously analyzed files?

Analyze only new files
Analyze all files

OK   Cancel

2. R will prompt you for which files you would like to analyze. You can select all, a subset (more than one but less than all), or just one file. The "Just one file" option includes "A different file" to load a file that's not currently in your raw data folder.

T... — ☐ ✕

There are 68 files for analysis.
Which files do you want to analyze?

All files
A subset of files
Just one file

OK   Cancel

Pleas... — ☐ ✕

Please select a single file to analyze

raw_data/BN1/25006BN1A.txt
raw_data/BN1/25023BN1A.txt
raw_data/BN1/25051BN1A.txt
raw_data/BN1/25058BN1A.txt
raw_data/BN2/25023BN2A.txt
raw_data/BN2/25051BN2A.txt
raw_data/BN2/25058BN2A.txt
raw_data/BN2/25310BN2A.txt
raw_data/BO1/25043BO1A.txt
raw_data/BO1/25048BO1A.txt
raw_data/BO1/25051BO1A.txt
raw_data/BO1/25058BO1A.txt
raw_data/BO1/25310BO1A.txt
raw_data/BO2/25023BO2A.txt
raw_data/BO2/25043BO2A.txt
raw_data/BO2/25051BO2A.txt
raw_data/BO2/25058BO2A.txt
raw_data/BO2/25310BO2A.txt
raw_data/BS1/25023BS1A.txt
raw_data/BS1/25043BS1A.txt
raw_data/BS1/25051BS1A.txt
raw_data/BS1/25058BS1A.txt
raw_data/BS1/25310BS1A.txt
raw_data/BS2/25023BS2A.txt
raw_data/BS2/25043BS2A.txt
raw_data/BS2/25051BS2A.txt

OK   Cancel

3. If you select either a subset of files or just one file, R will open a list of files it has detected. You can select one (just one file) or multiple (a subset of files) to analyze.

If you select all files, R will skip this dialogue box.

This is a screenshot of the RStudio application. The code editor (Run_Me_First.R) contains:

```r
26
27  # Run this set of lines the first time you run this script, and any time that
28  # you update R. This set of lines checks for and installs necessary packages.
29  # It doesn't hurt to re-run it, but you don't need to once they are installed.
30  # THE CODE WILL NOT RUN WITHOUT THESE PACKAGES
31      all_packages <- installed.packages()
32      if (!"khroma" %in% all_packages) {
33          install.packages("khroma")
34      }
35      if (!"viridis" %in% all_packages) {
36          install.packages("viridis")
37      }
38
39
40  # Run the line below to conduct data management and QAQC.
41  # Highlight the line below and select "Run" from above ---------------------^
42      source("qaqc_front_end.R")
43
44
45
46  # Run the line below to conduct data assembly.
47  # THIS SHOULD ONLY BE DONE AFTER HUMAN QAQC OF INCIDENT CALLS FROM R!!!
48  # GO BACK TO RESULTS IN /interm_data/ AND VERIFY THE CALLS ARE GOOD!!!
49  # DANGER DO NOT PROCEED WITHOUT HUMAN QAQC!!!
50  # Highlight the line below and select "Run" from above ---------------------^
51      source("assembly_front_end.R")
52
53
54
55
56
```

Console output:

```
Loading required package: stats
Loading required package: utils

Welcome at Wed Apr 16 14:03:14 2025
> # Run the line below to conduct data management and QAQC.
> # Highlight the line below and select "Run" from above --------------------^
>     source("qaqc_front_end.R")
Loading required package: viridisLite

Ping and incident detection in progress!
If this is the first run, or the first run in a while,
the process could take an hour or two.
```

Callout text box:

Once the code starts, it will print out text periodically to update you on how it's going. It starts by letting you know that if this is your first run of files, or if you've let files accumulate for a while, that it could take a bit to crunch everything.

Right now the code is identifying series of pings which are separated by >300 seconds, applying filters, constructing the output csv files, and building QAQC figures.

As the script completes a file, it will output some helpful information.

This includes the file number in the series, the file name, and where various outputs have been saved. It will also print a warning if it detects nonsensical dates or times.

Big files take more time. Sometimes a lot more time. The processing cost is exponential, so regular downloads will make this easier on you.

Script outputs part 1 QAQC tools

Run_Me_First.R creates several different files for QAQC and assessment of potential observations. If this is the first time you have run the script, you'll notice a new folder pop up (/interm_data), where some QAQC tools are stored.
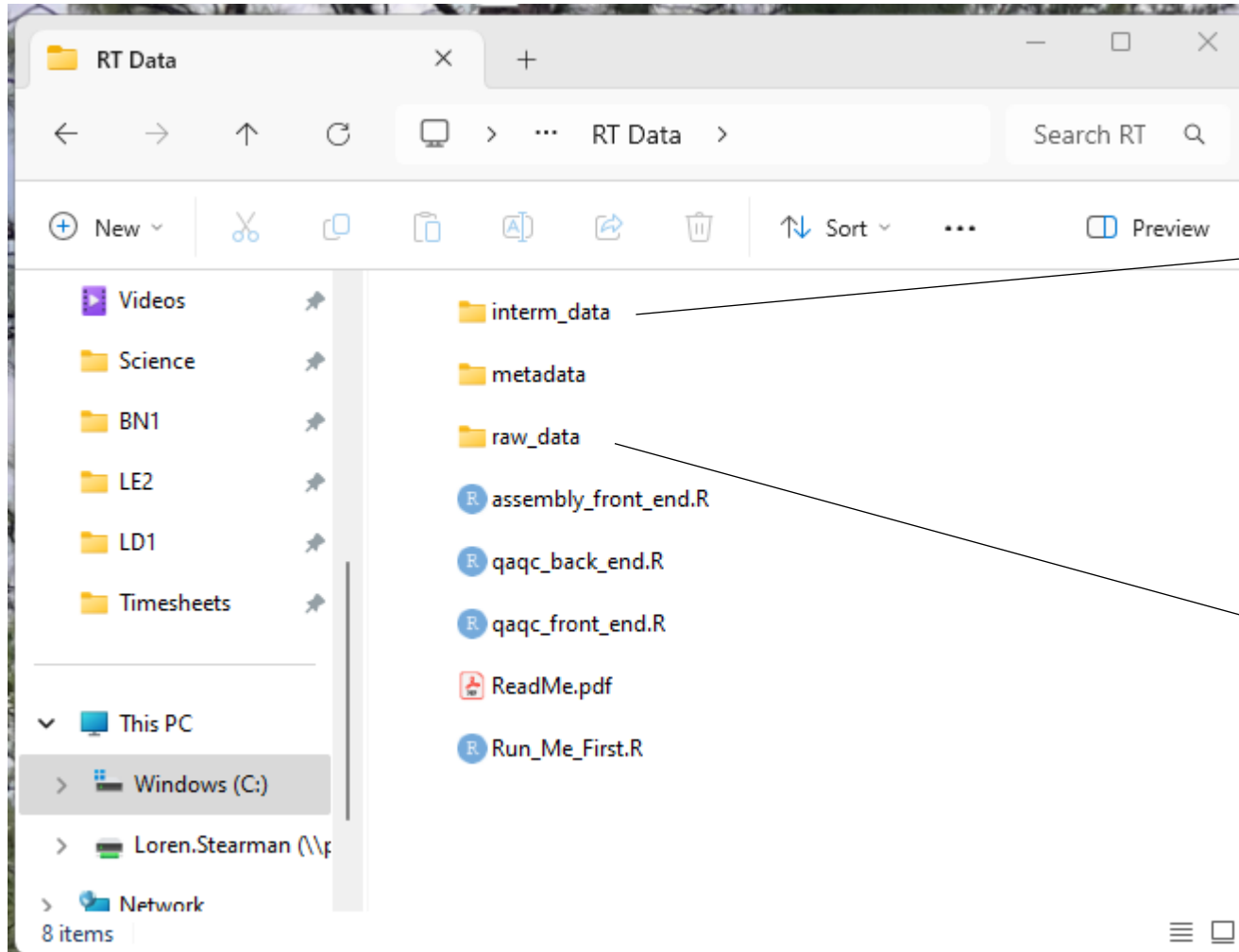
For each file:

This folder contains:
- A csv file of incidents, along with important data used in the filtering process
- A plot of all incidents, with noise and timer tag overlayed, and with suggested observations highlighted
- Individual plots of incidences with > a minimum number of pings

This folder now also includes:
- The same csv of incidents (backup copy)
- A processing receipt which confirms that a file was analyzed, plus some key metadata from the analysis
- An analysis history file which documents every time the file has been analyzed or reanalyzed, along with the filter settings used at that time (in case you want to replicate something again)

Why two copies of the incidents csv files?

- You will be making changes to the incidents csv files in /interm_data to let R know whether you agree or disagree with each suggested observation

- Because you'll probably be running this script again and again, and because accidents happen, I wanted to write something that protected changes you have already made. It does this by requiring you to *manually* copy a new version of the incidents csv files over to /interm_data to start over from scratch. R will write the file to /interm_data the first time, but not after that.

- So say you've gone through 500 incidents and found you disagree with R on 25 of them. You change their status in the incidents.csv files in /interm_data. You then accidentally re-run all files and overwrite the results. R will not overwrite the incidents.csv files which you have edited. You won't have to go back and re-examine each incident.

- But if you want to start over on a file (like say with new filter settings), you can re-run the code, and then manually copy the incidents.csv file from /raw_data over to /interm_data. Then you can re-examine each incident.

New files in /raw_data
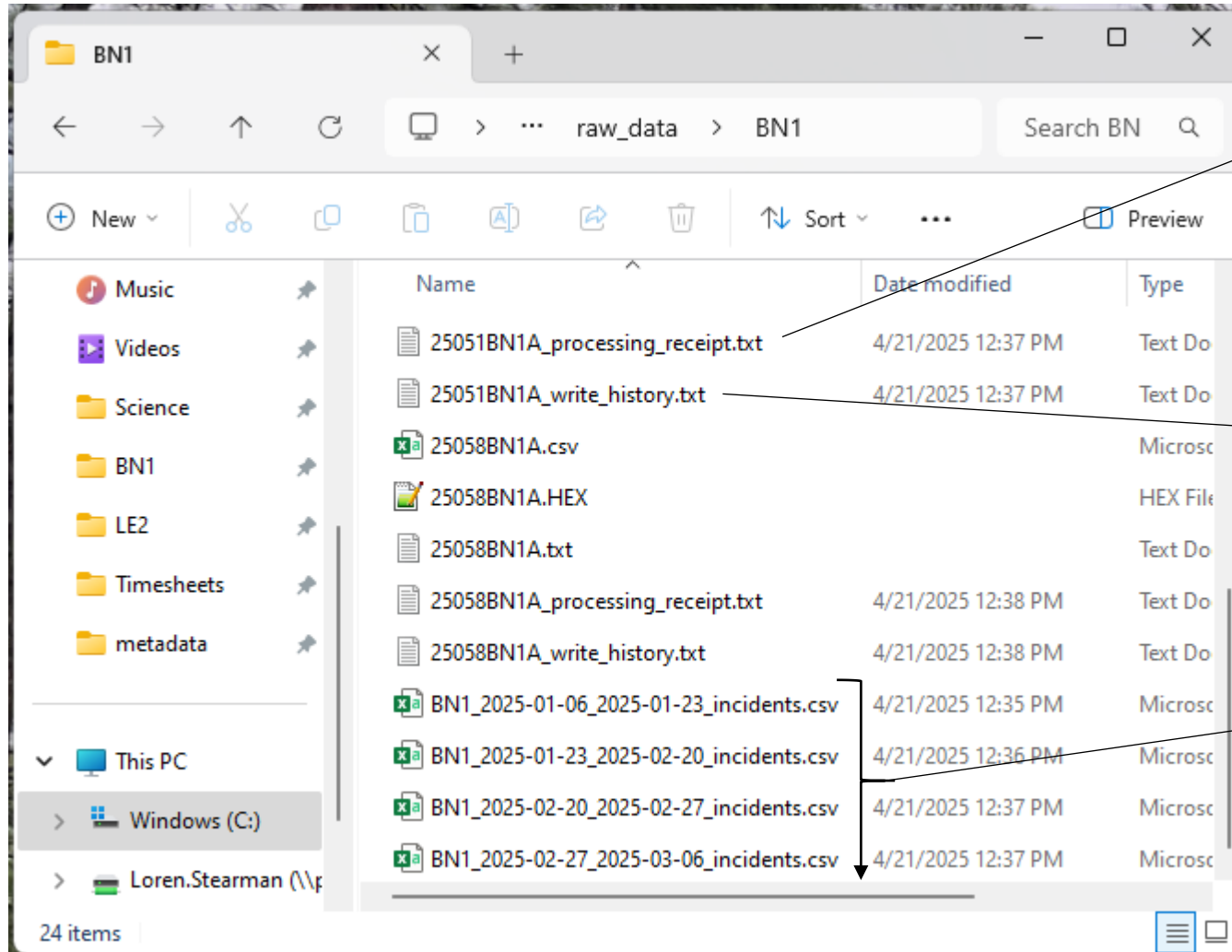
The processing_receipt.txt file
- This file records metadata about the QAQC run
- It also records whether or not the corresponding incidents.csv file in /interm_data was overwritten
- This file is overwritten with every run

The write_history.txt file
- Every time a file is reanalyzed, the date, time, and filter settings are recorded here. This is so that previous runs can be replicated.
- Changes append to this file – they aren't deleted.

The incidents.csv file(s)
- If this is the first run for the file, this file is identical to the version in /interm_data (see next slides)
- If the file was previously analyzed, this file is the new version
- To replace the old version in /interm_data, copy this one over

New files in /interm_data
Note the file name structure:
sensor_startdate_enddate_filetype

The incidents.csv file
- Individual incidents detected by R (a ping or series of pings with intervals of <300 seconds)
- Metadata associated with each ping
- Metadata associated with background noise
- Filter values and pass status
- A column called "Recommend" which is where R specifies (and you verify) that an incident is an observation (a fish)

A primary diagnostic plot which shows all incidents during the time period, overlayed on noise and timer tags, with suggested fish highlighted

Individual plots for each incident with a minimum number of pings:
- <100 incidents: 1 pings
- 100 – 500 incidents: 2 pings
- >500 incidents: 5 pings

Note that this is because it takes FOREVER to plot 500+ plots

File Explorer window:

BO1_2025-02-20_2025-02-27

← → ↑ C 🖥 > ⋯ BO1_2025-02-20_2025-02-27   Search BC 🔍

⊕ New ∨   ✂   ⧉   📋   🅰   ↗   🗑   ↑↓ Sort ∨   ⋯   ⬚ Preview

- ▶ Videos 📌
- 📁 Science 📌
- 📁 BN1 📌
- 📁 LE2 📌
- 📁 LD1 📌
- 📁 Timesheets 📌

- ∨ 💻 This PC
  - 〉 💾 Windows (C:)
  - 〉 💾 Loren.Stearman (\\
  - 〉 💾 Network

17 items

- 📊 BO1_2025-02-20_2025-02-27_incidents.csv
- 📄 BO1_2025-02-20_2025-02-27_initial_diagnosti...
- 📄 BO1_2060_2025-02-21_17-41-48_2025-02-21_1...
- 📄 BO1_2080_2025-02-23_01-58-12_2025-02-23_0...
- 📄 BO1_2080_2025-02-27_03-43-41_2025-02-27_0...
- 📄 BO1_2100_2025-02-25_18-13-44_2025-02-25_1...
- 📄 BO1_2100_2025-02-26_05-48-58_2025-02-26_0...
- 📄 BO1_2160_2025-02-24_05-42-20_2025-02-24_0...
- 📄 BO1_2200_2025-02-23_14-55-36_2025-02-23_1...
- 📄 BO1_2260_2025-02-24_04-55-33_2025-02-24_0...
- 📄 BO1_2340_2025-02-25_05-21-22_2025-02-25_0...

Structure of the incidents.csv file

- Sensor: the sensor
- Tag_Code: the inferred tag code
- Time_Start: the start time of the incident
- Time_Stop: the stop time of the incident
- Elapsed_Seconds: How many seconds the incident lasted
- Tag_Pings: How many pings were detected from the tag
- Noise_Pings: How many noise pings occurred at the same time
- Noise_Ping_Rate: The number of noise pings per second
- Tag_Noise_Ratio: The ratio of tag pings to noise pings

- Max_Tag_Strength: The strongest reading from the tag during the interval
- Mean_Tag_Strength: The mean tag reading during the interval
- Mean_Noise_Strength: The mean noise reading during the interval
- Diff_dB: The difference (in decibels) between mean tag and mean noise strength
- Filter_Pass: The number of filters passed by the incident
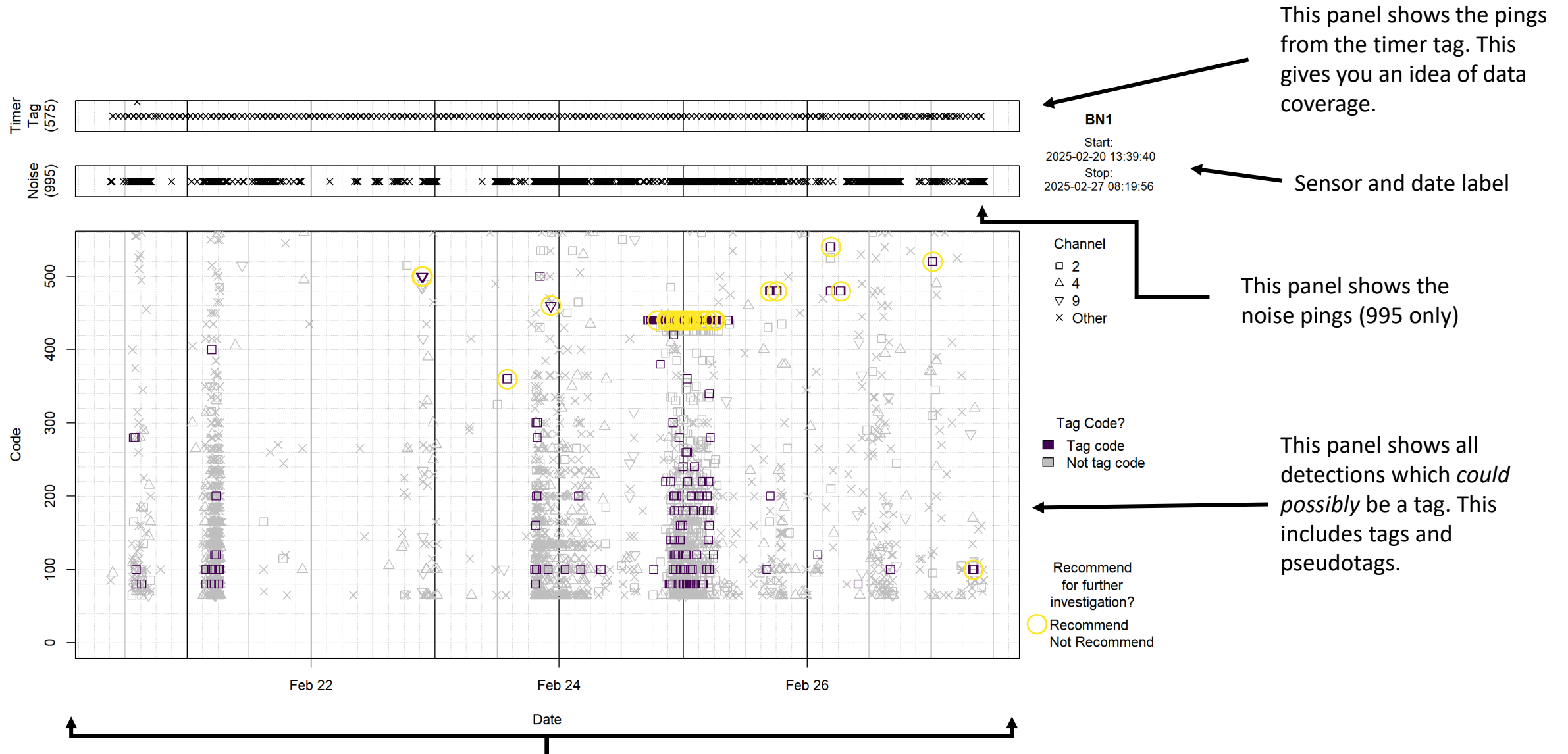- Recommend: R's call on whether the incident represents a real observation of a fish. 0 = no, 1 = yes

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Sensor	Tag_Code	Time_Start	Time_Stop	Elapsed_Seconds	Tag_Pings	Noise_Pings	Noise_Rate	Tag_Noise_Ratio	Max_Tag_Strength	Mean_Tag_Strength	Mean_Noise_Strength	Diff_dB	Filter_Pass	Recommend
2	BN1	2240	2/11/2025 17:01	2/11/2025 20:14	11584	4681	43	0.003712017	108.8604651	-77	-101.4029054	-95.86046512	-5.542440246	3	1
3	BN1	2440	2/12/2025 21:29	2/13/2025 0:26	10618	1986	28	0.002637031	70.92857143	-77	-108.4969789	-106.7142857	-1.782693138	3	1
4	BN1	2480	2/8/2025 22:24	2/9/2025 0:34	7846	1534	14	0.001784349	109.5714286	-73	-107.2451108	-95.57142857	-11.67368225	3	1
5	BN1	2440	2/13/2025 17:06	2/13/2025 18:32	5158	1379	8	0.001550989	172.375	-76	-100.5264685	-81	-19.52646846	3	1
6	BN1	2240	2/9/2025 23:38	2/10/2025 0:21	2550	758	6	0.002352941	126.3333333	-77	-98.8707124	-93.5	-5.370712401	3	1
7	BN1	2480	2/8/2025 20:55	2/8/2025 21:47	3170	635	3	0.000946372	211.6666667	-105	-108.0110236	-95.33333333	-12.67769029	3	1
8	BN1	2240	2/7/2025 19:02	2/7/2025 20:15	4359	182	1	0.00022941	182	-108	-112.2582418	-113	0.741758242	3	1
9	BN1	2480	2/8/2025 21:57	2/8/2025 22:17	1159	112	1	0.000862813	112	-104	-108.9285714	-109	0.071428571	3	1
10	BN1	2440	2/13/2025 23:20	2/13/2025 23:45	1463	109	3	0.002050581	36.33333333	-76	-110.559633	-110.6666667	0.107033639	3	1
11	BN1	2440	2/13/2025 1:02	2/13/2025 1:16	815	97	7	0.008588957	13.85714286	-104	-109.5876289	-108	-1.587628866	3	1
12	BN1	2240	2/9/2025 2:22	2/9/2025 2:38	953	89	0	0	Inf	-110	-112.4044944	0	-112.4044944	3	1
13	BN1	2440	2/12/2025 18:56	2/12/2025 19:18	1376	85	4	0.002906977	21.25	-110	-113.1411765	-114	0.858823529	3	1
14	BN1	2440	2/12/2025 20:57	2/12/2025 21:14	1011	59	1	0.00098912	59	-110	-112.7457627	-112	-0.745762712	3	1

Structure of the incidents.csv file

When you are examining incidents, this is the column where you will either go with R's recommendations or change them. Type a 1 if you want the incident to be considered in further analyses, or a 0 if you don't. R makes a suggestion, but you don't have to go with it!

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Sensor	Tag_Code	Time_Start	Time_Stop	Elapsed_Seconds	Tag_Pings	Noise_Pings	Noise_Rate	Tag_Noise_Ratio	Max_Tag_Strength	Mean_Tag_Strength	Mean_Noise_Strength	Diff_dB	Filter_Pas	Recommend
2	BN1	2240	2/11/2025 17:01	2/11/2025 20:14	11584	4681	43	0.003712017	108.8604651	-77	-101.4029054	-95.86046512	-5.542440246	3	1
3	BN1	2440	2/12/2025 21:29	2/13/2025 0:26	10618	1986	28	0.002637031	70.92857143	-77	-108.4969789	-106.7142857	-1.782693138	3	1
4	BN1	2480	2/8/2025 22:24	2/9/2025 0:34	7846	1534	14	0.001784349	109.5714286	-73	-107.2451108	-95.57142857	-11.67368225	3	1
5	BN1	2440	2/13/2025 17:06	2/13/2025 18:32	5158	1379	8	0.001550989	172.375	-76	-100.5264685	-81	-19.52646846	3	1
6	BN1	2240	2/9/2025 23:38	2/10/2025 0:21	2550	758	6	0.002352941	126.3333333	-77	-98.8707124	-93.5	-5.370712401	3	1
7	BN1	2480	2/8/2025 20:55	2/8/2025 21:47	3170	635	3	0.000946372	211.6666667	-105	-108.0110236	-95.33333333	-12.67769029	3	1
8	BN1	2240	2/7/2025 19:02	2/7/2025 20:15	4359	182	1	0.00022941	182	-108	-112.2582418	-113	0.741758242	3	1
9	BN1	2480	2/8/2025 21:57	2/8/2025 22:17	1159	112	1	0.000862813	112	-104	-108.9285714	-109	0.071428571	3	1
10	BN1	2440	2/13/2025 23:20	2/13/2025 23:45	1463	109	3	0.002050581	36.33333333	-76	-110.559633	-110.6666667	0.107033639	3	1
11	BN1	2440	2/13/2025 1:02	2/13/2025 1:16	815	97	7	0.008588957	13.85714286	-104	-109.5876289	-108	-1.587628866	3	1
12	BN1	2240	2/9/2025 2:22	2/9/2025 2:38	953	89	0	0	Inf	-110	-112.4044944	0	-112.4044944	3	1
13	BN1	2440	2/12/2025 18:56	2/12/2025 19:18	1376	85	4	0.002906977	21.25	-110	-113.1411765	-114	0.858823529	3	1
14	BN1	2440	2/12/2025 20:57	2/12/2025 21:14	1011	59	1	0.00098912	59	-110	-112.7457627	-112	-0.745762712	3	1

The primary diagnostic plot

This panel shows the pings from the timer tag. This gives you an idea of data coverage.

BN1

Start:
2025-02-20 13:39:40
Stop:
2025-02-27 08:19:56

Sensor and date label

This panel shows the noise pings (995 only)

This panel shows all detections which *could possibly* be a tag. This includes tags and pseudotags.
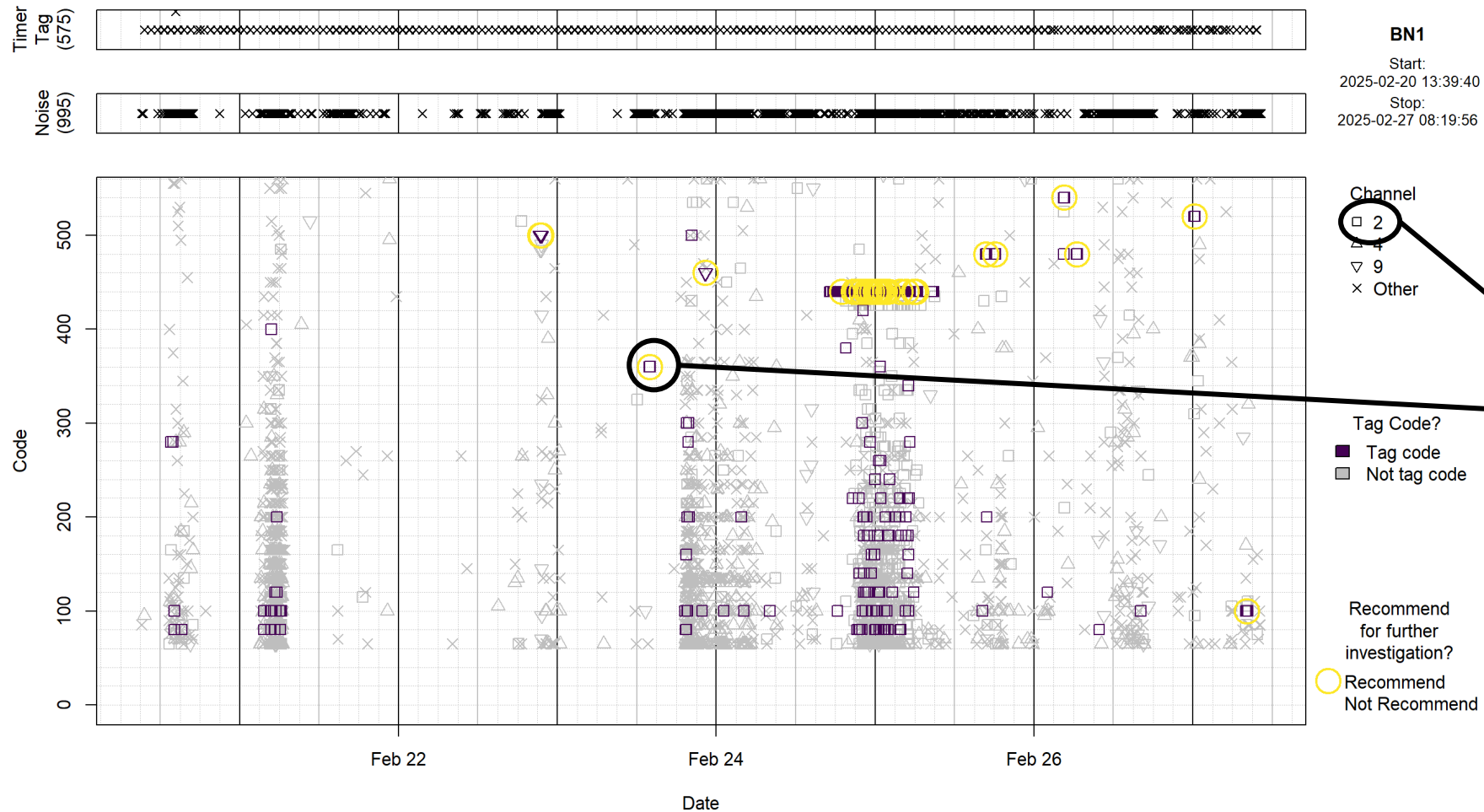
Dates and times are aligned across all plots. Note solid black lines at 0000 hours and solid grey lines at 1200 hours

The primary diagnostic plot
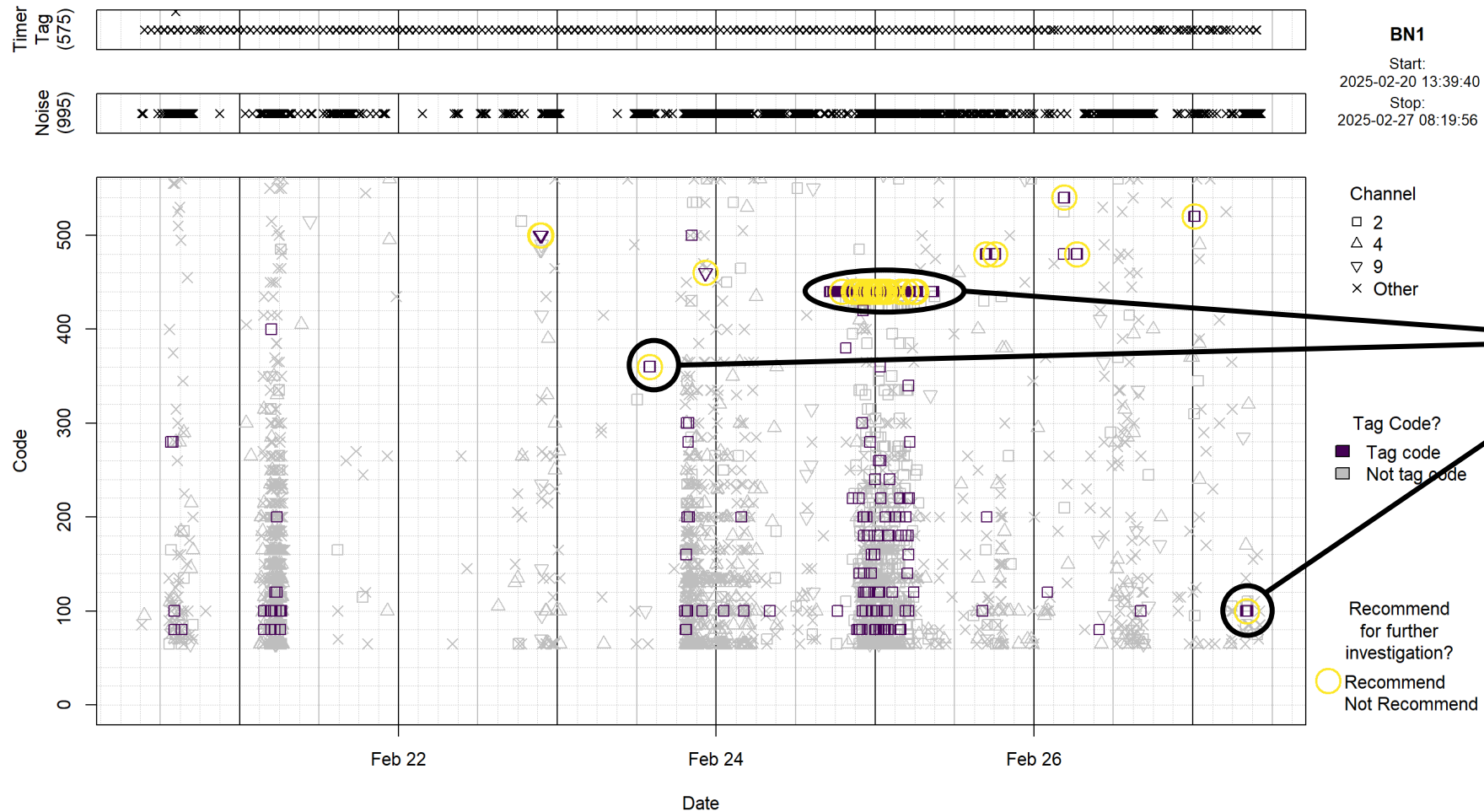
Real tag numbers are coded as dark purple points. Suggested incidents are highlighted with yellow circles.

The primary diagnostic plot

Tag codes on the Y axis represent the ping interval. The channel portion of the code is shown by the symbol in the legend. So this highlighted tag is channel 2 (square), interval 360 – tag code 2360.

Individual diagnostic plots
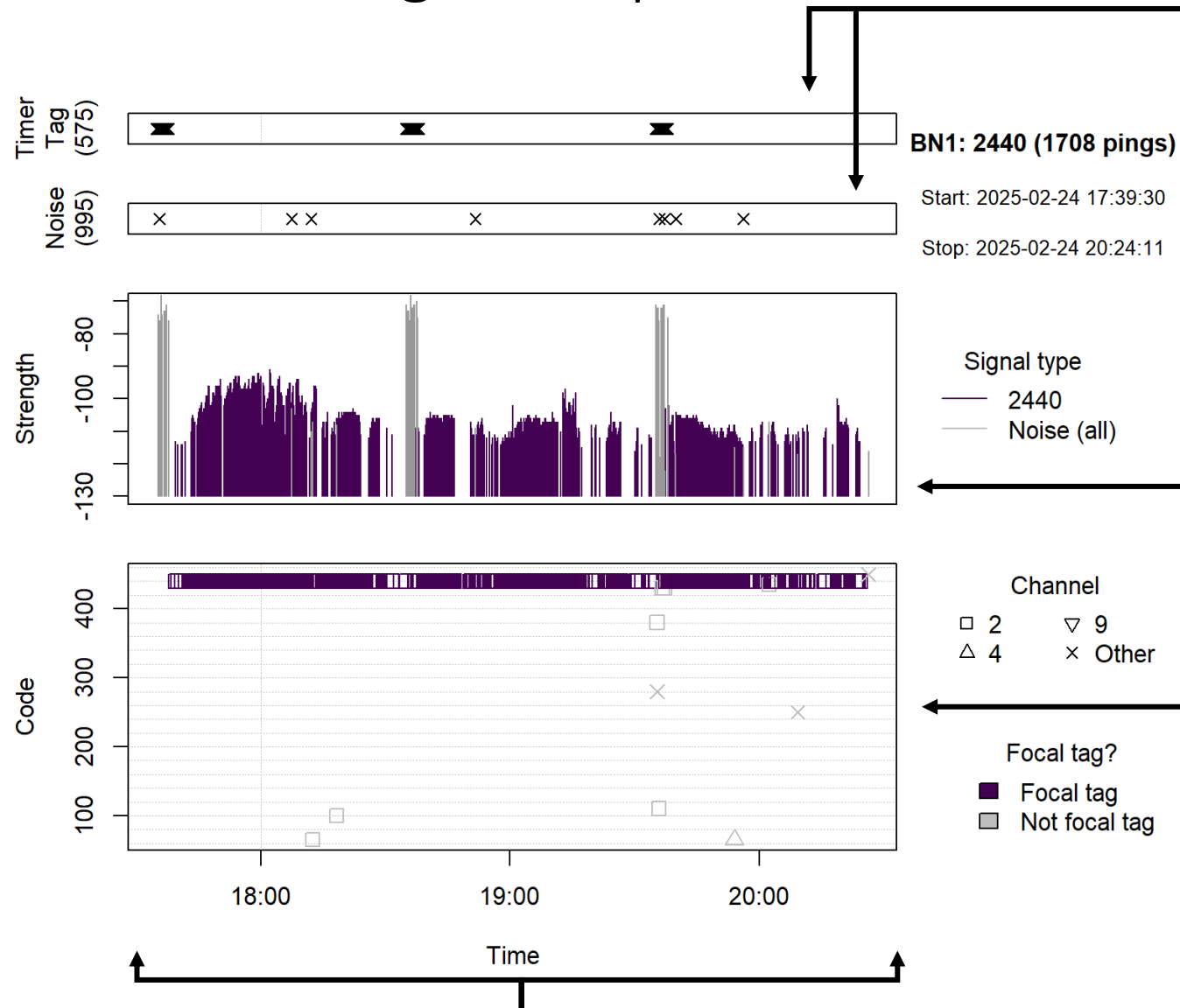
Just like the primary plot, the timer tag and noise are plotted in the top two panels

Sensor, tag, and date label with total ping count

BN1: 2440 (1708 pings)

Start: 2025-02-24 17:39:30

Stop: 2025-02-24 20:24:11

This panel shows the strength of both tag (dark purple) and noise (light grey) pings. Noise includes timer tag, other tags, and pseudotags. Noise is plotted on top of the tag, and will obscure the tag if it is louder than the tag. This is by design. If you see grey when purple should be there, noise is too loud.

This panel shows detections of the tag, plus detections of other potential tag numbers. Unlike the primary plot, only the focal tag is highlighted in purple.

Dates/times are still aligned, but as most detections are short interval, they are plotted as times

How filtering works

- R uses four filters to suggest whether an incident should be an observation (default values below):

  - Number of pings (minimum 5)

  - Rate of noise pings (maximum 0.1/second)

  - # tag pings : # noise pings ratio (minimum 0.5, or one tag ping per two noise pings)

  - Tag – noise strength difference (minimum 3dB)

- If there are fewer than 5 pings, the incident receives an automatic fail. The incident must also pass 2 of the other 3 filters to be suggested.
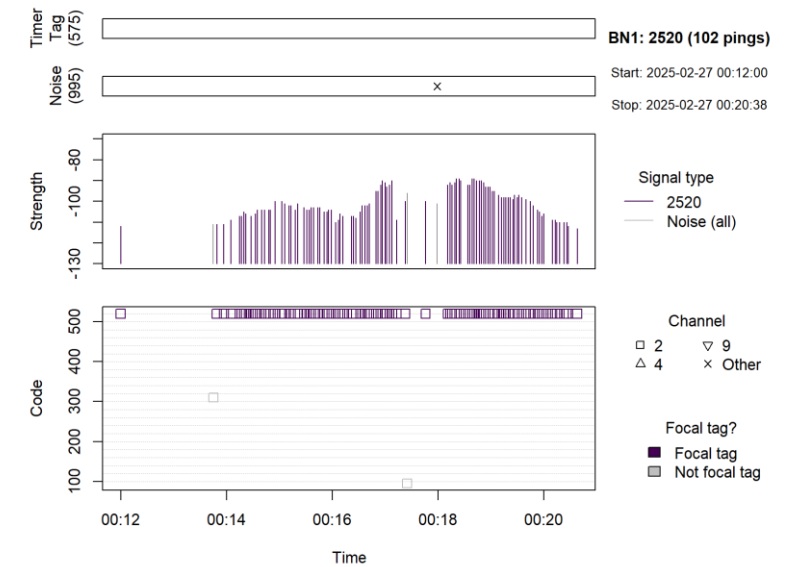

A good detection – lots of tag pings, little noise
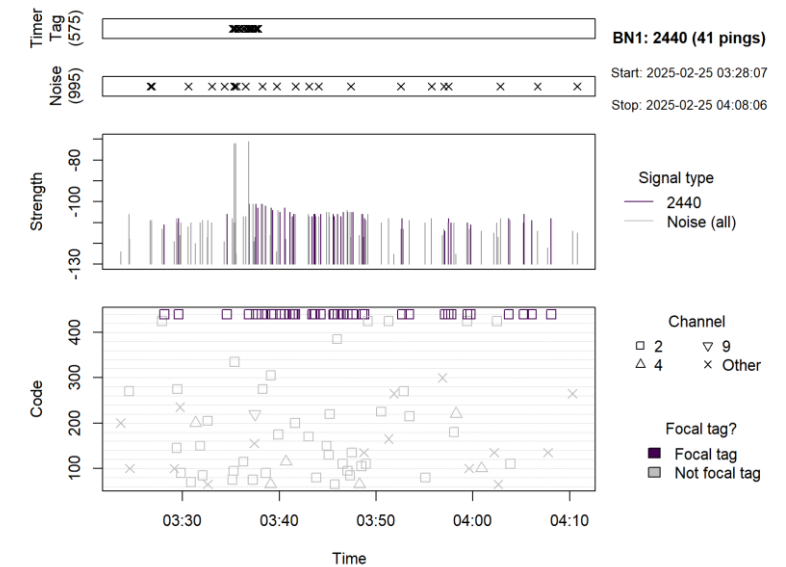
More questionable – lots of tag pings, but high noise rate and poor tag – noise strength difference

Next steps: validating incidents as observations

- This is done by hand after reviewing QAQC tools, using the incidents.csv file **in the /interm_data folder**

- If you like what R decided, keep it. If you don't, change it.

- If you are constantly changing calls, we can tweak the filters

- Most obvious cases will be good calls, but cases near the filter values may not be

A questionable call… remember R isn't perfect and nature is weird

BN1: 2440 (6 pings)

Start: 2025-02-24 23:26:20

Stop: 2025-02-24 23:34:01
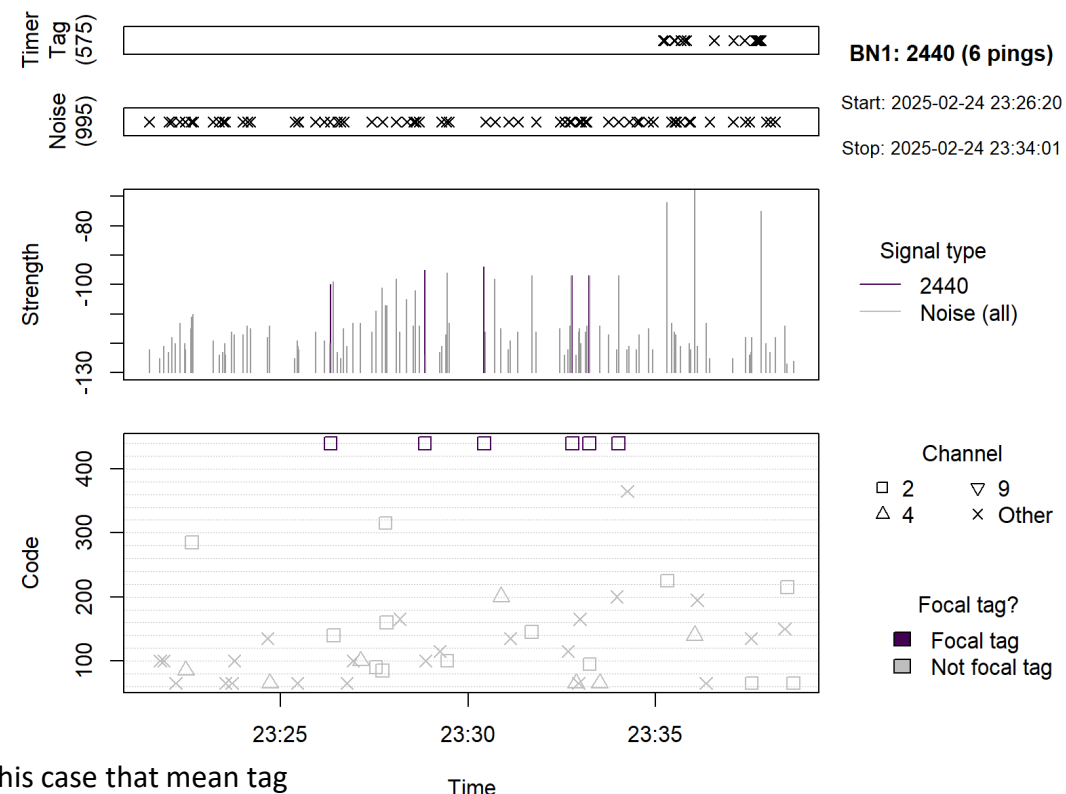
Note in this case that mean tag strength – mean noise strength looks good, but "tag" detections are occurring in periods of higher noise strength

Lots of noise
Tag and noise strength correlate
Lots of dubious "tags"

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Sensor	Tag_Code	Time_Start	Time_Stop	Elapsed_Seconds	Tag_Pings	Noise_Pings	Tag_Noise_Ratio	Max_Tag_Strength	Mean_Tag_Strength	Mean_Noise_Strength	Diff_dB	Filter_Pass	Recommend
23	BN1	2440	2/25/2025 1:54	2/25/2025 1:59	298	7	26	0.269230769	-95	-100.4285714	-114.1538462	13.72527473	3	1
24	BN1	2440	2/24/2025 22:34	2/24/2025 22:40	371	7	38	0.184210526	-73	-97.57142857	-109.9473684	12.37593985	3	1
25	BN1	2440	2/24/2025 23:26	2/24/2025 23:34	461	6	48	0.125	-94	-96.66666667	-113.9166667	17.25	3	1

This is where you would change the value

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)
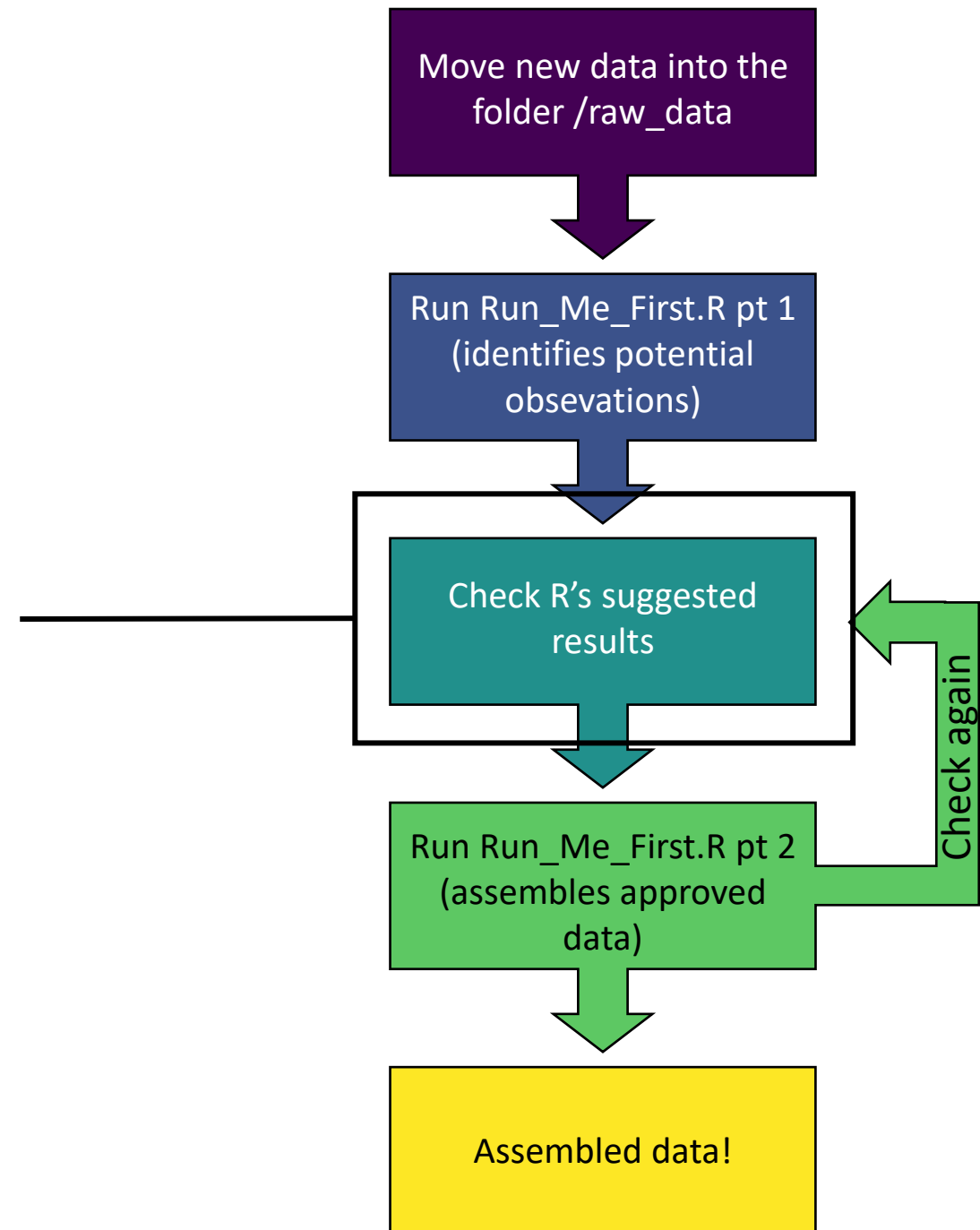
Run_Me_First.R   qaqc_back_end.R   qaqc_front_end.R

Source on Save   → Run   → Source

```
26
27   # Run this set of lines the first time you run this script, and any time that
28   # you update R. This set of lines checks for and installs necessary packages.
29   # It doesn't hurt to re-run it, but you don't need to once they are installed.
30   # THE CODE WILL NOT RUN WITHOUT THESE PACKAGES
31       all_packages <- installed.packages()
32       if (!"khroma" %in% all_packages) {
33           install.packages("khroma")
34       }
35       if (!"viridis" %in% all_packages) {
36           install.packages("viridis")
37       }
38
39
40   # Run the line below to conduct data management and QAQC.
41   # Highlight the line below and select "Run" from above ---------------------^
42       source("qaqc_front_end.R")
43
44
45
46   # Run the line below to conduct data assembly.
47   # THIS SHOULD ONLY BE DONE AFTER HUMAN QAQC OF INCIDENT CALLS FROM R!!!
48   # GO BACK TO RESULTS IN /interm_data/ AND VERIFY THE CALLS ARE GOOD!!!
49   # DANGER DO NOT PROCEED WITHOUT HUMAN QAQC!!!
50   # Highlight the line below and select "Run" from above ---------------------^
51       source("assembly_front_end.R")
52
53
54
55
56
```

52:5   (Top Level)   R Script

Console   Terminal   Background Jobs

R 4.4.1 · C:/Science/lower_granite/RT Data/

```
All files successfully analyzed!
Stop time at 2025-04-21 09:37:33.028659

STOP BEFORE YOU PROCEED!
You need to review the output files in the folders of /interm_data!!!
Check for edge cases where you disagree with R!!!
There are multiple plots and a csv to help you review each case.
If you disagree, change the value (0 = no, 1 = yes) in the 'Recommend' column
of the incident csv file in the /interm_data folder.

Warning message:
package 'viridis' was built under R version 4.4.3
>
```

Environment   History   Connections   Tutorial

metadata		
qaqc_back_end.R	32.6 KB	Apr 16, 2025, 10:45 AM
qaqc_front_end.R	7.5 KB	Apr 16, 2025, 1:11 PM
raw_data		
ReadMe.pdf	452.5 KB	Apr 16, 2025, 1:13 PM
Run_Me_First.R	2.8 KB	Apr 16, 2025, 1:11 PM
interm_data		

Activate Windows
Go to Settings to activate Windows.

Callout:

Now you can run the second line to assemble validated incidents (observations)

Remember, ONLY DO THIS IF YOU HAVE ALREADY VALIDATED THE INCIDENTS YOU WANT TO KEEP AS OBSERVATIONS

If you only go with what R has filtered, you are almost certainly going to have some wrong calls. Please, please double check R!!!
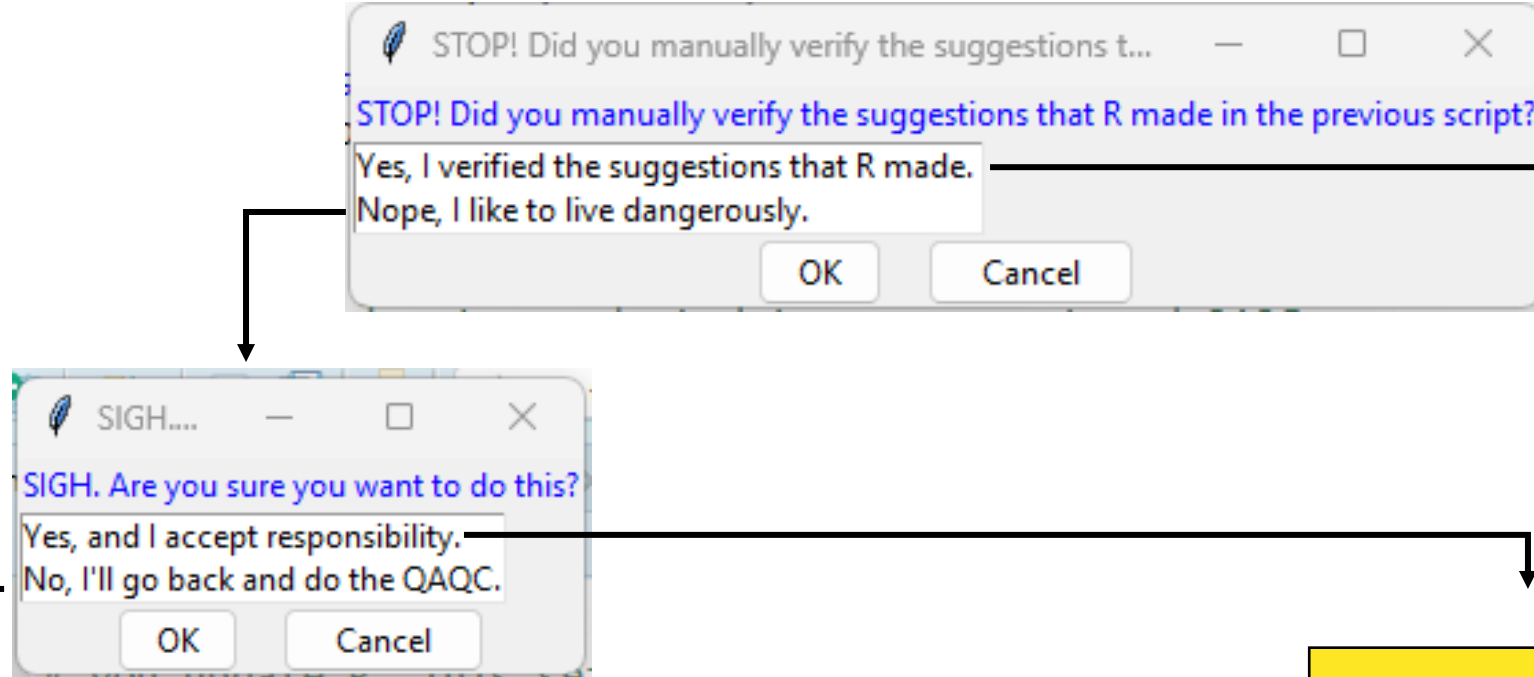
This part runs very fast (a few seconds to a minute at most)

User validation verification dialogues

These open when you run the script "assembly_front_end.R", which is automatically called by "Run_Me_first.R"

These dialogues are to help impress upon the user the importance of verifying R's calls. They are designed for when you have new helpers.

STOP! Did you manually verify the suggestions t...

STOP! Did you manually verify the suggestions that R made in the previous script?

Yes, I verified the suggestions that R made.
Nope, I like to live dangerously.

OK   Cancel

SIGH....

SIGH. Are you sure you want to do this?

Yes, and I accept responsibility.
No, I'll go back and do the QAQC.

OK   Cancel

```
package 'viridis' was built under R version 4.4.3
> # Run the line below to conduct data assembly.
> # THIS SHOULD ONLY BE DONE AFTER HUMAN QAQC OF INCIDENT CALLS FROM R!!!
> # GO BACK TO RESULTS IN /interm_data/ AND VERIFY THE CALLS ARE GOOD!!!
> # DANGER DO NOT PROCEED WITHOUT HUMAN QAQC!!!
> # Highlight the line below and select "Run" from above ----------------------^
>     source("assembly_front_end.R")

Good call! I'll be here ready to run again when you are finished!
```

Assembled data!

Remember, assembled data doesn't always equal good data.
VALIDATE YOUR INCIDENTS!!

New files in /assembled_data
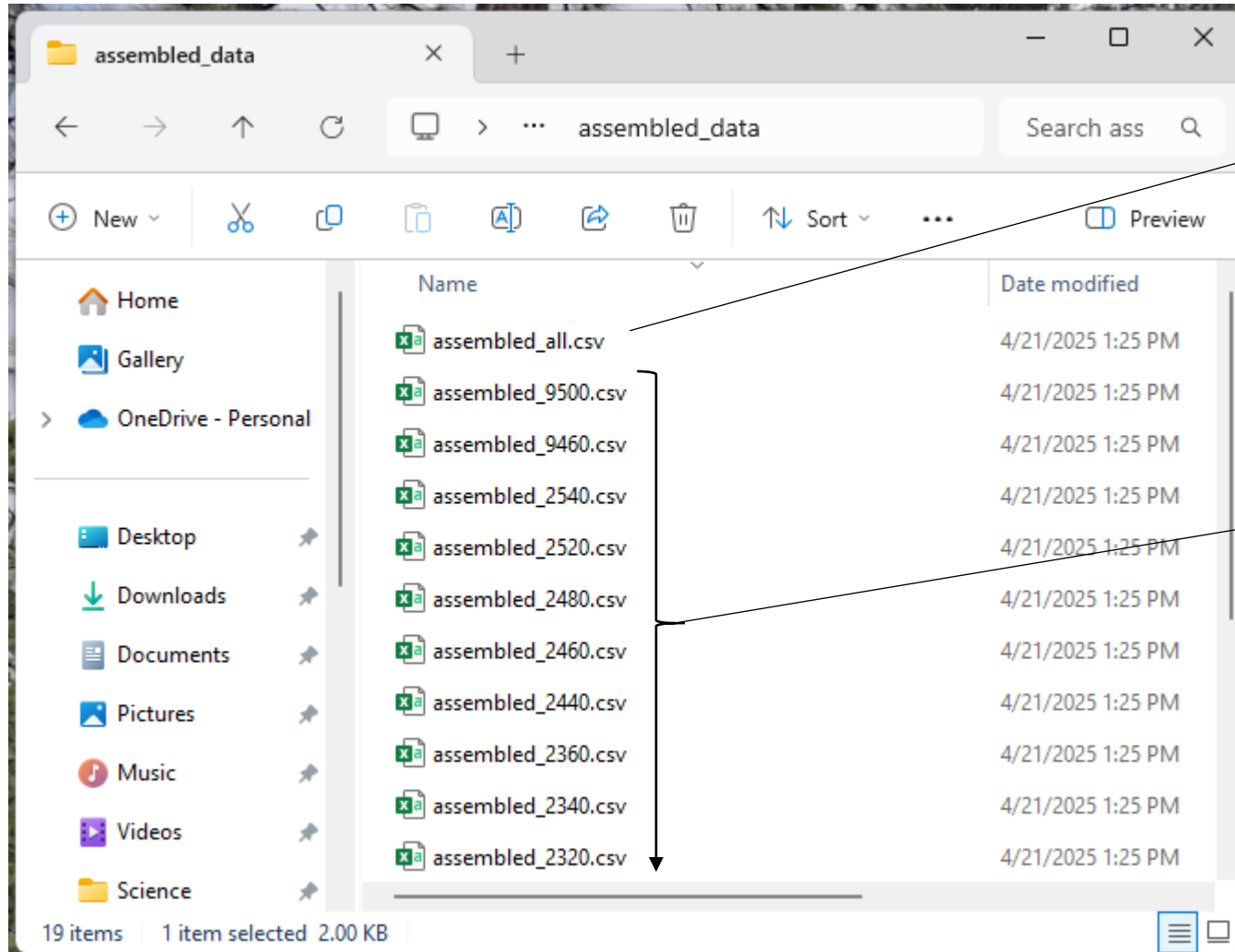
The assembled data file
- This file records validated observations
- The format is identical to the incident.csv files
- This file contains data for all tags

Individual tag assembled data files
- This file records validated observations for one tag
- The format is identical to the incident.csv files
- It is a subset of the main assembled data file

You may need to do another round of QAQC by looking at where tags show up. For example, are there simultaneous detections at two receivers far away from one another?

If you find incidents that you want to add or remove at this stage, you can re-run the last chunk of the R code (the assembly front end line) after editing the appropriate incidents.csv file in /interm_data

Move new data into the folder /raw_data

Run Run_Me_First.R pt 1 (identifies potential obsevations)

Check R's suggested results

Run Run_Me_First.R pt 2 (assembles approved data)

Check again

Assembled data!