PULSAR Otago Input and Output formats

Required input file format

1. Input files must be .csv
2. First cell of first row contains experiment name (can be any string)
3. Second row contains column headers (can be any strings; see column definitions below)
4. Rows 3 to 500 contain data values in multiple columns. Data columns will typically correspond to individual animals in a single experiment, and column headers will be each animal's ID. However this is not a requirement. The only formal restriction is that each column header must be unique.
5. In rows 3 to 500, column values are:
    a. Column 1 must be ordinal sample number. All data series in a file use the same sample numbers.
    b. Column 2 must be time in minutes from start of session. All data series in a file must use the same time values.
    c. Columns 3 to n are individual data series, one series per column. There is no limit on the number of data columns in a single file.[1]
6. Example: In this file, each column is an individual (simulated) animal. The column headers are animal ids.

	A	B	C	D	E	F	G	H
1	Sim Six							
2	sample	time	wufhl	ogndp	tjxgq	gpfdo	ruypp	arszm
3	1	0	0	1.937935	1.187298	1.508838	0.656911	1.259434
4	2	5	0.427438	0.512648	1.2772	1.604273	1.043127	1.812398
5	3	10	1.506043	4.449308	1.141557	0.934403	1.963995	0
6	4	15	1.114234	1.454254	1.840361	4.517327	1.155115	0.71719
7	5	20	0.437685	0.563906	1.660121	0.747696	2.275511	1.028442
8	6	25	0.720989	1.06228	1.324477	0.749747	0.373916	0.618342
9	7	30	1.778086	1.244067	2.039745	0.905767	2.226389	0.622235
10	8	35	0.26568	0.964429	0.550703	2.043873	0.933668	0.625648
11	9	40	0.871585	1.717319	1.694299	1.04076	4.315315	0.187265

Batch Processing

Multiple input files can be processed simultaneously (use function read_pulsar_input_folder in file pulsar_utilities.R). For batch processing:

1. All files must be contained in a single folder
2. Each file must have only one data series
3. All files must follow the input file format described above.

A single multi-column input file is constructed from the folder contents, and then normal processing can be performed. For batch processing:

1. The "experiment name" is set to the name of the input folder
2. The column header for each data series column is set to the input file name (minus the file type suffix) for that data series.

[1] In the R shiny version of PULSAR Otago, only the first nine columns are displayed.

Output Files Generated:

In these descriptions, for simplicity, we refer to each data series as “an animal”.

For each input file, the following are produced:

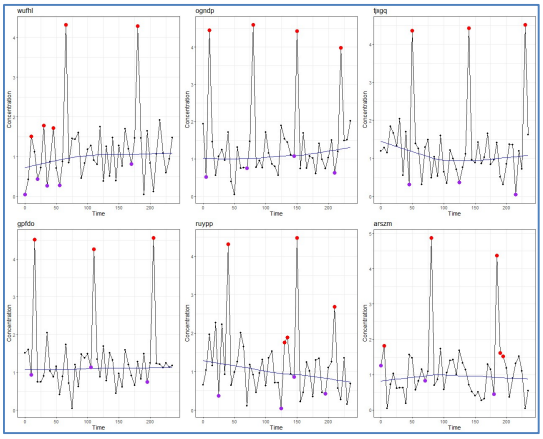
1. A .csv file containing descriptors for each pulse identified in each animal’s data.
    a. File name format is: <experiment name>_peak_desc_<time stamp>.csv
    b. Columns for each identified peak are:
        i. Animal ID (column header), start time (first point in pulse)
        ii. Peak time (point in pulse with maximum raw concentration)
        iii. Maximum concentration in pulse
        iv. Amplitude of pulse (maximum - nadir concentration)
    c. Example:

	A	B	C	D	E	F	G	H
1		AnimalID	StartTime	MaxTime	MaxConc	Amplitude	PeakLength	
2	1	wufhl	10	10	1.506042905	0.749123747	2	
3	2	wufhl	30	30	1.778086117	0.948331926	1	
4	3	wufhl	45	45	1.717160194	0.835167433	1	
5	4	wufhl	65	65	4.315713203	3.368899806	1	
6	5	wufhl	75	85	1.606715748	0.608256793	3	
7	6	wufhl	120	120	1.747445795	0.695807702	1	
8	7	wufhl	140	140	1.478637236	0.420960568	1	
9	8	wufhl	160	160	1.703385831	0.651411531	1	
10	9	wufhl	175	180	4.286979126	3.226869125	3	
11	10	wufhl	195	195	1.649577669	0.581433129	1	
12	11	wufhl	215	215	1.922016692	0.844766624	1	
13	12	ogndp	10	10	4.449308492	3.438889384	2	
14	13	ogndp	40	40	1.717319489	0.714386206	1	
15	14	ogndp	55	55	1.317907745	0.313528843	1	
16	15	ogndp	75	80	4.59405095	3.581628533	2	
17	16	ogndp	100	100	1.716181697	0.673502342	1	
18	17	ogndp	125	125	1.890410879	0.813688603	3	
19	18	ogndp	150	150	4.433539472	3.343693972	1	
20	19	ogndp	160	160	1.68909464	0.584103656	1	
21	20	ogndp	220	220	3.97743278	2.71420835	1	
22	21	tjxgq	15	15	1.840361	0.47878461	2	
23	22	tjxgq	30	30	2.039744585	0.759843508	1	

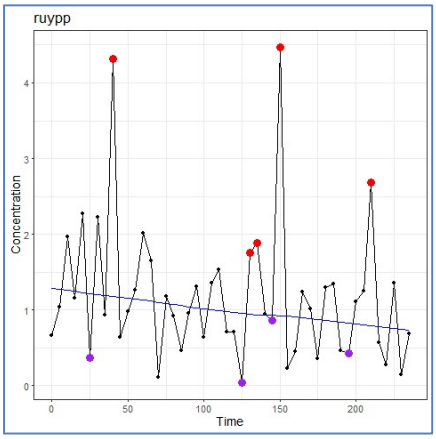
2. A .csv file containing summary values for the peaks identified within each animal.
    a. File name format is: <experiment name>_peak_summ_<time stamp>.csv
    b. Columns for each animal are:
        i. Animal ID
        ii. Average amplitude of all peaks
        iii. Standard deviation of amplitude of all peaks
        iv. Average peak interval length (in minutes)
        v. Standard deviation of peak interval length
        vi. Peak frequency in peaks per hour as measured from the time of the initial peak.
    c. Example:

	A	B	C	D	E	F	G
1		AnimalID	AvgAmp	SdAmp	AvgPeakInterval	SdPeakInterval	PeakFrequencyHour
2	1	wufhl	1.175548035	1.059552468	20.5	5.502524673	2.808510638
3	2	ogndp	1.797514432	1.422464231	26.25	15.05940617	2.29787234
4	3	tjxgq	1.16304686	1.270138169	17.91666667	6.200562047	3.319148936
5	4	gpfdo	1.30985583	1.320377332	20	9.428090416	2.808510638
6	5	ruypp	1.18910353	1.042463564	17.91666667	7.525210155	3.319148936
7	6	arszm	1.317872118	1.346196128	26.875	13.34634782	2.29787234
8							
9							

3. A .csv file containing average concentration for each data series.
    a. File name format is <experiment name>_conc_summ_<time stamp>.csv
    b. Columns for each data series are:
        i. Animal ID
        ii. Average concentration (i.e. average of all values in series)
        iii. Standard deviation of concentration (i.e. standard deviation of all values in series)
4. An image file (.png) containing the PULSAR Otago plots for each animal in the data file combined into a single image.
    a. File name format is: <experiment name>_all_<time stamp>.png
    b. Example:



5. Images containing individual plot images for each animal in the data file.
    c. Folder name is: <experiment name> Images
    d. File name format is: <experiment name>_<animal ID>_<time stamp>.png
    e. Example:



6. A .csv file containing the algorithm parameter values (e.g. G values, cutoffs, smoothing fraction) for the analysis.
    a. File name format is: <experiment name>_parameters_<time stamp>.csv\
    b. Example:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		smoothing_fraction	g1	g2	g3	g4	g5	extinction_threshold	peak_split_depth	sdrQuad	sdrLin	sdrConst	n_steps
2	1	0.75	2.5	2	1.9	1.7	1.6	0	2.7	2	3	4	6
3													
4													