

From Noldus file to Visualisation part 2 : prepare for wakefish

Clair Chaigne

8 novembre 2021

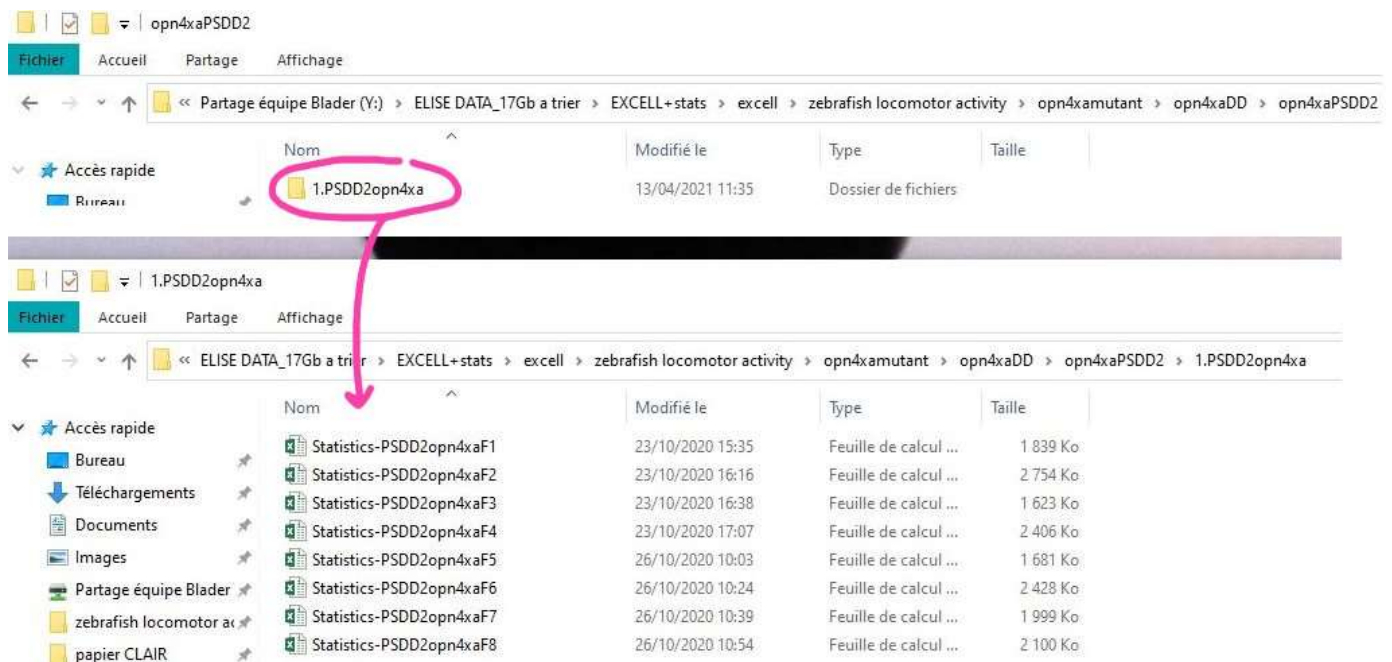
1. INSTRUCTIONS

This is a follow up from From Noldus File to Visualisation part 1 : Noldus

For this example I used the PSDD2opn4xa experiment (8 movies labelled PSDD2opn4xF1 to F8).

2. STEP BY STEP EXAMPLE

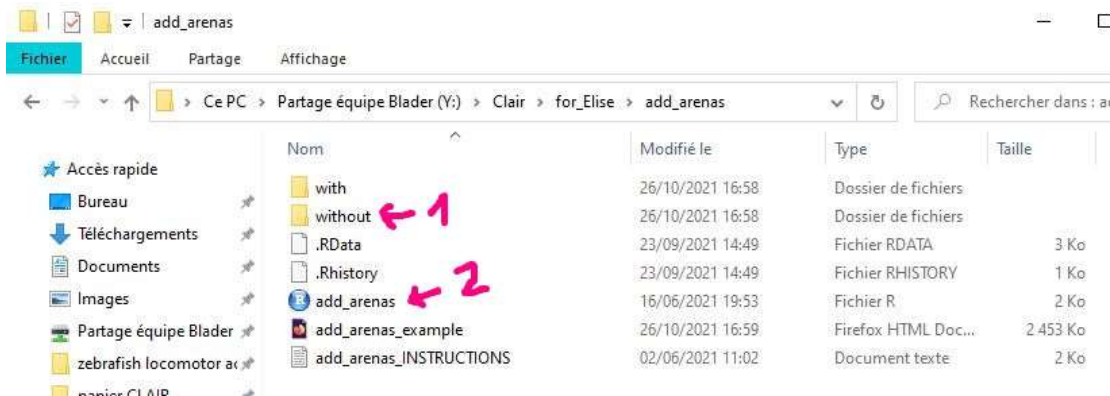
1. At the end of the Noldus processing, you should end up with one excel file for each one of your movies (here 8 excel files for "PSDD2opn4xa"). For each type of experiment (*opn4xaDD*), I create a folder for each one of my experiments (*opn4xaPSDD1*, *opn4xaPSDD2*, *opn4xaPSDD3*, etc...). In the folder corresponding to my experiment (*opn4xaPSDD2*), I like to create a folder called *1.*nameofmyexperiment** where I put the original excel files I directly extracted from the Noldus.



2. (if you have more than 24 fishes). Prepare the files for wakefish processing : to be able to process the files, the wakefish program (program written by L.Sanchou to extract the distance moved per min over x minutes from several files into one final file) needs to detect several column thanks to specific keywords in the excel files, so make sure you have these : **Distance moved** for the column containing the distance moved, **< User-defined 1 >** for the column containing the genotypes, **Arena** for the column containing the Arena names, **Start** for the column containing the time. With 96 arenas, the Noldus files contain "1,2,3,..." instead of "Arena 1, Arena 2, Arena 3,..." , this is why you need an extra step before being able to process them with the wakefish program.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1			1		Independ	Distance moved								
2					Independ	center-point								
3					<User-def	Total								
4						mm								
5	Result 1	Trial	1	1	Start-0:01:b		87,5914							
6	Result 1	Trial	1	1	0:01:00-0:b		78,0216							
7	Result 1	Trial	1	1	0:02:00-0:b		21,4394							
8	Result 1	Trial	1	1	0:03:00-0:b		1,16375							
9	Result 1	Trial	1	1	0:04:00-0:b		-							

2.a : (1) in `Y://Clair//for_Elise//add_arenas` , copy your excel files in the `without` folder, (2) open the `add_arenas` R file (you need to have R installed in your computer). You can follow the `add_arenas_example` if you have any issues or check `add_arenas_INSTRUCTIONS`.



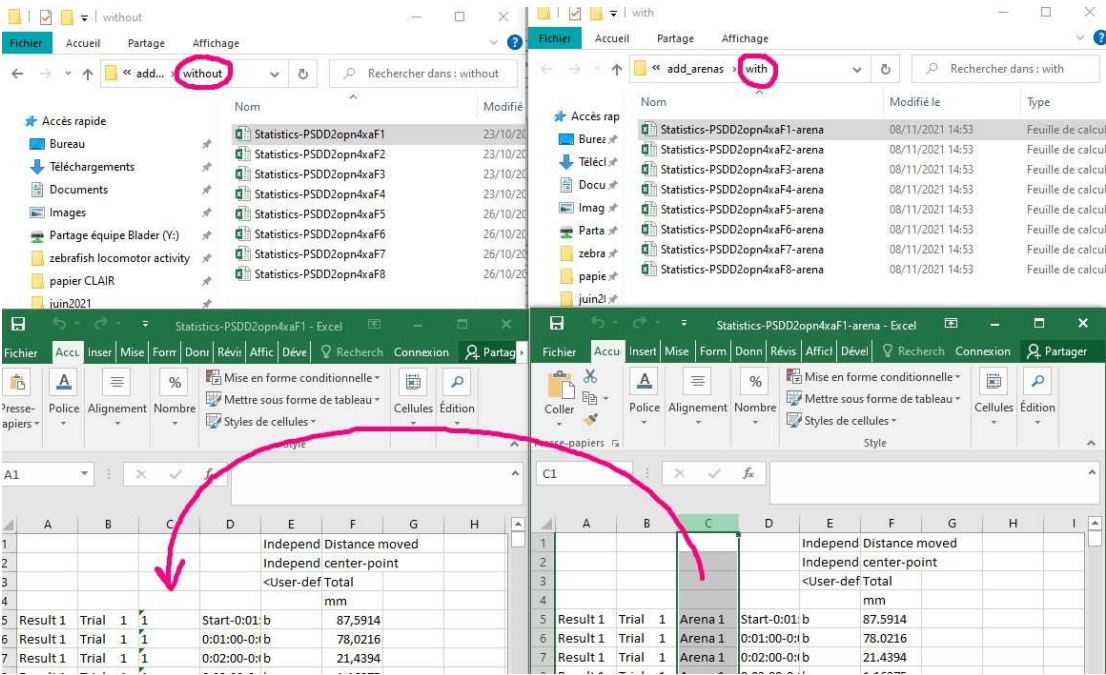
2.b Run the `add_arenas` R file (select all the code with **CTRL+A** and run it with **CTRL+ENTER**)

```

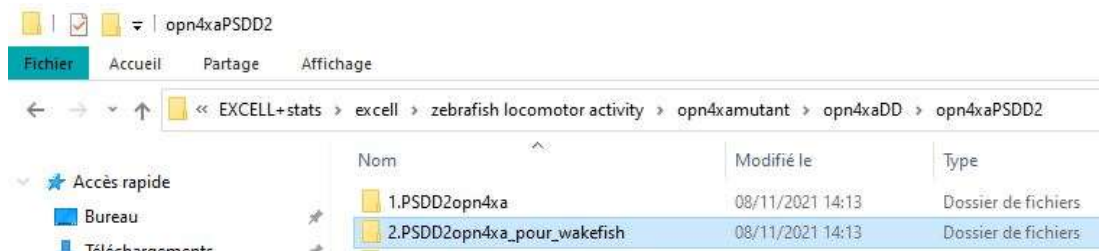
1 # BE SURE TO EMPTY THE 'WITHOUT' FOLDER BEFORE RUNNING THIS SCRIPT#
2 # CHECK THAT THE N CORRESPOND TO THE ARENA COLUMN #
3
4 rm(list=ls()) # clean R environment
5
6 # 0. LIBRAIRIES ----
7 library(readxl)
8 library(writexl)
9
10 n <- 3 # define the column arena
11 m <- 96 # number of arenas
12
13 # 1. DEFINE FOLDERS ----
14 setwd("Y:/clair//for_Elise//add_arenas//without") # input files
15 dossier_sortie <- "Y:/clair//for_Elise//add_arenas//with" # output files
16
17 # 2. LOAD FILES ----
18 nom_fichiers <- list.files()
19 fichiers <- list()
20 for (i in 1:length(nom_fichiers))
21 {
22   fichiers[[i]] <- read_excel(nom_fichiers[i], col_names = FALSE)
23 }
24 names(fichiers) <- sub('\\.xlsx$', '', nom_fichiers)
25
26 # 3. ADD ARENAS ----
27 for (i in 1:length(nom_fichiers))
28 {
29   j <- 1
30   for (j in 1:m)
31   {
32     fichiers[[i]][, n][fichiers[[i]][, n]==j] <- paste("Arena", j, sep = " ")
33   }
34 }
35
36 # 4. EXPORT FILES ----
37 nom_fichiers <- sub('\\.xlsx$', '', nom_fichiers)
38 for (i in 1:length(nom_fichiers))
39 {
40   write_xlsx(fichiers[[i]], path = paste(dossier_sortie, "/", nom_fichiers[i], "-arena.xlsx", sep = ""), col_names = FALSE)
41 }
42
43 # 5. MANUALLY MODIFY THE WITHOUT FILES ----
44
45 # open the files in the without folder and manually copy/paste the corresponding column arena from files with
46 # in the 'with' folder
47 # use the modified files from the without folder for wakefish

```

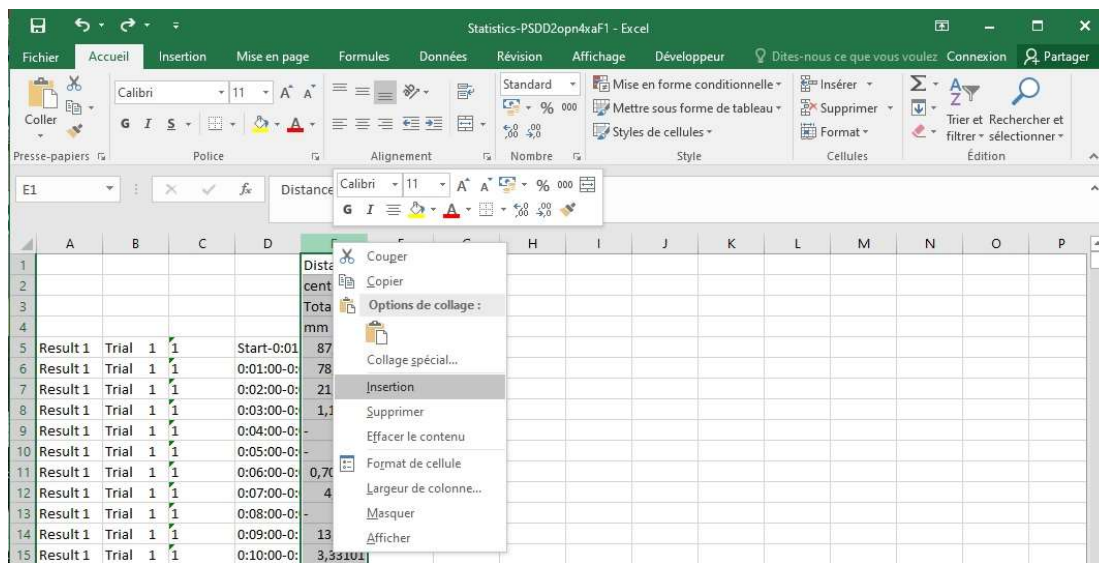
2.c for each file in the with folder (*Statistics-PSDD2opn4xaF1-arena*), copy and paste the Arena column in the corresponding file in the without folder (*Statistics-PSDD2opn4xaF1*).



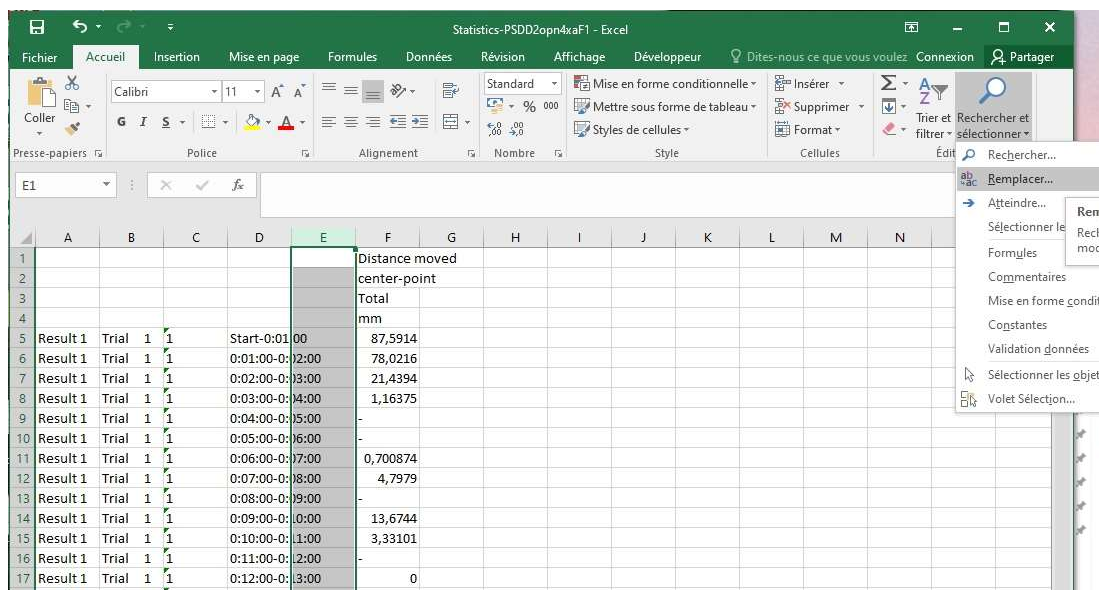
2.d I like to put the modified files ready to go to be processed in another folder called
2.*nameofmyexperiment*_pour_wakefish



3.a If you forgot to put a genotype column, the wakefish program will not like it. But do not worry, you can add it manually ! Open each file, do a right click on a column name (E for example) and click on Insertion (Insert a column)



3.b with the column still selected, click on Rechercher et sélectionner (Search and Select), Remplacer (Replace)



3.c keep the Rechercher box empty, and put a random letter in the Remplacer par box, click on Remplacer tout

Statistics-PSDD2opn4xF1 - Excel

Fichier Accueil Insertion Mise en page Formules Données Révision Affichage Développeur Dites-nous ce que vous voulez Connexion Partager

Calibri 11 A⁺ A⁻ G I S % 000 Mise en forme conditionnelle Mettre sous forme de tableau Styles de cellules Insérer Supprimer Trier et Rechercher et filtrer sélectionner Édition

Presse-papiers Police Alignement Nombre Style Cellules Édition

b

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1					b	Distance moved										
2					b	center-point										
3					b	Total										
4					b	mm										
5	Result 1	Trial 1	1	1	Start-0:01	b	87,5914									
6	Result 1	Trial 1	1	1	0:01:00-0:	b	78,0216									
7	Result 1	Trial 1	1	1	0:02:00-0:	b	21,4394									
8	Result 1	Trial 1	1	1	0:03:00-0:	b	1,16375									
9	Result 1	Trial 1	1	1	0:04:00-0:	b	-									
10	Result 1	Trial 1	1	1	0:05:00-0:	b	-									
11	Result 1	Trial 1	1	1	0:06:00-0:	b	0,700874									
12	Result 1	Trial 1	1	1	0:07:00-0:	b	4,7979									
13	Result 1	Trial 1	1	1	0:08:00-0:	b	-									
14	Result 1	Trial 1	1	1	0:09:00-0:	b	13,6744									
15	Result 1	Trial 1	1	1	0:10:00-0:	b	3,33101									
16	Result 1	Trial 1	1	1	0:11:00-0:	b	-									
17	Result 1	Trial 1	1	1	0:12:00-0:	b	0									
18	Result 1	Trial 1	1	1	0:13:00-0:	b	0									
19	Result 1	Trial 1	1	1	0:14:00-0:	b	0,910493									
20	Result 1	Trial 1	1	1	0:15:00-0:	b	1,44161									
21	Result 1	Trial 1	1	1	0:16:00-0:	b	0,531121									
22	Result 1	Trial 1	1	1	0:17:00-0:	b	1,87528									
23	Result 1	Trial 1	1	1	0:18:00-0:	b	0									
24	Result 1	Trial 1	1	1	0:19:00-0:	b	22,6859									
25	Result 1	Trial 1	1	1	0:20:00-0:	b	5,15667									
26	Result 1	Trial 1	1	1	0:21:00-0:	b	5,87252									
27	Result 1	Trial 1	1	1	0:22:00-0:	b	8,56805									
28	Result 1	Trial 1	1	1	0:23:00-0:	b	5,19839									

Rechercher et remplacer

Rechercher Remplacer

Rechercher:

Remplacer par: b

Options >>

Remplacer tout Remplacer Rechercher tout Suivant Fermer

Microsoft Excel

Cette action est terminée, 58000 remplacements ont été effectués.

OK

3.d copy and past the first 4 lines of a file with a genotype in your file, and voila !

Statistics-PSDD2opn4x... - Excel

Fichier Accu Inser Mise Form Don Révis Affic Déve Recherche Connexion

Calibri 11 A⁺ A⁻ G I S % 000 Mise en forme conditionnelle Mettre sous forme de tableau Styles de cellules Insérer Supprimer Trier et Rechercher et filtrer sélectionner Édition

Presse-papiers Police Alignement Nombre Style Cellules Édition

E1

b

	A	B	C	D	E	F	G
1					b	Distance moved	
2					b	center-point	
3					b	Total	
4					b	mm	
5	Result 1	Trial 1	1	1	Start-0:01	b	87,5914
6	Result 1	Trial 1	1	1	0:01:00-0:	b	78,0216
7	Result 1	Trial 1	1	1	0:02:00-0:	b	21,4394
8	Result 1	Trial 1	1	1	0:03:00-0:	b	1,16375
9	Result 1	Trial 1	1	1	0:04:00-0:	b	-
10	Result 1	Trial 1	1	1	0:05:00-0:	b	-
11	Result 1	Trial 1	1	1	0:06:00-0:	b	0,700874
12	Result 1	Trial 1	1	1	0:07:00-0:	b	4,7979
13	Result 1	Trial 1	1	1	0:08:00-0:	b	-
14	Result 1	Trial 1	1	1	0:09:00-0:	b	13,6744
15	Result 1	Trial 1	1	1	0:10:00-0:	b	3,33101
16	Result 1	Trial 1	1	1	0:11:00-0:	b	-

Independent Variable

	A	B	C	D	E	F	G
1					Independent	Distance moved	
2					Independent	center-point	
3					<User-def	Total	
4						mm	
5	Result 1	Trial 1	1	1	Start-0:01	b	87,5914
6	Result 1	Trial 1	1	1	0:01:00-0:	b	78,0216
7	Result 1	Trial 1	1	1	0:02:00-0:	b	21,4394
8	Result 1	Trial 1	1	1	0:03:00-0:	b	1,16375
9	Result 1	Trial 1	1	1	0:04:00-0:	b	-
10	Result 1	Trial 1	1	1	0:05:00-0:	b	-
11	Result 1	Trial 1	1	1	0:06:00-0:	b	0,700874
12	Result 1	Trial 1	1	1	0:07:00-0:	b	4,7979
13	Result 1	Trial 1	1	1	0:08:00-0:	b	-
14	Result 1	Trial 1	1	1	0:09:00-0:	b	13,6744
15	Result 1	Trial 1	1	1	0:10:00-0:	b	3,33101
16	Result 1	Trial 1	1	1	0:11:00-0:	b	-