

From Noldus file to Visualisation part 3 : wakefish

Clair Chaigne

8 novembre 2021

1. INSTRUCTIONS

This is a follow up from From Noldus File to Visualisation part 2 : prepare for wakefish

For this example I used the PSDD2opn4xa experiment (8 movies labelled PSDD2opn4xaF1 to F8).

2. STEP BY STEP EXAMPLE

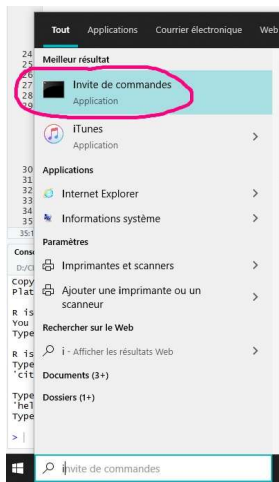
1. For processing to wakefish, your files should all look like this (make sure your Arenas are labelled "Arena 1, Arena 2, ..." and not just "1, 2, ..." and that you have a genotype column). If not check the From Noldus File to Visualisation part 2 : prepare for wakefish

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1					Independence moved									
2					Independence point									
3					<User-def total									
4					mm									
5	Result 1	Trial 1	1	Arena 1	Start-0:01:b	87,5914								
6	Result 1	Trial 1	1	Arena 1	0:01:00-0:b	78,0216								
7	Result 1	Trial 1	1	Arena 1	0:02:00-0:b	21,4394								
8	Result 1	Trial 1	1	Arena 1	0:03:00-0:b	11,6375								

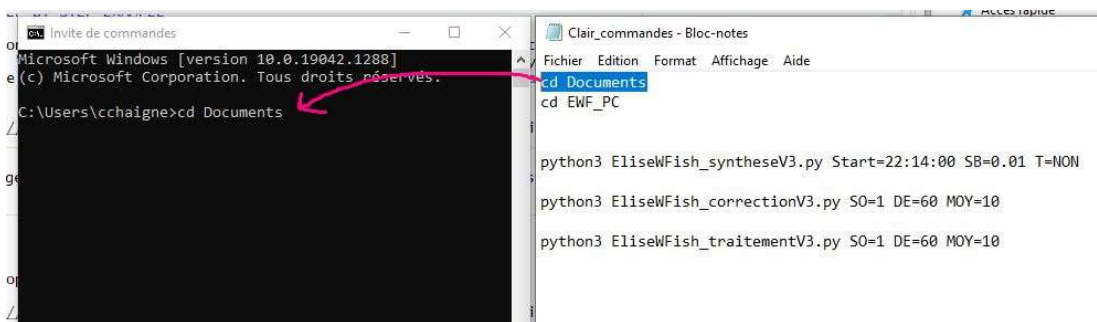
2. copy your files in your EWF_PC//ori folder (be sure this folder is empty before your copy your files).

Nom	Modifié le	Type	Taille
Statistics-PSDD2opn4xaF1	26/10/2020 14:59	Feuille de calcul ...	1 882 Ko
Statistics-PSDD2opn4xaF2	26/10/2020 14:59	Feuille de calcul ...	2 771 Ko
Statistics-PSDD2opn4xaF3	26/10/2020 14:59	Feuille de calcul ...	1 658 Ko
Statistics-PSDD2opn4xaF4	26/10/2020 14:59	Feuille de calcul ...	2 418 Ko
Statistics-PSDD2opn4xaF5	26/10/2020 15:00	Feuille de calcul ...	1 709 Ko
Statistics-PSDD2opn4xaF6	26/10/2020 15:00	Feuille de calcul ...	2 418 Ko
Statistics-PSDD2opn4xaF7	26/10/2020 15:00	Feuille de calcul ...	2 021 Ko
Statistics-PSDD2opn4xaF8	26/10/2020 15:00	Feuille de calcul ...	2 104 Ko

3. open the "terminal"

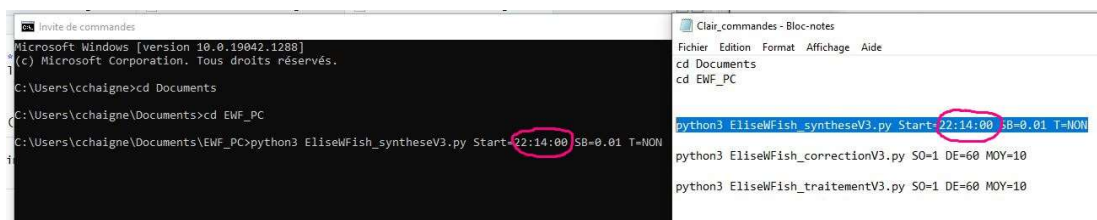


4. copy and paste the coding lignes `cd Documents` and `cd EWF_PC` (or directly `cd Documents\EWF_PC`) in the terminal to go in the correct folder



5. **Synthese part** : copy and paste the synthese part with the correct parameters

`python3 EliseWfish_syntheseV3.py Start=22:14:00 SB=0.01 T=NON`, put the correct Starting Time (**Start=HH:MM:SS**) : 23:00:00 correspond to 8am, beware do not put anything over 23:00:00 ; **SB=** (binarisation threshold in mm to consider that the larva is moving), **T=OUI ou T=NON** to sort out by genotype or not



This part takes 2-3min. For each file, it should say something like this if it's able to detect every column it needs:

```

Invite de commandes - python3 EliseWFish_syntheseV3.py Start=22:14:00 SB=0.01 T=NON
fin de colonne
Partition of a set of 1228285 objects. Total size = 99900361 bytes.
Index Count % Size % Cumulative % Kind (class / dict of class)
0 308200 25 32052800 32 32052800 32 openpyxl.cell.cell.Cell
1 334865 27 19184848 19 51237648 51 tuple
2 212341 17 14946647 15 66184295 66 str
3 1519 0 12285904 12 78470199 79 dict (no owner)
4 277856 23 7787856 8 86258055 86 int
5 34650 3 4019400 4 90277455 90 openpyxl.styles.cell_style.StyleArray
6 8222 1 1466560 1 91744015 92 types.CodeType
7 15955 1 1188938 1 92932953 93 bytes
8 7588 1 1031968 1 93964921 94 function
9 1075 0 949696 1 94914617 95 type
<499 more rows. Type e.g. '._more' to view.>

Analyse de :
C:\Users\cchaig\Documents\EFW_PC\ori\Statistics-PSDD2opn4xaF3.xlsx
Chargement du classeur en memoire...
Classeur charge en memoire
Selection de la feuille
Feuille chargee en memoire
51796
6
Identification des cellules d'interet fichier 2
Trouve 'Distance moved' cellule F1
Trouve '<User-defined 1>' cellule E3
Trouve 'Arena' cellule C5
Trouve 'Start' cellule D5
Mise a jour de la liste des Arena Globale

```

For example if you forgot to put the genotype column, it will look like that and it would work :

```

Invite de commandes
({'Start': '22:14:00', 'SB': '0.01', 'T': 'NON'})

Liste des fichiers a traiter :
['Statistics-PSDD2opn4xaF1.xlsx']
C:\Users\cchaig\Documents\EFW_PC\EliseWFish_syntheseV3.py:187: DeprecationWarning: Call to deprecated function get_sheet_by_name (Use wb[sheetname]).
ws = ClasseurOut.get_sheet_by_name('Sheet')
C:\Users\cchaig\Documents\EFW_PC\EliseWFish_syntheseV3.py:188: DeprecationWarning: Call to deprecated function remove_sheet (Use wb.remove(worksheet) or del wb[sheetname]).
ClasseurOut.remove_sheet(ws)

Analyse de :
C:\Users\cchaig\Documents\EFW_PC\ori\Statistics-PSDD2opn4xaF1.xlsx
Chargement du classeur en memoire...
Classeur charge en memoire
Selection de la feuille
Feuille chargee en memoire
58000
5
Identification des cellules d'interet fichier 0
Trouve 'Distance moved' cellule E1
Trouve 'Arena' cellule C5
Trouve 'Start' cellule D5
Mise a jour de la liste des Arena Globale
Traceback (most recent call last):
  File "C:\Users\cchaig\Documents\EFW_PC\EliseWFish_syntheseV3.py", line 271, in <module>
    CellIdx2= ColonneType+str(i)
NameError: name 'ColonneType' is not defined

```

6. Correction part When the Synthese part is over (you should get a single line saying C:\Users\your_login\Documents\EFW_PC\ori), copy and paste the synthese part with the correct parameters python3 EliseWFish_correctionV3.py SO=1 DE=60 MOY=10 , put the correct threshold time in minutes where you consider that the larva is sleeping (**SO=1** for 1min), binning time in the Noldus file seconds (**DE=60** for distance moved per minute), mean to do (**MOY=10** to have to distance moved by minute over 10min). This should also take 2-3 minutes.

```

Invite de commandes
Derniere ligne avt remplissage : 1
1 day, 8:20:00
duree du process : 1.3477725982666016 sec
duree depuis debut : 10.774916172027588 sec
Sauvegarde du fichier
C:\Users\cchaig\Documents\EFW_PC\tmp\Result_Statistics-PSDD2opn4xaF1163647
6178.31743 Start22-14-00_SB0.01_Tri-False.xlsx
fin de colonne
Partition of a set of 545601 objects. Total size = 51055038 bytes.
Index Count % Size % Cumulative % Kind (class / dict of class)
0 121400 22 12625600 25 12625600 25 openpyxl.cell.cell.Cell
1 118043 22 9753006 19 22378606 44 str
2 148065 27 8724048 17 31102654 61 tuple
3 1519 0 7006160 14 38108814 75 dict (no owner)
4 91090 17 2558548 5 40667362 80 int
5 8222 2 1465466 3 42132828 83 types.CodeType
6 15955 3 1188938 2 43321766 85 bytes
7 7588 1 1031968 2 44353734 87 function
8 1075 0 949696 2 45303430 89 type
9 6614 1 767224 2 46070654 90 openpyxl.styles.cell_style.StyleArray
<499 more rows. Type e.g. '._more' to view.>
duree totale : 13.381743669500888 sec
C:\Users\cchaig\Documents\EFW_PC>python3 EliseWFish_correctionV3.py SO=1 DE=60 MOY=10

```

```

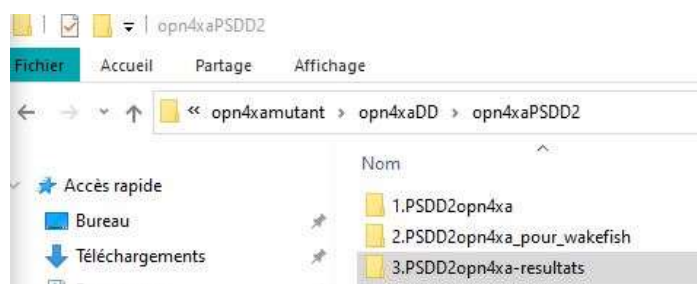
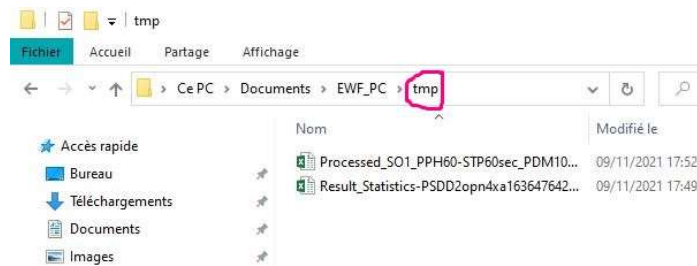
Clair commandes - Bloc-notes
Fichier Edition Format Affichage Aide
cd Documents
cd EFW_PC

python3 EliseWFish_syntheseV3.py Start=22:14:00 SB=0.01 T=NON
python3 EliseWFish_correctionV3.py SO=1 DE=60 MOY=10
python3 EliseWFish_traitementV3.py SO=1 DE=60 MOY=10

```

6. Traitement part when the Correction part is over, copy and paste the traitement part with the correct paramters python3 EliseWFish_traitementV3.py SO=1 DE=60 MOY=10 , ut the correct threshold time in minutes where you consider that the larva is sleeping (**SO=1** for 1min), bining time in the Noldus file seconds (**DE=60** for distance moved per minute), mean to do (**MOY=10** to have to distance moved by minute over 10min). This part can take several hours.

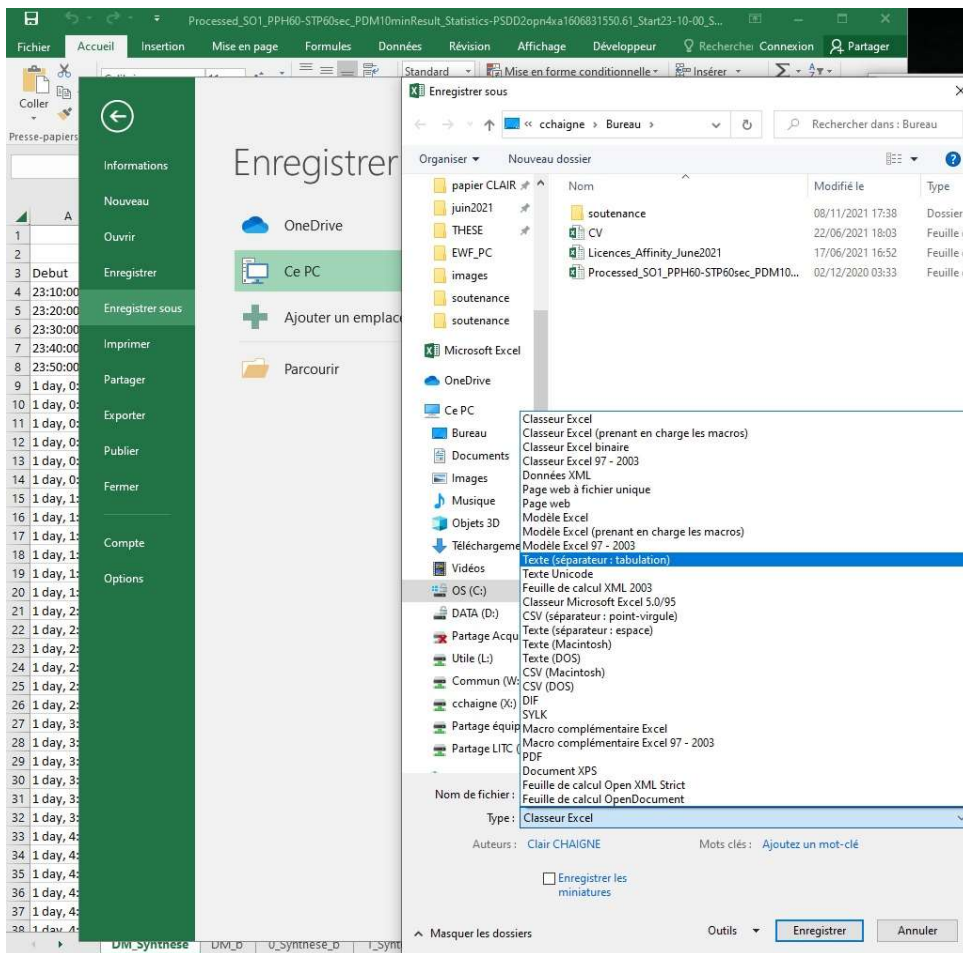
7. when the traitement part is over, I like to copy and paste the Processed and Result files in th e tmp folder in my 3.*nameofmyexperiment*_resultats folder



8. open the Processed file , it should look like this :

	A	B	C	D	E	F	G	H	I	J	
1			Arena_01_b	Arena_02_b	Arena_03_b	Arena_04_b	Arena_05_b	Arena_06_b	Arena_07_b	Arena_08_b	Are
2			DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM
3	Debut	Fin	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moy
4	23:10:00	23:20:00	20,7389324	20,8838	12,3896767	15,3930736	29,7546384	15,2132946	26,016535	9,7946506	22,9
5	23:20:00	23:30:00	3,0775414	5,2123396	6,707683	2,7457238	8,23572	3,5315957	9,324965	9,514283	5,74
6	23:30:00	23:40:00	10,697048	6,1175173	9,7817669	9,171691	6,386368	7,1887406	10,222173	41,820196	9,00
7	23:40:00	23:50:00	9,6918123	13,049718	6,450029	10,060099	19,343485	9,089252	9,5005	9,3142853	7,03
8	23:50:00	1 day, 0:00:00	4,743328	11,173036	9,806365	6,5600788	26,210091	10,409376	11,009551	3,0767981	5,65
9	1 day, 0:00:00	1 day, 0:10:00	8,701932	10,961208	18,819223	6,635242	5,279905	17,299382	18,615756	6,311505	14,0
10	1 day, 0:10:00	1 day, 0:20:00	8,9258116	13,45065	6,6997821	9,798785	2,4112136	9,6462099	45,81597	4,2847948	13,5
11	1 day, 0:20:00	1 day, 0:30:00	9,1973224	18,913278	2,597509	13,021831	32,338403	8,7278452	1,88145	11,039979	13,6
12	1 day, 0:30:00	1 day, 0:40:00	8,391229	9,9544947	6,917111	15,675043	41,42234	7,3658726	15,942835	10,6953	19,2
13	1 day, 0:40:00	1 day, 0:50:00	15,862407	15,991271	10,559773	11,4145736	51,26011	12,166603	61,44782	21,43824	3,50

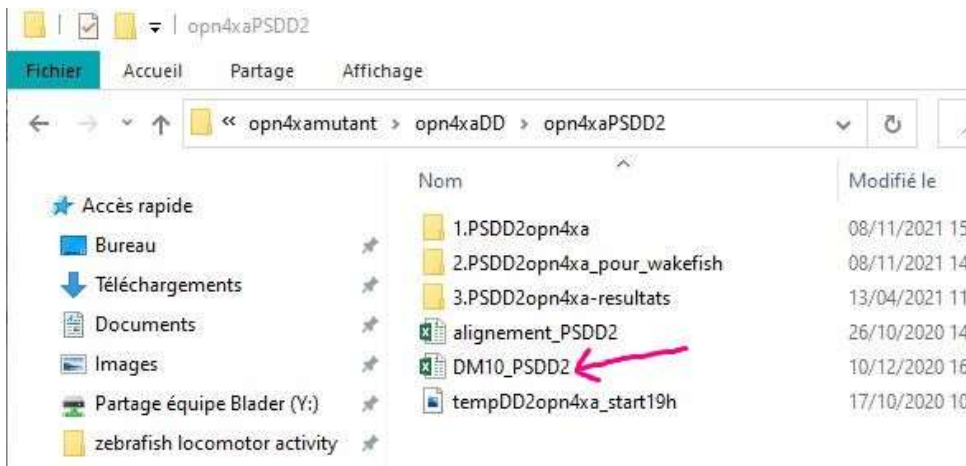
9. copy and paste the DM_Synthese sheet in another excel file (CTRL+A to select everything, CTRL+C to copy and CTRL+V to paste). Save this new file as a Texte (separateur : tabulation) file.



10. Open this new file in Excel, it should now look like this :

	A	B	C	D	E	F	G	H	I	J
1			Arena_01_b	Arena_02_b	Arena_03_b	Arena_04_b	Arena_05_b	Arena_06_b	Arena_07_b	Arena_08_b
2			DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn	DM moyenn
3	Debut	Fin	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne
4	23:10:00	23:20:00	20,7389324	20,8838	12,3896767	15,3930736	29,7546384	15,2132946	26,016535	9,7946506
5	23:20:00	23:30:00	3,0775414	5,2123396	6,707683	2,7457238	8,23572	3,5315957	9,324965	9,514283
6	23:30:00	23:40:00	10,697048	6,1175173	9,7817669	9,171691	6,386368	7,1887406	10,222173	41,820196
7	23:40:00	23:50:00	9,6918123	13,049718	6,450029	10,060099	19,343485	9,089252	9,5005	9,3142853
8	23:50:00	1 day, 0:00:00	4,743328	11,173036	9,806365	6,5600788	26,210091	10,409376	11,009551	3,0767981
9	1 day, 0:00:00	1 day, 0:10:00	8,701932	10,961208	18,819223	6,635242	5,279905	17,299382	18,615756	6,311505
10	1 day, 0:10:00	1 day, 0:20:00	8,9258116	13,45065	6,6997821	9,798785	2,4112136	9,6462099	45,81597	4,2847948
11	1 day, 0:20:00	1 day, 0:30:00	9,1973224	18,913278	2,597509	13,021831	32,338403	8,7278452	1,88145	11,039979
12	1 day, 0:30:00	1 day, 0:40:00	8,391229	9,9544947	6,917111	15,675043	41,42234	7,3658726	15,942835	10,6953
13	1 day, 0:40:00	1 day, 0:50:00	15,862407	15,991271	10,559773	11,4145736	51,26011	12,166603	61,44782	21,43824
14	1 day, 0:50:00	1 day, 1:00:00	12,232086	15,49647	39,812288	15,465826	45,54479	19,474097	60,79685	66,44069
15	1 day, 1:00:00	1 day, 1:10:00	6,2611355	12,228721	27,1519	16,743718	38,94404	13,208473	52,41269	77,27603

11. Save this file as an excel file (classeur Excel). I like to save name this file DM10_nameofmyexperiment for 10min means.



3. Issues you can have

- Do not put a starting time over 23:00:00
- If you have too many different genotypes, the wakefish won't be able to generate the final excel files (too many sheets to generate)
- Check that there is not any very very small data (the wakefish program should then give an error message with the line number where this very very small data is such as $e^{-0.006}$), replace it with 0
- If when you try to open the processed file, it says the name is too long to be opened, you can trim its name and/or try to copy it on your desk and open it from it.