

Manual for RTRIM-shell_2.0

This manual is intended to guide RTRIM-shell_2.0 users through the index calculation process.

If interested in detailed description of RTRIM-shell inputs and outputs, see the RTRIM-shell tutorial at the PECBMS webpage <https://pecbms.info/methods/software/trim/>, please. Although the tutorial was created as a guide through an older RTRIM-shell version, the description of RTRIM-shell inputs and outputs is still valid.

RTRIM-shell_2.0 is a version for 2024 Update created by Javier Rivas Salvador. Main intention is to make the tool as much user friendly as possible. On the other hand, the national coordinators experienced with R, who need to modify the program, have an opportunity in the script 'RTRIM-shell_2.0\01_Scripts\DO_NOT_MODIFY\ **x3_functions_Shell_Rtrim.R**'

THERE ARE NO CHANGES IN THE CALCULATION METHODS BETWEEN VERSIONS RTRIM-shell_v1.7 AND RTRIM-shell_2.0

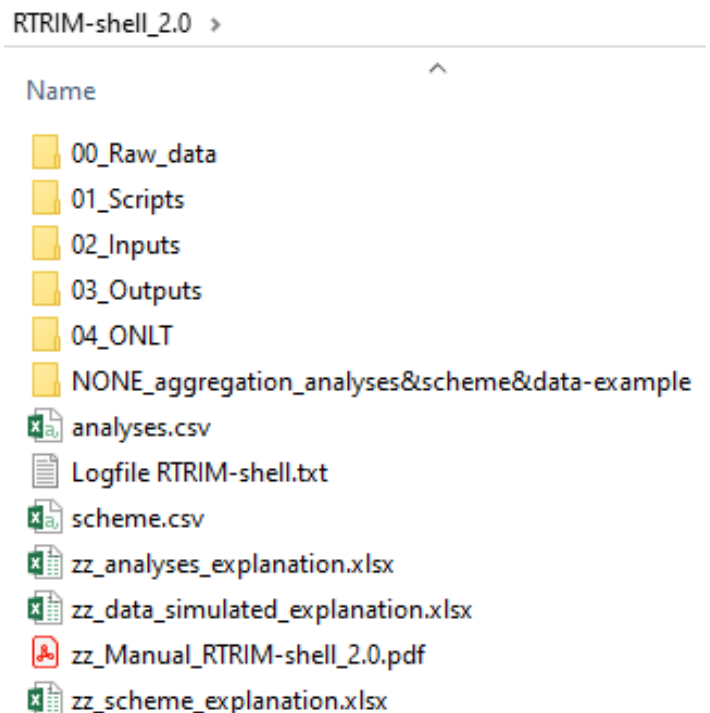
There are big changes in design. All the previously used scripts (for process from RTRIM-shell input preparation to zipping files required for On-line tool upload (ONLT)) are now merged into one script under the name **x3_functions_Shell_Rtrim.R**.

User has only to change 2 tables defining setting of calculation in RTRIM-shell (analyses.csv & scheme.csv) and modify path in the script RTRIM-shell_2.0.R.

RTRIM-shell_2.0 contains example data, which can be easily modified to your own data and outputs.

Figure 1: RTRIM-shell_2.0 contains:

- 5 folders (00_ - 04_) used for whole process from data preparation to zipping files requested for upload into the On-line tool
- folder "NONE_aggregation_analyses&scheme&data-example" with sources for aggregation NONE,
- files "analyses.csv" & "scheme.csv" for RTRIM-shell settings
- "Logfile RTRIM-shell.txt" documenting whole process from data preparation to zipping files requested for upload into On-line tool
- files marked by "zz_" explaining tables used in RTRIM-shell_2.0 and this Manual.



Tables used for index calculation in RTRIM-shell_2.0

Both tables have to be saved in the main RTRIM-shell folder.

analyses.csv

- Source table with the information for RTRIM-shell. National coordinator fills this table with information on her/his country, input dataset and formatting of the files.
- *Figure 2: Table “zz_analyses_explanation.xlsx” with explanation (orange colour). This table can be also found in the RTRIM-shell_2.0 folder. The data in red frame has to be filled in the table called “analyses.csv”.*

zz_analyses_explanation.xlsx - Excel			
	A	B	C
1			Explanation
2	country	Czechia	Enter name of your country.
3	country_code	4	Enter your country code used in PECBMS. If you wish to upload data in the PECBMS On-line tool, you must keep country code set by PECBMS. Other codes can be used only for private index calculation.
4	changepoints	all	We recommend to enter "all", so all the years will be automatically selected as changepoints. "all" works also for the situation of having some years with completely missing counts (e.g. due to foot&mouth disease or covid restrictions) because RTRIM-shell automatically recognizes and discards such years from the changepoints. However, if you wish to discard some years for your own reason, you must list all the years as numbers separated by dashes and omit the numbers belonging to the years to be discarded (e.g. if monitoring lasted 8 years and years 3 & 6 could not be monitored, write: 1-2-4-5-7). The last year of monitoring is NOT entered in any case.
5	data_file_name	data_simulated.csv	Name of your datafile including extension (.csv => only .csv files are used as input for RTRIM-shell).
6	monitoring_prefix	BMP	RTRIM-shell automatically creates prefix BMP during calculations. Enter your own prefix in case you use a different one.
7	decimal	.	Enter decimal used in your computer. => If comma "," is used, write comma here. However, the RTRIM-shell outputs will be saved with decimal dot "." used in PECBMS => Do not open the .zip folder, because your computer might change decimals.
8	separator	;	Enter separator of columns used in your computer. => If comma "," is used, write comma here. However, the outputs will be saved with semicolon ";" used in PECBMS => Do not open the .zip folder, because your computer might change separator.

scheme.csv

- Source of information for Monitoring_scheme.csv table and RTRIM-shell calculation.
- It is extremely important to fill the table precisely and double-check!!!
- Species_nr: can be written in 2 ways:
 - “all”: this row will contain info for the major part of species
 - “60-70”: species 60 AND 70 are treated as an exemption for the general rule and RTRIM-shell will treat them according to the information in their row. List all the species with dashes. It is possible to add various exemptions by adding rows.
- *Figure 3: Table “zz_scheme_explanation.xlsx” with explanation (orange colour). This table can be also found in the RTRIM-shell_2.0 folder. The data in red frame has to be filled in the table called “scheme.csv”. Here is the table divided into 3 parts for the Manual purposes. 1. Part*

zz_scheme_explanation.xlsx - Excel - 1. part						
	A	B	C	D	E	F
	If the calculation setting is used for all species, write "all". If you use different setting for a species, enter the species Euring code and complete the setting. If the other setting is valid for more species, all species must be listed. Use a dash (-) as a divider between species Euring codes (e.g. 70-90-5820).		Enter a general first year of monitoring scheme in this species. Only informative function (real data can start later, which is automatically identified by RTRIM-shell).	Enter number of monitoring visits per breeding season.	Specify if you use point count, line transect or territory mapping: 1=point count, 2=line transect, 3=territory mapping.	If Point count is used, enter number of points used per square or transect. Otherwise enter NA.
1						
2						
3	Species_nr	Monitoring_scheme	Start_year	Number_visits_spring	Census_method	Number_of_points_per_square_or_transect
4	all	ABC	2018	2	2	NA
5	60-70	EFG	2020	2	1	
6	11870-12000-14640	KLM	2019	2	2	NA
7	15910	OPQ	2021	2	1	
8	10990	XYZ	1982	2	3	NA

Table scheme.csv divided in 3 parts with added explanations: 2. Part:

zz_scheme_explanation.xlsx - Excel - 2. Part						
	G	H	I	J	K	L
	<p>Base year = 100%. Set Base_year_last_year to Set the year from which the first year of the a short term slope monitoring program. (trend) shall be calculated. This is for your private use. Usually last 10 years are used.</p>					
	If Line transect is used, enter its length in metres.	If Territory mapping is used, enter the area surveyed in hectares.	Specify if you use an individual, pair or territory as Count unit: 1=1 individual, 2=pair, 3=territory.	Base year = 100%. Set Base_year_first_year to the first year of the monitoring program.	Otherwise a period instead of 1 year will be used for setting Base year.	
1	Otherwise enter "NA".	Otherwise fill in "NA".				
2						
3	Transect_length	Mapping_area	Count_unit	Base_year_first_year	Base_year_last_year	year_from
4	1000	NA	2	2015	2015	2016
5	NA	NA	1	2020	2020	2021
6	1000	NA	1	1982	1985	1990
7	NA	NA	1	2000	2000	2010
8	NA		2	1982	1982	2014

Table scheme.csv divided in 3 parts with added explanations: 3. Part

zz_scheme_explanation.xlsx - Excel - 3. Part					
	M	N	O	P	Q
	<p>Specify if you set Serial correlation ON (write TRUE) or not (write FALSE). RTRIM-shell uses this column for its calculations. Serial correlation is recommended to be TRUE, because the monitoring data is correlated.</p>				
	If sites are visited more than once in the season, enter how you aggregate the data annually, e.g. MAX=maximum count, AVE_WEIGHTS or AVE_DIRECT=average count, FIRST=count in first visit from more visits, LAST=count in late visit from more visits. Otherwise use the MAX aggregation.	Enter type of species recorded in the census. 1=all species, 2=only breeders, 3=subset of breeding species.	Specify if you monitor 1=all the localities every year, 0= some localities monitored one year and rest in other year, i.e. each locality is monitored e.g. 1*/2 years.	Specify if you set Overdispersion ON (write TRUE) or not (write FALSE). RTRIM-shell uses this column for its calculations. Overdispersion is recommended to be TRUE because overdispersion appears very often in the monitoring data.	
1					
2					
3	Count_aggregation	Recorded_species	All_localities_used_every_year	Serial_correlation	Overdispersion
4	MAX	1	1	TRUE	TRUE
5	AVE_WEIGHTS	1	1	TRUE	TRUE
6	AVE_DIRECT	1	1	FALSE	TRUE
7	FIRST	1	1	TRUE	FALSE
8	LAST	1	1	TRUE	TRUE

Table 1: Detailed explanation of the Aggregations set in the table scheme.csv:

Type of Aggregation	General description	Use of Weights	Aggregation details
MAX	Selects the maximum count for each location/year combination.	No	Visit with maximal number of birds/pairs in the breeding season is used for index calculation
AVE_DIRECT	Provides mean number of observed counts among all the visits for each location/year combination.	No	Average number of birds detected in a locality is calculated already during data preparation and used for index calculation. Index calculated by using aggregation AVE_DIRECT differs slightly from index gained by AVE_WEIGHTS.
AVE_WEIGHTS	Calculates the sum of counts for each location/year combination, which is used during the index calculation.	Yes	Weight is used in the form $1/\text{number of visits per year}$. Sum per location/year is multiplied by weights during the calculation process. Index differs slightly from index gained by AVE_DIRECT. AWE_WEIGHTS corresponds with method used in TRIM/BirdStats.
FIRST	Selects the earliest visit for each location/year combination.	No	Data from first visit in the breeding season is used.
LAST	Selects the latest visit for each location/year combination.	No	Data from last visit in the breeding season is used. This aggregation is often used in the South Europe for some species because of migration during first visit. Use of last visit guarantees monitoring of birds breeding in a locality.
NONE	Intended for skilled coordinators who have their own scripts for aggregations/weights.	Depends on monitoring scheme.	No weights are used and the table scheme.csv must be prepared precisely for the needs of your monitoring scheme. Let's contact Javier Rivas <rivassalvador@birdlife.cz> if you are interested in this aggregation.

Folders used for index calculation in RTRIM-shell_2.0

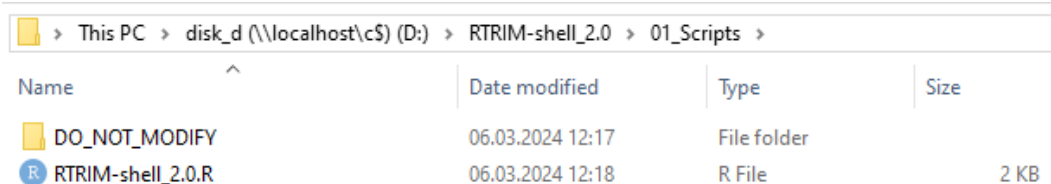
00_Raw_data

- Contains table with all raw data. RTRIM-shell_2.0 script is able to pick proper data according to the option stated in table scheme.csv.
- Column "date" is date of visit. It is only necessary to add this column for the aggregation FIRST/LAST visit. This column does not have to be added for other aggregations.
- We try to reduce the amount of pre-processing that some coordinators were conducting before running RTRIM. Therefore, the program itself will select the necessary data for the calculation according to the aggregation methods implemented in RTRIM. Let's just set aggregation method in the table "scheme.csv".
- *Figure 4: Data table serving as source of data for RTRIM-shell. Enter only zeros and positive data, RTRIM-shell will add NAs automatically. In case of a year without monitoring (due to e.g. Covid restrictions) do NOT enter this year at all. RTRIM-shell will automatically add this year with NAs and later impute the values.*

zz_data_simulated_explanation.xlsx - Excel						
	A	B	C	D	E	F
1	Latin species name	Euring code (see Species list in the On-line tool, "Taxonomy"	Site code used in your monitoring system. PECBMS does not receive this information. Combination of numbers and letters are permitted.	Count used in your system. It is important to be interannually consistent in using individual, pair or territory. If the data is not consistent, RTRIM-shell will calculate different values than in former years. Only zero and positive numbers can appear (RTRIM-shell will add NAs automatically).	Year of monitoring. Each visit in each location has its own raw, so RTRIM-shell can calculate maximum or average.	Date in any format. This column is needed only if Aggregation "FIRST" OR "LAST" is used. Otherwise is this column redundant. RTRIM-shell will work properly without date.
2	species	euring	site	count	year	date
3	Tachybaptus ruficollis	70	16	0	2020	11.05.2020
4	Tachybaptus ruficollis	70	16	0	2021	07.05.2021
5	Tachybaptus ruficollis	70	16	0	2023	03.05.2023
6	Tachybaptus ruficollis	70	16	0	2023	12.06.2023
7	Tachybaptus ruficollis	70	214	1	2021	25.05.2021
8	Tachybaptus ruficollis	70	298	0	2021	03.06.2021
9	Tachybaptus ruficollis	70	306	0	2021	18.04.2021
10	Tachybaptus ruficollis	70	426	3	2021	06.05.2021
11	Tachybaptus ruficollis	70	438	0	2022	06.05.2022
12	Tachybaptus ruficollis	70	442	0	2020	13.06.2020

01_Scripts

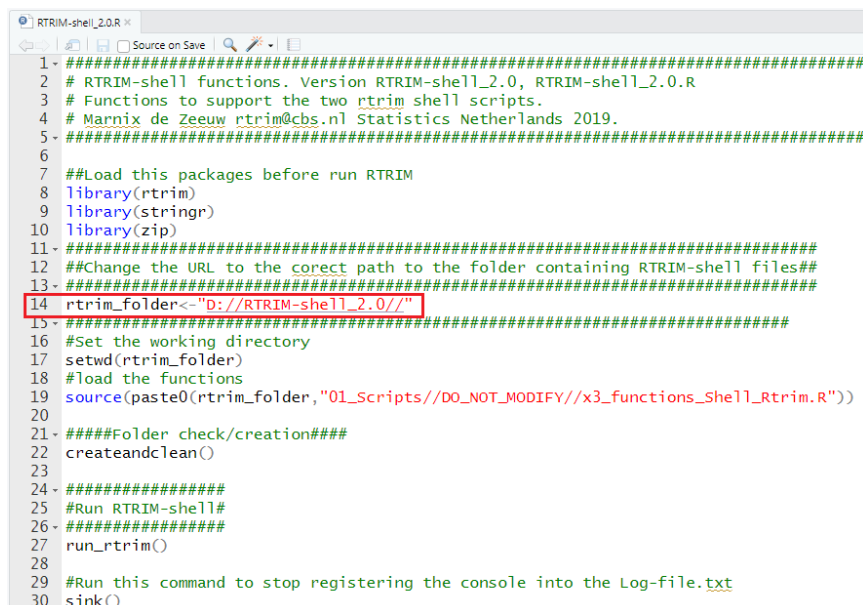
- This folder contains a script, which has to be run to calculate indices and then it contains a folder called DO_NOT_MODIFY
- *Figure 5: Folder 01_Scripts contains script RTRIM-shell_2.0, which automatically runs the script with functions and folder DO_NOT_MODIFY, which includes script with functions and Species list.*



Name	Date modified	Type	Size
DO_NOT_MODIFY	06.03.2024 12:17	File folder	
RTRIM-shell_2.0.R	06.03.2024 12:18	R File	2 KB

Script RTRIM-shell_2.0.R

- This is the script, which has to be run by you and which will automatically run the second script with functions.
- The whole RTRIM-shell calculation is automated and the only task that has to be done is to exchange path marked by red frame in the picture for the path in your own computer. Then let's run the script.
- "Createandclean()" function will create the folders 02,03 and 04 if they do not exist. If the folder RTRIM-shell_2.0 already contains these folders, their content will be deleted prior a new rtrim run.
- *Figure 6: Script RTRIM-shell_2.0; path to be modified to the path in your computer is highlighted in red frame.*



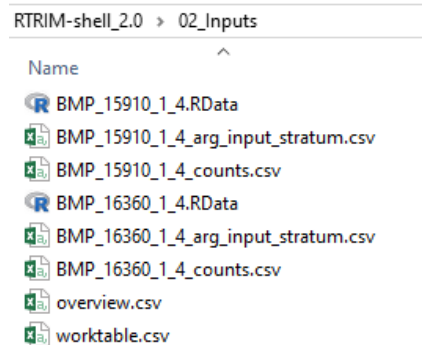
```
1 #####
2 # RTRIM-shell functions. Version RTRIM-shell_2.0, RTRIM-shell_2.0.R
3 # Functions to support the two rtrim shell scripts.
4 # Marnix de Zeeuw rtrim@cbs.nl Statistics Netherlands 2019.
5 #####
6
7 ##Load this packages before run RTRIM
8 library(rtrim)
9 library(stringr)
10 library(zip)
11 #####
12 ##Change the URL to the correct path to the folder containing RTRIM-shell files##
13 #####
14 rtrim_folder<-"D://RTRIM-shell_2.0//"
15 #####
16 #Set the working directory
17 setwd(rtrim_folder)
18 #load the functions
19 source(paste0(rtrim_folder,"01_Scripts//DO_NOT_MODIFY//x3_functions_Shell_Rtrim.R"))
20
21 #####Folder check/creation###
22 createandclean()
23
24 #####
25 #Run RTRIM-shell#
26 #####
27 run_rtrim()
28
29 #Run this command to stop registering the console into the Log-file.txt
30 sink()
```

Folder DO_NOT_MODIFY

- Species_list.csv
- This file contains a PECBMS list of all European species. The list is used also for On-line tool (left tool bar/ Taxonomy). It is recommended to check if your euring codes match with the PECBMS ones before the calculation.
- x3_functions_Shell_Rtrim.R
- Original scripts 01_shell_rtrim_strata.r& 02_processingOutputSpeciesStratumCombination_SumTab.R have been turned into functions and are now part of this script.
- If you intend to do your own modifications, let's inform Eva Šílarová silarova@birdlife.cz and Javier Rivas rivasalvador@birdlife.cz, please.

02_Inputs

- This folder is automatically created before index calculations (if not present in RTRIM-shell_2.0) or emptied (if present from former RTRIM-shell runs) and will be filled by the RTRIM-shell inputs when you run the script RTRIM-shell_2.0.R
- *Figure 7: PrintScreen with example of 02_Inputs filled with the RTRIM-shell inputs from which indices will be calculated. RTRIM-shell fills the folder “02_Inputs” with _arg_input_stratum.csv, _counts.csv, overview.csv and worktable.csv. For detailed description of these files watch RTRIM-shell tutorial at the PECBMS. Webpage <https://pecbms.info/methods/software/trim/>.*



- Let's check the file **overview.csv** after each run (see example below). Overview.csv shows all species and announces successful or unsuccessful index calculations (column “success”). RTRIM-shell does not calculate species with very poor data. If the calculation was unsuccessful, read details in the columns “error”. Check the counts.csv table of an unsuccessful species and assess, if this species is needed for calculations. If you need to calculate index anyway, it is possible to truncate the time series to series of years with positive values (if there is e.g. only one positive value in 10 years of NAs, let's delete these 10 years from the table with raw data and run the calculations again). If index is calculated, use your expert knowledge to assess, if the Time Totals and Indices are real. If the calculation was successful in attempt_4, RTRIM-shell used the simplest model, which may be unrealistic. Therefore, we recommend to decide, if such indices are reliable. If such a species is uploaded into On-line tool, let's comment it in the bookmark Data controls/Specific comments window.
- *Figure 8: Table “overview.csv” enables to see what species index was successfully calculated (column “success”) and also what model was used for successful calculation (columns “attempt_1-4”). See column “error1-4” acknowledge potential problem. Here is the table divided into 2 parts for the Manual purposes. Table “overview.csv” divided in 3 parts with added explanations: 1. Part and 2. Part*

	A	B	C	D	E	F	G	H
1	ss_combinations	species_group	species_number	stratum_type	stratum_number	first_year	last_year	success
2	BMP_60_1_4	BMP	60	1	4	2020	2023	yes
3	BMP_70_1_4	BMP	70	1	4	2020	2021	no
4	BMP_6680_1_4	BMP	6680	1	4	2015	2023	yes
5	BMP_8760_1_4	BMP	8760	1	4	2015	2023	yes
6	BMP_10200_1_4	BMP	10200	1	4	2015	2023	yes
7	BMP_10990_1_4	BMP	10990	1	4	1982	2023	yes
8	BMP_11870_1_4	BMP	11870	1	4	1982	2023	yes
9	BMP_12000_1_4	BMP	12000	1	4	1982	2023	yes
10	BMP_14640_1_4	BMP	14640	1	4	1982	2023	yes
11	BMP_15910_1_4	BMP	15910	1	4	2000	2023	yes
12	BMP_16360_1_4	BMP	16360	1	4	2015	2023	yes

	I	J	K	L	M	N	O	P	Q
1	attempt_1	attempt_2	attempt_3	attempt_4	error_1	error_2	error_3	error_4	
2	error	success	n.a.	n.a.	Serial correlation is negative	n.a.	n.a.	n.a.	
3	error	error	error	error	Removed 6 sites without pos	Removed 6	Removed 6	Removed 6 sites	
4	error	success	n.a.	n.a.	Serial correlation is negative	n.a.	n.a.	n.a.	
5	error	success	n.a.	n.a.	Serial correlation is negative	n.a.	n.a.	n.a.	
6	error	success	n.a.	n.a.	Serial correlation is negative	n.a.	n.a.	n.a.	
7	error	error	error	success	Zero expected value at year	Zero	Serial corri	n.a.	
8	success	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	
9	success	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	
10	success	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	
11	error	success	n.a.	n.a.	Serial correlation is very low	n.a.	n.a.	n.a.	
12	error	success	n.a.	n.a.	Serial correlation is negative	n.a.	n.a.	n.a.	

LogFile

- More detailed information on the preparation and calculation processes can be retrieved from the LogFile, which is automatically created and saved in the RTRIM-shell_2.0 folder (see Figure 1).
- This file is new in this RTRIM-shell version. It documents whole process from RTRIM-shell input preparation to zipping files required for On-line tool upload.
- It is easy to find out, where did the calculation stop (in case of a problem) and what is cause of the problem.

Figure 9: LogFile shows in detail all the processes running in RTRIM-shell. First part of LogFile with summary of NAs, Zero values and Counts different from zero in species 6680.

```
#####
Tue Mar 12 11:09:51 2024

Start TRIM-shell v 2.0 in R

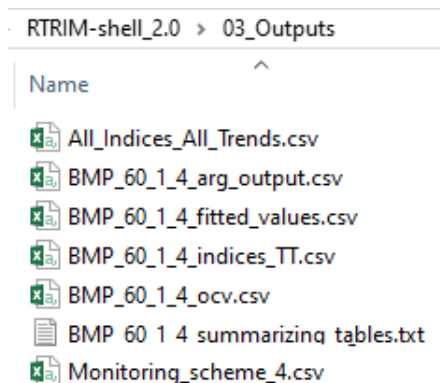
Analyses start for Czechia

RTRIM will run for:
all species with MAX aggregation
60-70 species with AVE_WEIGHTS aggregation
11870-12000-14640 species with AVE_DIRECT aggregation
15910 species with FIRST aggregation
10990 species with LAST aggregation

#####
#Data preparation function:
#####
The total number of locations is: 351
Worktable was successfully created
BMP_6680_1_4_counts.csv was successfully created and contains:
NAs: 49
Zero values: 47
Counts different from zero: 21
```

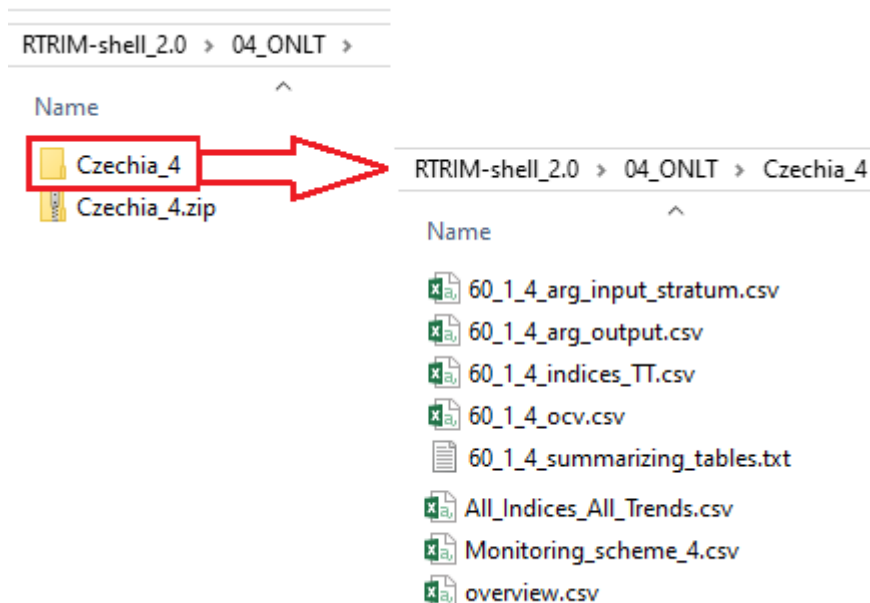
03_Outputs

- This folder is automatically created before index calculations (if not present in RTRIM-shell_2.0) or emptied (if present from former RTRIM-shell runs) and will be filled by the RTRIM-shell outputs when you run the script RTRIM-shell_2.0.R
- Table Monitoring_table_{country_code}.csv is automatically created by the program and saved here. Sometimes table Monitoring_table_{country_code}.csv is not visible in the Windows Commander, but Total commander always shows this table.
- Figure 10: RTRIM-shell fills the folder "03_Outputs" with files All_Indices_All_Trends.csv, _arg_output_stratum.csv, _fitted_values.csv, _indices_TT.csv, _ocv.csv, _summarizing_tables.txt and Monitoring_scheme_{country-code}.csv tables. For detailed description of these files watch RTRIM-shell tutorial at the PECBMS webpage <https://pecbms.info/methods/software/trim/>.



04_ONLT

- This folder is automatically created (if not present in RTRIM-shell_2.0) or emptied (if present from former RTRIM-shell runs) before index calculations.
- RTRIM-shell automatically copies the files from the folders 02_Inputs and 03_Outputs and zips them for the On-line tool (ONLT). The zipped folder is ready to be uploaded to the On-line tool. => Check the unzipped folder but DON'T TOUCH the .zip folder "country_{country-code}"
- RTRIM-shell selects only files requested by the ONLT => PECBMS does not require _counts.csv and _fitted_values.csv. These tables are under your private ownership.
- *Figure 11: RTRIM-shell automatically copies the files from the folders 02_Inputs and 03_Outputs. The zipped folder is ready to be uploaded to the On-line tool. => Check the unzipped folder but DON'T TOUCH the .zip folder "country_{country-code}", here "Czechia_4.zip".*



RTRIM-shell_2.0 Step-by-Step

1. Create table containing your own data and save it in the folder 00_Raw_data
 - This data will be used for creating RTRIM-shell inputs and further index and Time Total calculations
 - Remember to add the "date" column if you intend to use aggregation methods "FIRST" or "LAST"
2. Fill the table analyses.csv with information on your dataset
3. Fill the table scheme.csv with information on your dataset
4. Open a script RTRIM-shell_2.0.R in the folder '01_Scripts' and change path to the main folder with RTRIM-shell on line #14.
5. Run the script RTRIM-shell_2.0