**Documentation MITERRA-Portugal**

# 1. Get\_data

The main goal of this module is for the user to select all the activity data required for each step of the calculations. All the data is being obtained from the folder “Activity data”, which contains the data for gaseous N emissions (N-NH3, N-N2O, N-NOx), runoff, leaching and GIS data.

Function *define\_main\_path*

As the name indicates, this defines the main path as specified by the user. If not specified, it uses my default activity data main folder.

## all\_datamodules

This function prints all the subfiles/subfolders within a specific folder.

all\_datamodules(path)

# Activity data # Documentation.docx # Get\_data.R # GNB\_module

## 1.2 check\_folder

This checks the path as a string and performs a check to the last two indexes (i.e. “\\”) and returns it corrected.

check\_folder(“C:\\Users\\Serra”) # “C:\\Users\\Serra\\”

## 1.3 read\_folder

This function reads all the main activity data subfolder (N-NH3, N-N2O, etc) and allows the user to specific a submodule through the “data\_pattern” variable. This variable receives as an input a specific pattern, which should focus any of the environmental losses already identified.

Example:

nh3 <- read\_folder(main\_path, “NH3”)

## 1.4 list\_subfolders

This prints all the subfolders within the specified activity data main folder for any environmental losses.

Example:

list\_subfolders(nh3)

# Application #Manure\_housing #Manure\_storage

## 1.5 loop\_files

This function is useful to loop and print any existing spreadsheet within a specified folder. The main goal of this function was to be applied in the following function “1.4 read\_subfolders” to allow the user to easily visualize the existing spreadsheets.

Example:

loop\_files(folder\_to\_loop)

## 1.6 nh3\_read\_subfolders

As the name implies, this function reads the subfolders of any specific environmental loss and allows the user to specify any step of the lifecycle of environmental losses. Therefore, this is also closely linked to manure management systems, particularly housing, storage and application.

Example:

read\_subfolders(nh3, “Application”)

## 1.7 print\_return\_datafiles

This function can either print the files within a subfolder or it can return a vector with the filenames, which can then be used to read the files.

Example:

Print\_return\_datafiles(path, print) # ‘lol.csv’

a <- Print\_return\_dafiles(path) # a[1] #’lol.csv’

## 1.8 collate\_sub\_year

This receives the year as a parameter (in numeric form only!!!) and returns the year identified. The identifier concerns the last two numbers as strings, i.e. 99 or 09.

Example:

collate\_sub\_year(2009)

# “09”

## 1.9 data\_by\_year

This function disaggregates the baseline dataset per specified year. There are 3 different parameters, but only the first 2 are mandatory. The primary parameter concerns function “1.4 read\_subfolders”, followed by the year (in numeric; which then calls function 1.6). The remaining parameter is optional: if left unchecked, the function simply returns the required dataset; otherwise, if the parameter equals “**print**”, it **ONLY** prints the dataset for that year

Example:

data\_by\_year(read\_subfolder(nh3, “Application”), 1999)

# "fertiliser99.csv" "Grazing99.csv" "manure99.csv" "sludge99.csv"

## 1.10 read\_lists

This is used to read vectors stored in lists. This is particularly useful to read the GNB files, where a list stores the inputs and outputs

## 1.11 subset\_data

This subsets a specific file based on either column numbers or column names. The main goal of this function is to be used in the merge\_data.

Example:

subset\_data(datafile, c(1,2,3, 5,6)) or subset\_data(datafile, c(‘GIS\_ID’, ‘ID’, ‘Muni’, ‘GNB’)