Package: remulator (via r-universe)

August 24, 2024

Type Package

Title R emulator

Version 1.22.0

Date 2023-10-19

Description A collection of R tools for fitting model results.

License LGPL-3 | file LICENSE

Depends R(>= 2.10.0)

Imports gms, ggplot2, luplot, lusweave, madrat, magclass(>= 4.40), magpie4,

URL https://github.com/pik-piam/remulator, https://doi.org/10.5281/zenodo.2546517

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Suggests covr, knitr, rmarkdown

VignetteBuilder knitr

Repository https://pik-piam.r-universe.dev

RemoteUrl https://github.com/pik-piam/remulator

RemoteRef HEAD

RemoteSha 43f6734ee27d27ebbb223e149112d2da0b9e8888

Contents

---

bisect   *Use bisection to find x so that myform(param, x) is close to approx_this*

Description

Use bisection to find x so that myform(param, x) is close to approx_this

Usage

```
bisect(param, myform, approx_this, lower, upper, eps)
```

Arguments

param	Parameters applied in `myform`
myform	User defined function used in approximation
approx_this	y-value for which x-value should be approximated
lower	Lower limit of interval to search in
upper	Upper limit of interval to search in
eps	Maximal distance of x to approx_this

Value

x-value that is within the eps distance to approx_this

Author(s)

David Klein

---

calculate_fit | *Calculate fit by minimizing sum of error squares*

Description

This is an internal function of the remulator package. It fits a curve to given data. For each region and year it minimizes the function sum((y − (a + b * x ^d))^2 x and y are the data provided by the user, and a, b, and d are the coefficients to be found. So, the fit is y = a+b*x^d.

Usage

```
calculate_fit(data, initial_values = c(1, 1, 1), form, ...)
```

Arguments

data	MAgPIE object with data to fit.
initial_values	Initial values for the parameters to be optimized over.
form	Fit function for which least squares will be calaculated using the optim function.
...	Arguments passed on to the optim function. Useful to define bounds on fit coefficients.

Value

MAgPIE object with fit coefficients a, b, and d

Author(s)

David Klein

---

calc_supplycurve | *Calculate points of fitted curve for plotting using fit coefficients*

Description

This function uses the raw data on which you performed the fit (calculate_fit) and the resulting fit coefficients to calculate a couple of points on the fitted curve that can be used to plot the curve (plot_curve).

Usage

```
calc_supplycurve(data_in, fitcoef, myform, ylimit = "common")
```

Arguments

`data_in`	MAgPIE object containing the same data that was used to calculate the fitcoefficients.
`fitcoef`	MAgPIE object containing the fitcoefficients (output of [`calculate_fit`](#))
`myform`	Function that was fitted and is used here to calculate curve
`ylimit`	Choose the method to calculate the upper limit for the y values. Options: "individual", the maximal value up to which the y values of the supplycurve are calculated is done individualy for each reagion and year. "common" (default), the upper limit of the y values of the supplycurve for all regions and years is the same and is the maximum of the raw data. This is useful for plotting, because it expands "short" supplycurves beyond the maximal demand of the respective region and year and makes them better comparable.

Value

MAgPIE object containing the points on the curve.

Author(s)

David Klein

See Also

[`calculate_fit`](#) [`plot_curve`](#)

---

emudata	*emudata*

Description

Example dataset for a regional MAgPIE object

Value

example data as read from MAgPIE report with reduced set of variables

Author(s)

David Klein

---

emulator	*Calculate and plot fit.*

Description

This function fits a function of the form y = a + b * x ^ d for the given data.

Usage

```
emulator(
  data,
  name_x,
  name_y,
  name_modelstat = NULL,
  treat_as_feasible = c(2, 7),
  userfun = function(param, x) return(param[[1]] + param[[2]] * x^param[[3]]),
  initial_values = c(0, 0, 1),
  outlier_range = 1.5,
  n_suff = 1,
  fill = FALSE,
  output_path = "emulator",
  fitname = "linear",
  create_pdf = TRUE,
  ...
)
```

Arguments

data	MAgPIE object containing at least two variables and the modelstatus
name_x	Name of the variable in data that will be treated as x in the fit
name_y	Name of the variable in data that will be treated as y in the fit
name_modelstat	Name of the variable that contains the modelstatus
treat_as_feasible	
	GAMS model status codes that will be regarded feasible. See https://www.gams.com/24.8/docs/userguides/mccarl/modelstat_tmodstat.htm
userfun	Function to fit. User can provide a functional form using the following syntax: function(param,x)return(param[[1]] + param[[2]] * x ^param[[3]]). This function is the default.
initial_values	Vector with initial values of the fit coefficients.
outlier_range	Before the actual fit a linear pre-fit is performed. Based on their distance to this pre-fit the data points are allocated to quartiles. A data point is considered an outlier if it is more than outlier_range times of the interquartile range away from the the upper or lower end of the interquartile range http://colingorrie.github.io/outlier-detection.html.

n_suff	Minial number (default=1) of data points in a specific year and region that will be regarded as sufficient to perform a fit. If the number of available data points is less no fit will be generated for this year and region.
fill	Logical (default=FALSE) indicating whether data will be copied from subsequent year if in the current year not enough data points are avaialbe.
output_path	Path to save the output to
fitname	Name that describes the fit (default: linear) and will be used for naming the output folders.
create_pdf	Logical indicating whether a pdf should be produced that compiles all figures.
...	Arguments passed on to the optim function in calcualte_fit. Useful to define bounds on fit coefficients.

Value

MAgPIE object containning fit coefficients

Author(s)

David Klein

---

emulator_runs_complete

Checks whether set of given runs is complete.

Description

Checks whether set of given runs is complete.

Usage

```
emulator_runs_complete(list_of_directories, runnumbers = 1:73)
```

Arguments

list_of_directories	
	Vector of strings containing path names. Must have "-xy" at the end of each name, with xy beinga one or two digit number.
runnumbers	Vector of integers providing the number that have to be present in list_of_directories.

Value

Logical. TRUE if set of numbers at the end of the path names is idential to runnumbers.

Author(s)

David Klein

---

fill_missing_years *Find fit coefficients for years where no fit could be calculated*

Description

Find fit coefficients for years where no fit could be calculated

Usage

```
fill_missing_years(fitcoef, nodata, method = 1)
```

Arguments

fitcoef	MAgPIE object containing the fit coefficients
nodata	MAgPIE object of the same shape as fitcoef containing only logicals that indicate where data is not available (TRUE)
method	Choose method that finds fit coefficients for years with no fit. Currently there is only one method availalbe that in a first step takes the fit from the subsequent year. If all subsequent years have no fit it takes the fit from the preceding year.

Value

Magpie object with the updated fit coefficients. Has an attribute attached providing the years missing fits were taken from.

Author(s)

David Klein

---

minimize_least_squares

Minimizes the sum of the error squares

Description

This is an internal function of the remulator package. It fits a user defined function to data minimizing the sum of the error squares.

Usage

```
minimize_least_squares(dat, initial_values, userform, ...)
```

Arguments

`dat`	Array with data points to fit containing one column with sample numbers, one column with x values, and one column with y values
`initial_values`	Initial values for the parameters to be optimized over.
`userform`	Fit function for which least squares will be calaculated using the `optim` function.
`...`	Arguments passed on to the `optim` function. Useful to define bounds on fit coefficients.

Value

list containing fit coefficients and messages

Author(s)

David Klein

---

mute_duplicated	*Set duplicated samples to NA*

Description

This function sets duplicated samples (per region, year, model, scenario) to NA. This is useful to clean your data before fitting.

Usage

```
mute_duplicated(data)
```

Arguments

`data`	MAgPIE object containing the samples to remove the duplicates from.

Value

Magpie object with duplicated samples set to NA.

Author(s)

David Klein

---

mute_infes	*Set data of infeasible years and their successors to NA*

Description

This function sets data of infeasible years and all years after the first infeasible year to NA because results after an infeasible year are not meaningful. The modelstatus has to be provided as one of the variables in the data. This function will look for the modelstatus variable with the name given in `name`. If `return_infes` is TRUE this function return the matrix with infeasible years instead of the data.

Usage

```
mute_infes(data, name = "Modelstatus (-)", feasible = 2)
```

Arguments

data	MAgPIE object containing the results of a MAgPIE run and the modelstatus.
name	String providing the name of the variable that holds the modelstatus.
feasible	Integer vector defining which modelstatus will be treated as feasible.

Value

Magpie object with either filtered model data or the modelstatus.

Author(s)

David Klein

---

mute_insufficient	*In years and regions with only an unsufficient number of data points set all data points to NA*

Description

This function counts data points (per region, year, scenario) and sets them to NA if the count is less than given in `n_suff`. This is useful to clean your data before fitting. No fit will be calculated where no data is available.

Usage

```
mute_insufficient(data, n_suff)
```

Arguments

`data`	MAgPIE object containing the samples to remove the duplicates from.
`n_suff`	Minial number (default=1) of data points that will be regarded as sufficient to perform a fit.

Value

Magpie object with duplicated samples set to NA.

Author(s)

David Klein

`mute_outliers`	*Set outlier samples to NA*

Description

This function sets outlier samples (per region, year, model, scenario) to NA. This is useful to clean your data before fitting. A data point is considered an outlier if it is more than `range` times of the interquartile range away from the the upper or lower end of the interquartile range http://colingorrie.github.io/outlier-detection.html

Usage

```
mute_outliers(data, range = 1.5)
```

Arguments

`data`	MAgPIE object containing the samples to remove the duplicates from.
`range`	Multiplied with the interquartile range this is the maximal distance a data point may have to not be considered an outlier.

Value

Magpie object with duplicated samples set to NA.

Author(s)

David Klein

See Also

[boxplot](#)

plot_compare_supplycurves
Calculate and plot multiple supplycurves to one plot

Description

This function calculates supplycurves of multiple scenarios and plots them into one figure to compare them.

Usage

```
plot_compare_supplycurves(folders, pdfname = NULL)
```

Arguments

folders	Vector giving the paths to the fitted data (usually within the 'output/emulator' folder)
pdfname	If you want the figures to be compiled in a pdf provide the name of the file here

Author(s)

David Klein

See Also

emulator

plot_curve
Plot fitted curve to png files and additionally compile them in a pdf file.

Description

This function produces plot showing the raw data and the fittet curves. It saves it to a folder named after the scenario and it produces a pdf containing all the figures.

Usage

```
plot_curve(
  raw,
  supplycurve_commonY,
  supplycurve_indiviY,
  infes,
  emu_path = "emulator",
  fitname = "linear",
  create_pdf = TRUE
)
```

Arguments

raw	MAgPIE object containing the same raw data that was used to calculate the fitcoefficients.
supplycurve_commonY	
	MAgPIE object containing the points of the curve (with common y limit) (output of [calc_supplycurve](#))
supplycurve_indiviY	
	MAgPIE object containing the points of the curve (with individual y limit) (output of [calc_supplycurve](#))
infes	MAgPIE object containing the modelstatus (optional output of [mute_infes](#))
emu_path	Name of the folder the figures and pdf will be saved to.
fitname	Name that describes the fit (default: linear) and will be used for naming the output folders.
create_pdf	Logical indicating whether a pdf should be produced that compiles all figures.

Author(s)

David Klein

See Also

[calc_supplycurve](#) [mute_infes](#)

---

read_and_combine	*Read multiple report mif files from REMIND or MAgPIE and combine into single magpie object*

Description

This function reads model output from mif files from the provided list of folders. It also reads the GAMS modelstatus from the fulldata.gdx and appends it to the report. If more than one folder is given it combines the contents of the multiple reports into a single magpie object and returns it. If the name of an output file is provided the object is additionally stored in this file. In each of the folders there must be a report_scenname.mif, a fulldata.gdx, a spatial_header.rda, and config.Rdata.

Usage

```
read_and_combine(list_of_directories, outfile = NULL)
```

Arguments

list_of_directories	
	Vector of strings providing the folder names from which reports will be read.
outfile	Path to an output file which the output of the function can optionally be stored to.

Author(s)

David Klein

---

replace_flat_fits	*Replace fits in years where they are flat with fits from other years*

Description

This function considers fits with a slope less than defined in threshold flat and tries to replace them with non-flat fits from other years.

Usage

```
replace_flat_fits(
  path_to_postfit_rdata,
  flat = NULL,
  threshold = 0.01,
  plot = FALSE
)
```

Arguments

path_to_postfit_rdata

        Path to the Rdata file that contains raw data, fitted data and fitcoefficients saved by the emulator function

flat	If you want this function to replace further fits provide their region and year here using a vector of the form c("LAM:2020,2025", "IND:2005")
threshold	Defines the value of the slope (b) below which fits are considered flat
plot	logical. If TRUE supply curves will be saved to png files

Author(s)

David Klein

See Also

emulator fill_missing_years

Index