

# Package ‘mbRes’

January 29, 2022

**Type** Package

**Title** Integrating Multiple Biomarker Responses in Aquatic Organisms using Effect Size, Statistical Uncertainty, and Ecological Relevance

**Version** 0.1.3

**Description** Compute and visualize the ps-index, a new integrated index for multiple biomarker responses, as described in Pham & Sokolova (2022, unpublished).

**Depends** R (>= 4.1.0)

**Imports** cowplot (>= 1.1.1), data.table (>= 1.14.2), dplyr (>= 1.0.7), forcats (>= 0.5.1), ggforce (>= 0.3.3), ggplot2 (>= 3.3.5), grid, magrittr (>= 2.0.1), purrr (>= 0.3.4), rlang (>= 0.4.12), stats, tidyr (>= 1.1.4), tibble (>= 3.1.6), utils

**Suggests** RProbSup (>= 3.0)

**BugReports** <https://github.com/phamdn/mbRes/issues>

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.2

**NeedsCompilation** no

**Author** Duy Nghia Pham [aut, cre] (<<https://orcid.org/0000-0003-1349-1710>>),  
Inna M. Sokolova [aut] (<<https://orcid.org/0000-0002-2068-4302>>)

**Maintainer** Duy Nghia Pham <nghiapham@yandex.com>

**Repository** CRAN

**Date/Publication** 2022-01-28 23:00:02 UTC

## R topics documented:

mbRes-package . . . . .	2
cliff . . . . .	3

compare . . . . .	4
ggbar . . . . .	4
ggdot . . . . .	5
ggheat . . . . .	6
mbr . . . . .	7
perch . . . . .	8
resampling . . . . .	9
sokolova2021 . . . . .	10

<b>Index</b>	<b>13</b>
--------------	-----------

---

mbRes-package	<i>mbRes: Integrating Multiple Biomarker Responses in Aquatic Organisms using Effect Size, Statistical Uncertainty, and Ecological Relevance</i>
---------------	--

---

## Description

Compute and visualize the ps-index, a new integrated index for multiple biomarker responses, as described in Pham & Sokolova (2021, unpublished).

## Guidelines

ps-index is a new integrated index for multiple biomarker responses. `mbr` is the main function to compute and visualize the ps-index. `sokolova2021` is provided as a sample dataset. `compare` simulates a hypothetical dataset and compare the results of ps-index and two other integrated indices published earlier. The others are helper functions and are not meant to be called directly by users.

## Copyright

mbRes: Integrating Multiple Biomarker Responses in Aquatic Organisms using Effect Size, Statistical Uncertainty, and Ecological Relevance. Copyright (C) 2021 Duy Nghia Pham & Inna M. Sokolova

mbRes is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

mbRes is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with mbRes. If not, see <https://www.gnu.org/licenses/>.

## Author(s)

Duy Nghia Pham & Inna M. Sokolova

---

cliff	<i>Compute Effect Size</i>
-------	----------------------------

---

**Description**

cliff calculates Cliff's delta statistic using the rank sum method.

**Usage**

```
cliff(v1, v0)
```

**Arguments**

v1                    a vector, biomarker values from the treatment group.  
v0                    a vector, biomarker values from the control group.

**Value**

cliff returns a numeric that is the Cliff's delta of the treatment group.

**References**

Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, 114(3), 494–509. doi: [10.1037/00332909.114.3.494](https://doi.org/10.1037/00332909.114.3.494).

Vargha, A., & Delaney, H. D. (2000). A Critique and Improvement of the CL Common Language Effect Size Statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics*, 25(2), 101–132. doi: [10.3102/10769986025002101](https://doi.org/10.3102/10769986025002101).

Ruscio, J., & Mullen, T. (2012). Confidence Intervals for the Probability of Superiority Effect Size Measure and the Area Under a Receiver Operating Characteristic Curve. *Multivariate Behavioral Research*, 47(2), 201–223. doi: [10.1080/00273171.2012.658329](https://doi.org/10.1080/00273171.2012.658329).

**See Also**

[CalcA1](#).

**Examples**

```
cliff(unlist(sokolova2021[7:12, 2]), unlist(sokolova2021[1:6, 2]))
```

---

compare	<i>Hypothetical Case Study Analysis</i>
---------	---

---

### Description

compare generates a hypothetical dataset of biomarker responses and analyzes the dataset using three different strategies.

### Usage

```
compare()
```

### Value

compare returns a list of length 3:

tab	a list of length 9
	pop the true population means of normal distributions
	pop_long the true population means in long format
	sam the sample dataset of biomarker responses
	sam_long the sample dataset in long format
	m mean biomarker responses by exposure group
	blaise2002 integrated index from an ordinalization method
	beliaeff2002 integrated index from a standardization method
	eco assigned ecological relevance of biomarkers
	pham2021 ps-index from our estimation method
fig	a list of 2 ggplot objects
	fig_sam hypothetical dataset
	fig_compare result comparison of three methods
est	full results of our estimation method given by mbr

---

ggbar	<i>Make Bar Plot</i>
-------	----------------------

---

### Description

ggbar creates bar plot of the ecological relevance. This is not meant to be called directly.

### Usage

```
ggbar(dat, hax, vax, sub, env = parent.frame())
```

**Arguments**

dat	a data frame with at least two columns.
hax	a character, name of the column to be used as the horizontal axis.
vax	a character, name of the column to be used as the vertical axis.
sub	a numeric, mean ecological relevance.
env	an environment, to access outer scope variables.

**Value**

ggbar returns a ggplot object.

---

ggdot *Make Dot Plot*

---

**Description**

ggdot creates dot plot of the ps-index. This is not meant to be called directly.

**Usage**

```
ggdot(dat, hax, vax)
```

**Arguments**

dat	a data frame with at least two columns.
hax	a character, name of the column to be used as the horizontal axis.
vax	a character, name of the column to be used as the vertical axis.

**Value**

ggdot returns a ggplot object.

---

`ggheat`*Make Heatmap*

---

## Description

`ggheat` creates heatmaps of the Cliff's delta, surprisal, confidence interval width, and weighting factor. This is not meant to be called directly.

## Usage

```
ggheat(  
  dat,  
  hax,  
  vax,  
  cell,  
  nm,  
  lim,  
  lo,  
  hi,  
  diverging = FALSE,  
  env = parent.frame()  
)
```

## Arguments

<code>dat</code>	a data frame with at least three columns.
<code>hax</code>	a character, name of the column to be used as the horizontal axis.
<code>vax</code>	a character, name of the column to be used as the vertical axis.
<code>cell</code>	a character, name of the column to be used as the cells.
<code>nm</code>	a character, name of the heatmap.
<code>lim</code>	a numeric vector, limits of the color scale.
<code>lo</code>	a character, color of the color scale low end.
<code>hi</code>	a character, color of the color scale high end.
<code>diverging</code>	a logical, whether to use diverging color gradient.
<code>env</code>	an environment, to access outer scope variables.

## Value

`ggheat` returns a `ggplot` object.

---

mbr	<i>Compute and Visualize ps-index</i>
-----	---------------------------------------

---

### Description

mbr calculates the ps-index and visualizes the whole process.

### Usage

```
mbr(df, df2 = NULL, rotate = FALSE, display = TRUE)
```

### Arguments

df	a data frame with the name of experimental groups or biomonitoring sites as the first column and the measurement of biomarkers as the remaining columns.
df2	a data frame with the name of biomarkers as the first column and the ecological relevance as the second column.
rotate	a logical, whether to rotate the biomarker labels in figures.
display	a logical, whether to display cell values in heatmaps.

### Details

The header of the first column can be any character, for example, 'group' or 'site'. The first name appearing in the first column will determine the control group or the reference site. The other names will be treatment groups or test sites. The header of the remaining columns will define the list of biomarkers.

The row order of biomarkers in df2 must match the column order of biomarkers in df. If df2 is missing, all biomarkers automatically have ecological relevance of only 1.

### Value

mbr returns a list of length 4:

input	a list of length 3
	main the input biomarker data frame
	extra a data frame with 2 columns:
	biomarker individual biomarker
	eco the input ecological relevance
	eco.mean the mean ecological relevance
es	a data frame with 14 columns:
	treatment treatment groups or test sites
	control control group or reference site
	tm_size the sample size of treatment group
	ct_size the sample size of control group

	biomarker	individual biomarker
	delta	the Cliff's delta of treatment group
	delta.abs	the absolute Cliff's delta
	pval	the P-Value
	sval	the surprisal or S-Value
	se	the standard error of Cliff's delta
	ci.lower	the lower bound of the confidence interval
	ci.upper	the upper bound of the confidence interval
	ciw	the width of the confidence interval
	eco	the ecological relevance of the biomarker
	sval.cdf	the ascending rank of S-value by CDF method, see <a href="#">cume_dist</a>
	ciw.cdf	the descending rank of confidence interval width by CDF method
	eco.cdf	the ascending rank of ecological relevance by CDF method
	weight	the total weighting factor
idx		a data frame with 2 columns:
	treatment	treatment groups or test sites
	ps	the ps-index
fig		a list of 6 ggplot objects
	delta	the Cliff's delta
	sval	the surprisal or S-value
	ciw	confidence interval width
	eco	the ecological relevance
	weight	the weighting factor
	ps	the ps-index

### Examples

```
mbr(sokolova2021, ecorelevance, rotate = TRUE) #might take more than 5s in some machines
```

---

perch

*Empirical Case Study Analysis*

---

### Description

perch analyses the biomarker results of Hansson et al. (2014) <doi:10.1007/s00244-013-9974-5>.

### Usage

```
perch()
```

**Value**

perch returns a list of length 3:

tab	a list of length 2 hansson2014 biomarker data reported by Hansson et al. (2014) percheco assigned ecological relevance of biomarkers
fig	a list of 5 ggplot objects SG length-corrected somatic growth SCI somatic condition index GSI time-corrected gonadosomatic index LSI time-corrected liver somatic index EROD liver ethoxyresorufin-O-deethylase activity
est	full results of our estimation method given by mbr

**References**

Hansson, T., Hansen, W., Tjärnlund, U., Balk, L., & Bengtsson, B.-E. (2014). Biomarker Investigations in Adult Female Perch (*Perca fluviatilis*) From Industrialised Areas in Northern Sweden in 2003. *Archives of Environmental Contamination and Toxicology*, 66(2), 237–247. doi: [10.1007/s0024401399745](https://doi.org/10.1007/s0024401399745).

---

 resampling

*Measure Statistical Uncertainty*


---

**Description**

resampling performs randomization test and bootstrapping to calculate P-value and percentile bootstrap confidence interval of Cliff's delta.

**Usage**

```
resampling(v1, v0, nrand = 1999, nboot = 1999, conf.level = 0.95, seed = 1)
```

**Arguments**

v1	a vector, biomarker values from the treatment group.
v0	a vector, biomarker values from the control group.
nrand	an integer, the number of randomization samples. The default value is 1999.
nboot	an integer, the number of bootstrap samples. The default value is 1999.
conf.level	a numeric, the confidence level to calculate percentile bootstrap confidence interval. The default value is 0.95.
seed	an integer, the seed for random number generation. Setting a seed ensures the reproducibility of the result. See <a href="#">set.seed</a> for more details.

**Value**

resampling returns a one-row data frame with four numerics:

delta	the Cliff's delta of the treatment group.
pval	the observed P-value p under the null hypothesis.
sval	the S-value s calculated from P-value p.
se	the standard error of Cliff's delta.
ci.lower	the lower bound of the confidence interval.
ci.upper	the upper bound of the confidence interval.
ciw	the width of the confidence interval.

**References**

Greenland, S. (2019). Valid P-Values Behave Exactly as They Should: Some Misleading Criticisms of P-Values and Their Resolution With S-Values. *The American Statistician*, 73(sup1), 106–114. doi: [10.1080/00031305.2018.1529625](https://doi.org/10.1080/00031305.2018.1529625).

Phipson, B., & Smyth, G. K. (2010). Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn. *Statistical Applications in Genetics and Molecular Biology*, 9(1). doi: [10.2202/15446115.1585](https://doi.org/10.2202/15446115.1585).

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Chapman & Hall.

Ruscio, J., & Mullen, T. (2012). Confidence Intervals for the Probability of Superiority Effect Size Measure and the Area Under a Receiver Operating Characteristic Curve. *Multivariate Behavioral Research*, 47(2), 201–223. doi: [10.1080/00273171.2012.658329](https://doi.org/10.1080/00273171.2012.658329).

**See Also**

[A1](#).

**Examples**

```
resampling(unlist(sokolova2021[7:12, 2]), unlist(sokolova2021[1:6, 2]))
```

---

sokolova2021

*Biomarker Responses of the Blue Mussels to Organic UV Filters*

---

**Description**

sokolova2021 contains the biomarker responses of the blue mussel *Mytilus edulis* to organic UV filters ensulizole and octocrylene. ecorelevance contains the ecological relevance of the biomarkers.

**Usage**

sokolova2021

ecorelevance

**Format**

sokolova2021 is a data frame with 30 rows and 31 variables:

**Exposure** 5 experimental groups:

**CTRL** control group

**EN10** 10 ug/L of ensulizole

**EN100** 100 ug/L of ensulizole

**OC10** 10 ug/L of octocrylene

**OC100** 100 ug/L of octocrylene

**NRR** lysosomal membrane stability

**ROS** reactive oxygen species generation

**TBARSd** lipid peroxidation in digestive gland

**TBARSg** lipid peroxidation in gills

**PCd** protein carbonylation in digestive gland

**PCg** protein carbonylation in gills

**CPRd** NADPH-P450 reductase activity in digestive gland

**CPRg** NADPH-P450 reductase activity in gills

**ERODd** 7-ethoxyresorufin-O-deethylase activity in digestive gland

**ERODg** 7-ethoxyresorufin-O-deethylase activity in gills

**CEd** carboxylesterase activity in digestive gland

**CEg** carboxylesterase activity in gills

**GSTd** glutathione-S-transferase activity in digestive gland

**GSTg** glutathione-S-transferase activity in gills

**GRd** glutathione reductase activity in digestive gland

**GRg** glutathione reductase activity in gills

**CTSDTd** total cathepsin D activity in digestive gland

**CTSDTg** total cathepsin D activity in gills

**CTSDFd** free cathepsin D activity in digestive gland

**CTSDFg** free cathepsin D activity in gills

**Cas2** caspase 2

**Cas3** caspase 3

**BAX** Bcl-2-associated X protein

**Bcl-2** B-cell lymphoma 2

**p53** tumor protein 53

**GADD45** growth arrest and DNA-damage-inducible protein 45

**NF- $\kappa$ B** nuclear factor  $\kappa$ B

**IL-17** interleukin 17

**COX-2** cyclooxygenase 2

**ACC** acetyl-CoA carboxylase

ecorelevance is a data frame with 30 rows and 2 variables:

**Biomarker** 30 endpoints as documented in sokolova2021

**Eco** ecological relevance

### Source

Sokolova, I. M., Falfushynska, H., & Sokolov, E. P. (2021). Biomarker responses of the blue mussels to organic UV filters [Data set]. Zenodo. doi: [10.5281/zenodo.5176087](https://doi.org/10.5281/zenodo.5176087).

# Index

## \* datasets

sokolova2021, [10](#)

A1, [10](#)

CalcA1, [3](#)

cliff, [3](#)

compare, [2](#), [4](#)

cume\_dist, [8](#)

ecorelevance (sokolova2021), [10](#)

ggbar, [4](#)

ggdot, [5](#)

ggheat, [6](#)

mbr, [2](#), [7](#)

mbRes-package, [2](#)

perch, [8](#)

resampling, [9](#)

set.seed, [9](#)

sokolova2021, [2](#), [10](#)