

Package ‘diathor’

May 18, 2021

Type Package

Title Calculate Ecological Information and Diatom Based Indices

Version 0.0.6

Maintainer Joaquín Cochero <jcochero@ilpla.edu.ar>

Description Calculates ecological information and biotic indices for diatoms in a sample.

It includes number/shape of chloroplasts diversity indices, size classes, ecological guilds, and multiple biotic indices.

It outputs both a dataframe with all the results and plots of all the obtained data in a defined output folder.

Sample data was taken from Nicolosi Ge-

lis, Cochero & Gómez (2020, <doi:10.1016/j.ecolind.2019.105951>).

The package uses the 'Diat.Barcode' database to calculate morphological and ecological information by Rimet & Couchez (2012, <doi:10.1051/kmae/2012018>),

and calculates the DES index by Descy (1979, <http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=PASCAL8060205402>),

the EPID index by Dell'Uomo (1996, ISBN: 3950009002),

the IDAP index by Prygiel & Coste (1993, <doi:10.1007/BF00028033>),

the ID-CH index by Hürlimann & Nieder-

hauser (2007, <https://www.bafu.admin.ch/bafu/fr/home/themes/eaux/publications/publications-eaux/methodes-analyse-appreciation-cours-eau-diatomees.html>),

the IDP index by Gómez & Licursi (2001, <doi:10.1023/A:1011415209445>),

the ILM index by Leclercq & Maquet (1987, <http://www.vliz.be/imisdocs/publications/286641.pdf>),

the IPS index by Coste (1982, <https://www.oieau.org/eaudoc/notice/ETUDE-DES-METHODES-BIOLOGIQUES-DAPPRECIATION-QUANTITATIVE-DE-LA-QUALITE-DES-EAUX>),

the LOBO index by Lobo, Callegaro, & Bender (2002, ISBN:9788585869908),

the SLA by Sládeček (1986, <doi:10.1002/ahch.19860140519>),

the TDI index by Kelly, & Whitton (1995, <doi:10.1007/BF00003802>),

and the SPEAR(herbicide) index by Wood, Mitrovic, Lim, Warne, Dunlop, & Kefford (2019, <doi:10.1016/j.ecolind.2018.12.035>).

License GNU General Public License

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Depends R (>= 2.10), stringdist, vegan, ggplot2, tidy

Imports data.table, purrr, stringr, tibble

Suggests knitr, rmarkdown

NeedsCompilation no

Author María Mercedes Nicolosi Gelis [aut]

(<<https://orcid.org/0000-0001-6324-7930>>),

María Belén Sathicq [aut] (<<https://orcid.org/0000-0002-3534-8950>>),

Joaquín Cochero [cre] (<<https://orcid.org/0000-0003-3957-6819>>)

Repository CRAN

Date/Publication 2021-05-18 15:40:19 UTC

R topics documented:

| | |
|---------------------------|----|
| dbc_offline | 3 |
| des | 3 |
| diaThor | 4 |
| diaThorAll | 6 |
| diat_des | 7 |
| diat_diversity | 8 |
| diat_epid | 9 |
| diat_guilds | 10 |
| diat_idap | 11 |
| diat_idch | 12 |
| diat_idp | 13 |
| diat_ilm | 14 |
| diat_ips | 15 |
| diat_loadData | 16 |
| diat_lobo | 17 |
| diat_morpho | 18 |
| diat_sampleData | 19 |
| diat_size | 19 |
| diat_sla | 20 |
| diat_spear | 21 |
| diat_tdi | 22 |
| diat_vandam | 23 |
| epid | 24 |
| idap | 24 |
| idch | 25 |
| idp | 25 |
| ilm | 26 |
| ips | 26 |
| lobo | 27 |
| sla | 27 |
| spear | 28 |
| tdi | 28 |

dbc_offline

3

Index

29

dbc_offline

DBC (offline)

Description

Diatom database from the 'Diat.Barcode' project V9.0

Usage

```
data(dbc_offline)
```

Format

A data frame with ecological and morphological information for 8066 diatoms

Source

<https://www.kmae-journal.org/articles/kmae/abs/2012/03/kmae120025/kmae120025.html>

References

Rimet F. & Bouchez A., 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. Knowledge and management of aquatic ecosystems, 406: 1-14. DOI:10.1051/kmae/2012018

des

DES

Description

Index values for diatom species included in the DES index

Usage

```
data(des)
```

Format

A data frame with the ecological values for 622 species

Source

<http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=PASCAL8060205402>

References

Descy, J. P. 1979. A new approach to water quality estimation using diatom. Beih. Nov Hedw. 64:305-323

Description

The package calculates multiple biotic indices using diatoms from environmental samples. Diatom species are recognized by their species' name using a heuristic search, and their ecological data is retrieved from multiple sources. Morphological information about the species is retrieved from the 'Diat.Barcode' project:

- Rimet F., Gusev E., Kahlert M., Kelly M., Kulikovskiy M., Maltsev Y., Mann D., Pfannkuchen M., Trobajo R., Vasselon V., Zimmermann J., Bouchez A., 2019. Diat.barcode, an open-access curated barcode library for diatoms. *Scientific Reports*. <https://www.nature.com/articles/s41598-019-51500-6>

Size class classification is obtained from:

- Rimet F. & Bouchez A., 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. *Knowledge and management of aquatic ecosystems*, 406: 1-14. DOI:10.1051/kmae/2012018

Guild classification is obtained from:

- Rimet F. & Bouchez A., 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. *Knowledge and management of aquatic ecosystems*, 406: 1-14. DOI:10.1051/kmae/2012018

Van Dam classification is obtained from:

- Van Dam, H., Mertens, A., & Sinkeldam, J. (1994). A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. *Netherlands Journal of Aquatic Ecology*, 28(1), 117-133.

Diversity index (Shannon's H') is calculated using the vegan package, following:

- Shannon, C. E., and Weaver, W. (1949). 'The Mathematical Theory of Communication.' (University of Illinois Press: Urbana, IL, USA.)

Species tolerance and their ecological information to calculate each biotic index is retrieved from their original sources:

- IPS: Coste, M. (1982). Étude des méthodes biologiques d'appréciation quantitative de la qualité des eaux. Rapport Cemagref QE Lyon-AF Bassin Rhône Méditerranée Corse.
- TDI: Kelly, M. G., & Whitton, B. A. (1995). The trophic diatom index: a new index for monitoring eutrophication in rivers. *Journal of Applied Phycology*, 7(4), 433-444.
- IDP: Gómez, N., & Licursi, M. (2001). The Pampean Diatom Index (IDP) for assessment of rivers and streams in Argentina. *Aquatic Ecology*, 35(2), 173-181.
- DES: Descy, J. P. 1979. A new approach to water quality estimation using diatom. *Beih. Nov Hedw.* 64:305-323

- EPID: Dell’Uomo, A. (1996). Assessment of water quality of an Apennine river as a pilot study for diatom-based monitoring of Italian watercourses. Use of algae for monitoring rivers, 65-72.
- IDAP: Prygiel, J., & Coste, M. (1993). The assessment of water quality in the Artois-Picardie water basin (France) by the use of diatom indices. *Hydrobiologia*, 269(1), 343-349.
- ID-CH: Hürlimann J., Niederhauser P. 2007: Méthodes d’analyse et d’appréciation des cours d’eau. Diatomées Niveau R (région). État de l’environnement n° 0740. Office fédéral de l’environnement, Berne. 132 p
- ILM: Leclercq, L., & Maquet, B. (1987). Deux nouveaux indices diatomique et de qualité chimique des eaux courantes. Comparaison avec différents indices existants. *Cahier de Biologie Marine*, 28, 303-310.
- LOBO: Lobo, E. A., Callegaro, V. L. M., & Bender, E. P. (2002). Utilização de algas diatomáceas epilíticas como indicadoras da qualidade da água em rios e arroios da Região Hidrográfica do Guaíba, RS, Brasil. *Edunisc*.
- LOBO: Lobo, E. A., Bes, D., Tudesque, L., & Ector, L. (2004). Water quality assessment of the Pardino River, RS, Brazil, using epilithic diatom assemblages and faecal coliforms as biological indicators. *Vie et Milieu*, 54(2-3), 115-126.
- SLA: Sládeček, V. (1986). Diatoms as indicators of organic pollution. *Acta hydrochimica et hydrobiologica*, 14(5), 555-566.
- SPEAR(herbicides): Wood, R. J., Mitrovic, S. M., Lim, R. P., Warne, M. S. J., Dunlop, J., & Kefford, B. J. (2019). Benthic diatoms as indicators of herbicide toxicity in rivers—A new SPEcies At Risk (SPEARherbicides) index. *Ecological Indicators*, 99, 203-213.

Sample data included in the package is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951.

Functions

diat_loadData() diat_morpho() diat_size() diat_diversity() diat_guilds() diat_vandam() diat_loadData()
 diat_ips() diat_tdi() diat_idp() diat_des() diat_epid() diat_idch() diat_ilm() diat_lobo() diat_sla()
 diat_spear()

`diaThorAll`*Runs all the DiaThor functions in a pipeline*

Description

The `diaThorAll` function is the master function of the package. It calculates all outputs from the data, and places them in the Output folder. The input file for the package is a dataframe or an external CSV file. Species should be listed as rows, with species' names in column 1 (column name should be "species"). The other columns (samples) have to contain the abundance of each species (relative or absolute) in each sample. The first row of the file has to contain the headers with the sample names. Remember that a column named "species" is mandatory, containing the species' names. If a dataframe is not specified as a parameter (`species_df`), the package will show a dialog box to search for the CSV file. A second dialog box will help set up an Output folder, where all outputs from the package will be exported to (dataframes, CSV files, plots in PDF). The package also downloads and installs a wrapper for the 'Diat.Barcode' project. Besides citing the DiaThor package, the Diat.Barcode project should also be cited, as follows:

- Rimet, Frederic; Gusev, Evgeny; Kahlert, Maria; Kelly, Martyn; Kulikovskiy, Maxim; Maltsev, Yevhen; Mann, David; Pfannkuchen, Martin; Trobajo, Rosa; Vasselon, Valentin; Zimmermann, Jonas; Bouchez, Agnès. 2018. "Diat.barcode, an open-access barcode library for diatoms". *Scientific Reports*, 9, 15116. <https://doi.org/10.15454/TOMBYZ>

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diaThorAll(  
  species_df,  
  isRelAb = FALSE,  
  maxDistTaxa = 2,  
  resultsPath,  
  calculateguilds = TRUE,  
  vandam = TRUE,  
  vandamReports = TRUE,  
  singleResult = TRUE,  
  exportFormat = 3,  
  exportName = "DiaThor_results",  
  plotAll = TRUE,  
  color = "#0073C2"  
)
```

Arguments

| | |
|-----------------|---|
| species_df | The data frame with your species data. Species as rows, Sites as columns. If empty, a dialog will prompt for a CSV file |
| isRelAb | Boolean. If set to 'TRUE' it means that your species' data is the relative abundance of each species per site. If FALSE, it means that it the data corresponds to absolute densities. Default = FALSE |
| maxDistTaxa | Integer. Number of characters that can differ in the species' names when compared to the internal database's name in the heuristic search. Default = 2 |
| resultsPath | String. Path to the output folder. If none specified (default), a dialog box will prompt to select it |
| calculateguilds | Boolean. If set to 'TRUE' the percentage of abundance of each diatom guild will be calculated. Default = TRUE |
| vandam | Boolean. If set to 'TRUE' the Van Dam classifications will be calculated in the Output. Default = TRUE |
| vandamReports | Boolean. If set to 'TRUE' the detailed reports for the Van Dam classifications will be reported in the Output. Default = TRUE |
| singleResult | Boolean. If set to 'TRUE' all results will go into a single output file. If FALSE, separate output files will be generated. Default = TRUE |
| exportFormat | Integer. If = 1: only a CSV (external file) will be generated with the output matrices; 2: only an internal R dataframe will be generated; 3: both a CSV and an internal R dataframe are generated. Default = 3 |
| exportName | String. Prefix for the CSV exported file. Default = "DiaThor_results" |
| plotAll | Boolean. If set to 'TRUE', plots will be generated for each Output in a PDF file. Default = TRUE |
| color | Color code (hex). Default color for bar charts and lolipop plots. Default = "#0073C2" |

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# In the example, a temporary directory will be used in resultsPath
allResults <- diaThorAll(diat_sampleData, resultsPath = tempdir())
```

Description

The input for all of these functions is the resulting dataframe (resultLoad) obtained from the diat_loadData() function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the diat_loadData() function so the acronyms and species' names are recognized. References for the index:

- Descy, J. P. 1979. A new approach to water quality estimation using diatom. *Beih. Nov Hedw.* 64:305-323

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_des(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
desResults <- diat_des(df)
```

diat_diversity

Calculate diversity parameters for diatoms using the vegan package

Description

The input for these functions is the resulting dataframe obtained from the diat_loadData() function, to calculate diversity data using the vegan package. Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Diversity index (Shannons H') is calculated using the vegan package, following:

- Shannon, C. E., and Weaver, W. (1949). 'The Mathematical Theory of Communication.' (University of Illinois Press: Urbana, IL, USA.)

Usage

```
diat_diversity(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

| | |
|-----------|---|
| diat_epid | <i>Calculates the EPID index (EPID)</i> |
|-----------|---|

Description

The input for all of these functions is the resulting dataframe (resultLoad) obtained from the diat_loadData() function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the diat_loadData() function so the acronyms and species' names are recognized. References for the index:

- Dell'Uomo, A. (1996). Assessment of water quality of an Apennine river as a pilot study for diatom-based monitoring of Italian watercourses. Use of algae for monitoring rivers, 65-72.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". Ecological Indicators, 110, 105951. <<https://doi.org/10.1016/j.ecolind.2019.105951>>

Usage

```
diat_epid(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
```

```
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
epidResults <- diat_epid(df)
```

diat_guilds

Calculate ecological guilds for diatoms

Description

The input for these functions is the resulting dataframe obtained from the `diat_loadData()` function, to calculate the ecological guilds for the diatoms. Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Guild classification is obtained from:

- Rimet F. & Bouchez A., 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. *Knowledge and management of aquatic ecosystems*, 406: 1-14. DOI:10.1051/kmae/2012018

Usage

```
diat_guilds(resultLoad)
```

Arguments

`resultLoad` The resulting list obtained from the `diat_loadData()` function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
guildsResults <- diat_guilds(df)
```

`diat_idap`*Calculates the Indice Diatomique Artois-Picardie (IDAP)*

Description

The input for all of these functions is the resulting dataframe (`resultLoad`) obtained from the `diat_loadData()` function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the `diat_loadData()` function so the acronyms and species' names are recognized. References for the index:

- Prygiel, J., & Coste, M. (1993). The assessment of water quality in the Artois-Picardie water basin (France) by the use of diatom indices. *Hydrobiologia*, 269(1), 343-349.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_idap(resultLoad)
```

Arguments

`resultLoad` The resulting list obtained from the `diat_loadData()` function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
idapResults <- diat_idap(df)
```

`diat_idch`*Calculates the Swiss Diatom Index (IDCH)*

Description

The input for all of these functions is the resulting dataframe (`resultLoad`) obtained from the `diat_loadData()` function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the `diat_loadData()` function so the acronyms and species' names are recognized. References for the index:

- Hürlimann J., Niederhauser P. 2007: Méthodes d'analyse et d'appréciation des cours d'eau. Diatomées Niveau R (région). État de l'environnement n° 0740. Office fédéral de l'environnement, Berne. 132 p

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_idch(resultLoad)
```

Arguments

`resultLoad` The resulting list obtained from the `diat_loadData()` function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
idchResults <- diat_idch(df)
```

| | |
|----------|--|
| diat_idp | <i>Calculates the Pampean Diatom Index (IDP)</i> |
|----------|--|

Description

The input for all of these functions is the resulting dataframe (resultLoad) obtained from the diat_loadData() function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the diat_loadData() function so the acronyms and species' names are recognized. References for the index:

- Gómez, N., & Licursi, M. (2001). The Pampean Diatom Index (IDP) for assessment of rivers and streams in Argentina. *Aquatic Ecology*, 35(2), 173-181.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_idp(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
idpResults <- diat_idp(df)
```

| | |
|----------|---------------------------------------|
| diat_ilm | <i>Calculates the ILM Index (ILM)</i> |
|----------|---------------------------------------|

Description

The input for all of these functions is the resulting dataframe (resultLoad) obtained from the diat_loadData() function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the diat_loadData() function so the acronyms and species' names are recognized. References for the index:

- Leclercq, L., & Maquet, B. (1987). Deux nouveaux indices diatomique et de qualité chimique des eaux courantes. Comparaison avec différents indices existants. Cahier de Biology Marine, 28, 303-310.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". Ecological Indicators, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_ilm(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
ilmResults <- diat_ilm(df)
```

`diat_ips`*Calculates the Specific Polluosensitivity Index (IPS) index*

Description

The input for all of these functions is the resulting dataframe (`resultLoad`) obtained from the `diat_loadData()` function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the `diat_loadData()` function so the acronyms and species' names are recognized. References for the index:

- Coste, M. (1982). Étude des méthodes biologiques d'appréciation quantitative de la qualité des eaux. Rapport Cemagref QE Lyon-AF Bassin Rhône Méditerranée Corse.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_ips(resultLoad)
```

Arguments

`resultLoad` The resulting list obtained from the `diat_loadData()` function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
ipsResults <- diat_ips(df)
```

diat_loadData

*Loads the Data into DiaThor in the correct format***Description**

Loads the CSV or dataframe file, sets the Output folder for the package, and conducts both an exact and an heuristic search of the species' names.

The input file for the package is a dataframe or an external CSV file. Species should be listed as rows, with species' names in column 1 (column name should be "species") The other columns (samples) have to contain the abundance of each species (relative or absolute) in each sample. The first row of the file has to contain the headers with the sample names. Remember that a column named "species" is mandatory, containing the species' names If a dataframe is not specified as a parameter (species_df), the package will show a dialog box to search for the CSV file A second dialog box will help set up an Output folder, where all outputs from the package will be exported to (dataframes, CSV files, plots in PDF) The package also downloads and installs a wrapper for the 'Diat.Barcode' project. Besides citing the DiaThor package, the Diat.Barcode project should also be cited, as follows:

- Rimet F., Gusev E., Kahlert M., Kelly M., Kulikovskiy M., Maltsev Y., Mann D., Pfannkuchen M., Trobajo R., Vasselon V., Zimmermann J., Bouchez A., 2019. Diat.barcode, an open-access curated barcode library for diatoms. Scientific Reports. <https://www.nature.com/articles/s41598-019-51500-6>

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". Ecological Indicators, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_loadData(species_df, isRelAb = FALSE, maxDistTaxa = 2, resultsPath)
```

Arguments

| | |
|-------------|---|
| species_df | The data frame with your species data. Species as rows, Samples as columns. If empty, a dialog box will prompt to import a CSV file |
| isRelAb | Boolean. If set to 'TRUE' it means that your species' data is the relative abundance of each species per site. If FALSE, it means that it the data corresponds to absolute densities. Default = FALSE |
| maxDistTaxa | Integer. Number of characters that can differ in the species' names when compared to the internal database's name in the heuristic search. Default = 2 |
| resultsPath | String. Path for the output data. If empty (default), it will prompt a dialog box to select an output folder |

`diat_lobo`*Calculates the Lobo Index (LOBO)*

Description

The input for all of these functions is the resulting dataframe (`resultLoad`) obtained from the `diat_loadData()` function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the `diat_loadData()` function so the acronyms and species' names are recognized. References for the index:

- Lobo, E. A., Callegaro, V. L. M., & Bender, E. P. (2002). Utilização de algas diatomáceas epilíticas como indicadoras da qualidade da água em rios e arroios da Região Hidrográfica do Guaíba, RS, Brasil. *Edunisc*.
- Lobo, E. A., Bes, D., Tudesque, L., & Ector, L. (2004). Water quality assessment of the Pardinho River, RS, Brazil, using epilithic diatom assemblages and faecal coliforms as biological indicators. *Vie et Milieu*, 54(2-3), 115-126.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_lobo(resultLoad)
```

Arguments

`resultLoad` The resulting list obtained from the `diat_loadData()` function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
loboResults <- diat_lobo(df)
```

diat_morpho

Calculate morphological parameters for diatoms

Description

The input for these functions is the resulting dataframe obtained from the `diat_loadData()` function to calculate morphological parameters. The morphological data (size classes, chloroplasts) is obtained from the 'Diat.Barcode' project. Besides citing DiaThor, the Diat.Barcode project should also be cited if the package is used, as follows:

- Rimet F., Gusev E., Kahlert M., Kelly M., Kulikovskiy M., Maltsev Y., Mann D., Pfannkuchen M., Trobajo R., Vasselon V., Zimmermann J., Bouchez A., 2019. Diat.barcode, an open-access curated barcode library for diatoms. Scientific Reports. <https://www.nature.com/articles/s41598-019-51500-6>

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". Ecological Indicators, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_morpho(resultLoad, isRelAb = FALSE)
```

Arguments

| | |
|-------------------------|--|
| <code>resultLoad</code> | The resulting list obtained from the <code>diat_loadData()</code> function |
| <code>isRelAb</code> | Boolean. If set to 'TRUE' it means that your species' data is the relative abundance of each species per site. If FALSE, it means that the data corresponds to absolute densities. Default = FALSE |

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
morphoResults <- diat_morpho(df)
```

| | |
|-----------------|--------------------|
| diat_sampleData | <i>Sample Data</i> |
|-----------------|--------------------|

Description

This sample data is a dataset used in: Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipellic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951.

Usage

```
data(diat_sampleData)
```

Format

A data frame with the abundance of 164 diatoms in 108 sampled sites

Source

<https://www.sciencedirect.com/science/article/abs/pii/S1470160X1930946X>

References

Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipellic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951.

| | |
|-----------|---|
| diat_size | <i>Calculate size classes for diatoms</i> |
|-----------|---|

Description

The input for these functions is the resulting dataframe obtained from the `diat_loadData()` function to calculate size classes for diatoms. Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipellic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Size class classification is obtained from:

- Rimet F. & Bouchez A., 2012. Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. *Knowledge and management of aquatic ecosystems*, 406: 1-14. DOI:10.1051/kmae/2012018

Usage

```
diat_size(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
sizeResults <- diat_size(df)
```

diat_sla

Calculates the Sladeczek Index (SLA)

Description

The input for all of these functions is the resulting dataframe (resultLoad) obtained from the diat_loadData() function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the diat_loadData() function so the acronyms and species' names are recognized. References for the index:

- Sládeček, V. (1986). Diatoms as indicators of organic pollution. Acta hydrochimica et hydrobiologica, 14(5), 555-566.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". Ecological Indicators, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_sla(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
slaResults <- diat_sla(df)
```

diat_spear

Calculates the SPEAR(herbicides) Index (SPEAR)

Description

The input for all of these functions is the resulting dataframe (resultLoad) obtained from the diat_loadData() function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the diat_loadData() function so the acronyms and species' names are recognized. References for the index:

- Wood, R. J., Mitrovic, S. M., Lim, R. P., Warne, M. S. J., Dunlop, J., & Kefford, B. J. (2019). Benthic diatoms as indicators of herbicide toxicity in rivers—A new SPEcies At Risk (SPEARherbicides) index. *Ecological Indicators*, 99, 203-213.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_spear(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
```

```
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
spearResults <- diat_spear(df)
```

diat_tdi

Calculates the Trophic (TDI) index

Description

The input for all of these functions is the resulting dataframe (resultLoad) obtained from the diat_loadData() function. A CSV or dataframe cannot be used directly with these functions, they have to be loaded first with the diat_loadData() function so the acronyms and species' names are recognized. References for the index:

- Kelly, M. G., & Whitton, B. A. (1995). The trophic diatom index: a new index for monitoring eutrophication in rivers. *Journal of Applied Phycology*, 7(4), 433-444.

Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cocheró, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Usage

```
diat_tdi(resultLoad)
```

Arguments

resultLoad The resulting list obtained from the diat_loadData() function

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
tdiResults <- diat_tdi(df)
```

| | |
|-------------|--|
| diat_vandam | <i>Calculates ecological information for diatoms based on the Van Dam classification</i> |
|-------------|--|

Description

The input for these functions is the resulting dataframe obtained from the `diat_loadData()` function, to calculate ecological information for diatoms based on the Van Dam classification. Sample data in the examples is taken from:

- Nicolosi Gelis, María Mercedes; Cochero, Joaquín; Donadelli, Jorge; Gómez, Nora. 2020. "Exploring the use of nuclear alterations, motility and ecological guilds in epipelagic diatoms as biomonitoring tools for water quality improvement in urban impacted lowland streams". *Ecological Indicators*, 110, 105951. <https://doi.org/10.1016/j.ecolind.2019.105951>

Van Dam classification is obtained from:

- Van Dam, H., Mertens, A., & Sinkeldam, J. (1994). A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. *Netherlands Journal of Aquatic Ecology*, 28(1), 117-133.

Usage

```
diat_vandam(resultLoad, vandamReports = TRUE)
```

Arguments

| | |
|----------------------------|---|
| <code>resultLoad</code> | The resulting list obtained from the <code>diat_loadData()</code> function |
| <code>vandamReports</code> | Boolean. If set to 'TRUE' the detailed reports for the Van Dam classifications will be reported in the Output. Default = TRUE |

Examples

```
# Example using sample data included in the package (sampleData):
data("diat_sampleData")
# First, the diat_loadData() function has to be called to read the data
# The data will be stored into a list (loadedData)
# And an output folder will be selected through a dialog box if resultsPath is empty
# In the example, a temporary directory will be used in resultsPath
df <- diat_loadData(diat_sampleData, resultsPath = tempdir())
vandamResults <- diat_vandam(df)
```

epid

EPID

Description

Index values for diatom species included in the EPID index

Usage

`data(epid)`

Format

A data frame with the ecological values for 1038 species

Source

<https://www.tib.eu/en/search/id/BLCP:CN034949165/Use-of-algae-for-monitoring-rivers-in-the-Czech?chash=fdd9e0b1bf812a31ec0f692a273cab04>

References

Dell'Uomo, A. (1996). Assessment of water quality of an Apennine river as a pilot study for diatom-based monitoring of Italian watercourses. Use of algae for monitoring rivers, 65-72.

idap

IDAP

Description

Index values for diatom species included in the IDAP index

Usage

`data(idap)`

Format

A data frame with the ecological values for 194 species

Source

<https://link.springer.com/article/10.1007/BF00028033>

References

Prygiel, J., & Coste, M. (1993). The assessment of water quality in the Artois-Picardie water basin (France) by the use of diatom indices. *Hydrobiologia*, 269(1), 343-349.

idch

ID-CH

Description

Index values for diatom species included in the IC-CH index

Usage

```
data(idch)
```

Format

A data frame with the ecological values for 550 species

Source

<https://www.bafu.admin.ch/bafu/fr/home/themes/eaux/publications/publications-eaux/methodes-analyse-appreciation-cours-eau-diatomees.html>

References

Hürlimann J., Niederhauser P. (2007). Méthodes d'analyse et d'appréciation des cours d'eau. Diatomées Niveau R (région). État de l'environnement n° 0740. Office fédéral de l'environnement, Berne. 132 p

idp

IDP

Description

Index values for diatom species included in the IDP index

Usage

```
data(idp)
```

Format

A data frame with the ecological values for 475 species

Source

<https://link.springer.com/article/10.1023/A:1011415209445>

References

Gómez, N., & Licursi, M. (2001). The Pampean Diatom Index (IDP) for assessment of rivers and streams in Argentina. *Aquatic Ecology*, 35(2), 173-181.

ilm

ILM

Description

Index values for diatom species included in the ILM index

Usage

`data(ilm)`

Format

A data frame with the ecological values for 667 species

Source

<http://www.vliz.be/imisdocs/publications/286641.pdf>

References

Leclercq, L., & Maquet, B. (1987). Deux nouveaux indices diatomique et de qualité chimique des eaux courantes. Comparaison avec différents indices existants. *Cahier de Biology Marine*, 28, 303-310

ips

IPS

Description

Index values for diatom species included in the IPS index

Usage

`data(ips)`

Format

A data frame with the ecological values for 6881 species

Source

<https://www.oieau.org/eaudoc/notice/ETUDE-DES-METHODES-BIOLOGIQUES-DAPPRECIATION-QUANTITATIVE-DE-LA>

References

Coste, M. (1982). Étude des méthodes biologiques d'appréciation quantitative de la qualité des eaux. Rapport Cemagref QE Lyon-AF Bassin Rhône Méditerranée Corse.

lobo

LOBO

Description

Index values for diatom species included in the LOBO index

Usage

data(lobo)

Format

A data frame with the ecological values for 297 species

References

Lobo, E. A., Callegaro, V. L. M., & Bender, E. P. (2002). Utilização de algas diatomáceas epilíticas como indicadoras da qualidade da água em rios e arroios da Região Hidrográfica do Guaíba, RS, Brasil. Edunisc.

sla

SLA

Description

Index values for diatom species included in the SLA index

Usage

data(sla)

Format

A data frame with the ecological values for 976 species

Source

<https://onlinelibrary.wiley.com/doi/abs/10.1002/ahch.19860140519>

References

Sládeček, V. (1986). Diatoms as indicators of organic pollution. Acta hydrochimica et hydrobiologica, 14(5), 555-566.

| | |
|-------|-----------------|
| spear | <i>SPEAR(h)</i> |
|-------|-----------------|

Description

Index values for diatom species included in the SPEAR(h) index

Usage

data(spear)

Format

A data frame with the ecological values for 285 species

Source

<https://www.sciencedirect.com/science/article/abs/pii/S1470160X18309683>

References

Wood, R. J., Mitrovic, S. M., Lim, R. P., Warne, M. S. J., Dunlop, J., & Kefford, B. J. (2019). Benthic diatoms as indicators of herbicide toxicity in rivers—A new SPEcies At Risk (SPEARherbicides) index. *Ecological Indicators*, 99, 203-213.

| | |
|-----|------------|
| tdi | <i>TDI</i> |
|-----|------------|

Description

Index values for diatom species included in the TDI index

Usage

data(tdi)

Format

A data frame with the ecological values for 3445 species

Source

<https://link.springer.com/article/10.1007/BF00003802>

References

Kelly, M. G., & Whitton, B. A. (1995). The trophic diatom index: a new index for monitoring eutrophication in rivers. *Journal of Applied Phycology*, 7(4), 433-444.

Index

* bioindicator

- diat_des, 7
- diat_diversity, 8
- diat_epid, 9
- diat_guilds, 10
- diat_idap, 11
- diat_idch, 12
- diat_idp, 13
- diat_ilm, 14
- diat_ips, 15
- diat_loadData, 16
- diat_lobo, 17
- diat_morpho, 18
- diat_size, 19
- diat_sla, 20
- diat_spear, 21
- diat_tdi, 22
- diat_vandam, 23
- diaThorAll, 6

* biotic

- diat_des, 7
- diat_diversity, 8
- diat_epid, 9
- diat_guilds, 10
- diat_idap, 11
- diat_idch, 12
- diat_idp, 13
- diat_ilm, 14
- diat_ips, 15
- diat_loadData, 16
- diat_lobo, 17
- diat_morpho, 18
- diat_size, 19
- diat_sla, 20
- diat_spear, 21
- diat_tdi, 22
- diat_vandam, 23
- diaThorAll, 6

* datasets

- dbc_offline, 3
- des, 3
- diat_sampleData, 19
- epid, 24
- idap, 24
- idch, 25
- idp, 25
- ilm, 26
- ips, 26
- lobo, 27
- sla, 27
- spear, 28
- tdi, 28

* diatom

- diat_des, 7
- diat_diversity, 8
- diat_epid, 9
- diat_guilds, 10
- diat_idap, 11
- diat_idch, 12
- diat_idp, 13
- diat_ilm, 14
- diat_ips, 15
- diat_loadData, 16
- diat_lobo, 17
- diat_morpho, 18
- diat_size, 19
- diat_sla, 20
- diat_spear, 21
- diat_tdi, 22
- diat_vandam, 23
- diaThorAll, 6

* ecology

- diat_des, 7
- diat_diversity, 8
- diat_epid, 9
- diat_guilds, 10
- diat_idap, 11
- diat_idch, 12

- diat_idp, 13
- diat_ilm, 14
- diat_ips, 15
- diat_loadData, 16
- diat_lobo, 17
- diat_morpho, 18
- diat_size, 19
- diat_sla, 20
- diat_spear, 21
- diat_tdi, 22
- diat_vandam, 23
- diaThorAll, 6

- dbc_offline, 3
- des, 3
- diat_des, 7
- diat_diversity, 8
- diat_epid, 9
- diat_guilds, 10
- diat_idap, 11
- diat_idch, 12
- diat_idp, 13
- diat_ilm, 14
- diat_ips, 15
- diat_loadData, 16
- diat_lobo, 17
- diat_morpho, 18
- diat_sampleData, 19
- diat_size, 19
- diat_sla, 20
- diat_spear, 21
- diat_tdi, 22
- diat_vandam, 23
- diaThor, 4
- diaThorAll, 6

- epid, 24

- idap, 24
- idch, 25
- idp, 25
- ilm, 26
- ips, 26

- lobo, 27

- sla, 27
- spear, 28

- tdi, 28