

Package: andurinha (via r-universe)

August 25, 2024

Type Package

Title Make Spectroscopic Data Processing Easier

Version 0.0.2

Description The goal of 'andurinha' is provide a fast and friendly way to process spectroscopic data. It is intended for processing several spectra of samples with similar composition (tens to hundreds of spectra). It compiles spectroscopy data files, produces standardized and second derivative spectra, finds peaks and allows to select the most significant ones based on the second derivative/absorbance sum spectrum. It also provides functions for graphic evaluation of the outputs.

Depends R (>= 3.5.0)

License GPL-2 | file LICENSE

URL <https://github.com/noemiallefs/andurinha>

BugReports <https://github.com/noemiallefs/andurinha/issues>

Encoding UTF-8

LazyData true

Imports signal, tidyr, ggplot2, cowplot, rlang, utils, plyr

Suggests extrafont, dplyr, knitr, rmarkdown, testthat, MASS

RoxygenNote 7.1.1

VignetteBuilder knitr

Repository <https://noemiallefs.r-universe.dev>

RemoteUrl <https://github.com/noemiallefs/andurinha>

RemoteRef HEAD

RemoteSha 6bbf8b2a1ea1bca378bcdeb6b9adac50e6cdd9c0

Contents

andurinha	2
andurinhaData	2
findPeaks	3
gOverview	4
importSpectra	5
plotPeaks	6

Index	8
--------------	----------

andurinha	<i>andurinha: Tools to make spectroscopic data processing easier</i>
-----------	--

Description

This package contains a set of functions that makes spectroscopic data processing easier and faster. It is intended for processing several spectra (tens to hundreds) of samples with similar composition. It compiles spectroscopy data files, produces standardised and second derivative spectra, finds peaks and allows to select the most significant ones based on the second derivative/absorbance sum spectrum. It also provides functions for graphic evaluation of the outputs.

andurinha functions

1. **importSpectra**: in case you have your spectra in separated files (.csv) this function imports and bind them in a single data frame.
2. **findPeaks**: finds peaks and allows to select the most relevant based on the second derivative sum spectrum.
3. **gOverview**: generates a graphic overview of the spectroscopic data.
4. **plotPeaks**: makes a graphic representation of the peaks over the second derivative/absorbance sum spectrum.

andurinhaData	<i>Andurinha data set</i>
---------------	---------------------------

Description

A collection of FTIR-ATR mid-infrared spectra of peat samples with increasing degree of peat humification.

Usage

andurinhaData

Format

A data frame with 1736 observations of 3 peat samples.

- *WN*: wave numbers of the spectra.
- *A-C*: absorbances of a FTIR spectra of three peat samples.

Source

EcoPast research group, Universidade de Santiago de Compostela (Spain) <https://ecopast.es>

See Also

[importSpectra](#), [findPeaks](#), [plotPeaks](#) and [gOverview](#)

findPeaks

findPeaks

Description

This function finds peaks and allows to the most relevant based on the second derivative/absorbance sum spectrum.

Usage

```
findPeaks(  
  data,  
  resolution = 4,  
  minAbs = 0.1,  
  cutOff = NULL,  
  scale = TRUE,  
  ndd = TRUE  
)
```

Arguments

<code>data</code>	A data frame object, which contains in the first column the wave numbers and in the following columns the samples absorbances.
<code>resolution</code>	The equipment measurement resolution; by default 4 cm-1.
<code>minAbs</code>	The cut off value to check spectra quality; by default 0.1.
<code>cutOff</code>	The second derivative/absorbance sum spectrum cut off to reduce the raw peaks table; by default NULL.
<code>scale</code>	By default (TRUE) the data is scaled by Z-scores. Use FALSE in case you do not want to scale it.
<code>ndd</code>	By default (TRUE) the peaks are searched based on the second derivative sum spectrum. Use FALSE in case you want to search them based on the absorbance sum spectrum.

Value

A list with a collection of data frames which contains:

1. *dataZ*: the standardised data by Z-scores.
2. *secondDerivative*: the second derivative values of the data.
3. *sumSpectrum_peaksTable*: the peaks wave numbers and their second derivative/absorbance sum spectrum values.
4. *peaksTable*: the selected peaks wave numbers and their absorbance for each spectrum.

See Also

[importSpectra](#), [gOverview](#) and [plotPeaks](#)

Examples

```
# Find Peaks based on the absorbance sum spectrum
fp.abs <- findPeaks(andurinhaData, ndd = FALSE)

# See the peaks table of the absorbance sum spectrum
fp.abs$sumSpectrum_peaksTable

# Find Peaks based on the second derivative sum spectrum
fp.ndd <- findPeaks(andurinhaData)

# See the peaks table of the second derivative sum spectrum
fp.ndd$sumSpectrum_peaksTable

# Select a cutOff to reduce the number of peaks in the table
# (i.e. select the most relevant)
# fp.ndd$sumSpectrum_peaksTable %>%
#   arrange(desc(sumSpectrum))
# Run findPeaks() with the new cutOff
fp.ndd2 <- findPeaks(andurinhaData, cutOff = 0.25)
```

gOverview

gOverview

Description

This function generates a graphic overview of the spectroscopic data.

Usage

```
gOverview(data_abs, data_ndd, fontFamily = NULL)
```

Arguments

data_abs	A data frame, which contains in the first column the wave numbers and in the following columns the samples absorbances.
data_ndd	A data frame, which contains in the first column the wave numbers and in the following columns the samples second derivative values.
fontFamily	The desired graphic font family.

Value

If data_ndd is provided: It returns a grid with three plots:

- The ensemble of all samples spectra.
- The ensemble of the second derivative spectra of all samples.
- The average and standard deviation spectra.

If data_ndd is omitted: It returns a grid with two plots:

- The ensemble of all samples spectra.
- The average and standard deviation spectra.

See Also

[importSpectra](#), [findPeaks](#) and [plotPeaks](#)

Examples

```
# Grapic overview of your raw data
gOverview(andurinhaData)

# Graphic overview of your processed data by findPeaks()
fp <- findPeaks(andurinhaData)
gOverview(fp$dataZ, fp$secondDerivative)
```

importSpectra	<i>importSpectra</i>
---------------	----------------------

Description

In case you have your spectra in separated files (.csv) this function imports and binds them in a single data frame. The files directory must contain only the samples files.

Usage

```
importSpectra(path, sep = ";")
```

Arguments

path	A character vector with the full path to the data directory; by default corresponds to the working directory, getwd .
sep	The field separator character; by default sep = ";".

Value

A data frame with the structure:

- *First column (WN)*: wave numbers of the spectra.
- *1-n*: samples spectra (the column names correspond to the files names).

See Also

[findPeaks](#), [gOverview](#) and [plotPeaks](#)

Examples

```
# Create an empty directory
# Now create some spectra separate files
A <- andurinhaData[, 1:2]
B <- andurinhaData[, c(1, 3)]
C <- andurinhaData[, c(1, 4)]

MASS::write.matrix(A, file = tempfile(pattern = "A.csv"), sep = ";")
MASS::write.matrix(B, file = tempfile(pattern = "B.csv"), sep = ";")
MASS::write.matrix(C, file = tempfile(pattern = "C.csv"), sep = ";")

# Try
importSpectra(path = paste0(tempdir(), "/"), ";")
```

plotPeaks

plotPeaks

Description

This function makes a graphic representation of the peaks over the second derivative and/or absorbance sum spectra.

Usage

```
plotPeaks(peaksWN, data_abs, data_ndd, fontFamily = NULL)
```

Arguments

<code>peaksWN</code>	A vector with the peaks wave numbers.
<code>data_abs</code>	A data frame, which contains in the first column the wave numbers and in the following columns the samples absorbances.
<code>data_ndd</code>	A data frame, which contains in the first column the wave numbers and in the following columns the samples second derivative values.
<code>fontFamily</code>	The desired graphic font family.

See Also

[importSpectra](#), [findPeaks](#) and [gOverview](#)

Examples

```
# Plot the peaks found by findPeaks()
# 1. Based on absorbance sum spectrum
fp.abs <- findPeaks(andurinhaData, ndd = FALSE)
plotPeaks(fp.abs[[3]]$WN, fp.abs$dataZ)

# 2. Based on second derivative spectrum
fp.ndd <- findPeaks(andurinhaData, cutOff = 0.25)
plotPeaks(fp.ndd[[4]]$WN, fp.ndd$dataZ, fp.ndd$secondDerivative)
```

Index

* datasets

andurinhaData, 2

andurinha, 2

andurinhaData, 2

findPeaks, 2, 3, 3, 5–7

getwd, 6

gOverview, 2–4, 4, 6, 7

importSpectra, 2–5, 5, 7

plotPeaks, 2–6, 6