

Package ‘ColonyTrack’

January 31, 2024

Type Package

Title Analysis of Multi-Subject Tracking Data from Interconnected Cage Networks

Version 1.0.5

Maintainer Rupert Overall <rtrack@rupertoverall.net>

Description

This package prepares raw timestamped Radio Frequency ID data into a standardised storage format, infers the position of each subject at each time point and extracts behavioural metrics.

License GPL-3

Encoding UTF-8

LazyData true

ByteCompile true

RoxygenNote 7.2.3

biocViews behaviour

Imports crayon,

igraph,
IRanges,
pbapply,
Rfast,
fastmatch

R topics documented:

calculate_metrics	2
ColonyTrack	3
combine_metrics	4
get_metrics	4
plot_ethogram	5
plot_metric	6
read_data	7

Index	9
--------------	----------

calculate_metrics *Calculation of tracking metrics.*

Description

Calculates metrics from cage location data.

Usage

```
calculate_metrics(
  data,
  days = "all",
  drop.days = NULL,
  subjects = "all",
  drop.subjects = NULL,
  trim = c(0, 0),
  estimate = FALSE,
  log = FALSE,
  n.cores = parallel::detectCores() - 1
)
```

Arguments

data	A colonytrack_data object, generated by read_data .
days	Specify the days to be included is the metrics calculation. This may be in any of the standard formats (names, indices, or a logical vector). Default is to include all days present in the input data file.
drop.days	Specify the days to be <code>_excluded_</code> is the metrics calculation. This takes precedence over the days parameter. This may be in any of the standard formats (names, indices, or a logical vector). Default is not to exclude any days.
subjects	Specify the subjects to be included is the metrics calculation. This may be in any of the standard formats (names, indices, or a logical vector). Default is to include all subjects present in the input data file.
drop.subjects	Specify the subjects to be <code>_excluded_</code> is the metrics calculation. This takes precedence over the subjects parameter. This may be in any of the standard formats (names, indices, or a logical vector). Default is not to exclude any subjects.
trim	An integer vector of length 2, specifying the number of hours (in Zeitgeber time) to be trimmed off the ends of the metrics calculation period. This may be useful to allow focus on a core segment of nightly activity. Default is no trimming.
estimate	Logical (default FALSE) Whether an estimate of total runtime should be shown as the calculation proceeds. The calculation of metrics on a large dataset can take many hours, so thus can be a useful indicator. Performance is, however, better when this option is set to false.
log	Logical (default FALSE) indicating whether a file should be created for each job. This is useful for monitoring large experiments and for estimating total run time. Note that a progress bar is not provided as the computational overhead is too high.

`n.cores` The number of CPU cores to use. The default tries to estimate an optimal value but, especially on Linux, it may be better to set this manually based on the available hardware. If `n.cores` is set higher than the number of days to be calculated, then the cluster will be created with only one node per day.

Details

This function takes the pre-processed data from [read_data](#) and calculates a range of behavioural metrics. Metrics are calculated daily, with the exception of the ethogram data which are calculated hourly.

Value

An `colonytrack_metrics` object. The `calls` element contains the called strategy/strategies as well as confidence scores for all possible strategies.

See Also

[read_data](#).

ColonyTrack	<i>Analysis of Multi-Subject Tracking Data from Interconnected Cage Networks.</i>
-------------	---

Description

This package prepares raw timestamped Radio Frequency ID data into a standardised storage format, infers the position of each subject at each time point and extracts behavioural metrics.

Functions provided

[read_data](#)

[calculate_metrics](#)

[combine_metrics](#)

[get_metrics](#)

[plot_metric](#)

[plot_ethogram](#)

combine_metrics	<i>Combine metrics objects.</i>
-----------------	---------------------------------

Description

Joins metrics results into one contiguous metrics object.

Usage

```
combine_metrics(metrics.list)
```

Arguments

`metrics.list` A list containing any number of `colonytrack_metrics` objects in any order. These will be merged into one combined `colonytrack_metrics` object sorted by date. This parameter may also be a character list of filenames pointing to saved RData archives containing the `colonytrack_metrics` objects.

Details

This function takes processed results from several `calculate_metrics` runs and merges them into one `colonytrack_metrics` object. This allows metrics to be calculated in batches to limit memory usage.

Value

A `colonytrack_metrics` object.

See Also

[calculate_metrics](#).

get_metrics	<i>Retrieve one or more metrics.</i>
-------------	--------------------------------------

Description

Retrieves data for all individual metrics for a single subject.

Usage

```
get_metrics(subject, metrics, days = "all")
```

Arguments

`subject` The name of the subject for which the data is to be retrieved.
`metrics` A `colonytrack_metrics` object, as returned by [calculate_metrics](#).
`days` Specify the days for which metrics should be retrieved. This may be in any of the standard formats (names, indices, or a logical vector). Default is to include all days present in the metrics object.

Value

Returns a matrix with days/nights as rows and metrics as columns.

plot_ethogram	<i>Plot an ethogram.</i>
---------------	--------------------------

Description

Plots a heatmap representation of the high-level ethogram metrics.

Usage

```
plot_ethogram(
  metrics,
  days = "all",
  subjects = "all",
  plot = "all",
  scheme = "yrb",
  file = NULL,
  width = NULL,
  height = NULL,
  cex.axis = 0.75,
  las = 2,
  mar = c(5, 8, 0, 0)
)
```

Arguments

metrics	A colonytrack_metrics object, as returned by calculate_metrics .
days	Specify the days to be included in the metrics calculation. This may be in any of the standard formats (names, indices, or a logical vector). Default is to include all days present in the input data file.
subjects	Specify the subjects to be included in the metrics calculation. This may be in any of the standard formats (names, indices, or a logical vector). Default is to include all subjects present in the input data file. If a vector of subjects is supplied, the resulting plot will respect the ordering of this vector.
plot	A character vector describing which behavioural categories ('activity', 'sociality' and 'exploration') should be plotted. Abbreviations may also be used. The default, 'all', plots all three categories mixed using a colour model determined by the parameter 'scheme'. Note that if 'all' is present in any of the elements of this argument, then it will override any other elements. If plot is FALSE, then all plotting is suppressed and only the ethogram colours and components are returned (see below).
scheme	A code specifying the colour scheme to be used. See Details for information on the available schemes. Default is 'yrb'.
file	An optional filename for PDF output. If supplied, the plot will be written to a PDF document with this name. The file is saved relative to the working directory. The default is to plot to the active graphics device.
width	The width of the PDF (only if file is specified) in mm.

height	The height of the PDF (only if file is specified) in mm.
cex.axis	The scaling factor for the axis labels (see par). Default (0.75) can be overridden for different sized plot devices.
las	The label aspect (see par). Can be overridden based on the user's preferences.
mar	The margins surrounding the plot (see par). Can be overridden for different length ID labels or to match different sized plot devices.

Details

This function uses the three high-level metrics 'activity', 'sociality' and 'exploration' and produces a heatmap representation for all subjects over the given time period.

Available colour schemes are 'rgb' ('activity' in the red channel, 'sociality' in the green channel and 'exploration' in the blue channel). Scheme variants beginning with 'd' use discrete colouring such that the most prevalent activity in any time block is taken as the colour (the default is to mix the components). The scheme family 'ymc' uses yellow for 'activity', magenta for 'sociality' and cyan for 'exploration'. The discrete and scaled versions are also available for this scheme. The (default) scheme family 'yrb' uses yellow for 'activity', red for 'sociality' and blue for 'exploration' in a subtractive colour model. The activity and sociality scores are scaled and thresholded to fall within the range 0-1. The scheme 'basic' uses a discrete colour mapping based on a subtractive colouring so that yellow = 'activity', red = 'sociality', blue = "exploration", orange = 'activity' and 'sociality', green = 'activity' and 'exploration', purple = 'sociality' and 'exploration', brown indicates the presence of all three features and white the absence of all features.

Value

Invisibly returns a list with the components 'ethogram' (a matrix of the colours of each cell), and 'activity', 'sociality' and 'exploration' (matrices with the values for each component respectively).

plot_metric	<i>Plot one or more metrics.</i>
-------------	----------------------------------

Description

Plots a line graph of the individual metrics over time.

Usage

```
plot_metric(
  metric,
  metrics,
  days = "all",
  subjects = "all",
  scheme = "rainbow",
  col = NULL,
  file = NULL,
  cex.axis = 1,
  las = 2,
  mar = c(8, 8, 4, 0),
  plot = TRUE
)
```

Arguments

metric	The name of the metric that is to be plotted.
metrics	A colonytrack_metrics object, as returned by calculate_metrics .
days	Specify the days to be included in the metrics calculation. This may be in any of the standard formats (names, indices, or a logical vector). Default is to include all days present in the input data file.
subjects	Specify the subjects to be included in the metrics calculation. This may be in any of the standard formats (names, indices, or a logical vector). Default is to include all subjects present in the input data file. If a vector of subjects is supplied, the resulting plot will respect the ordering of this vector.
scheme	A code specifying the colour scheme to be used. See Details for information on the available schemes. Default is 'rainbow'.
col	A named vector of colours. This vector must contain the SubjectIDs as names. Any subjects not present in this vector will be drawn in black. If not NULL, then this parameter overrides scheme. Default is NULL.
file	An optional filename for PDF output. If supplied, the plot will be written to a PDF document with this name. The file is saved relative to the working directory. The default is to plot to the active graphics device.
cex.axis	The scaling factor for the axis labels (see par). Default (0.75) can be overridden for different sized plot devices.
las	The label aspect (see par). Can be overridden based on the user's preferences.
mar	The margins surrounding the plot (see par). Can be overridden for different length ID labels or to match different sized plot devices.
plot	If FALSE, then the plot is not drawn and only the plotting data are returned.

Details

The data used in plotting (the plotted values for each subject and the colours) are returned invisibly to enable further plot customisation.

read_data	<i>Preparation of tracking data.</i>
-----------	--------------------------------------

Description

Reads raw antenna contact data and metadata files.

Usage

```
read_data(
  dataFiles,
  subjectFile,
  networkFile,
  eventsFile,
  cageQualityFile = NULL
)
```

Arguments

dataFiles	A vector of filenames. Each file is a CSV dump of antenna contact data. These files should cover a contiguous period encompassing the whole experiment. Files can be given as full pathnames so that the often large data volumes can be stored outside the analysis directory (for example on an external fileserver). Any file format readable by read.table can be used.
subjectFile	A tab-delimited file describing the subjects/animals. This information is required to link RFID tags to variables/factors such as experimental group, age, weight etc.
networkFile	A tab-delimited file describing the layout of the cage system. This is currently in the the form of an unweighted undirected network.
eventsFile	A tab-delimited file describing events. This must minimally include details of the light cycle. See Details for more information.
cageQualityFile	An optional tab-delimited file describing the content of the cages (food, bedding, toys etc.). See Details for more information.

Details

This function requires raw CSV data from the Colony Cage as well as a subject/animal description file, a cage layout description file and a log of events. An optional metadata file allows finer description of the cages.

The processed data are padded to start at "lights on" of the first day to "lights off" of the last day (as specified in the 'Events' file). Nevertheless, the first and last days (and any interruptions in recording) will result in missing data that may need to be handled during analysis. The actogram (produced by [plot](#) of the colonytrack_data object) can be used to identify such missing blocks.

Value

A colonytrack_data object.

Index

`calculate_metrics`, 2, 3–5, 7

`ColonyTrack`, 3

`combine_metrics`, 3, 4

`get_metrics`, 3, 4

`par`, 6, 7

`plot`, 8

`plot_ethogram`, 3, 5

`plot_metric`, 3, 6

`read.table`, 8

`read_data`, 2, 3, 7