

The metrics object

The `colonytrack_metrics` object contains features/metrics calculated from the tracking data. The data are segmented by day (technically by ‘window’; see the following section) and processed in parallel to extract a range of metrics describing the activity of the subjects and their interactions with each other. Details of the resulting `colonytrack_metrics` object are given below.

Calculation of the metrics

Time windows

The metrics are calculated for blocks of time, referred to as ‘windows’. In the current version of the software, these windows default to a 12-hour period during the dark phase (ZT12–ZT24). Support for calculation of light phase metrics (and limited support for customised window intervals) will be added to an upcoming version. An additional parameter `trim` allows for hours to be trimmed off the start and end of the metrics calculation period. This makes it possible to remove the first and last hours of data—in the situation that the light cycle is not strictly maintained for example. This feature could also enable the user to focus on a particular period of interest—although it is recommended to perform such trimming during later analysis of the metrics results. The `trim` parameter will only accept integer values (rounded hours) and care should be taken in interpreting the results if this feature is used. The windows are described in the `windows` section of the `info` component and consist of an ID (by default the date in YYYY-MM-DD format) and a start and end time as UNIX timestamps.

Cage layout

The cage layout is described in the `cage.layout` section of the `info` component as a list of `igraph` objects (see the [data object description](#) for additional background). This list, with one cage layout per day, allows for the flexibility to include modifications of the cage layout during an experiment. The `igraph` objects are passed through from the `colonytrack_data` object and contain the same information.

Components of the metrics object

Features vs. individual metrics/variables

The `calculate_metrics` function generates a list of 109 variables, many of which are derivatives (e.g. mean, median, variance) of the same information (see the detailed description below for a list of all metrics that are calculated). All of these values describe the underlying data in subtly different ways and are thus useful for certain applications such as machine learning. Reflecting this intended purpose, these variables are referred to in the metrics object as ‘features’. A list of the feature names can also be obtained computationally by querying a `colonytrack_metrics` object and inspecting the `metrics$info$feature.names` component.

For general use, however, the large collection of partially redundant variables is not helpful, so a select subset is returned in the `individual` component of the metrics object. The names of the individual metrics can be recalled from the `metrics$info$var.names` component of a `colonytrack_metrics` object. These individual metrics are accessible for plotting by the `plot_metric` function. For a detailed description of each of the individual metrics, please refer to the separate [metrics details vignette](#).

Cage use

The `cage.use` component holds a list for each day of tracking which is further subdivided into each cage; where the proportion of time spent in each cage by each subject is recorded. Thus the values for the first

subject over all cages in one day should sum to 1. This enables the time spent in a cage of interest (e.g. a cage with water access) to be easily calculated and can provide the basis for a customised analysis of your cage set-up.

Ethogram

The ethogram is calculated hourly and is used as the basis for the built-in `plot_ethogram` function. The ethogram uses 3 summary metrics to encapsulate the ‘activity’, ‘exploration’ and ‘sociality’ dimensions of the subjects behaviour. Although this does not make use of the full multi-parametric power of ColonyTrack, it provides a rapid and intuitive overview (for users with full colour vision) of the behavioural patterns for very many subjects over long periods of time. The `ethogram` component makes the numeric data underlying the ethogram plot accessible to users—for example to enable custom plots to be generated. Split into a list by day, the three dimensions (‘activity’, ‘exploration’ and ‘sociality’) are available as a numeric matrix with 12 columns (one for each hour of the day) and a row for each subject. Users can, for example, build a custom one-colour heatmap for each dimension separately using an appropriate palette to provide colourblind-friendly output.

Clustering

The group-housing nature of the ColonyRack enables interactions between subjects to be recorded. Different to the individual metrics, these social interactions are reported as a subject x subject matrix. The group metrics that are calculated are `interaction.time` (the amount of time that each pair of subjects spend in the same cage), `social.distance` (the average number of cages separating each pair of subjects; technically, the mean shortest path in the cage network between the two subjects over the course of the day¹) and `following` (the number of times one subject [rows] follows another [columns]; see the section on ‘following’ below). With the exception of `following`, these interaction matrices are symmetrical (in graph theory parlance, they are undirected networks). The `following` network, on the other hand, is directed—the subject doing the following (the ‘aggressor’) is in the rows with the subject being followed (the ‘victim’) in the columns².

Following and dominance calculation

A follow event is defined as two subjects passing through the same tunnel in the same direction in rapid (less than 1 second delay) succession. The second subject passing through the tunnel is understood to be pursuing the first (this interpretation has been initially confirmed using video observation of the cage system). The directed nature of this interaction means that a hierarchy can be calculated by summing the ‘wins’ (the number of times a subject was the pursuer) for each pair of subjects. A subject with more ‘wins’ over another is taken to be dominant in the social hierarchy. The results are presented using this terminology for practical convenience—the actual dominance relationships identified by this method are interpretations from the automated data. The `dominance` score for each subject is pre-calculated for convenience from the follow network for each day. It is provided as a raw sum of ‘wins’ and ‘losses’ and can be converted to ranks by the user if desired³.

Structure of the metrics object

Below is an overview of the hierarchy of the `colonytrack_metrics` object together with the names and classes of each component. Where the class is not part of the R base package, the package providing it is given in square brackets after the class name.

```
metrics : colonytrack_metrics [ColonyTrack]
  info : list
```

¹The calculation of this component changed in version 1.0.4 to use weighted means—See the description of the `social.distance` metric for more information.

²Note that this has changed in v1.0.2 and was transposed in earlier ColonyTrack versions.

³Other scores, such as the commonly-used David’s score, can be calculated from the `following` tables in the `clustering` component. Future versions of ColonyTrack may offer multiple variants of the dominance score.

```

windows : colonytrack_windows [ColonyTrack]
  id : character
  start : numeric
  end : numeric
subjects : character
subject.info : data.frame
  SubjectID : character
  Tag : character
  (optional additional user-defined columns)
feature.names : character
var.names : character
cage.layout : list
  (window) : igraph [igraph]
processed : POSIXct
version : character
features : list
  (subject) : data.frame
    path.length : numeric
    max.time.in.cage : numeric
    mean.time.in.cage : numeric
    min.time.in.cage : numeric
    sd.time.in.cage : numeric
    upper.time.in.cage : numeric
    median.time.in.cage : numeric
    lower.time.in.cage : numeric
    time.high.activity : numeric
    time.sustained.activity : numeric
    time.active : numeric
    time.inactive : numeric
    total.activity.burst.time : numeric
    max.activity.burst.length : numeric
    max.activity.burst.time : numeric
    max.cage.use : numeric
    mean.cage.use : numeric
    min.cage.use : numeric
    sd.cage.use : numeric
    upper.cage.use : numeric
    median.cage.use : numeric
    lower.cage.use : numeric
    cage.time.entropy : numeric
    max.adj.cage.use : numeric
    mean.adj.cage.use : numeric
    min.adj.cage.use : numeric
    sd.adj.cage.use : numeric
    upper.adj.cage.use : numeric
    median.adj.cage.use : numeric
    lower.adj.cage.use : numeric
    adjusted.cage.time.entropy : numeric
    max.cage.transitions : numeric
    mean.cage.transitions : numeric
    min.cage.transitions : numeric
    sd.cage.transitions : numeric
    upper.cage.transitions : numeric

```

median.cage.transitions : numeric
lower.cage.transitions : numeric
cage.location.entropy : numeric
max.adj.cage.transitions : numeric
mean.adj.cage.transitions : numeric
min.adj.cage.transitions : numeric
sd.adj.cage.transitions : numeric
upper.adj.cage.transitions : numeric
median.adj.cage.transitions : numeric
lower.adj.cage.transitions : numeric
adjusted.cage.location.entropy : numeric
max.revisit.time : numeric
mean.revisit.time : numeric
min.revisit.time : numeric
sd.revisit.time : numeric
upper.revisit.time : numeric
median.revisit.time : numeric
lower.revisit.time : numeric
max.revisit.length : numeric
mean.revisit.length : numeric
min.revisit.length : numeric
sd.revisit.length : numeric
upper.revisit.length : numeric
median.revisit.length : numeric
lower.revisit.length : numeric
peak.inactive.timepoint : numeric
peak.active.timepoint : numeric
number.activity.blocks : numeric
max.cage.sharing : numeric
mean.cage.sharing : numeric
upper.cage.sharing : numeric
median.cage.sharing : numeric
lower.cage.sharing : numeric
time.alone : numeric
max.social.interaction.time : numeric
mean.social.interaction.time : numeric
min.social.interaction.time : numeric
sd.social.interaction.time : numeric
upper.social.interaction.time : numeric
median.social.interaction.time : numeric
lower.social.interaction.time : numeric
social.exposure : numeric
max.distance.from.all : numeric
mean.distance.from.all : numeric
min.distance.from.all : numeric
sd.distance.from.all : numeric
upper.distance.from.all : numeric
median.distance.from.all : numeric
lower.distance.from.all : numeric
max.distance.from.each : numeric
min.distance.from.each : numeric
sd.distance.from.each : numeric
upper.distance.from.each : numeric

```

median.distance.from.each : numeric
lower.distance.from.each : numeric
max.sharing.change : numeric
mean.sharing.change : numeric
min.sharing.change : numeric
sd.sharing.change : numeric
upper.sharing.change : numeric
median.sharing.change : numeric
lower.sharing.change : numeric
max.influence : numeric
mean.influence : numeric
min.influence : numeric
sd.influence : numeric
upper.influence : numeric
median.influence : numeric
lower.influence : numeric
number.follow.events : numeric
sd.follow.events : numeric
mean.follow.wins : numeric
sd.follow.wins : numeric
follow.sds : numeric
individual : list
  (subject) : data.frame
    distance.moved : numeric
    time.per.cage : numeric
    high.activity : numeric
    sustained.activity : numeric
    cage.variability : numeric
    cage.time.entropy : numeric
    adjusted.cage.time.entropy : numeric
    cage.location.entropy : numeric
    revisit.time : numeric
    revisit.length : numeric
    peak.inactive : numeric
    peak.active : numeric
    activity.blocks : numeric
    cage.sharing : numeric
    time.alone : numeric
    social.interaction : numeric
    social.distance : numeric
    social.gradient : numeric
    social.influence : numeric
    follow.events : numeric
    follow.dominance : numeric
cage.use : list
  (window) : data.frame
  (cage) : numeric
ethogram : list
  (window) : list
    activity : numeric
    exploration : numeric
    sociality : numeric
clustering : list

```

```
interaction.time : list
  (window) : numeric matrix
social.distance : list
  (window) : numeric matrix
following : list
  (window) : numeric matrix
dominance : list
  (window) : numeric
follow.events : list
  (window) : list
  (subject) : list
    (subject) : data.frame
      Timestamp : numeric
      Transition : character
development : list
  elapsed.time : numeric
  threads : list
    start.time : numeric
    end.time : numeric
    elapsed.time : numeric
```