Description of available datasets.

David Simons

2022-06-15

Data

Data has been obtained from a systematic literature search of studies reporting rodent trapping across West Africa. Studies have been performed for various reasons and the quality of data obtained from each study differs. Where studies have been conducted to investigate zoonoses data on pathogens have also been extracted.

The data has been divided into three datasets for you. First, a detection-non-detection dataset, second, an **abundance** dataset (not true ecological abundance), third a long-format pathogen dataset. In the following sections I will explain the variables.

Detection-non-detection

If a study identified a rodent species at any of the trap sites this has been labelled a detection. If the species was not identified at another trap-site but was detected within the study (i.e. if it was there they would have reported it) it is termed a non-detection.

```
## # A tibble: 5 x 13
##   unique_id      year_trapping month_trapping country iso3c region  town_village
##   <chr>          <chr>         <chr>          <chr>   <chr> <chr>   <chr>
## 1 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 2 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 3 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 4 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 5 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## # ... with 6 more variables: habitat <chr>, classification <chr>,
## #   trap_nights <dbl>, longitude <dbl>, latitude <dbl>, detection <chr>
```

Variable name	Interpretation
unique_id	This is the identifier of the study in which the data came from. This can be crosschecked against the studies page here
year_trapping	Where provided this is the period in which trapping occurred, this may have important implications due to land use change, population change, invasive rodents etc.
month_trapping	Where provided this is the period in which trapping occurred, this may have important implications due to seasonality

Variable name	Interpretation
country & iso3c	The country in which trapping occurred with harmonised ISO3C codes
region	Specified region in which trapping occurred
town_village	If given the locations of the study site
habitat	The described habitat type in which trapping occurred
classification	The descriptor of the trapped rodent at the lowest taxonomic level available
trap_nights	The number of trap nights that were surveyed at this location. This value gives a measure of the trapping effort. For some studies this has been imputed where not presented based on the number of trapped rodents within a habitat
longitude and latitude	The GPS coordinates of the trapping location at the highest geographical resolution provided
detection	Whether that species was detected or not-detected at that trapping location

Abundance

This reports the number of individuals from a single species that were trapped at the highest geographical and temporal resolution. If you are not interested in temporal changes these can be grouped by habitat type to give the total numbers of individuals trapped across all sessions within a study.

```
## # A tibble: 5 x 13
##   unique_id      year_trapping month_trapping country iso3c region  town_village
##   <chr>          <chr>         <chr>          <chr>   <chr> <chr>   <chr>
## 1 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 2 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 3 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 4 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## 5 fc_2010_guinea 2003          May            Guinea  GIN   Faranah Gbetaya
## # ... with 6 more variables: habitat <chr>, classification <chr>, number <dbl>,
## #   trap_nights <dbl>, longitude <dbl>, latitude <dbl>
```

Only additional data to above is the number (abundance) variable

Pathogens

I think you are less interested in this but perhaps this may help to focus on which rodent species you are interested in looking at in more depth. This data is in long format with each host-pathogen pair tested for separated out by study, geography, habitat and time.

```
## # A tibble: 5 x 15
##   unique_id       year_trapping month_trapping country iso3c region town_village
##   <chr>           <chr>         <chr>          <chr>   <chr> <chr>  <chr>
## 1 az_2020_burkina 2005-2006     Nov-Oct        Burkin~ BFA   <NA>   Ouagadougou
## 2 az_2020_burkina 2005-2006     Nov-Oct        Burkin~ BFA   <NA>   Ouagadougou
## 3 az_2020_burkina 2005-2006     Nov-Oct        Burkin~ BFA   <NA>   Ouagadougou
## 4 az_2020_burkina 2005-2006     Nov-Oct        Burkin~ BFA   <NA>   Ouagadougou
## 5 az_2020_burkina 2005-2006     Nov-Oct        Burkin~ BFA   <NA>   Ouagadougou
## # ... with 8 more variables: classification <chr>, record_id <dbl>,
```

```
## #   pathogen_tested <chr>, assay <chr>, number <dbl>, geometry <chr>,
## #   longitude <dbl>, latitude <dbl>
```

Variable name	Interpretation
unique_id	This is the identifier of the study in which the data came from. This can be crosschecked against the studies page here
year_trapping	Where provided this is the period in which trapping occurred, this may have important implications due to land use change, population change, invasive rodents etc.
month_trapping	Where provided this is the period in which trapping occurred, this may have important implications due to seasonality
country & iso3c	The country in which trapping occurred with harmonised ISO3C codes
region	Specified region in which trapping occurred
town_village	If given the locations of the study site
habitat	The described habitat type in which trapping occurred
classification	The descriptor of the trapped rodent at the lowest taxonomic level available
record_id	As several tests can be performed on the same rodent this value links all the tests for that same rodent species within the study (geography, habitat and trapping session) with one row for the assay and one row for the result for each pathogen
pathogen_tested	The lowest taxonomic name available for the pathogen tested as reported by the study
assay	The type of assay used for the pathogen (i.e. PCR or Ab/Ag)
number	The number of individuals tested or positive based on the value in the assay variable
longitude and latitude	The GPS coordinates of the trapping location at the highest geographical resolution provided