

that it will become a central knowledge and data-sharing centre for the camera trapping community.

## 4.6 The future: more repositories, better data management and analytical services

We envision a future for camera trapping where there are more and more camera trap repositories and databases being managed by a wide variety of individuals and organisations. As a result the software solutions mentioned here (as well as potentially many others not described here) will provide the community with the tools it needs to manage and analyse camera trap data. The creation of repositories that are shareable with efforts such as WI and other data-discovery solutions (e.g. Global Biodiversity Information Facility, <http://www.gbif.org>, and DataOne, <https://www.dataone.org>) will ensure that valuable data is not lost and can be maximally utilised by the broader community.

Video also presents an additional challenge for camera trap tools, both in terms of increased storage demands as well as the transformation of video into data points. Fortunately many camera traps currently enable researchers to take sequences of images (from a camera trap trigger) and the analytical approaches being developed account for these sequences. These approaches will be very relevant to accurately handling video and turning this into actionable data.

As a parallel to the growth of camera trap repositories, we envision better web and data services around camera trap data. For example, data services could provide climate information to complement camera trap data for a specific location. The capability to embed automated image identification and individual identification into camera trap data management tools and/or online repositories can make species/individual identification more efficient and provide an extra layer of quality control (Yu *et al.* 2013).

## References

- Fegraus, E., Lin, K., Ahumada, J., Baru, C., Chandra, S. and Youn, C. (2011) Data acquisition and management software for camera trap data: a case study from the TEAM Network. *Ecological Informatics* 6: 345–353.
- Harris, G., Thompson, R., Childs, J. and Sanderson, J. (2010) Automatic storage and analysis of camera trap data. *Bulletin of the Ecological Society of America* 91: 352–360.
- Yu, X., Wang, J., Kays, R., Jansen, P., Wang, T. and Huang, T. (2013) Automated identification of animal species in camera trap images. *EURASIP Journal on Image and Video Processing* 2013: 52.

# 5. Presence/absence and species inventory

Francesco Rovero and Daniel Spitale

## 5.1 Introduction

This chapter addresses descriptive analysis from camera trapping surveys aimed at assessing the presence of target species, groups of species, or the whole community of terrestrial vertebrates (hence effectively compiling a faunal inventory for the study area). According to a scheme that will be followed in all the analytical chapters of this book, the first part of the chapter provides the theoretical background. We first introduce the raw metrics that are derived from faunal surveys, namely the observed species richness, the number of photographic events, and the relative abundance index (RAI) for each species. These are useful descriptors for an initial assessment of data obtained from camera trapping surveys. We then provide a detailed overview of sampling design, pointing the reader, in particular, to the fundamental differences between opportunistic and systematic designs. This section includes indications on sampling effort, i.e. number of camera trap stations and survey duration, as well as means to assess sampling completeness.

The second part of the chapter provides a case study with real data whereby the reader is guided step-by-step through data import, data screening and a set of fundamental analytical routines to derive the metrics above and related summaries such as temporal and spatial distribution of trapping events at the species level. As this chapter is the first on data analysis, it also briefly introduces the preparatory steps that are necessary to use R (<http://www.r-project.org/>), the main analytical software used here (R editor, setting directories, loading libraries, sourcing functions, etc.), and hence this knowledge is taken for granted in subsequent chapters. The analyses proposed assume basic general confidence in the use of R and Excel.

The use of camera trapping for faunal surveys is the simplest and perhaps most common use of this tool, and one that is instrumental to more detailed analysis. Indeed a clear definition of study objectives, beyond inventory compilation, is a prerequisite to define the correct sampling design. If the study is limited to assessing the presence of species, then the sampling design, sampling duration and selection of camera trap sites can be relatively flexible. For the purpose of maximising trapping success, it might be possible to use baits placed in front of the camera trap, or targeting particular sites of intense use such as feeding or drinking sites, nests, rubbing trees, etc.

However, we anticipate here that an opportunistic sampling design, i.e. one that does not adhere to the fundamental principles of randomisation and replication of samples, will fail to provide representative observations of the studied population, and as such, cannot be used to *infer* parameters of the whole target population (Williams *et al.* 2002). This means that the data obtained from opportunistic sampling are usually not suited to test hypotheses and make predictions on these populations. However, if the objective of the study is to provide only a species list of a target area, then an array of camera traps arranged in an opportunistic way may represent a reasonable and affordable strategy. It is likely that opportunistic sampling will bias the results towards one or a pool of species that are more likely to use the selected sites relative to other species (i.e. not all species will have the same chances of being detected at all sites), and/or because no consideration is given to the spacing of camera traps in relation to the home range of the target species (see below and Chapter 6). These are only two of the several potential sources of variation, which are unrecognisable, and whose influence is addressed through randomisation and replication (Williams *et al.* 2002).

If, in contrast, the inventory is conducted as a baseline assessment for future studies, for replication to other areas, or for modelling species' distribution, then a systematic sampling design is required. Similarly, such design is recommended for robust assessment of activity pattern (see Chapter 8). Hence, sites should be selected at *random* relative to animal movements and therefore relative to the chances of animals encountering the cameras. In this way we can assume that species have the same chances of being detected across all sites, i.e. there is no bias towards selected species by deploying variable setups (e.g. camera traps on nests or drinking sites at some sites and along wildlife trails at others). In such cases, the study design could follow the one recommended for single-season occupancy studies (Chapter 6) that adopts a regular grid of cameras, or could follow a stratified random design, as well as a full randomisation approach (Rowcliffe *et al.* 2011). Moreover, if the survey targets the whole community of terrestrial vertebrates, a robust sampling for analysis in occupancy framework allows the estimation of true species richness using the Bayesian approach described in Chapter 9.

## 5.2 Raw descriptors: naïve occupancy and detection rate as a relative abundance index

Besides the checklist, i.e. the list of captured species, data from faunal surveys at the species level can be analysed to derive two basic descriptors of species' presence: (1) naïve occupancy and (2) camera trapping detection rate, often called the RAI.

Naïve (or observed) occupancy is simply the proportion of camera trap stations where the species has been captured relative to the total number of camera stations sampled. It therefore has a value ranging from 0 to 1 that provides an indication of the extent of species' presence across the area sampled. The closer the value is to 1, the larger the proportion of sites where the species occur. This may in turn indicate wider distribution of a species, especially if the camera sites are sufficient in number to cover a representative sample of the target area. These simplistic presence-absence data may be biased by detection error – i.e. a recorded absence of a species at a site may in fact be a non-detection, and not a true absence. Using such data with naïve estimators will thus result in underestimates of true occupancy (MacKenzie *et al.* 2002; see Chapter 6 for details). When imperfect detection is not accounted for, the metric we are estimating is not true species occupancy, as occupancy and detection are confounded. Naïve

occupancy provides information on where the species is more or less likely to be detected rather than an estimate of true occupancy. Since detection probability can change even for the same species inhabiting different regions (e.g. as a result of different behaviours or habitat characteristics), comparison of naïve occupancy among different surveys could lead to erroneous conclusions.

The camera trapping rate is the photographic 'event' rate, i.e. the number of photographic events at which a species is trapped during the sampling. It is therefore calculated as the ratio events/sampling effort. Survey effort is typically measured as camera days, i.e. the number of camera traps used multiplied by the number of days they were in operation (taking into account malfunctioning cameras if necessary). Hence, the unit of the camera trapping rate is events per day of sampling.

An interval between consecutive images is typically used to separate out single, passing animal events from repeated images of the same event. Intervals of 30 min to 1 h are common (e.g. O'Brien *et al.* 2003; Bowkett *et al.* 2008; Rovero *et al.* 2014) although Kays and Parsons (2014) found that temporal autocorrelation dropped off after 1 min, and thus chose this as their independence interval. Indeed, Yasuda (2004) addressed this issue by analysing the number of species' appearances with variable intermission lengths and found that 39–52% of all the photographs were taken within a 1 min interval, depending on target species, and the number levelled off at a 30 min intermission length. Hence, beyond this interval, the probability that new images are independent events of passage increases. However, this author used baited camera trap stations, and hence it is doubtful if these results can be generalised.

The camera trapping rate is also called RAI as this metric may provide information on population abundance and indeed be considered an index of relative abundance (O'Brien 2011). On the general concepts of indices of abundance we refer the reader to general essays (e.g. Williams *et al.* 2002) as well as specific applications to camera trapping (O'Brien *et al.* 2003; Rovero and Marshall 2009; O'Brien 2011). While the camera trapping rate will, in principle, be related to true abundance, it remains an index based on observed data (frequency of captures) and does not account for time- and space-related factors that will affect detectability (e.g. Yoccoz *et al.* 2001; Pollock *et al.* 2002). Species-specific factors such as body size (Kelly and Holub 2008; Tobler *et al.* 2008; Anile and Devillard 2015; but see Rowcliffe *et al.* 2011 on how this can be taken into account), trail (Trolle and Kéry 2003), daily range (Rowcliffe *et al.* 2008; Tobler *et al.* 2008) and behaviour (Steenweg 2012) will also affect detectability. In addition, camera trap model, temperature and moisture (affecting the performance of the sensor, see Chapter 2) and habitat features at the camera site (e.g. vegetation density, steepness and other features that affect the camera's field of view) will bias the camera trapping rate according to a species- and site-specific process (Rowcliffe *et al.* 2011; Rovero *et al.* 2014). As a result, comparisons of camera trapping rates between species, and within species over time and space, that do not take these factors into account may to some extent lead to erroneous conclusions (Sollmann *et al.* 2013).

However, a few studies that have calibrated RAI to density estimates (Carbone *et al.* 2001; O'Brien *et al.* 2003; Rovero and Marshall 2009; Rowcliffe *et al.* 2008) show that there may be a robust, linear relationship between density and RAI hence supporting its use as a relative abundance index. Nevertheless, calibrations are difficult, as they require an independently derived density estimate, and would ideally need to be performed at each study area and over time. Therefore, these studies partially support the use of RAI only for comparisons within populations (Rovero and Marshall 2009; Sollmann *et al.* 2013).

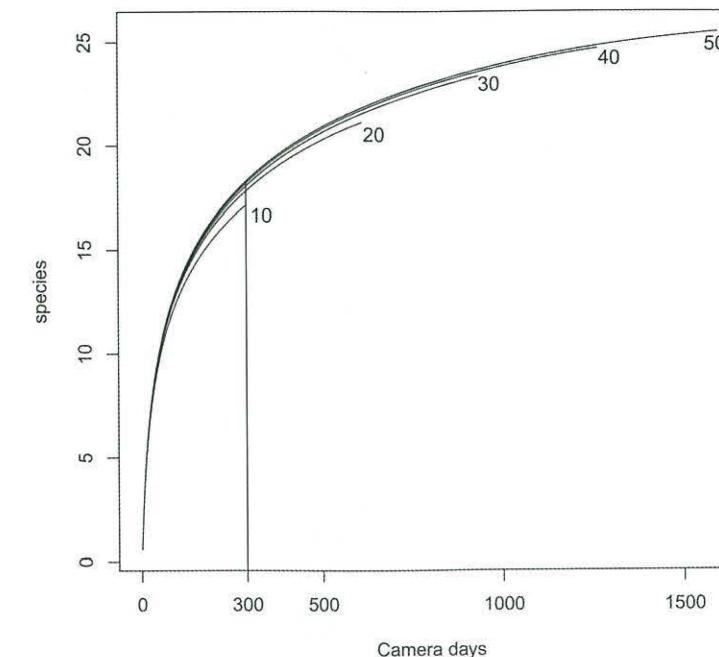
In summary, when data are collected opportunistically from faunal surveys they are not suitable for abundance estimation that accounts for detectability (such as occupancy, see Chapter 6). In this case, the use of RAI as a raw index of event rate, and naïve occupancy as an indication of occurrence, will be helpful for descriptive comparisons among populations and hence for making preliminary observations and hypotheses. These in turn will help in designing more robust studies suitable for rigorous analysis and hypothesis testing. In addition to deriving the species' list, species' accumulation and the raw metrics of relative abundance, in the following example we will also show simple, additional descriptive analysis such as plotting the activity pattern of each species, using the RAI at the camera trap sites to plot distribution maps, make simple assessments on habitat associations of focal species and on possible interactions between potentially competing species.

### 5.3 Sampling design

If the study objective is simply to assess which species are present within a given area, the sampling design implies setting single camera traps across the study area, preferably stratifying according to the habitat types targeted. Tobler *et al.* (2008) indicated that the area covered by the camera traps may have little impact on the number of species detected so long as the sampled area is representative of the general study area and habitat type. There are no strict indications on the number of camera trap stations to be sampled, which will broadly depend on the size of the area. When facing a trade-off between number of sampling sites and sampling days, however, perhaps because only a few camera traps are available, we recommend maximising the number of sites at the expense of sampling days per site, hence moving camera traps around the area (Si *et al.* 2014). This is intuitive to some extent, as the larger the number of cameras used and spatial coverage of sampling, the greater the chances of maximising the number of species captured.

As an illustration of this suggestion we created a simulation using the TEAM Network data provided for the case study presented further below. We set a target of 300 camera days to be deployed in the study area, and wanted to assess how many species we would detect with increasing numbers of camera trap sites sampled (and accordingly, decreasing number of days for each). We considered 10, 20, 30, 40 and 50 camera trap sites. We then randomly selected (out of an overall sample of 60 camera trap sites) the camera traps and determined the number of species detected using 10, 20, 30, 40, 50 cameras once 300 camera days had been reached; we repeated this computation 1,000 times for each set of cameras and determined an average number of species detected. We found that the number of species detected increases with the number of camera stations (Figure 5.1). That is, with the same sampling effort, we found more species using more cameras working for less time than using few cameras working longer. This indicates that when the number of camera traps available is small (10 or fewer), it is better to move camera traps around to sample a larger area when the aim is to maximise the number of species. The overall duration needed to carry out a survey will be inversely proportional to the number of camera trap stations sampled. While there may be no strict limits on the duration of a one-off survey, we recommend keeping it within a season, or a year, i.e. within a period that can be assumed closed to changes in species' richness, and species' occupancy state (Rovero *et al.* 2013).

The spatial arrangement of camera traps should reflect the study objective and, in particular, the target species. Spacing will determine the probability of which species will



Effort	Number of cameras (camera days)	Average number of species ( $\pm$ SD)
300	10 (working 30 days each)	17.1 $\pm$ 2.0
	20 (15 days)	17.8 $\pm$ 1.3
	30 (10 days)	18.0 $\pm$ 1.0
	40 (7.5 days)	18.2 $\pm$ 0.7
	50 (6 days)	18.3 $\pm$ 0.5
600	20 (30 days)	21.1 $\pm$ 1.7
	30 (20 days)	21.5 $\pm$ 1.2
	40 (15 days)	21.7 $\pm$ 0.9
	50 (12 days)	21.8 $\pm$ 0.5
900	30 (30 days)	23.2 $\pm$ 1.4
	40 (22.5 days)	23.4 $\pm$ 0.9
	50 (18 days)	23.5 $\pm$ 0.6
1,200	40 (30 days)	24.5 $\pm$ 1.0
	50 (24 days)	24.6 $\pm$ 0.6

**Figure 5.1** Results of simulations of observed species' richness determined by a constant survey effort of 300 camera days with varying combinations of number of cameras used (10–50) and sampling duration. The number of species detected (values shown on the table) increases with the number of camera sites sampled. See text for details.

be captured based on the species' home range. In general terms, if the spacing between camera stations is larger than a species' home range, this species will have less chance of being detected relative to a species with a home range that is larger than the spacing of camera stations. This is because in the latter case animals will have, on average, more sampling points within their home ranges. An adequate spacing for detecting as many species as possible would therefore be a compromise between missing species with small home ranges by spacing camera traps too far apart, and missing species with large home ranges by spacing camera traps within a very small area (TEAM Network 2011; see also Chapter 6).

Unless researchers can identify prior to the study a set of optimal sites (drinking sites, nests, etc.), the approach generally taken is to design a regular grid (the size of which will depend on the total number of camera stations and the area targeted) and broadly set one camera station at each node. We also recommend the adoption of a consistent criterion to identify the camera trap site at the node, which would simply imply finding a wildlife trail within a given distance from the node, e.g. 50 or 100 m maximum (TEAM Network 2011). Similarly, the actual setting of camera trap units on supports, usually trees, should follow a standardised procedure in terms of camera height and inclination, orientation relative to trails, etc., as described in Chapter 3 (see also TEAM Network 2011). The use of bait at camera sites is not recommended for this systematic approach. Bait increases the chance that animals passing nearby modify their trajectory towards the camera trap detection zone, and will likely bias the results towards certain sites and/or species (see Chapter 7 on the use of baits and lures).

Finally, we recommend not using different camera trap models when undertaking a survey as this will likely introduce a potential source of bias due mainly to varying camera sensitivity (see Chapter 2). We are aware that this may not always be possible as camera units within camera sets may be progressively replaced. In Chapter 6 we mention how this potentially confounding factor can be taken into account in data analysis.

## 5.4 Sampling completeness

The completeness of sampling effort for a faunal inventory is assessed by building species accumulation curves and looking at the levelling off, as exemplified in detail in the next section. In principle, the more individuals are sampled, the more species will be recorded (Gotelli and Colwell 2001). This sampling curve rises relatively rapidly at first, then much more slowly in later samples as increasingly rare taxa are added. A species accumulation curve records the total number of collected species as additional camera days are added. This has important practical implications, as researchers can use this curve to judge when sampling is adequate and adjust the study design and duration accordingly. For diversified communities of terrestrial vertebrates, such as in tropical forests where up to 40–50 species may occur at one particular area, there is convincing evidence that 1,000–2,000 camera trap days may be enough to detect up to 70–80% of the species (Tobler *et al.* 2008; Ahumada *et al.* 2011; Rovero *et al.* 2014). This of course will partly depend on site-specific features and species-specific detection rates. Nevertheless, many thousand camera trap days may be required to obtain a more complete species list, and the rarest or most elusive species may not be captured even when deploying a very large sampling effort (Srbek-Araujo and Chiarello 2005; Tobler *et al.* 2008). When the species accumulation curve reaches the asymptote, we can be quite confident that the species community has been sampled exhaustively. However, when this is not the case

(either because the community is very species rich, or because the sampling effort was not sufficient), we could be still interested in knowing the 'real' species richness. For this goal, statistical studies have suggested a large number of estimators of the asymptotic number of species (Magurran 2003).  if not all the estimators, ignore the imperfect detectability of species. One of the most promising methods to estimate species richness while accounting for imperfect detection of individual species is provided by Dorazio *et al.* (2006). An application of this method is provided in Chapter 9 (see also Rovero *et al.* 2014 for an example).

## 5.5 Case study

### 5.5.1 Raw data format (.CSV file)

In this section we provide an example using a data set of mammal images taken in the Udzungwa Mountains of Tanzania by the TEAM Network project during 2009–2013. According to a standardised protocol designed to assess and monitor communities of terrestrial vertebrates in tropical forests (TEAM Network 2011), this project deploys 60 camera stations through three sequential arrays of 20 camera traps each. Camera trapping was conducted from July to November each year from 2009 to 2013, and we use the 2009 data in this example as the baseline sampling of the long-term programme (see Rovero *et al.* 2014 for more details).

Images were entered into Wild.ID (see Chapter 4) and the output data set is a .CSV file for which we report in Figure 5.2 the columns that are relevant to the analysis with the first 10 records (see Chapter 4 for the full format of the output file). Each record is an image stored on the memory card. The complete dataset file can be downloaded as Appendix 5.1.

	Sampling Unit.Name	Latitude	Longitude	Project.Name	Sampling.Event	Photo.Date	Photo.Time	Genus	Species	Number.of.Animals	Camera.Start.Date.and.Time	Camera.End.Date.and.Time
1	Sampling Unit.Name	Latitude	Longitude	Project.Name	Sampling.Event	Photo.Date	Photo.Time	Genus	Species	Number.of.Animals	Camera.Start.Date.and.Time	Camera.End.Date.and.Time
2	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-07-28	00:26:25	Crictomys	gambianus	1	2009-07-27 13:16:17	2009-08-28 15:39:15
3	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-01	05:17:13	Allax	paludinosus	1	2009-07-27 13:16:17	2009-08-28 15:39:15
4	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-01	05:17:19	Allax	paludinosus	1	2009-07-27 13:16:17	2009-08-28 15:39:15
5	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-01	05:17:20	Allax	paludinosus	1	2009-07-27 13:16:17	2009-08-28 15:39:15
6	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-04	00:08:21	Boleogale	crassicauda	1	2009-07-27 13:16:17	2009-08-28 15:39:15
7	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-04	00:08:23	Boleogale	crassicauda	1	2009-07-27 13:16:17	2009-08-28 15:39:15
8	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-04	00:08:24	Boleogale	crassicauda	1	2009-07-27 13:16:17	2009-08-28 15:39:15
9	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-06	16:12:27	Cercocetus	sanjai	1	2009-07-27 13:16:17	2009-08-28 15:39:15
10	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-06	16:12:28	Cercocetus	sanjai	1	2009-07-27 13:16:17	2009-08-28 15:39:15
11	CT-UDZ-1-01	7.84697	36.82797	Udzungwa	2009.01	2009-08-06	16:12:29	Cercocetus	sanjai	1	2009-07-27 13:16:17	2009-08-28 15:39:15

Figure 5.2 Screenshot of the first 10 rows of the data set used for analysis, a .CSV file produced by Wild.ID and opened in Excel.

This is the data format on which all subsequent analysis are based. The columns contain the following information:

- Sampling Unit Name: in this example, the standard TEAM coding of camera trap sites is used (Team Network 2011), whereby CT = Camera Trap, UDZ = Udzungwa, 1 = array 1, i.e. the first array of the three sequential arrays of 20 camera traps that are deployed each year, 01 = the first camera trap of 20 set for each array.
- Longitude, Latitude: coordinates of the sampling unit. The default coordinate system is in decimal degrees (format WGS 84). In the example provided, these are converted into UTM format.
- Sampling Event: the year of sampling (this is especially important for multiple-year studies). As described in Chapter 4, this field can be customised: we recommend the format adopted here (i.e. 2009.01 with the decimals indicating the season, in case there are multiple seasons sampled within the same year).
- Photo Date: the date the image was taken. It is fundamental to maintain the exact output format of Wild.ID which is yyyy-mm-dd. Since spreadsheet software such as Excel will modify this format when the .CSV file is opened depending on the computer settings, we recommend opening the file with a text editor to see the correct format.
- Photo Time: the time the image is taken using the format hh:mm:ss. The same observation as for the Photo Date applies.
- Genus and Species: the scientific binomial species' name (in two columns).
- Number of Animals: the number of individuals captured in the image.
- Camera Start Date and Time, Camera End Date and Time: these are the date of beginning and ending of the sampling, respectively. These are critical information and the .CSV output format is 'yyyy-mm-dd hh:mm:ss' (note the space between the date and time). Here, too, we note that opening the file in Excel is likely to modify these settings.

Note that the first steps described in the analysis will involve adding Class, Order and Family to the dataset using the IUCN database as the reference taxonomy. This will allow, for example, the analysis to be restricted to mammals only.

### 5.5.2 Importing data in R

To run the analysis, readers should first learn the basics of using the open-source software R (<http://www.r-project.org/>), for which a wealth of books and online references exist (see <http://www.r-project.org/doc/bib/R-books.html>, Venables and Smith 2009; Adler 2010; Kabacoff 2011; Crawley 2012). Since R receives instructions through the command line, we can code directly in the R console but often it is more convenient to write scripts in a separate file instead of typing them directly in the console. This requires a text editor. R comes with a default editor (which is accessed from the menu 'File' and then 'New script'), but it is very basic and lacks a lot of useful utilities. Among the available R editors, we recommend RStudio (available at <http://www.rstudio.com>). A typical screenshot of RStudio is given in Figure 5.3, showing four windows (that can be reduced or arranged as one prefers):

- The R script (top left in the example) which is where the code is written and saved.

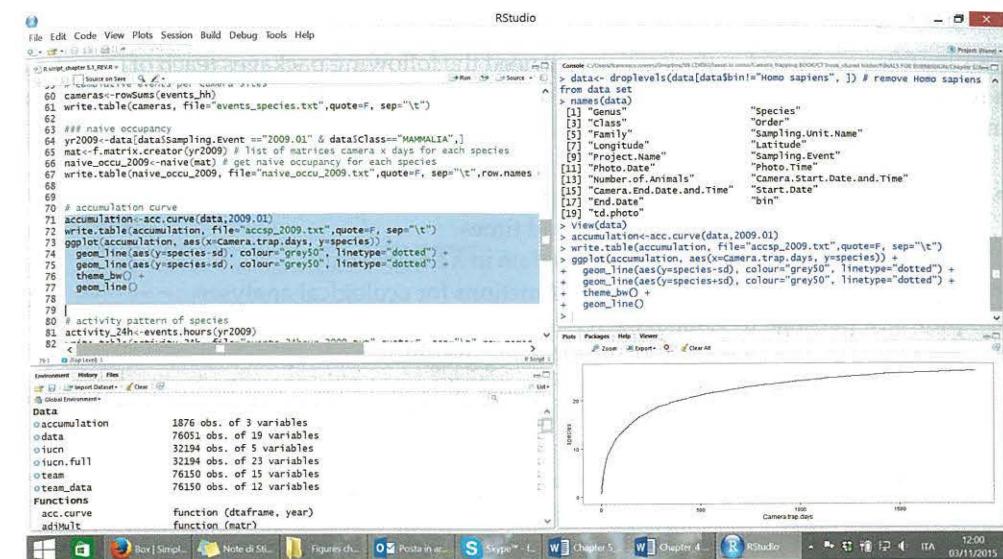


Figure 5.3 Example of a screenshot of RStudio.

- The R console (top right) which is the R environment exactly as it would be if one runs analysis in R without using an editor. Here, the lines of the script are run and the outputs are displayed.
- Two additional panels are available, each with different attributes; in the example are shown the 'environment' (bottom left), where data and functions are listed and can be visualised by clicking on the corresponding icon on the right, and the plot(s) generated (bottom right).

Before starting, we need to set the working directory. It is recommended to create a folder where the .CSV data file, all scripts sourced and the R project are saved. This is the directory where all the output files will be saved. The directory can be set by selecting 'Set Working Directory' in the menu 'Session'.

Before importing data in R, we need to source the script called 'TEAM library 1.7' where all the functions needed to prepare and analyse the data are found:

```
source("TEAM library 1.7.R")
```

The sourced script contains many useful functions that can be called directly from this file without any need to report them in the R console. In practice, when we want to perform a specific task, we call the functions stored in that script. These sourced files will remain unchanged.

Next, we need to install several packages needed for the following tasks. There is a large number of available packages, for graphics, for performing statistical tests in most disciplines, etc. (see <http://cran.r-project.org/web/packages/> to get a sense of the available packages). To use a package in R, we need to install it into a local library (i.e. a folder on our computer where all packages are stored). Next, we need to load the package into the current working session. Upon installing R for the first time, it comes with a number of different packages depending on the R version, and we can type the following command line to obtain a list of packages available by default in the version installed:

```
packages(all.available=TRUE)
```

The analyses presented in this section will need the following packages (each of them has a reference page on the web from which a manual in .PDF can be freely downloaded, as indicated for the package 'chron'):

- **chron** (<http://cran.r-project.org/web/packages/chron/chron.pdf>): handles chronological objects such as dates and times.
- **reshape**: restructures and aggregates data in a flexible way.
- **vegan**: a package containing a lot of functions for ecological analyses.
- **plotrix**: a package to visualize circular data.
- **ggplot2**: a powerful package for producing graphics.
- **maptools**: a package for combining spatial data and hence producing maps using GIS layers.

Information on each package (with a manual and examples) can be found on the R website. In order to install these packages type:

```
install.packages(c("chron", "reshape", "vegan", "plotrix",
  "ggplot2"))
```

If all went well, the packages are now ready to be loaded in the current session of R.

```
library(chron)
library(reshape)
library(vegan)
library(plotrix)
library(ggplot2)
library(maptools)
```

We are now ready to load the data into R. A simple way is to use the command below that uses the function `file.choose()` which opens a window to search the data file manually:

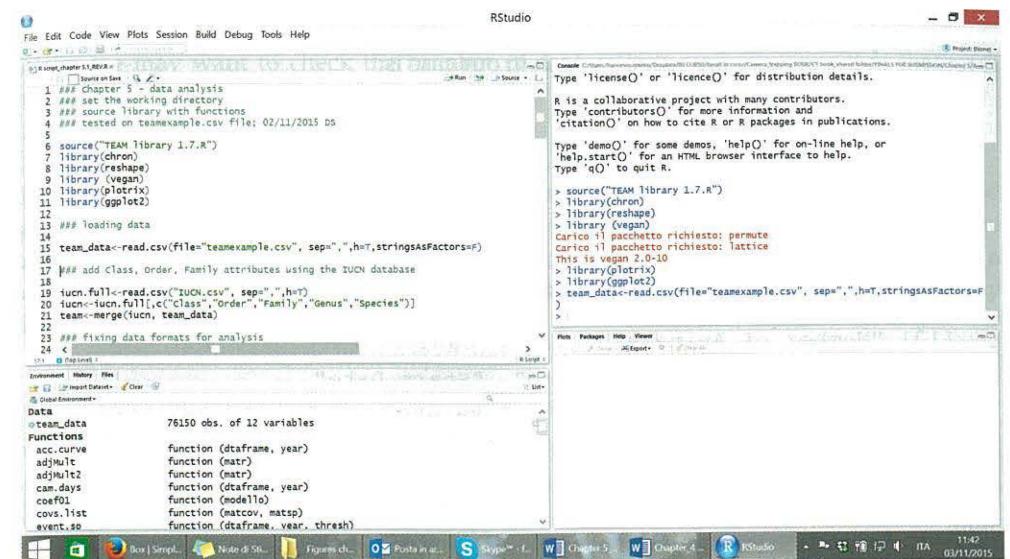
```
team_data<-read.csv(file.choose(), sep=",", h=T)
```

Alternatively, if the data file is stored in the working directory where the R script runs, we can use the following command that states the exact name of the .CSV file:

```
team_data<-read.csv(file="teamexample.csv", sep=",", h=T,
  stringsAsFactors=F)
```

In both cases, it is important to make sure that the separator of the .CSV file is correctly identified in the command. In this example we have a comma-separated file, as it can be viewed by opening the data in a text editor (e.g. Notebook). We recommend *not* opening the .CSV file in Excel as the date and time formats will change depending on the local formats; a text editor will always display the correct format of the file. The command includes `h=T` because the data set contains the headings in the first row.

If we have installed RStudio, we can now observe in our 'Environment' window (bottom left in Figure 5.4) that our `team_data` object is being listed as containing 76,150 observations (obs.) of 12 variables. Additional datasets created by the analysis will be progressively stored in this window under the 'Data' heading.



**Figure 5.4** Screenshot of RStudio showing the data set object `team_data` loaded in the 'Environment' window (bottom left).

As mentioned above, we now add taxonomic attributes (Class, Order, Family), sourcing these from the IUCN database, which is a .CSV file provided in Appendix 5.5.<sup>1</sup>

```
iucn.full<-read.csv("IUCN.csv", sep=",", h=T)
iucn<-iucn.full[,c("Class", "Order", "Family", "Genus", "Species")]
team<-merge(iucn, team_data, all.y=T)
```

We then run a function which elaborates our raw data as imported (object called `team_data`) to derive our 'clean' data ('data' object) ready for analysis:

```
data<-fix.dta(team)
```

By running this, we end up with a database that contains 19 columns from the 12 that were found in the original `team_data`. We had already included the taxonomy attributes (Class, Order, Family), moreover the function `fix.dta` corrects the format of dates (check, for example, the variable `Photo.Time` in `team` and `data`), creates two new time objects (camera start date and time; camera end date and time), an ID for each picture (date and time) and a binomial name of species (`bin`).

Our 'data' object is ultimately the database on which subsequent analysis is done.

### 5.5.2.1 Checking data

It is good rule first to look at the data to ensure that everything is in order and consistent with the raw data being imported. Hence, we run this command to view the column headings of the 19 columns in the object `data`:

<sup>1</sup> These attributes (Class, Family, Order) may be automatically integrated in future versions of Wild.ID, in such cases the `fix.dta` function can be applied directly to the uploaded data set, and the three script lines needed to source them from the IUCN database will not be required.

```
names(data)
```

R returns in the console the output, which in our case is:

```
[1] "Genus"           "Species"
[3] "Class"          "Order"
[5] "Family"         "Sampling.Unit.Name"
[7] "Longitude"      "Latitude"
[9] "Project.Name"   "Sampling.Event"
[11] "Photo.Date"     "Photo.Time"
[13] "Number.of.Animals" "Camera.Start.Date.and.Time"
[15] "Camera.End.Date.and.Time" "Start.Date"
[17] "End.Date"       "bin"
[19] "td.photo"
```

Next, we extract the first year, 2009, from the whole dataset. We also extract the mammals only:

```
yr2009<-data[data$Sampling.Event == "2009.01" &
  data$Class=="MAMMALIA", ]
```

By using the function `unique`, which returns values in a dataset without duplicates, we can then check the content of a number of key columns to simply check if things are in order. For example, we can type:

```
unique(yr2009$bin)
```

This returns the binary name of species contained in the database simply by listing them only once for every name found. We obtain the following list:

```
[1] Cricetomys gambianus      Atilax paludinosus
[3] Bdeogale crassicauda    Cercopithecus sanjei
[5] Cephalophus harveyi      Cephalophus spadix
[7] Panthera pardus         Hystrix africaeaustralis
[9] Civettictis civetta      Potamochoerus larvatus
[11] Nesotragus moschatus    Homo sapiens
[13] Papio cynocephalus     Loxodonta africana
[15] Colobus angolensis      Nandinia binotata
[17] Paraxerus vexillarius   Genetta servalina
[19] Cercopithecus mitis      Mellivora capensis
[21] Procolobus gordonorum   Dendrohyrax validus
[23] Rhynchocyon cirnei      Mungos mungo
[25] Petrodromus tetradactylus Syncerus caffer
[27] Rhynchocyon udzungwensis
```

This is a quick way to check the species reported in the database, in this case 27 mammals from the TEAM Network site in the Udzungwa Mountains. Since in this case we are not interested in the images of humans, we remove this species from the data:

```
data<- droplevels(yr2009 [data$bin!="Homo sapiens", ])
```

Next, we may want to check the number of camera trap stations for which there are records in the database:

```
unique(yr2009$Sampling.Unit.Name)
```

In our example, we obtain a list of 58 camera stations, which means that out of the 60 camera stations originally sampled two camera stations did not produce any image record that is contained in the database. This could be due to a total malfunction of a camera trap, a camera trap stolen or a camera trap damaged by an animal without any image being recorded in the memory.

Another useful check is on the start and end date of sampling:

```
unique(yr2009$Camera.Start.Date.and.Time)
```

We obtain (see first row below) a list of 58 start dates and times which are useful to check for consistency with the beginning and ending of our sampling as we will know this from the fieldwork:

```
[1] (09-07-27 13:16:17) (09-07-25 15:52:52) (09-07-26 11:46:24)
```

### 5.5.3 Deriving sampling effort, events and species' list

We derive the survey effort (camera days) by using the function `cam.days` sourced in the libraries. The function needs as arguments the data and the sampling period, or sampling year (in this case 2009.01) which is in the data column `Sampling.Event`. This is to identify the sampling year in studies that involve repeated sampling in multiple years (see Chapter 6).

This function yields a table with a list of cameras, start and end dates, and the number of days the camera traps have worked. In the case that a camera trap did not work for the all duration of sampling (e.g. because of memory saturation or malfunctioning), Wild.ID adjusts the end date to the last image taken, so that the correct sampling effort can be calculated.

We therefore run the command:

```
camera_days<-cam.days(data, 2009.01)
```

The object `camera_days` is now stored in the Environment window of RStudio, and contains 58 observations (the number of camera trap stations sampled) of four variables, namely: `Sampling.Unit.Name`, `Start.Date`, `End.date` and `ndays`.

We then use the function `write.table` if we want to save this object in a text file that we call '`camera_days_2009.txt`', saved in the default directory:

```
write.table(camera_days,file="camera_days_2009.txt",quote=F,
  sep="\t")
```

The file can be opened in a text editor (such as Notepad) or in Excel and should appear as in Figure 5.5 (first ten rows):

We can then derive minimum, maximum, median and quartiles of dates and sampling effort:

Sampling.Unit.Name	Start.Date	End.Date	ndays
CT-UDZ-1-01	2009-07-27	2009-08-28	32
CT-UDZ-1-02	2009-07-25	2009-08-28	34
CT-UDZ-1-03	2009-07-26	2009-08-28	33
CT-UDZ-1-04	2009-07-24	2009-08-30	37
CT-UDZ-1-05	2009-07-27	2009-08-30	34
CT-UDZ-1-06	2009-07-28	2009-08-31	34
CT-UDZ-1-07	2009-07-25	2009-08-28	34
CT-UDZ-1-08	2009-07-25	2009-08-28	34
CT-UDZ-1-09	2009-07-28	2009-08-31	34
CT-UDZ-1-10	2009-07-28	2009-08-31	34

Figure 5.5 Output table for the calculation of the sampling effort.

```
summary(camera_days[,2:4])
  Start.Date      End.Date      ndays
  Min. :2009-07-24  Min. :2009-08-06  Min. : 7.00
  1st Qu.:2009-07-31 1st Qu.:2009-09-02  1st Qu.:30.00
  Median :2009-09-12 Median :2009-10-13  Median :31.00
  Mean   :2009-09-12 Mean   :2009-10-14  Mean   :31.34
  3rd Qu.:2009-10-28 3rd Qu.:2009-11-29  3rd Qu.:33.00
  Max.   :2009-11-03 Max.   :2009-12-04  Max.   :37.00
```

Such simple summaries can of course also easily be done in Excel.

The next step is to derive, for each species and camera trap station, the trapping events according to a user-defined time-interval criterion. We use the function `event.sp` that has as arguments: a data frame (i.e. the data object produced earlier), the sampling year (year as indicated above), and the time interval expressed in minutes (thresh). This function reshapes the object date, yielding a table with the cameras in rows and the species in columns. The entries of this table are the number of events separated by a specific threshold.

Below we run the commands and store away the results for an interval of 1 hour (thresh=60) and 1 day (thresh=1440), respectively:

```
events_hh<-event.sp(dataframe=data, year=2009.01, thresh=60)
events_dd<-event.sp(dataframe=data, year=2009.01, thresh=1440)
```

Comparing the two tables, `events_hh` and `events_dd`, we see that most of the species were not captured repeatedly in a day. Only the most detected species, as *Cephalophorus harveyi*, passed more times a day in front of the cameras.

We can store away the results (objects `events_hh` and `events_dd`) as above:

```
write.table(events_hh, file="events_hh.txt", quote=F, sep="\t")
write.table(events_dd, file="events_dd.txt", quote=F, sep="\t")
```

These tables are 58 × 28 in size, i.e. the number of camera trap stations by the number of species (plus the sampling unit name column). We can also produce a summary table of events by species:

```
events_hh_species<-colSums(events_hh)
```

```
write.table(species, file="events_hh_species.txt", sep="\t",
           quote=F)
```

We now have derived the event scores for each species (and, if needed, by camera trap site), which are used to calculate the RAI. We can first verify the differences between daily and hourly scores, the total number of events being 1,039 and 1,259, respectively.

species	events_hh	events_dd	Survey effort	RAI_hh	RAI_dd
<i>Atilax paludinosus</i>	3	3	1818	0.17	0.17
<i>Bdeogale crassicauda</i>	130	126	1818	7.15	6.93
<i>Cephalophorus harveyi</i>	367	281	1818	20.19	15.46
<i>Cephalophorus spadix</i>	60	58	1818	3.30	3.19
<i>Cercocebus sanjei</i>	73	71	1818	4.02	3.91
<i>Cercopithecus mitis</i>	22	22	1818	1.21	1.21
<i>Civettictis civetta</i>	1	1	1818	0.06	0.06
<i>Colobus angolensis</i>	1	1	1818	0.06	0.06
<i>Cricetomys gambianus</i>	276	215	1818	15.18	11.83
<i>Dendrohyrax validus</i>	23	23	1818	1.27	1.27
<i>Genetta servalina</i>	18	18	1818	0.99	0.99
<i>Homo sapiens</i>	2	2	1818	0.11	0.11
<i>Hystrix</i>					
<i>africanaaustralis</i>	11	10	1818	0.61	0.55
<i>Loxodonta africana</i>	11	10	1818	0.61	0.55
<i>Mellivora capensis</i>	7	6	1818	0.39	0.33
<i>Mungos mungo</i>	2	2	1818	0.11	0.11
<i>Nandinia binotata</i>	2	2	1818	0.11	0.11
<i>Nesotragus moschatus</i>	114	97	1818	6.27	5.34
<i>Panthera pardus</i>	8	8	1818	0.44	0.44
<i>Papio cynocephalus</i>	3	3	1818	0.17	0.17
<i>Paraxerus vexillarius</i>	46	46	1818	2.53	2.53
<i>Petrodromus</i>					
<i>tetradactylus</i>	3	2	1818	0.17	0.11
<i>Potamochoerus larvatus</i>	18	18	1818	0.99	0.99
<i>Procolobus gordonorum</i>	5	5	1818	0.28	0.28
<i>Rhynchocyon cirnei</i>	4	4	1818	0.22	0.22
<i>Rhynchocyon</i>					
<i>udzungwensis</i>	45	45	1818	2.48	2.48
<i>Syncerus caffer</i>	4	4	1818	0.22	0.22
	1259	1083			

Figure 5.6 List of species detected and calculation of the camera trapping rate, or relative abundance index (RAI).

This step can also be easily done in Excel (Figure 5.6). We simply need to copy into a spreadsheet the species' list and events, and divide the events by the sampling effort (in this case 1,818) to derive the RAI for each species, using both hourly and daily events. Note that we multiply the index by 100 as most values are <0.

An assessment of RAI across camera stations may be useful to explore if there are differences in species' relative abundance with possible environmental gradients (elevation, habitat types, distance to sources of disturbance, etc.) and below we show how to plot species' distribution to produce a descriptive assessment. However, we stress again that proper inference on abundance patterns of the population at large can only be done when using a state variable of abundance that accounts for imperfect detection (Chapter 6).

#### 5.5.4 Naïve occupancy

The calculation of naïve occupancy (camera trap sites positive to presence on the total number of sites, see above) is simple and can easily be done in Excel using the `events_dd` object as derived above (Figure 5.5). For each species, we need to calculate the number of camera trap sites where events are greater than zero and divide the value by 58 (the total number of camera trap sites). The function `COUNT.IF` in Excel does that, with '>0' being the criterion for counting the cells. As with any other function in Excel, we also need to provide the interval of cells to which the function is applied. We will then copy the resulting values somewhere else besides the species' list (see object `events_dd_species`). Note that we will derive the counts of positive camera trap sites arranged on a row while we have the species' list in column, which is more convenient; we therefore opt to transpose the counts using the option 'Paste special' > 'Transpose' under the 'Home' menu of the main Excel bar. Once this is done, we can calculate naïve occupancy (in Figure 5.7 the results are shown for the first five species).

Alternatively, we can run this analysis in R using two functions: `f.matrix.creator` and `naive`. The function `f.matrix.creator` is applied to the main data set (object `data`). This function creates, for each species, a table with cameras on rows and days on columns. The function `naive` in turns yields the naïve occupancy values for each species. As usual, we can save away the results.

```
mat<-f.matrix.creator(yr2009)
naive_occu_2009<-naive(mat)
write.table(naive_occu_2009, file="naive_occu_2009.txt",
sep="\t", quote=F, row.names = F)
```

Species	Num. positive cameras	Num. sites	RAI
<i>Atilax paludinosus</i>	3	58	0.0517
<i>Bdeogale crassicauda</i>	43	58	0.7414
<i>Cephalophorus harveyi</i>	50	58	0.8621
<i>Cephalophorus spadix</i>	27	58	0.4655
<i>Cercocebus sanjei</i>	30	58	0.5172

**Figure 5.7** Example of calculation of the naïve occupancy for the first five species of the checklist.

#### 5.5.5 Species accumulation

As explained earlier, the completeness of the faunal inventory can be assessed by creating a species accumulation curve and checking if it has reached an asymptote. We build a species accumulation curve by using the function `acc.curve`. The function needs as usual the object `data` and the year we intend to analyse. Then it retains only the class `Mammalia` (i.e. the other classes, if present, are dropped), and the table is arranged in the correct way as the internal function `specaccum` needs. The function `specaccum` is loaded from the package `vegan` and it is set to count the species richness as camera trap days increase. The function repeats the counts many times (we set it to 100), each time changing the order of days. The output of the function `acc.curve` is the average species richness and its standard deviation.

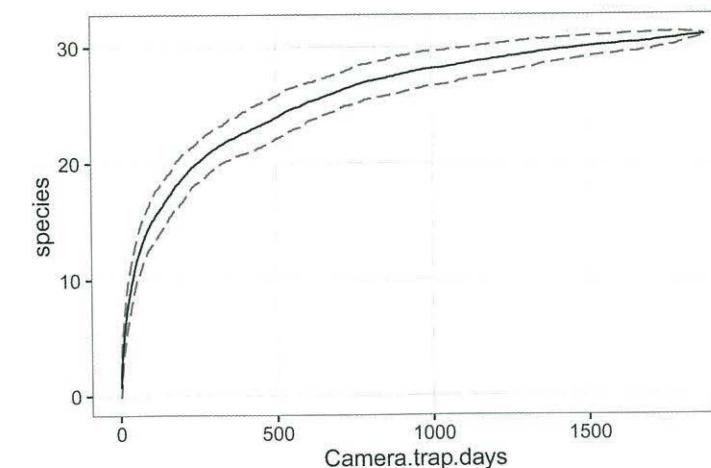
```
accumulation<-acc.curve(data,2009.01)
```

We then save the output object `accumulation` and observe that it contains three columns: the progressive number of camera days, the richness and the standard deviation of the richness:

```
write.table(accumulation, file="accsp_2009.txt", sep="\t",
quote=F)
```

Finally, we plot the average species richness, with standard deviation, and the camera trap days to obtain the species' accumulation curve (Figure 5.8):

```
ggplot(accumulation, aes(x=Camera trap.days, y=species)) +
  geom_line(aes(y=species-sd), colour="grey50",
  linetype="dotted") +
  geom_line(aes(y=species+sd), colour="grey50",
  linetype="dotted") +
  geom_line()
```



**Figure 5.8** Randomized species accumulation curve (solid line), and confidence intervals (dashed line).

### 5.5.6 Activity pattern

When surveys cumulate a robust survey effort (e.g. >500 camera days) and hence some species are recorded with enough events (>20) for focal analysis, species-specific activity pattern can be assessed by plotting all hour-separated events by the hour of the day.

This is done by the function `events.hours` that is fed, as usual, by the object `year2009`. It yields a table with hours (0–24) and species summing all the events, keeping 1 hour as the interval.

```
activity_24h<-events.hours(year2009)
```

This could also be calculated for all years data, as follows:

```
activity_24h<-events.hours(data)
```

We then save the resulting object `activity_24h`:

```
write.table(activity_24h, file="events_24hour_2009.txt", quote=F,
sep="\t", row.names = F)
```

An extract of this output file opened in Excel is shown in Figure 5.9 for the first four species and the first five 1 h intervals:

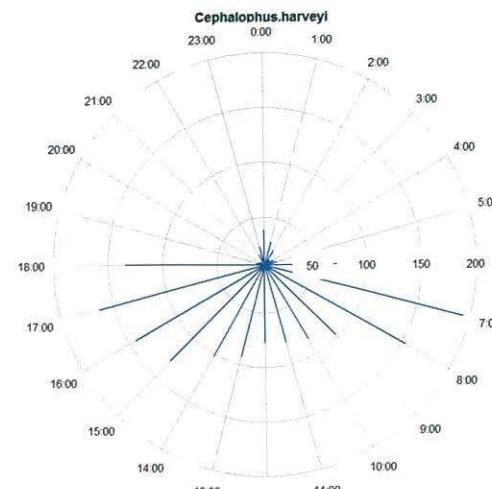
With this, we can make charts of events in the 24 h directly in Excel.

Alternatively, we can use the ad-hoc `clock24.plot` function in R (package `plotrix`, see <http://cran.r-project.org/web/packages/plotrix/index.html> or type `? clock24.plot` in the R console). We first define `clock` as a vector of 24 values, and we then build the plot for a focal species, in this example the red cuiker *Cephalophus harveyi*:

```
clock<-c(0:23)
clock24.plot(activity_24h$Cephalophus.harveyi,clock,show.
grid=T,lwd=2,line.col="blue", main="Cephalophus.harveyi",
cex.lab=0.5)
```

	Atilax. paludinosus	Bdeogale. crassicauda	Cephalophus. harveyi	Cephalophus. spadix
00:00:00	0	9	8	0
01:00:00	0	13	11	2
02:00:00	0	19	4	1
03:00:00	0	9	5	1
04:00:00	1	9	3	0
05:00:00	1	8	4	4
06:00:00	0	0	32	4
07:00:00	0	0	39	5
08:00:00	0	0	26	9
09:00:00	0	0	16	1
10:00:00	0	0	15	5

**Figure 5.9** Hourly counts of events per species, needed to derive charts of activity. The example shows only the first 10 h for four species.



**Figure 5.10** Radial plot of activity pattern based on hourly counts of events.

The resulting radial plot is shown in Figure 5.10.

With data from multiple years, such as in our case, the following code derives the hourly events for all available years:

```
activity_24h<-events.hours(data)
```

We refer to Chapter 8 for detailed background and analysis of activity pattern including quantitative assessment of activity overlap between species (see also below for a descriptive comparison of activity patterns).

### 5.5.7 Presentation and interpretation of results

As the species accumulation gives an indication of sampling completeness, we recommend showing this curve (see Rovero *et al.* 2014 for an example). In our example, the shape of the curve with its steep increase and tendency to reach a plateau indicates that sampling has captured a good portion of species in the community. Curves that do not flatten are generally an indication of insufficient sampling.

Next, we suggest presenting the full inventory results in terms of checklist of species camera trapped group by taxonomic attributes together with the raw indices we derived in the example (Table 5.1).

These results will in turn provide a description of the variation among species in terms of event rates (RAI) and naïve occupancy (the two indices may be broadly concordant) and, for non-systematic surveys where no further analysis is feasible, enable some conclusions to be drawn based on observed metrics about the relative abundance of species.

When the survey covers adequately the target area, and the spatial design is regular enough, it is useful to plot results on a map and inspect the distribution of species. This may also help assessing any potential, mutual distribution among two or more species for which an interaction is postulated. To do this we use the camera trap site coordinates from the Open DeskTEAM output file.

**Table 5.1** Checklist of mammals camera trapped in Mwanihana forest, Udzungwa Mountains, Tanzania during the baseline survey of the TEAM programme (adapted from Rovero *et al.* 2014)

Taxonomic group	Common name	Scientific name	Events per hour	RAI	Naïve occupancy
Afrotheria	tree hyrax	<i>Dendrohyrax arboreus</i>	23	1.27	0.241
	African elephant	<i>Loxodonta africana</i>	11	0.61	0.121
	four-toed sengi	<i>Petrodromus tetradactylus</i>	3	0.17	0.017
	chequered sengi	<i>Rhynchocyon cirnei</i>	4	0.22	0.052
	grey-faced sengi	<i>Rhynchocyon udzungwensis</i>	45	2.48	0.259
Carnivores	marsh mongoose	<i>Atilax paludinosus</i>	3	0.17	0.052
	bushy-tailed mongoose	<i>Bdeogale crassicauda</i>	130	7.15	0.741
	African civet	<i>Civettictis civetta</i>	1	0.06	0.017
	Lowe's servaline genet	<i>Genetta servalina lowei</i>	18	0.99	0.259
	honey badger	<i>Mellivora capensis</i>	7	0.39	0.103
	banded mongoose	<i>Mungos mungo</i>	2	0.11	0.034
	African palm civet	<i>Nandinia binotata</i>	2	0.11	0.034
	leopard	<i>Panthera pardus</i>	8	0.44	0.052
Primates	Sanje mangabey	<i>Cercocebus sanjei</i>	73	4.02	0.517
	Sykes' monkey	<i>Cercopithecus mitis</i>	22	1.21	0.241
	yellow baboon	<i>Papio cynocephalus</i>	3	0.17	0.052
	Udzungwa red colobus	<i>Procolobus gordoni</i>	5	0.28	0.069
	Angolan colobus	<i>Colobus angolensis</i>	1	0.06	0.017
Rodents	giant pouched-rat	<i>Cricetomys gambianus</i>	276	15.18	0.534
	Cape porcupine	<i>Hystrix africaeaustralis</i>	11	0.61	0.086
	Tanganyika mountain squirrel	<i>Paraxerus vexillarius</i>	46	2.53	0.328
Ungulates	Harvey's duiker	<i>Cephalophus harveyi</i>	367	20.19	0.862
	Abbott's duiker	<i>Cephalophus spadix</i>	60	3.30	0.466
	suni	<i>Neotragus moschatus</i>	114	6.27	0.448
	bush pig	<i>Potamochoerus larvatus</i>	18	0.99	0.190
	African buffalo	<i>Synacerus caffer</i>	4	0.22	0.052

Such maps can be done directly in R, with a basic graphic being simply a plot of geo-referenced camera locations represented as dots of varying colour or size according to the metric used. Ideally, however, we can use any geographical attribute that will help interpreting the map; these geographical 'layers' should be shapefiles, i.e. graphical objects that are geo-referenced so that they can be assembled using, for example, the package `maptools`. In the example below we simply use the shape of the forest, that we have as a shapefile, i.e. a geo-referenced polygon. Additional attributes could be road networks, rivers, digital elevation models, or even a geo-referenced topographical map.

The species' metric used for plotting can be simple presence/absence, number of events, or RAI.

Here, we plot as an example the cumulative daily trapping events scored for the two giant sengis, or elephant-shrew (genus *Rhynchocyon*). Since these species belong to the same genus, are similar in size and have similar diet, we are interested to see if their distributions overlap and whether any sign of competition is evident in the spatial distribution.

We first load in R our forest contour shape (file `park.shp`):

```
shape <- readShapeSpatial("park.shp", repair=T)
```

We also load the package `maptools` needed to convert the coordinate format from decimal degrees into UTM:

```
library(maptools)
```

We then create a table with coordinates of the camera traps locations and associated daily events for all species at first:

```
ev.dd.map<-merge(unique(data[,c("Sampling.Unit.Name",
  "Longitude", "Latitude")]), events_dd)
```

Because the default format of coordinates is decimal degrees, we transform the coordinates in UTM using the lines below (should the coordinates originally be included in UTM, then the reader can skip these lines). Note that these are referred to the specific UTM zone of the study area (UTM 37S), and therefore this code will need to be customised:

```
coord<-ev.dd.map[,c("Longitude", "Latitude")]
xy <- project(as.matrix(coord), "+proj=utm +zone=37
  +south +ellps=clrk80 +units=m +no_defs")
ev.dd.map$Longitude<-xy[,1]
ev.dd.map$Latitude<-xy[,2]
```

We tell R how to arrange the two maps that we are going to create in a panel, and we first just plot the forest contour:

```
par(mfcol=c(1,2), mar=c(0.5,0.5,0.5,0.5), oma=c(1,1,1,1))
plot(shape, axes=F)
```

We then plot the map for *R. cirnei*, specifying the title, font and size; since we want to have dots on a grey scale according to the number of events, we create a vector called `Rc`, standardising the values to the maximum.

```
mtext("Rhynchocyon cirnei", cex = 1.5, font =3 )
Rc<-ev.dd.map[,c("Rhynchocyon cirnei")]/max(ev.
  dd.map[,c("Rhynchocyon cirnei")])
points(ev.dd.map[, "Longitude"], ev.dd.map[, "Latitude"],
  pch = 21, bg=grey(1-Rc))
plot(shape, axes=F)
```

We do the same for the other species, *R. udzungwensis*:

```

mtext("Rhynchocyon udzungwensis", cex = 1.5, font =3)
Ru<-ev.dd.map[,c("Rhynchocyon udzungwensis")]/max(ev.
  dd.map[,c("Rhynchocyon udzungwensis")])
points(ev.dd.map[, "Longitude"], ev.dd.map[, "Latitude"],
  pch = 21, bg=grey(1-Ru))

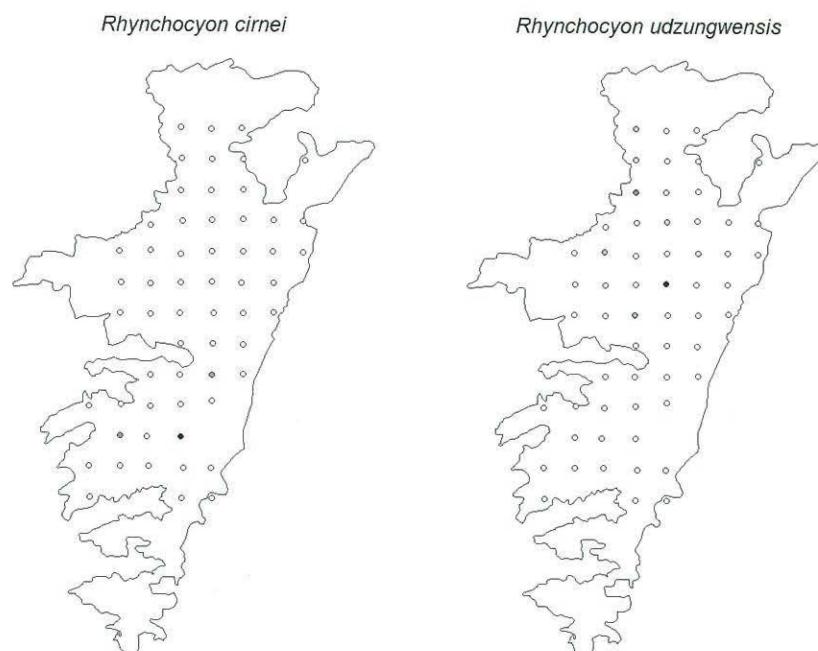
```

The resulting panel is shown in Figure 5.11.

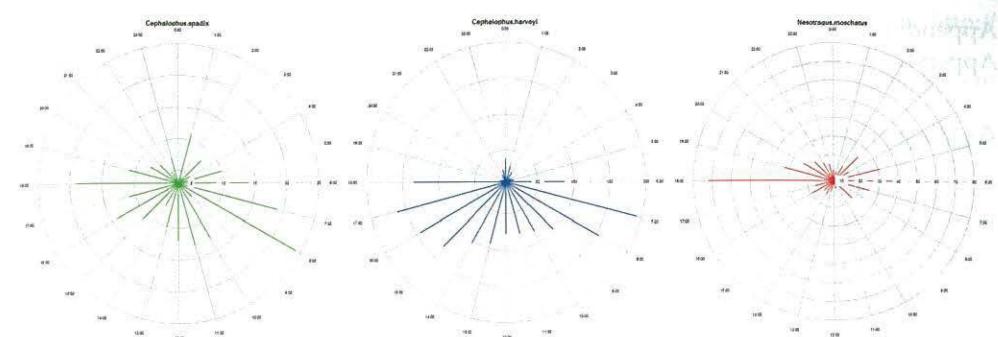
This map is very informative as we observe a typical, *parapatric* distribution, whereby the two species have contiguous but non-overlapping distributions across the forest. This may indeed be the result of competitive exclusion shown by species with similar ecological niches. Similar spatial plots of RAIs for target species or group of species can also be done to assess if any likely pattern of habitat associations exist – for example species that may prefer forest edge vs. interior habitat, species that prefer riverine habitats or species that react to sources of disturbance.

Similar inspections of results can be done along the temporal dimension using the results of the activity pattern. For example, we can plot side-by-side the activity pattern plots for species that may be under competition for resources and may therefore adjust their temporal pattern of activity. We can show this for three species of forest antelope (*Cephalophus spadix*, *C. harveyi* and *Nesotragus moschatus*) in the example below.

We first tell R to arrange the subsequent plots in a panel of  $1 \times 3$  columns, and then we reload the `clock` object (see above):



**Figure 5.11** Map of the camera trapping event score for the two species of giant sengi, or elephant-shrew in Mwanihana Forest, Udzungwa Mountains. The grey scale of the symbols represents the cumulative daily events standardised on the maximum number of events for that species.



**Figure 5.12** Panel with plots of activity patterns for three species of forest antelope.

```

par(mfrow=c(1,3), cex.lab=0.5, cex.axis=0.5)
clock<-c(0:23)

```

We then make the three plots sequentially using the same command as before and only changing the colour of the line (`line.col=""`):

```

clock24.plot(activity_24h$Cephalophus.spadix, clock,
  show.grid=T, lwd=2, line.col="green", main="Cephalophus.spadix")
clock24.plot(activity_24h$Cephalophus.harveyi, clock,
  show.grid=T, lwd=2, line.col="blue", main="Cephalophus.harveyi")
clock24.plot(activity_24h$Nesotragus.moschatus, clock,
  show.grid=T, lwd=2, line.col="red", main="Nesotragus.moschatus")

```

The resulting panel of plots is shown in Figure 5.12.

## 5.6 Conclusions

This chapter has addressed the use of camera trapping for faunal surveys. We have emphasised the importance of setting a clear study objective and how the difference between a systematic vs. an opportunistic study design will determine the type and quality of the ensuing analysis, with a systematic approach allowing for inferential analysis on the population at large. We have also cautioned against overreliance on an index such as the event rate (RAI) – which does not account for multiple sources of variation in detectability – for making inferences on abundance in space and time, and for comparison among populations and species. Through a case study, we have provided the essential tools for an analysis that can be generalised and customised to the readers' data. Finally, for an inferential assessment on focal species or communities based on occupancy, i.e. a state variable that accounts for imperfect detection, the reader is referred to Chapter 6.

## Appendices

Appendix 5.1 Case study data set: 'teamexample.csv'

Appendix 5.2 R script: 'R script\_chapter 5.R'

Appendix 5.3 R library with all functions: 'TEAM library 1.7.R'

Appendix 5.4 Shape file of forest contour: 'park.shp'  
 Appendix 5.5 IUCN species taxonomy: 'IUCN.csv'

## Acknowledgements

We are grateful to Roland Kays and Fridolin Zimmermann for valuable comments on previous versions of this chapter.

## References

- Adler, J. (2010) *R in a Nutshell*. Sebastopol, CA: O'Reilly Media.
- Ahumada, J.A., Silva, C.E.F., Gajapersad, K., Hallam, C., Hurtado, J., Martin, E., McWilliam, A., Mugerwa, B., O'Brien, T., Rovero, F., Sheil, D., Spironello, W.R., Winarni, N. and Andelman, S.J. (2011) Community structure and diversity of tropical forest mammals: data from a global camera trap network. *Philosophical Transactions of the Royal Society B Biological Sciences* 366: 2703–2711.
- Anile, S. and Devillard, S. (2015) Study design and body mass influence RAIs from camera trap studies: evidence from the Felidae. *Animal Conservation* 19: 35–45.
- Bowkett, A.E., Rovero, F. and Marshall, A.R. (2008) The use of camera-trap data to model habitat use by antelope species in the Udzungwa Mountain forests, Tanzania. *African Journal of Ecology* 46: 479–487.
- Carbone, C., Christie, S., Conforti, K., Coulson, T., Franklin, N., Ginsberg, J.R., Griffiths, M., Holden, J., Kawanishi, K., Kinnaird, M., Laidlaw, R., Lynam, A., Macdonald, D.W., Martyr, D., McDougal, C., Nath, L., O'Brien, T., Seidensticker, J., Smith, D.J.L., Sunquist, M., Tilson, R. and Wan Shahruddin W.N. (2001) The use of photographic rates to estimate densities of tigers and other cryptic mammals. *Animal Conservation* 4: 75–79.
- Crawley, M. (2012) *The R Book*. New York: Wiley.
- Dorazio, R.M., Royle, J.A., Söderström, B. and Glimskär, A. (2006) Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology* 87: 842–854.
- Gotelli, N.J. and Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4: 379–391.
- Kabacoff, R.I. (2011) *R in Action*. Greenwich, CT: Manning Publications Co.
- Kays, R. and Parsons, A.W. (2014) Mammals in and around suburban yards, and the attraction of chicken coops. *Urban Ecosystems* 17: 691–705.
- Kelly, J.M. and Holub, E.L. (2008) Camera trapping of carnivores: trap success among camera types and across species, and habitat selection by species on Salt Pond Mountain, Giles Country, Virginia. *Northeastern Naturalist* 15: 249–262.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew Royle, J. and Langtimm, C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83: 2248–2255.
- Magurran, A.E. (2003) *Measuring Biological Diversity*. Oxford: Wiley-Blackwell.
- O'Brien, T.G. (2011) Abundance, density and relative abundance: a conceptual framework. In: A.F. O'Connell, J.D. Nichols, U.K. Karanth (eds), *Camera Traps in Animal Ecology. Methods and Analyses*. Springer: New York. pp. 71–96.
- O'Brien, T.G., Kinnaird, M.F. and Wibisono, H.T. (2003) Crouching tigers, hidden prey: Sumatran tiger and prey populations in a tropical forest landscape. *Animal Conservation* 6: 131–139.
- Pollock, K.H., Nichols, J.D., Simons, T.R., Farnsworth, G.L., Bailey, L.L. and Sauer, J.R. (2002) Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics* 13: 105–119.
- Steenweg, R., Whittington, J. and Hebblewhite, M. (2012) Canadian Rockies Carnivore Monitoring Project: Examining Trends in Carnivore Populations and their Prey Using Remote Cameras. Year 1 Progress Report, 2011–2012. August 2012. University of Montana. <http://www.cfc.umt.edu/Heblab/ParksCamera.html> (accessed 5 February 2015).
- Rovero, F. and Marshall, A.R. (2009) Camera trapping photographic rate as an index of density in forest ungulates. *Journal of Applied Ecology* 46, 1011–1017.
- Rovero, F., Zimmermann, F., Berzi, D. and Meek, P. (2013) "Which camera trap type and how many do I need?" A review of camera features and study designs for a range of wildlife research applications. *Hystrix: the Italian Journal of Mammalogy* 24: 148–156.
- Rovero, F., Martin, E., Rosa, M., Ahumada, J.A. and Spitale, D. (2014) Estimating species richness and modelling habitat preferences of tropical forest mammals from camera trap data. *PLoS ONE* 9: e103300.
- Rowcliffe, J.M., Field, J., Turvey, S.T. and Carbone, C. (2008) Estimating animal density using camera traps without the need for individual recognition. *Journal of Applied Ecology* 45: 1228–1236.
- Rowcliffe, M., J., Carbone, C., Jansen, P.A., Kays, R. and Kranstauber, B. (2011) Quantifying the sensitivity of camera traps: an adapted distance sampling approach. *Methods in Ecology and Evolution* 2: 464–476.
- Si, X., Kays, R. and Ding, P. (2014) How long is enough to detect terrestrial animals? Estimating the minimum trapping effort on camera traps. *PeerJ* 2: e374.
- Sollmann, R., Mohamed, A., Samejima, H. and Wilting, A. (2013) Risky business or simple solution – relative abundance indices from camera-trapping. *Biological Conservation* 159: 405–412.
- Srbek-Araujo, A.C. and Chiarello, A.G. (2005) Is camera-trapping an efficient method for surveying mammals in neotropical forests? A case study in south-eastern Brazil. *Journal of Tropical Ecology* 21: 121–125.
- TEAM Network (2011) *Terrestrial Vertebrate Protocol Implementation Manual, v. 3.1*. Tropical Ecology, Assessment and Monitoring Network, Center for Applied Biodiversity Science, Conservation International, Arlington, VA, USA. <http://www.teamnetwork.org/files/protocols/terrestrial-vertebrate/TEAMTerrestrialVertebrate-PT-EN-3.1.pdf> (accessed 5 February 2015).
- Tobler, M.W., Carrillo-Percastegui, S.E., Leite Pitman, R., Mares, R. and Powell, G. (2008) An evaluation of camera traps for inventorying large- and medium-sized terrestrial rainforest mammals. *Animal Conservation* 11: 169–178.
- Trolle, M. and Kéry, M. (2003) Estimation of ocelot density in the Pantanal using capture-recapture analysis of camera-trapping data. *Journal of Mammalogy* 84: 607–614.
- Venables, W.N. and Smith, D.M. (2009) *An Introduction to R*. Network Theory Limited.
- Yasuda, M. (2004) Monitoring diversity and abundance of mammals with camera traps: a case study on Mount Tsukuba, central Japan. *Mammal Study* 29: 37–46.
- Yoccoz, N.G., Nichols, J.D. and Boulanger, T. (2001) Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution* 16: 446–453.
- Williams B.K., Nichols J.D. and Conroy M.J. (2002) *Analysis and Management of Animal Populations*. New York: Academic Press.