Browsing Movebank within R

Marco Smolla   Anne Scharf

Abstract

This vignette gives examples of how to browse the Movebank database within R and how to import data for the **move** package. It explains how to login, search for studies, get sensor, animal and study IDs, download location data, non-location data and data from the Movebank Data Repository.

Keywords: movebank, move, R.

Introduction

The provided functions allow you to search for studies stored in Movebank, browse them and download[1] them. A possible workflow could look like this:

First you always have to create an object containing your credentials of your Movebank account (www.movebank.org) in order to use the functions listed in this manual (except for Section 6.):
1. Login to Movebank

Than you can search for studies or get information about studies, animals, tags, etc.:
2. Search for a study name and ID
3. Get information about the used tags and animals

Once you know the study name and the kind of data you want to download you can:
4. Download the location data of a study as a `move/moveStack` object
5. Download the non-location data of a study as a `data.frame`

To download published data of the repository you do not need to have an account on Movebank:
6. Download data from the Movebank Data Repository as a `move/moveStack` object

1. Login to Movebank

There are two ways to login. Either you login every time you use the functions that are presented in this vignette. Or you use the `movebankLogin()` function to login to Movebank

---

[1] If you have the rights to do so.

and create an object that stores your login information. You can pass this object on to every function you use to skip the login process. Use your username and password which you use to login to the web based Movebank. Note that the password is stored in this object meaning that if you store the session or objects these are stored in plain text. If you want to hide the password while typing it might be worth to have a look at the package: "getPass".

```
> loginStored <- movebankLogin(username="user", password="password")
```

2. Search for a study name and ID

2.1. Search for a study name with keywords

You can use the `searchMovebankStudies()` function to search within the study names for a specific study. For example, if you want to find all studies that worked with goose try the following:

```
> searchMovebankStudies(x="oose", login=loginStored)
```

You may rather use the search term without the first letter, e.g. 'oose' instead of 'Goose' or 'goose', to find studies with both ways of writing.

2.2. Get the Movebank ID of a study

All of the here presented studies can work with the study ID or the study name to find information within the database. If you rather want to work with the short study ID instead of the longer study name use `getMovebankID()`. This number can also be found on the 'Study Details' page of the study on Movebank.

```
> getMovebankID("BCI Ocelot",login=loginStored)
> # 123413
```

3. Get information about the used tags and animals

3.1. Get general information of a study

You found a study you are interested in, let's say 'BCI Ocelot'. To get more information about this study, e.g. the authors of the study, licence type, citation and more, use:

```
> getMovebankStudy(study="BCI Ocelot",login=loginStored)
```

3.2. Get information about sensors

If you want to know, which sensor types were used for each tag in this study you can use:

```
> getMovebankSensors(study="BCI Ocelot",login=loginStored)
```

To see all available sensor types on Movebank use the same function, leaving the "study" argument empty.

```
> getMovebankSensors(login=loginStored)
```

To get a list of the attributes that are available for the sensors of a particular study use:

```
> getMovebankSensorsAttributes(study="BCI Ocelot",login=loginStored)
```

3.3. Get information about the animals of a study

A list of the animals names, their tag ids, sensors used and more information associated to each individual is returned with this command:

```
> getMovebankAnimals(study="BCI Ocelot",login=loginStored)
```

4. Download the location data of a study as a 'move/moveStack' object

NOTE: ***Agreement to license terms***: *to be able to download with R the data of a study on Movebank, first the license terms of the study have to be accepted. For this go to www.movebank.org, search for your study of interest, click on the download tab, and accept the license terms. **This only has to be done once per data set.***

If you know which coordinates of which animal you want to download you can either download the timestamp and coordinates, or directly create a move object.

4.1. Download location data of a study

An entire study can be downloaded:

```
> bci_ocelot <- getMovebankData(study="BCI Ocelot", login=loginStored)
```

4.2. Download location data for selected individuals of a study

Or the data for one or several individuals:

```
> # for one individual
> bobby <- getMovebankData(study="BCI Ocelot", animalName="Bobby",
+                          login=loginStored)

> # for several individuals
> ocelot2ind <- getMovebankData(study="BCI Ocelot", animalName=c("Bobby","Darlen"),
+                                login=loginStored)
```

4.3. Download location data for a selected time range

You can also limit your download to a given time range. The timestamp has to be in format 'yyyyMMddHHmmssSSS' or as a POSIXct, then it is converted to the character using the UTC timezone:

```
> # download all data between "2003-03-22 17:44:00.000" and "2003-04-22 17:44:00.000"
> bci_ocelot_range1 <- getMovebankData(study="BCI Ocelot", login=loginStored,
+                                       timestamp_start="20030322174400000",
+                                       timestamp_end="20030422174400000")
> # alternative:
> t<-strptime("20030322174400",format="%Y%m%d%H%M%S", tz='UTC')
> bci_ocelot_ranget <- getMovebankData(study="BCI Ocelot", login=login,
+                                       timestamp_start=t,
+                                       timestamp_end=t+as.difftime(31,units='days'))

> # download all data before "2003-07-24 20:00:00.000"
> bci_ocelot_range2 <- getMovebankData(study="BCI Ocelot", login=loginStored,
+                                       timestamp_end="20030724200000000")
>

> # download all data after "2003-07-01 20:00:00.000" for "Bobby"
> bobby_range <- getMovebankData(study="BCI Ocelot", login=loginStored,
+                                 animalName="Bobby", timestamp_start="20030701200000000")
```

4.4. Dealing with duplicated timestamps

In case the study contains duplicated timestamps, you can set the argument `removeDuplicatedTimestamps=TRUE`. It is strongly advised not to use this option because there is no control over which records are removed. It's better to edit the records in movebank and mark the appropriate records as outliers.

```
> bci_ocelot <- getMovebankData(study="BCI Ocelot", login=loginStored,
+                               removeDuplicatedTimestamps=TRUE)
```

Another possibility is to download the data directly from Movebank, or via the `getMovebank()` function, and then use the function `getDuplicatedTimestamps()` to find the duplicated timestamps and to decide which one to remove before creating a move or movestack object using the `move()` function.

5. Download the non-location data of a study as a 'data.frame'

NOTE: **Agreement to license terms**: *to be able to download with R the data of a study on Movebank, first the license terms of the study have to be accepted. For this go to www.movebank.org, search for your study of interest, click on the download tab, and accept the license terms. **This***

only has to be done once per data set.

To download the non location data from a study you can use the `getMovebankNonLocationData()` function. It returns a `data.frame`. Data from different sensors can be downloaded by specifying the sensor name or is in the "sensorID" argument. The valid names for the `sensorID` argument are those of the column 'name' or 'id' of the table returned by (`getMovebankSensors(login=loginloginStored)`. Non-location sensors (e.g. Acceleration, Magnetometer,...) are those marked as 'false' in the 'is_location_sensor' column of this table.

5.1. Download non-location data of a study

Download non location data for a specific individual for a specific sensor. A vector of several sensors can be stated in the argument `sensorID` and/or of several individuals in the argument `animalName`. If the `animalName` argument is left empty, the data of all individuals is downloaded. If the `sensorID` argument is left empty, the data of all available non location sensors is downloaded:

```
> getMovebankNonLocationData(study=74496970 , sensorID="Acceleration",
+                            animalName="DER AR439", login=loginStored)
```

5.2. Download non-location data for a selected time range

Non location data can be also downloaded for a specific time range (See Section 4. for more details):

```
> # get acceleration data for all individuals of the study between
> # the "2013-08-15 15:00:00.000" and "2013-08-15 15:01:00.000"
> getMovebankNonLocationData(study=74496970 , sensorID=2365683,
+                            login=loginStored, timestamp_start="20130815150000000",
+                            timestamp_end="20130815150100000")
```

5.3. Download non-location data with the 'getMovebankData' function

There is also the option to download non-location data with the `getMovebankData()` function. When the argument `includeExtraSensors=TRUE` than the data of all non-location sensors available for that study will be downloaded and stored in the 'unusedRecords' slot. With this option it is not possible to select specific sensors.

```
> mymove <- getMovebankData(study=74496970, login=loginStored,
+                           animalName="DER AR439",includeExtraSensors=TRUE)


> ## to get a data.frame containing the data for the non-location
> ## sensors use the "unUsedRecords" function
> nonlocation <- as.data.frame(unUsedRecords(mymove))
```

6. Download data from the Movebank Data Repository as a 'move/moveStack' object

You can also download data from the 'Movebank Data Repository'. This function returns a `move` object (if only one individual is included) or a `moveStack` (if several individuals are included). If data of non-location sensors are included in the data set, these will be stored in the 'unUsedRecords' slot of the move object.

```
> getDataRepositoryData("doi:10.5441/001/1.2k536j54")
```

Affiliation:

Marco Smolla
Max-Planck-Institute for Ornithology, Radolfzell, Germany
E-mail: marco.smolla@postgrad.manchester.ac.uk

Anne Scharf
Max-Planck-Institute for Ornithology, Radolfzell, Germany
E-mail: ascharf@orn.mpg.de