batident

Version *1.5*

*eco***Obs** GmbH

http://www.ecoobs.com

Legal

The application batIdent was developed under strict rules, nevertheless errors may be existing with the implementation. Please report such directly to us listing in addition your computer system, OS version and a detailed instruction on how to reproduce the error.

System requirements

The application batIdent was developed for Apple computers running 10.7 or newer OS versions. It runs on Intel Macs. We recommend main memory of at least 2 GB. You will also need the statistical tool R (version 3.x) as well as two statistical R packages (randomForest and kernlab).

Responsible

Volker Runkel

ecoObs GmbH

Tolstoistrasse 8

90475 Nuernberg, Germany

Inhaltsverzeichnis

1. Overview of batIdent

batIdent uses statistical methods to identify bat species based on measurements. These must be available as CSV files and can be analyzed on a per file base or in batch-mode. Results are logged to a console and can be viewed within the application. In addition result files (extension res) are stored at the same location and with the same name as the CSV input files.

At first start batIdent needs an internet connection to download supplemental files and R packages. You will need to have installed R prior to this step.

2. Installation

2.1. „Binary-Distribution“

As normal user you are downloading a ready made application which you should move to your Applications folder. The download contains an R 3.0.0 isntaller package which you have to run prior to running batIdent.

2.2. „Source-Distribution“

If you are interested in the app sources, please contact runkel@ecoobs.de

2.3. Supplemental R packages

batIdent uses two statistical tools implemented as R packages. These usually get installed automatically at first start. If for some reason you want to install them manually, you‘ll need to follow these steps. Some experience with R is recommended.

After starting R choose Packages and Data -> Package installation and install the packages kernlab and randomForest from your preferred R distribution server.

2.4. batIdent Installation

When starting batIdent the first time it downloads some files from our

server:

http://www.batident.eu/batIdent/stat-objects.zip

If your Mac has no internet connection, you can download these manually and supply them to batIdent. The download dialog will guide you.

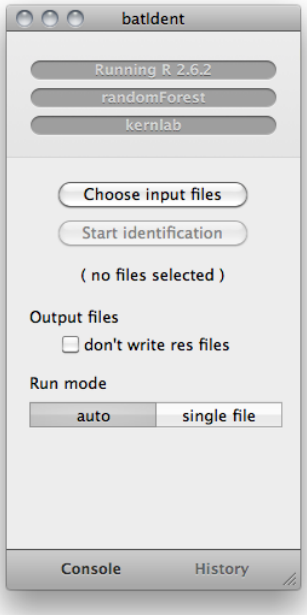
3. Using batIdent

3.1. Starting the application

After starting the application batIdent will load the basic R routines and checks if randomForest and kernlab can be started. This will take from 5 to 15 seconds. After success it is ready to use.

3.2. Interface concepts and elements

The application window shows the success of loading R and packages. The main part of the window contains a button to load CSV files, and a button to start identification.

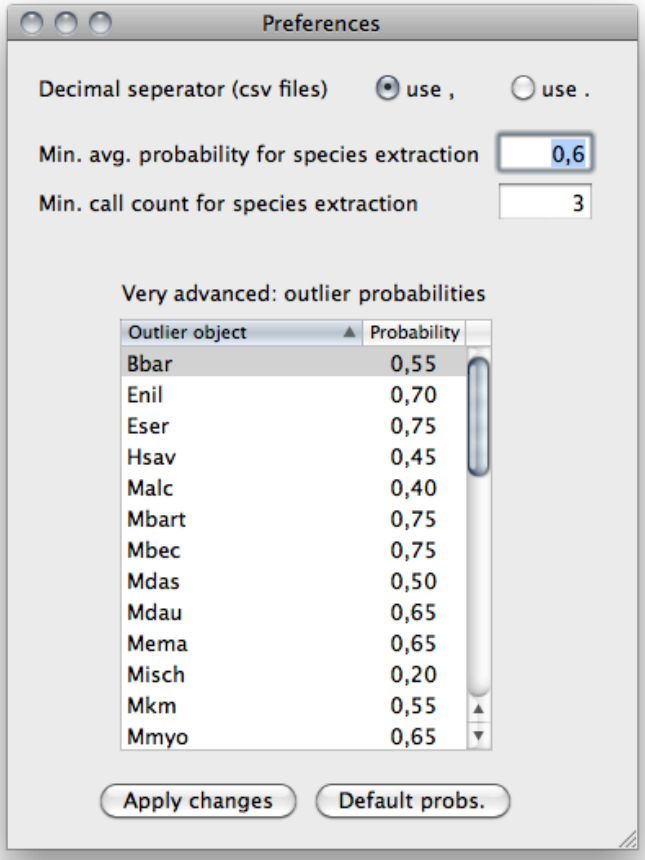


Checking „Don‘t write res files“ disables the creation of .res files. The run mode switch toggles between batch and single-file mode. Using the Console button in the lower left opens the runlog (last 1024 lines). There you will see the exact steps batIdent undertook to analyse your files as well as errors that may have occurred. The full logfile can be opened via the supplied button. The History button opens a table showing all analyzed files since the last run.

3.3. Preferences

You can influence how batIdent works by setting preferences.

The most important setting, the only you should touch, is the decimal separator. If your system works with . set it to . otherwise leave it at , .

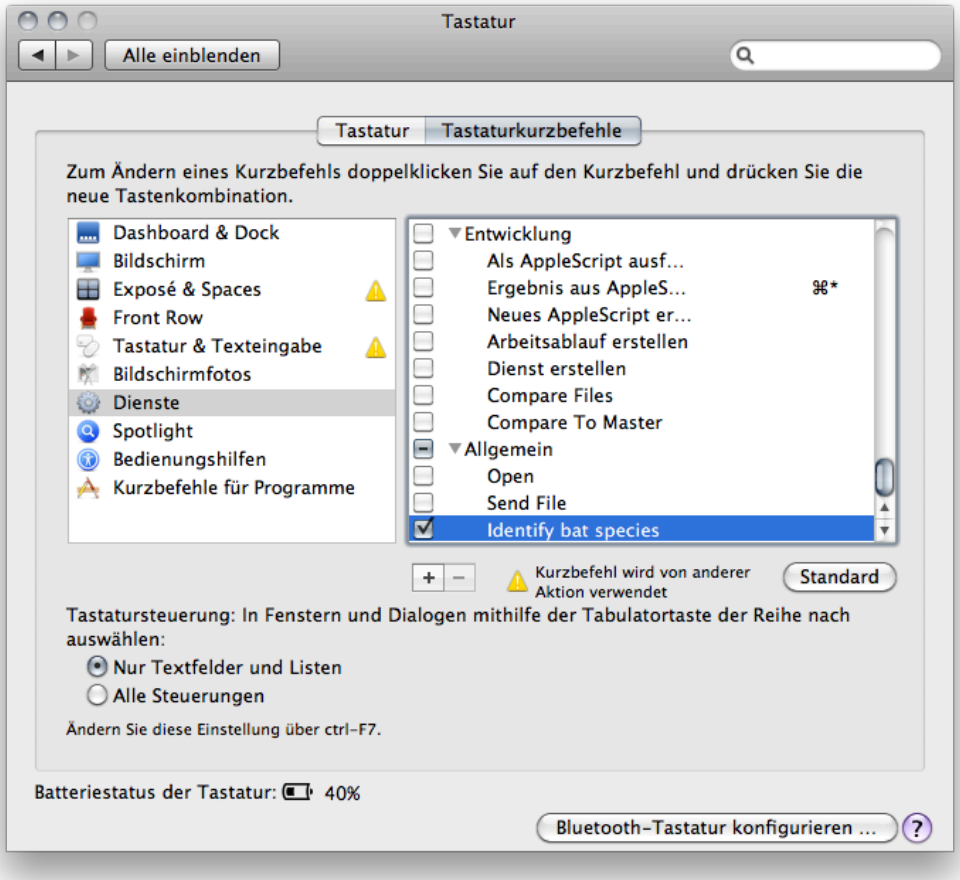


3.4. Playing nicely with bcAdmin

bcAdmin creates the necessary csv files in batch mode when measuring calls. Usually you then select all created csv files manually from within batIdent after bcAdmin finishes. Tedious and error prone, I commit. Thus, we added a new feature to batIdent, it will act as a system service and can be started from within bcAdmin for all selected recordings or from the Finder (therefore you have to select csv files).

You can use this feature after selecting recordings in bcAdmin or csv files in the Finder by using a right-click in the selection and from there choose

„Identify bat species“. To make it work you may have to activate batIdent as service first. In bcAdmin choose Services from the bcAdmin menu. There choose Services preferences and a window opens. Scroll down in the list until you find „Identify bat species“ and activate it.



4. How batIdent works

4.1. Steps in analysis

4.1.1. Identification and outlier detection

Species identification works on a per recording base and allows identification of up to three species per recording. The following table gives an overview of built-in species and groups:

Abbrev	Species	Abbrev	Species
Tten	*Tadarida teniotis*	Mema	*Myotis emarginatus*
Nnoc	*Nyctalus noctula*	Mdau	*Myotis daubentonii*
Nlei	*Nyctalus leisleri*	Mbec	*Myotis bechsteinii*
Enil	*Eptesicus nilssonii*	Ppyg	*Pipistrellus pygmaeus*
Eser	*Eptesicus serotinus*	Ppip	*Pipistrellus pipistrellus*
Vmur	*Vespertilio murinus*	Pnat	*Pipistrellus nathusii*
Mmyo	*Myotis myotis*	Pkuh	*Pipistrellus kuhlii*
Mnat	*Myotis nattereri*	Hsav	*Hypsugo savii*
Malc	*Myotis alcathoe*	Misch	Miniopterus schreibersii
Mbart	*Myotis brandtii/mystacinus*	Rfer	Rhinolophus ferrumequinum
Mdas	*Myotis dasycneme*	Bbar	*Barbastella barbastellus*

Additionally these groups are implemented

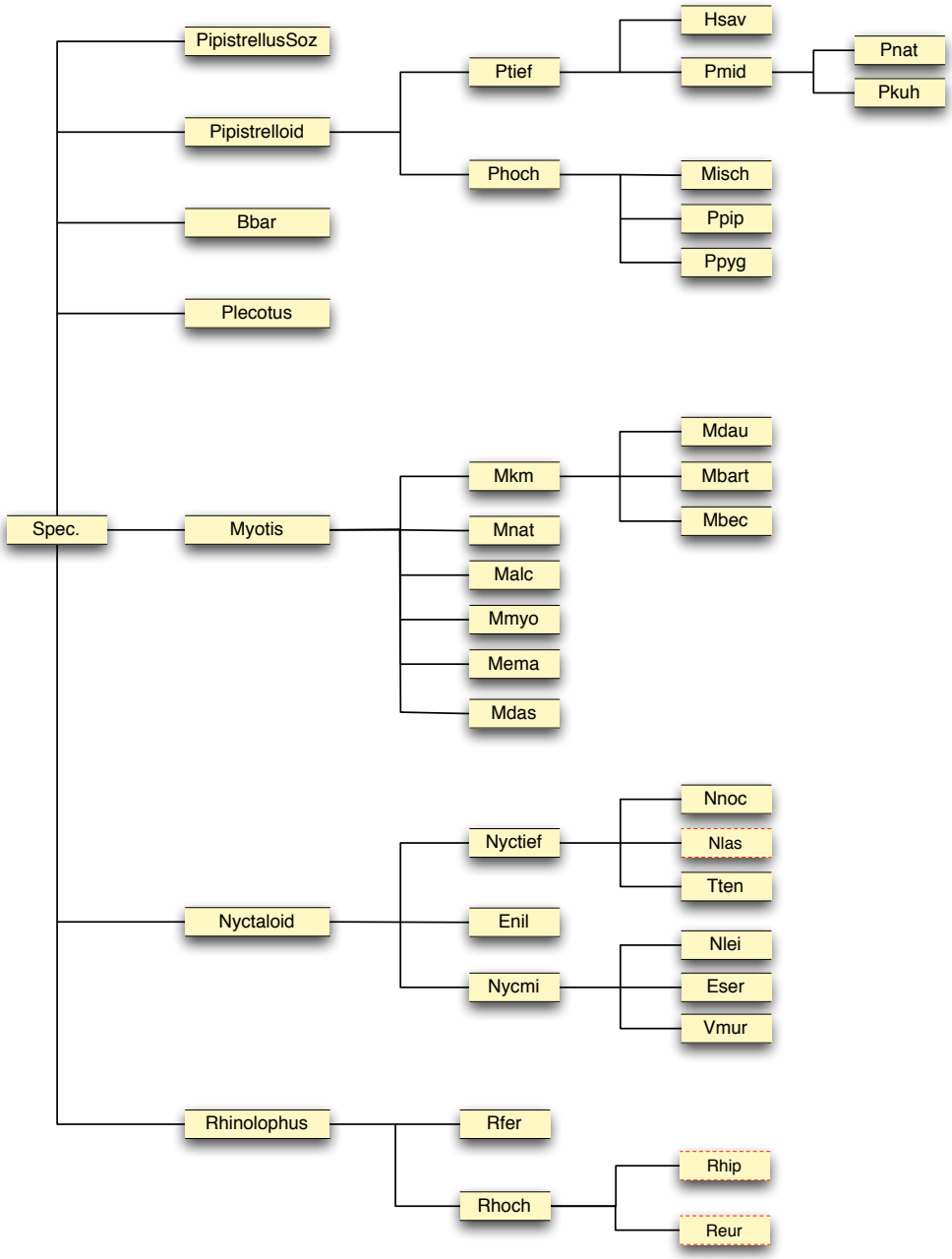
Abbrev	Genus /group
Rhinolophus	*Genus Rhinolophus*
Rhoch	*R. hipposideros or R. euryale*
Nyctaloid	*includes Nyctalus, Vespertilio, Eptesicus, Tadarida and Vespertilio*

Abbrev	Genus /group
Nyctief	*Nnoc, Tten and planned N. lasiopterus*
Nycmi	*Nlei, Eser and Vmur*
Myotis	*Genus Myotis*
Plecotus	*Genus Plecotus*
Pipistrelloid	*Genus Pipistrellus, Miniopterus and Hypsugo*
Phoch	*Ppip, Ppyg*
Ptief	*Pmid, Hsav*
Pmid	*Pnat, Pkuh*

Analysis always starts at genus/group niveau. After a first identification on per call level the probability for an outlier is calculated. Outliers are excluded from further steps in analysis. If the call was no outlier, and further steps are possible, batIdent checks on the next level to get an improved result. Again outlier detection is done. As soon as there is no further step or if an outlier was detected, analysis stops and reports the last result for each call.

4.1.2. Final result for the recording

batIdent will give one to three species for each recording. Since it is based on a single call analysis, it needs to summarize the per call results. Single mis identifications shouldn‘t then lead to bad results. Thus, batIdent uses a weighted identification probability per species and determines the best overall result. Criteria are the number of calls and the average probability after weighting. If both are healthy, the result will include this species.



Graph: Species tree

4.2 Measurements used by batIdent

batIdent requires certain measurements and a fixed input format. The csv file needs a header line followed by the actual measurements per call. Fields are separated by tabulators. The measurements are as follows:

Field	Description
Filename	Filename of the recording, not used
Species	Species name, not used
Call	call number, not used
Dur	Duration (ms)
Sfreq	Start frequenzy (kHz)
Efreq	End frequenzy (kHz)
Stime	Start time of the call, not used
NMod	see: description after the table
FMod	see: description after the table
FRmin	see: description after the table
Rmin	see: description after the table
tRmin	see: description after the table
Rlastms	see: description after the table
Flastms	see: description after the table
X10, X11 ... X60	see: description after the table
X62, X64 ... X148	see: description after the table

„FRmin“, „Rmin“ and „tRmin“ are measured at the location of lowest slope within the call. The values are frequency (kHz), slope and location within the call measured from the call end (ms). Frequency and slope are also extracted from the location with lowest slope from within the last millisecond of the call („Flastms“ and „Rlastms“.

Fields X10 to X148 are calculated as follows: For the whole call frequency measurements are taken every 100µs. This resembles the call curve. These measurements are combined in a histogram like way using a bin width of 1 kHz in the range from 10 to 60 kHz and a 2 kHz bin width from 60 to 150 kHz. The fields resemble these classes starting with the bin 10-11 kHz ending with the class 148 to 150 kHz. „NMod“ and „FMod“ are calculated from the histogram, they denote the best filled class with the number of measures „NMod“ and the lower frequency of that class („FMod“).

A pipistrelle call woudl gvie these values for example:

Field	Value		Field	Value
Dur	4,5		X49	48
Sfreq	92,361313		X50	31
Efreq	47,414757		X51	23,75
Stime	493,668640		X52	19
NMod	78		X53	16,25
FMod	48		X54	13,75
FRmin	47,598934		X55	11,75
Rmin	0,175629		X56	10,5
tRmin	-0,90		X57	9,5
Rlastms	0,031929		X58	8
Flastms	47,685329		X59	6
X10-X45	0		X60	8
X46	26,5		X62	6,5
X47	74		X64	5,5
X48	78		...	...