

# *bcDiscriminator*

Version 1.14 (20090510)

Runkel, Marckmann & Schuster GbR

**ecoObs**

<http://www.ecoobs.de>

## Rechtliches

### Umfang, Gewährleistung und Garantie

Die Software bcDiscriminator wurde nach bestem Wissen und Gewissen erstellt und getestet, dennoch können sich Fehler eingeschlichen haben. Bitte melden Sie solche direkt an uns unter Angabe ihres Rechnersystems und der Aktion, die zum Fehler geführt hat.

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA

Wir übernehmen keine Haftung die durch Fehler in der Software zu Schäden an Ihrem Rechner oder zu materiellen oder immateriellen Folgeschäden oder Gewinnausfällen bei Projekten führen, die Sie mit bcDiscriminator bearbeitet haben.

### Systemvoraussetzungen

Die Software bcDiscriminator ist für Apple Rechner mit dem Betriebssystem 10.4 (ab Version 10.4.6) geschrieben und läuft auf PowerPC- (PPC), sowie Intel Prozessoren. Empfohlen wird Hauptspeicher von wenigstens 512 MB. Benötigt werden die Statistikbibliothek R (Version 2.6.x, 2.7.x) und die R Pakete *randomForest* und *kernlab*.

### Kontakt, Redaktion

Verantwortlich für den Inhalt dieser Anleitung ist:

Volker Runkel

Runkel, Marckmann & Schuster GbR

ecoObs

Reindelstrasse 2

90402 Nürnberg

# Inhaltsverzeichnis

<b>1. Übersicht bcDiscriminator</b>	<b>4</b>
<b>2. Installation</b>	<b>4</b>
2.1. „Binary-Distribution“	4
2.2. „Source-Distribution“	4
2.3. Benötigte Ergänzungen	5
2.3.1. R	5
2.3.2. Benötigte R-Pakete	5
<b>3. Benutzung von bcDiscriminator</b>	<b>7</b>
3.1. Starten	7
3.2. Aufbau des Programmfensters	7
3.3. Voreinstellungen - Preferences	7
3.4. Zusammenspiel mit bcAdmin	10
<b>4. Funktionsweise von bcDiscriminator</b>	<b>10</b>
4.1. Analyseschritte	10
4.1.1. Diskriminierung und Ausreisser-Erkennung	10
4.1.2. Ermittlung des Ergebnisses / Artausgaben	11
4.2 Messwertdateien und Messwerte	12
<b>5. Anpassung für andere Rufparameter</b>	<b>15</b>

# 1. Übersicht bcDiscriminator

bcDiscriminator kann aus Rufmesswerten mittels statistischer Verfahren die zugehörigen Fledermausarten ermitteln. Die Messwerte müssen dazu als csv-Dateien vorliegen und werden einzeln oder im Stapelmodus analysiert. Die Ergebnisse werden im Programmfenster angezeigt und zusätzlich als Ergebnisdatei (.res) an der selben Stelle gespeichert, an der auch die Messwertdatei liegt.

Das Programm lädt beim Start automatisch alle benötigten R Pakete und die für die Diskriminierung verwendeten Objekte in den Speicher. Es setzt dafür eine korrekte Installation von R und der R Pakete voraus!

## 2. Installation

### 2.1. „Binary-Distribution“

Üblicherweise haben Sie als Nutzer die Binärdistribution von bcDiscriminator herunter geladen, d.h. die Version, die Sie mittels Doppelklick starten können. Um diese nutzen zu können, müssen sie auch die unter 2.3. erklärten Schritte durchführen, bevor das Programm sinnvoll anwendbar ist. Kopieren Sie bcDiscriminator außerdem vor der Verwendung in den Programme Ordner Ihres Rechners.

### 2.2. „Source-Distribution“

Sollten Sie Interesse haben, an der Weiterentwicklung von bcDiscriminator mitzuwirken oder Anpassungen vornehmen wollen, können Sie auch die Programmquellen herunter laden. Sie benötigen für die Weiterverarbeitung dieser die kostenlose Entwicklungsumgebung Xcode von Apple in der Version 2.4.1 oder neuer. Für Xcode 3 sind Anpassungen der Projekteinstellungen nötig. Diese können sie bei Bedarf von uns erfragen. Außerdem muss wie in 2.3. erklärt, das R.framework installiert sein, um die Quellen zu kompilieren.

## 2.3. Benötigte Ergänzungen

### 2.3.1. R

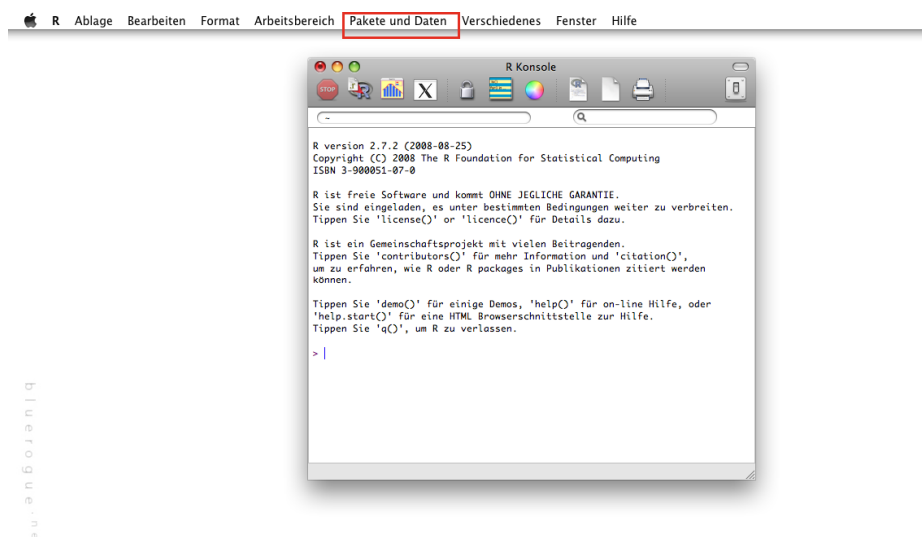
Die Erkennung von Fledermausrufen erfolgt anhand von Routinen, die im *open-source* Statistik Programm R implementiert sind. Das wiederum bedeutet, dass neben bcDiscriminator auch dieses Programm installiert sein muss. Zum Zeitpunkt der Erstellung dieses Dokuments sind die Versionen R 2.6.x und R 2.7.x kompatibel zu bcDiscriminator, diese Version erhalten Sie unter (alte Versionen innerhalb des Links old):

#### [R for Mac OS X](#)

Bitte installieren Sie R inklusive aller im Installationspaket enthaltenen Erweiterungen.

### 2.3.2. Benötigte R-Pakete

Innerhalb von R müssen zwei Pakete installiert werden, die für bcDiscriminator benötigt werden. Es handelt sich dabei um *randomForest* und *kernlab*.



Nach dem Starten von R können die beiden Pakete mittels des Menüpunkts „Pakete und Daten“ -> „Paketinstallation“ installiert werden (siehe Screenshot!). Stellen Sie sicher, dass im erscheinenden Fenster „CRAN (binaries)“ im Aufklappmenü ausgewählt ist. Nach Klick auf „Liste holen“

erscheint ein Dialog, der Sie bittet, eine Quelle für die Downloads auszuwählen. Als zuverlässig hat sich hier der Server in Österreich (Austria) erwiesen, der auch Mac Versionen der Pakete vorhält. Aus der Liste wählen Sie die beiden Pakete *kernlab* und *randomForest* und installieren diese, ohne andere Einstellungen zu verändern. Sie können daraufhin R beenden. Sollten keine Fehler aufgetreten sein, ist bcDiscriminator jetzt einsatzfähig.

## 3. Benutzung von bcDiscriminator

### 3.1. Starten

Nach dem Start von bcDiscriminator startet das Programm die R-Umgebung und lädt die benötigten R-Routinen. Dies nimmt in Abhängigkeit der Rechengeschwindigkeit zwischen 4 und 20 Sekunden in Anspruch. Erst wenn dieser Prozess abgeschlossen ist, kann bcDiscriminator benutzt werden.

### 3.2. Aufbau des Programmfensters

Die obere Hälfte des Programmfensters ist unterteilt in eine - zu Anfangs - leere Ergebnisanzeige (rechte Fensterseite), sowie einen Knopf zur Auswahl der zu diskriminierenden Dateien (linke Seite). In der unteren Programmhälfte befindet sich ein Textfeld, das aktuelle Mitteilungen des Programms ausgibt.

Sobald das Programm alle benötigten R-Pakete geladen hat, kann mittels des Knopfs „**Choose Input File**“ eine einzelne oder mehrere Dateien aus einem Verzeichnis der Festplatte ausgewählt werden. Wie in anderen Anwendungen, kann die Mehrfachauswahl durch Kombination mit den Tasten SHIFT oder ALT getroffen werden. Nach Auswahl der Dateien beginnt bcDiscriminator automatisch mit der Analyse der Dateien. Ist nur eine Datei ausgewählt, können die Ergebnisse im Anschluss an die Analyse betrachtet werden. Wurden mehrere Dateien gewählt, wird ohne Pause jede Eingabedatei abgearbeitet. Es erscheinen dann nur die Ergebnisse der letzten, analysierten Datei. Ist „**omit results file**“ ausgewählt, werden die Ergebnisse zu jeder Datei nicht gespeichert.

### 3.3. Voreinstellungen - Preferences

Seit der Version 1.13 kann die Funktionsweise von bcDiscriminator durch Voreinstellungen (Preferences) beeinflusst werden.

**Dezimaltrennzeichen:** Für Apple-System mit einem „.“ als Dezimaltrennzeichen wurde eine Voreinstellung ergänzt, da ansonsten die von bcAdmin erzeugten Messwertdateien nicht gelesen werden können.

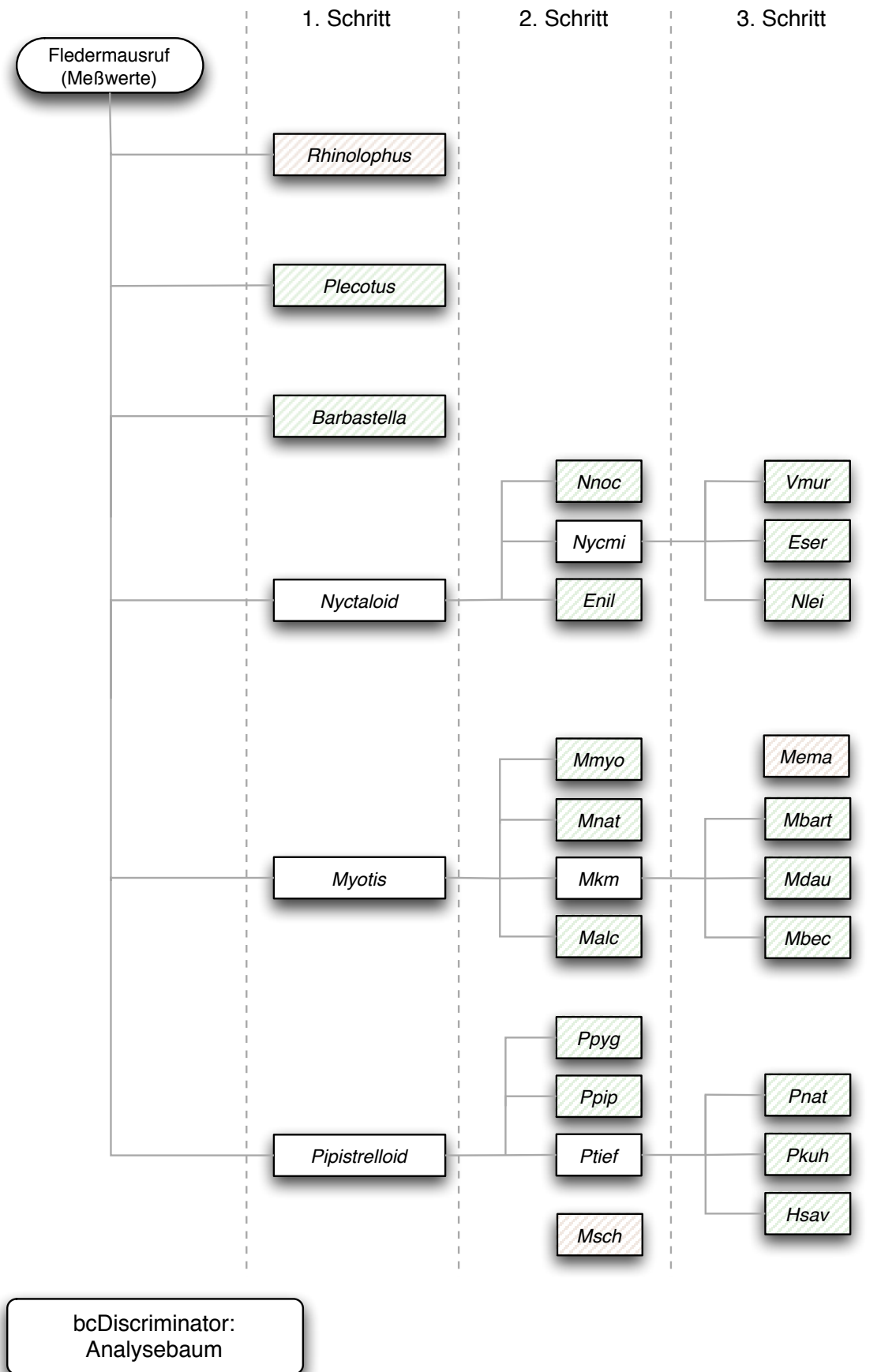


Abbildung 1: Gruppen der einzelnen Analyseschritte



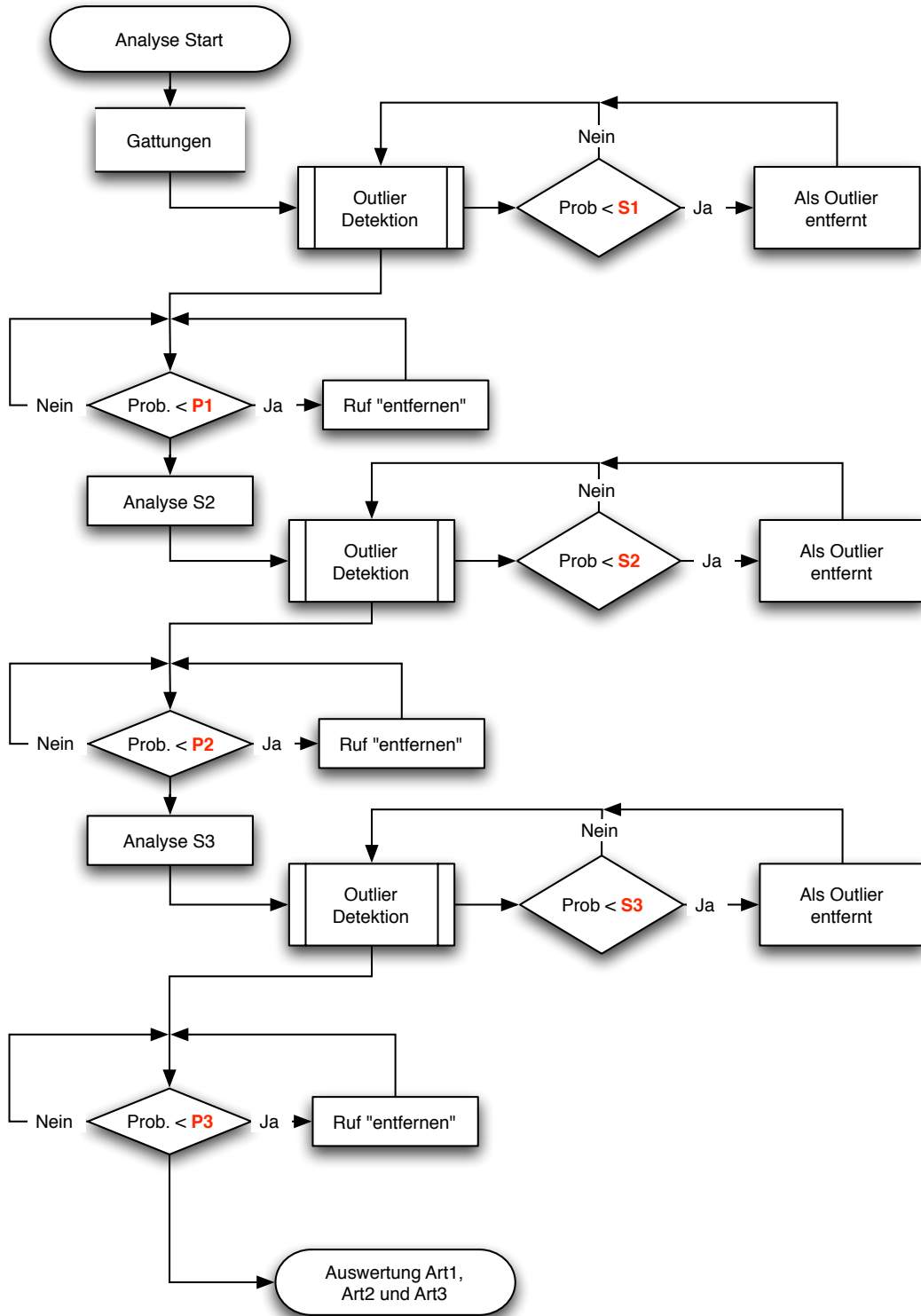


Abbildung 2: Entscheidungsbaum

**Artkriterien:** Die Kriterien für das Finden einer Art sind mittlerweile wählbar. Standardwerte sind 3 Rufe und eine Wahrscheinlichkeitsschwelle von 0,6 (=60%).

**Outlier-Schwelle:** Das Programm verwendet seit der Version 1.13 intern individuelle Werte je Gruppe/Art für die Erkennung von Ausreißern. Es kann jedoch auch eine eigene Schwelle, die dann für alle Arten gilt, gewählt werden.

**Neue Zusammenfassung:** Die Bildung des Mittelwerts der Wahrscheinlichkeiten aller Rufe einer Art wurde überarbeitet. Der neue Algorithmus berücksichtigt neben den Wahrscheinlichkeiten auch die Rufanzahl je Art bei der Mittelung. Es kann gewählt werden, ob das neue Verfahren (empfohlen) oder das herkömmliche verwendet werden soll.

### 3.4. Zusammenspiel mit bcAdmin

Das Programm bcAdmin liefert bei der Analyse von Rufaufnahmen Messwerte, die kompatibel zu bcDiscriminator sind. Es kann sowohl eine einzelne, als auch eine beliebige dieser csv-Dateien mittels „**Choose Input File**“ ausgewählt werden. Die Vorgehensweise im Zusammenspiel von bcAdmin und bcDiscriminator ist recht einfach. Nach Analyse und Messwerterzeugung in bcAdmin werden die entstandenen csv-Dateien in bcDiscriminator ausgewertet. Dabei werden Ergebnisdateien (.res) erzeugt, die wiederum in bcAdmin importiert werden können.

## 4. Funktionsweise von bcDiscriminator

### 4.1. Analyseschritte

#### 4.1.1. Diskriminierung und Ausreisser-Erkennung

Die Artanalyse wird je Sequenz Rufweise vorgenommen, um auch mehrere Arten innerhalb einer Aufnahme zu erkennen. Die unterschiedenen Gruppen/Arten sind in der Abbildung 1 gezeigt. Der Zeit fehlen die Gattung *Rhinolophus* (*R. hipposideros* und *R. ferrumequinum* werden jedoch

bereits von bcAdmin erkannt!), *Miniopterus* und *Tadarida*. Ebenso fehlen zur Zeit noch die Arten *Myotis dasycneme*, *M. emarginatus*, *M. cappaccini* und *Nyctalus lasiopterus*.

Die verwendeten Abkürzungen sind wie folgt:

Kürzel	Art	Kürzel	Art
Nnoc	<i>Nyctalus noctula</i>	Mdau	<i>Myotis daubentonii</i>
Nlei	<i>Nyctalus leisleri</i>	Mbec	<i>Myotis bechsteinii</i>
Enil	<i>Eptesicus nilssonii</i>	Ppyg	<i>Pipistrellus pygmaeus</i>
Eser	<i>Eptesicus serotinus</i>	Ppip	<i>Pipistrellus pipistrellus</i>
Vmur	<i>Vespertilio murinus</i>	Pnat	<i>Pipistrellus nathusii</i>
Mmyo	<i>Myotis myotis</i>	Pkuh	<i>Pipistrellus kuhlii</i>
Mnat	<i>Myotis nattereri</i>	Hsav	<i>Hypsugo savii</i>
Malc	<i>Myotis alcathoe</i>		
Mbart	<i>Myotis brandtii/mystacinus</i>		

Die Analyse beginnt immer auf Gattungsniveau. Nach einer Diskriminierung mittels *randomForest* wird für jeden Ruf überprüft, ob es sich um einen fälschlicherweise einer Gattung, Gruppe bzw. Art zugeteilten Ruf handelt (Methode: *svm*). Solche Ausreisser werden bei weiteren Analyse-schritten ausgenommen und auf den letzten sicheren Arteintrag gesetzt. Nach dem alle Rufe so analysiert sind, dass keine weiterführenden Schritte mehr möglich sind, werden die Artinformationen der Aufnahme ermittelt. Die Abbildung 2 zeigt den Analysevorgang, der für jeden Ruf Schrittweise (S1, S2 und S3) vorgenommen wird.

#### 4.1.2. Ermittlung des Ergebnisses / Artausgaben

Ziel von bcDiscriminator ist es, an Hand der Messwerte der Rufe einer Aufnahme, die aufgezeichnete Fledermausart zu ermitteln. Da regelmäßig auch mehrere Fledermausarten gleichzeitig aufgezeichnet werden, ist bcDiscriminator darauf ausgelegt, bis zu drei Arten zu erkennen. Um si-

cher eine Art zu bestimmen, werden die Anzahl der Rufe einer Art und die ermittelte, mittlere Wahrscheinlichkeit für die Art genutzt. Werden für beide Werte im Programm gesetzte Schwellen (0,6 für die Wahrscheinlichkeit und 3 für die Anzahl Rufe) nicht erreicht, da es sich z.B. um eine leise Aufnahme mit nur wenigen Rufen handelt, wird die Art ausgegeben, die die höchste Rufzahl und Wahrscheinlichkeit erreicht hat.

Ansonsten wird der Baum nach der Rufanzahl der Blätter, also Endpunkte der Diskriminierung, sortiert. Die Art mit den meisten Rufen (> 2 Rufe) bzw. der höchsten Wahrscheinlichkeit (>60%) wird als Art 1 gespeichert. Liegen noch weitere, potenziell gefundene Arteinträge vor, und erfüllen diese die genannten Kriterien, werden sie als weitere Arten (2 und 3) gespeichert.

Nach erfolgreicher Analyse stehen die Ergebnisse im Programmfenster in einer Baumdarstellung zur Verfügung und werden ebenso in Form einer einfachen Textdatei (.res) an die selbe Stelle geschrieben, an der sich die Messwerte auf der Festplatte befinden. Seit der Version 1.14 ergänzt bcDiscriminator in der .data (oder .bcCalls) Datei (=Rufverläufe) zusätzlich das Endergebnis (Art, Wahrscheinlichkeit) je Ruf. Diese Information kann dann z.B. von bcAdmin in der Rufdarstellung geplottet werden. Ist keine Rufverlaufsddatei vorhanden, oder weicht die Rufanzahl in den Dateien voneinander ab, unterbleibt der Eintrag.

#### **4.2 Messwertdateien und Messwerte**

bcDiscriminator ist so ausgelegt, dass es Messwertdateien im .csv Format einliest. Diese werden zum Zeitpunkt der Erstellung dieses Handbuchs nur von bcAdmin mit entsprechenden Variablen ausgegeben. Die csv Datei muss in der ersten Zeile Spaltenüberschriften tragen, ab der zweiten Zeile folgen die Messwerte. Die Spalten sind durch ein Tabulatorzeichen getrennt, Fließkommazahlen müssen mit Komma oder Punkt als Trennzeichen in den Spalten stehen. Die Spaltentitel lauten:

Feld	Beschreibung
Datei	Dateiname der Tonaufnahme, nicht benötigt
Art	Artnamen, nicht benötigt
Ruf	Rufnummer, nicht benötigt
Dur	Rufdauer (ms)
Sfreq	Startfrequenz (kHz)
Efreq	Endfrequenz (kHz)
Stime	Beginn des Rufes innerhalb der Datei (ms), nicht benötigt
NMod	siehe: im Anschluss der Tabelle
FMod	siehe: im Anschluss der Tabelle
FRmin	siehe: im Anschluss der Tabelle
Rmin	siehe: im Anschluss der Tabelle
tRmin	siehe: im Anschluss der Tabelle
Rlastms	siehe: im Anschluss der Tabelle
Flastms	siehe: im Anschluss der Tabelle
X10, X11 ... X60	siehe: im Anschluss der Tabelle
X62, X64 ... X148	siehe: im Anschluss der Tabelle

„FRmin“, „Rmin“ und „tRmin“ werden an der Stelle des Rufes mit der geringsten Steigung gemessen. Es handelt sich dabei um Werte der dort anliegenden Frequenz („FRmin“), der Steigung („Rmin“) und dem Zeitpunkt innerhalb des Rufes („tRmin“), gemessen vom Rufende. Innerhalb der letzten Millisekunde des Rufes werden ebenso Frequenz („Flastms“) und Steigung („Rlastms“) an der Stelle mit der geringsten Steigung gemessen.

Die Spalten X10 bis X148 ergeben sich wie folgt: Über den Rufverlauf wird

im Abstand von 100  $\mu$ s jeweils die Frequenz gemessen, daraus resultiert eine Abbildung des Rufverlaufs. Diese Messwerte werden in Frequenzklassen zusammengefasst, d.h. es wird ein Histogramm (*moving averages*) mit 1 kHz Klassenbreite im Bereich von 10 kHz bis 60 kHz und mit 2 kHz Klassenbreite von 60 kHz bis 150 kHz erstellt. Die Werte des Histogramms werden beginnend mit der Klasse 10-11 kHz (X10) bis zur Klasse 148-150 kHz (X148) gespeichert. „NMod“ und „FMod“ können direkt aus diesen Histogramm Daten bestimmt werden, sie geben die Anzahl an Messwerten der Frequenzklasse mit dem höchsten Wert („NMod“), sowie die Untergrenze dieser Frequenzklasse (in kHz) an („FMod“).

Ein Ruf der Zwergfledermaus ergibt z.B. die folgenden Werte:

Feld	Wert		Feld	Wert
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		...	...

## 5. Anpassung für andere Rufparameter

bcDiscriminator kann angepasst werden, um andere Rufparameter für die Diskriminierung zu verwenden. Dazu müssen entsprechende *randomForest*-Objekte erzeugt werden, die die selben Schritte abdecken, wie bisher von bcDiscriminator verwendet. Für diese Anpassung oder Erweiterung sind jedoch gute Kenntnisse in Statistik und R dringend nötig.

D.h. wenn sie eine umfangreiche Rufsammlung heimischer Fledermausarten haben und andere Messwerte nutzen, als die derzeit in bcDiscriminator implementierten, können Sie mittels Messwertdateien Ihrer Rufsammlung in R neue Diskriminierungsobjekte erzeugen. Neben den Diskriminierungsobjekten empfehlen wir auch die Erzeugung der Objekte für eine Ausreisser-Erkennung, die mittels *svm* erhalten werden. Mit diesen Objekten ersetzen Sie dann die in bcDiscriminator enthaltenen. Wir empfehlen, dass Sie die in bcDiscriminator mitgelieferten Objekte vorher gut untersuchen und sich mit den verwendeten Artnamen und dem Aufbau dieser Objekte vertraut machen.