

# Kurzanleitung für TURBOVEG 2.0 und JUICE 7.0 im Rahmen der Lehrveranstaltung „Auswertung und Analyse vegetationskundlicher Daten“, Version 2.4, September 2015

Christian Berg & Martin Magnes, Institut für Pflanzenwissenschaften, Universität Graz

## Inhaltsverzeichnis

1.	Arbeiten mit TURBOVEG .....	1
1.1	Computerkonfiguration und Voreinstellungen .....	1
1.2	Dateneingabe in TURBOVEG .....	3
1.3	Datenausgabe.....	6
2.	Arbeiten mit JUICE.....	7
2.1	Einführung.....	7
2.2	Daten einlesen in JUICE.....	7
2.3	Der Grundaufbau von JUICE .....	7
2.4	Erste Schritte der Tabellenarbeit .....	9
2.5	Ordnen der Tabelle .....	10
2.6	Klassifikation (pflanzensoziologische Tabellenarbeit).....	12
2.7	Arbeiten mit ökologischen Zeigerwerten .....	18
2.8	Multivariate Statistik: Detrended correspondence analysis mit JUICE und R .....	21
2.9	Verarbeiten von Pflanzen-Eigenschaften in Vegetationstabellen .....	25
2.10	Biodiversitätsuntersuchungen in Vegetationstabellen .....	26
3.	Weitere Tipps und Tricks zum Arbeiten mit JUICE .....	28
3.1	Datenvorbereitung .....	28
3.2	Klassifikation.....	31
3.3	Analyse.....	31
4.	References .....	34

## 1. Arbeiten mit TURBOVEG

TURBOVEG ist eine von Stephan Hennekens entwickelte Microsoft Windows Anwendung zum Speichern und Selektieren von großen Mengen vegetationskundlicher Daten (Vegetationsaufnahmen, relevés). Verschiedene Exportfunktionen ermöglichen die Weiterverarbeitung in anderen Programmen wie JUICE, ACCESS, EXCEL, ArcGIS, GoogleEarth und viele andere. Die Nutzung für Publikationen sollte mit Verweis auf HENNEKENS & SCHAMINEE (2001) immer angegeben werden.

### 1.1 Computerkonfiguration und Voreinstellungen

TURBOVEG läuft auf allen Windows Computern, nicht auf Apple. Die Einzeluserversion kostet 450 €, Studierende können eine freie Kopie über ihre Lehrkräfte erhalten, welche sorgfältig zu behandeln ist. Weitere Informationen und ein ausführliches Handbuch gibt es zum Download auf der Webseite <http://www.synbiosys.alterra.nl/turboveg/>.

Das Programm sollte im Stammverzeichnis installiert werden, nicht im Verzeichnis Programme.

#### 1.1.1 Generelle Struktur der Benutzeroberfläche

TURBOVEG verfügt über eine feste Menüleiste mit Pull-down-Menüs, welche den Zugriff auf alle Befehle ermöglicht. Häufige Befehle werden zusätzlich in einer darunter liegenden Button-Leiste verlinkt. Bei geöffneter Database zeigen sich drei Fenster in Tabellen-Struktur (**Abbildung 1**): eine Kopfdaten-Tabelle [header data], eine Arten-Tabelle ([species data] mit Informationen zur Abundanz [cover] und zur Schicht [layer] der Aufnahme, in der sich gerade der Cursor befindet) und eine Tabelle mit den Aufnahmeummern der letzten durchgeführten Selektion [selected relevés]. Ist keine Database geöffnet, ist nur das Selektions-Fenster sichtbar.

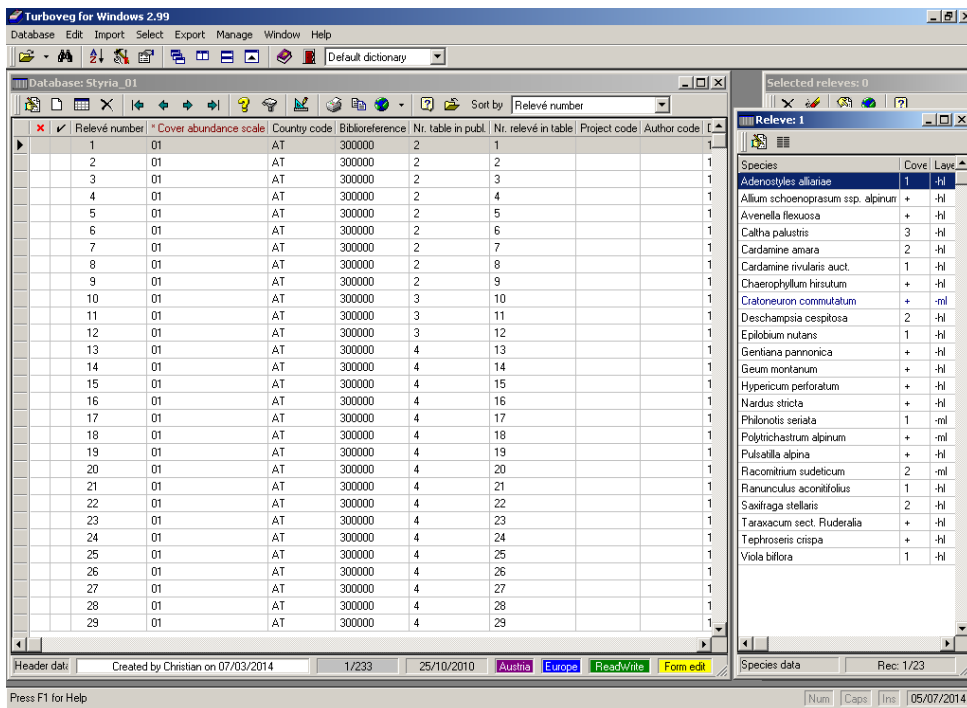


Abbildung 1: Aufbau von TURBOVEG

Das Programm ist für den Gebrauch der Maus optimiert, kann aber auch vollständig mit der Tastatur gesteuert werden (Tastaturbefehle sind im Handbuch nachzulesen).

### 1.1.2 Dateistruktur der Database

Alle Daten werden in einer relationalen Datenbank abgelegt, welche hier als „Database“ bezeichnet wird. Man kann sämtliche Daten in einer Database verwalten oder auch in mehreren; letzteres ist aber nicht notwendig, außer man arbeitet in Gebieten mit verschiedenem Artenpool und damit mit verschiedenen Artenlisten. (z. B. Österreich und Montenegro).

Jede Database baut sich aus 3 miteinander über Schlüsselfelder verknüpfte Datenbank-Dateien und drei Index-Dateien auf:

Datenbank-Dateien	Index Dateien
TVHABITA.DBF	TVHABITA.CDX (Tabelle mit header data)
TVABUND.DBF	TVABUND.CDX (Tabelle mit species data)
REMARKS.DBF	REMARKS.CDX (Tabelle mit Bemerkungen [field remarks])

Der Name der Database ist der Verzeichnisname, in dem diese Dateien liegen. Zum Öffnen von einer oder mehrerer Databases dient der Befehl Database → Open. Will man eine Datenbank weitergeben, brauchen nur die 3 DBF-Dateien weitergegeben werden, die Indexdateien bauen sich beim ersten Start neu auf. Für eine Weitergabe, insbesondere beim Versenden über Internet, empfiehlt es sich, die 3 Dateien in einer ZIP-Datei zu komprimieren, da Datenbankdateien grundsätzlich anfällig für Viren sind.

TURBOVEG arbeitet grundsätzlich mit vorgegebenen, stets gleichen Dateinamen. Zur Unterscheidung (z. B. verschiedene Databases) werden verschiedene Verzeichnisse (Ordner) verwendet.

### 1.1.3 Andere wichtige Dateien für die Arbeit mit TURBOVEG

Jede Database bekommt Informationen von zusätzlichen Tabellen, welche über Schlüsselfelder mit den Tabellen der Database verknüpft sind: die Artenliste [species list] und die Hilfslisten [popup lists]. Die Artenliste enthält alle Arten (Gefäßpflanzen, Moose, Flechten und einige Makroalgen einschließlich Synonymen) und stellt so die korrekte Schreibweise der Pflanzennamen sicher. Die popup lists bieten Hilfe bei der Eingabe der Kopfdaten. Bei Feldern, für die eine Hilfsliste verfügbar ist, erscheint während der Eingabe ein Fragezeichen links vom Eingabefeld. Klickt man darauf, erscheint die jeweilige popup list, z. B. für die Umrechnung der Exposition von Himmelsrichtungsangaben in Grad. Die Popup Listen können vom Nutzer geändert und mit eigenen Informationen gefüllt werden, z. B. die Vergabe verschiedenen Projekt-Codes.

Bitte nie die Artenliste ändern! Dies kann zu ernsthaften Problemen beim Datenaustausch führen. Ist eine Art nicht in der Liste, liegt es daran, dass der Nutzer sie nicht findet. Es gibt keine fehlenden Arten in der species list, höchstens mal fehlende Synonyme!

## 1.2 Dateneingabe in TURBOVEG

Daten können nur in eine vorhandene Database eingegeben werden. Deshalb müssen wir zuerst eine erstellen. Über den Befehl `Manage → Change database directories` können Sie das Verzeichnis festlegen, in das die Database geschrieben werden soll. Standardmäßig gibt TURBOVEG hier das Verzeichnis **Data** vor.

### 1.2.1 Erstellen einer Database

**Database → New**

Dieser Befehl führt zu einer einfachen Maske. Sie brauchen hier lediglich einen Namen [Database name] vergeben, die species list auswählen und den Rang der Aufnahmeummern für die Database vergeben. Letzteres ist von Bedeutung, wenn Sie mit mehreren Databases arbeiten, denn die Aufnahme-Nummern, welche das wichtigste key field in TURBOVEG sind, dürfen sich nicht überschneiden. Auch der Umfang der Database wird hier festgelegt (z. B. 30.000 Aufnahmen). Geben Sie einfach 1 bis 200 für Ihre erste Übungs-Database ein. Die beiden Fragen nach der Struktur der Database sollten auf „Standard“ stehen. Mit „Create“ erstellen Sie die Database. Diese ist zunächst leer und besitzt eine festgelegte Struktur bezüglich der Kopfdaten, die Sie zwar nicht ändern, wohl aber erweitern können.

### 1.2.2 Erweitern der Struktur einer Database

**Database → Modify structure**

Die Standardstruktur von TURBOVEG hat 2 Nachteile: die Koordinaten-Abfrage basiert auf dem UTM-System, und es gibt kein Feld, in das man verbale Informationen zur Lokalität eingeben kann. Diese beiden Nachteile wollen wir nun abstellen.

Ergänzen wir also zuerst zwei Felder für ein auf Dezimalgrad basiertes Koordinatensystem. Zuerst die geographische Länge [longitude]. Als Feldname [field name] nehmen wir `DEG_LON`, als Feldtyp [Type] `n` (=numeric), als Feldlänge [Len] 13 und als Dezimalwert [Dec] 7. Mit `Add` wird das Feld erstellt. Das wiederholen wir jetzt mit einem Feld für die geographische Breite [latitude], welches `DEG_LAT` heißen soll (diese Feldnamen korrespondieren mit einigen Funktionen von JUICE).

Jetzt noch ein Feld, in das wir die Ortsangabe schreiben können. Wir nennen dieses Feld `LOCATION`, der Feldtyp ist diesmal `c` (=character), die Feldlänge 99 und der Dezimalwert natürlich 0. Nach drücken von `Add` sollten nun die 3 neuen Felder am Ende der gezeigten Feldübersicht stehen. Angeschlossen wird das Ganze mit `Rebuild`.

Die Möglichkeit, eigene Kopfdaten hinzuzufügen ist sehr wertvoll für eigene, spezielle Fragestellungen, da damit alle erdenklichen Daten (z. B. Messwerte) den Aufnahmen zugeordnet werden können.

### 1.2.3 Organisieren der Kopfdatenanzeige

**Database → Organize header data**

Mit dieser nützlichen Funktion können Sie aus dem angebotenen Kopfdatensatz nur jene herausuchen, für die Sie in dem jeweiligen Projekt Daten haben. Beim Eingeben von Felddaten können Sie z. B. alle Angaben, die mit der Biblioreferenz zusammenhängen, ausblenden. Damit wird das Formular für die Kopfdateneingabe erheblich übersichtlicher. Die Felder werden nur ausgeblendet, nicht etwa physisch gelöscht.

### 1.2.4 Bearbeiten von Popup lists

Die Popup lists bieten Hilfe bei der Eingabe der Kopfdaten. Auch diese sollten wir also, bevor wir mit der Dateneingabe beginnen, noch bearbeiten. Mit dem Befehl

**Manage → Popup list → Edit → ...** (z. B. Project code)

können wir jenen Feldern, die über Popup listen codiert werden, eigene Codes hinzufügen. So können wir unseren Aufnahmen einen eigenen Project code geben, oder uns als Autor der Aufnahmen in den Autor code eintragen. Unter `cover scales` können Sie sich die Umrechnung der Abundanzcodes (z.B. Braun-Blanquet new) in Bedeckungsprozente ansehen (nicht ändern!). Diese Datei ist der Schlüssel zur Umrechnung verschiedener Aufnahmemethoden, so dass alle Aufnahmen miteinander verglichen werden können.

### 1.2.5 Einfaches Importieren von Daten aus xml-Dateien

Jetzt wollen wir unsere leere Database erst einmal mit ein paar Daten füllen:

**Import → Standard Turboveg XML-File**

Wir suchen in dem entsprechenden Verzeichnis die bereitgestellte Datei `tvexport.xml` und klicken auf `Öffnen`. In dem sich öffnenden Fenster tragen wir die zu verwendende species list (z. B. Austria) ein (nie das Feld `Update species list` aktivieren!) und aktivieren unten das Kästchen `Import releve and species data`. Im Message log zeigt das Programm einige Fehlermeldungen, die auf Unterschiede in den Popup Listen zurückzuführen sind, und die uns jetzt nicht interessieren sollen. Die Daten sind nun eingelesen.

Bei dieser Datei hatten wir die Arten schon an die Austria-Liste angepasst. Findet TURBOVEG einen Artnamen, der nicht mit der Liste übereinstimmt, kann diese Art über den 3-3-code manuell ausgewählt werden.

### 1.2.6 Dateneingabe

Vegetationsaufnahmen können als Einzelaufnahmen eingegeben werden (das ist sinnvoll bei handschriftlichen Aufnahmen wie in unserem Fall), oder auch als Tabelle (sinnvoll für die Eingabe bereits publizierter Vegetationstabellen). Für Einzelaufnahmen nutzen wir den Befehl

**Edit → Add a relevé**

Es erscheint eine Eingabemaske für die Kopfdaten. Form 1 enthält die Standard-Felder, Form 2 die 3 von uns ergänzten Felder. Rote Felder [Mandatory fields] müssen eingegeben werden: Das betrifft aber nur das Feld „Cover abundance scale“, alle anderen können theoretisch frei gelassen werden.

Geben Sie nun die Kopfdaten Ihrer Aufnahme ein. Eingabehilfen über die popup lists sind mit einem ? gekennzeichnet. So können Sie sehen, dass *Cover abundance scale* nach der Schätzskala fragt, nach der in Ihrer Aufnahme die Bedeckung der Arten geschätzt wurde. Im Fall der Skala mit 2m, 2a und 2b handelt es sich um „Braun-Blaquet new“ mit dem Code 02.

Die Fragen nach der *Biblioreference* und der *Nr. table in publ.* sowie *Nr. relevé in table* bezieht sich auf Literaturdaten, diese Felder bleiben bei der Eingabe von Original-Felddaten leer. In die *Remarks* können Sie alles eingeben, was Sie ansonsten nicht in den vorgegebenen Feldern unterbringen können.

Bestätigen Sie die Eingabe der Kopfdaten mit *Save*. Das Programm springt jetzt in eine Eingabemaske für die Art Daten. Die Arten werden mit einem 3-3-Code gefunden, d. h Sie geben 3 Buchstaben des Gattungsnamens und 3 Buchstaben des Artnamens ein und finden dann in dem linken Fenster die entsprechende Auswahl.

**Achtung:** Viele 3-3-Codes führen nicht zu einem eindeutigen Ergebnis, Sie müssen also nach der Eingabe noch den Cursor auf die richtige Art lenken. Suchen Sie *Carex hirta* und geben *Car-hir* ein, findet TURBOVEG auch *Cardamine hirsuta*! Also nicht blind dem Code vertrauen!

Steht der Cursor auf dem richtigen Namen (blau hinterlegt), kommen Sie mit Doppelklick oder einfach *Enter* zur Eingabe der Schicht (*layer*). Ein pull-down-Menü zeigt die Eingabemöglichkeiten (siehe Tabelle 1).

**Tabelle 1:** Vegetationsschichten und ihre Bezeichnung in TURBOVEG und JUICE

Schicht	Englisch	Numerisch	Abkürzung
Keine Schicht	No layer	0	-
Erste Baumschicht	Tree layer (high)	1	t1
Zweite Baumschicht	Tree layer (middle)	2	t2
Dritte Baumschicht	Tree layer (low)	3	t3
Erste Strauchschicht	Shrub layer (high)	4	s1
Zweite Strauchschicht	Shrub layer (low)	5	s2
Krautschicht	Herb layer	6	hl
Juvenile	Juvenile	7	jl
Sämlinge < 1 Jahr alt	Seedling (< 1 year)	8	sl
Moosschicht	Moss layer	9	ml

Wieder *Enter* führt zum nächsten Eingabefeld, wo der entsprechende Abundanz-Code (*cover*, also *r*, *+*, *1*, *2m* usw.) einzugeben ist. Klicken auf *Add* oder 2 mal *Enter* schreibt die Art dann in den Datensatz für diese Aufnahme. Diesen Vorgang muss man wiederholen, bis alle Arten mit korrekter Schichtangabe und Abundanz eingegeben sind. Komfortabler Weise bleibt der Schichteintrag bei der nächsten Art erhalten. Nach der letzten Art sollte man nochmal eine Überprüfung der Arteinträge (insbesondere der Schichten) vornehmen und mit *Save* schreibt TURBOVEG die Aufnahme in die Database. Geben Sie eine weitere Aufnahme ein, können Sie die Kopfdaten der letzten Aufnahme übernehmen und brauchen Sie nur noch jene Angaben eintragen, die sich geändert haben.

Der Vollständigkeit halber sei erwähnt, dass Sie auch Tabellen eingeben können, die eine Artenliste, aber mehrere Aufnahmen enthalten, und dass es komfortable Einlesemöglichkeiten für Aufnahmen aus anderen Programmen, (z. B. EXCEL) gibt.

### 1.2.7 Importieren von Daten aus EXCEL

Diese Funktion erlaubt es, alte Vegetationsaufnahmen, die in anderen Programmen wie *Tab*, *Hitab*, *Sort*, *EXCEL*, *ACCESS*, *DBASE* oder einfach als Text digital vorliegen, in TURBOVEG einzulesen. Am einfachsten geht es mit EXCEL, so dass Sie zuerst die Aufnahmen aus anderen Programmen in EXCEL überführen sollten. Wie es von dort weitergeht, soll hier kurz erläutert werden.

Es dürfte jedem klar sein, dass wenn man Texte in Datenbanken einlesen will, es sowohl auf die Anordnung des Textes, auf die einheitliche Trennung der einzelnen Textteile und auch auf das Format einzelner Textfelder ankommt. Normale Vegetationstabellen enthalten Kopf und Art Daten in einer Tabelle, häufig Leerzeilen oder Erläuterungen innerhalb der

Artenliste, fast nie eine eigene Spalte für die Schichtangaben und am Ende oft einen Block mit selten vorkommenden Arten, der von der Tabellenstruktur komplett abweicht. Deshalb erfordert der Import umfangreiche Vorarbeiten.

Die Aufnahmenummern werden aufsteigend durchnummeriert (mit 1 beginnend). Art- und Kopfdaten müssen nun getrennt werden, wobei die Aufnahmenummern und die Reihenfolge beizubehalten ist. Die Arttabelle bleibt in der üblichen Form erhalten, die Kopfdaten müssen transponiert werden, so dass die Aufnahmenummern bei den Arttdaten in der 1. Zeile, bei den Kopfdaten in der 1. Spalte stehen. Wenden wir uns zuerst der Arttabelle zu. Im Feld A1 muss „Relevé number“ stehen, rechts folgen ein Leerfeld und dann die Aufnahmenummern mit 1 beginnend. Im Feld A2 beginnen abwärts die Artnamen ohne Leerzeilen. Alle anderen Zeilen müssen gelöscht werden, und ein ggf. unter der Tabelle stehender Artenblock muss hier einsortiert werden (Abundanzwerte den Aufnahmespalten zuordnen!). In die B-Spalte kommen die Schichtenangaben (entweder 1 bis 9 oder tl, sl, hl, ml), ab der C-Spalte folgen die Abundanzwerte-Codes. Die Leerfelder können leer bleiben oder Punkte enthalten.

Wenden wir uns nun dem 2. Datenblatt mit den transponierten Kopfdaten zu. Im Feld A1 steht auch hier „Relevé number“, die Nummern folgen, mit 1 beginnend, in der A-Spalte darunter. In der ersten Zeile stehen die Bezeichnungen für die Kopfdaten (z. B. Datum, Größe der Aufnahmefläche usw.), die Werte stehen jeweils darunter. Keine Leerzeilen/-spalten lassen. Obwohl TURBOVEG auch falsche Formate erst einmal einliest, ist die Änderung der Formate in EXCEL oft viel einfacher zu bewerkstelligen. Die wichtigsten Formate, die dabei zu beachten sind, stehen in Tabelle 2, die wichtigste Änderung betrifft das Datum.

**Tabelle 2:** Umwandlung bestimmter Kopfdatenformate zum Einlesen in TURBOVEG

Kopfdaten-Typ	Übliche Schreibweise	TURBOVEG-Schreibweise
Datum	15.8.2010	20100815
Koordinaten (in WSG 84)	54° 16' 23''	54.0809722
Exposition (Aspect)	NW	315 [Grad] (Umrechnung siehe popup list)
Baumhöhe	600 [cm]	6 [m]
Höhe Krautschicht	0,8 [m]	80 [cm]

Nun können wir die Datei im EXCEL-Format abspeichern und mit dem Import beginnen. Dazu erstellen wir eine neue Database, deren Aufnahmenummern ebenfalls mit 1 beginnen. Mit dem Befehl

**Import → free format species data table**

öffnen wir die EXCEL-Datei, sämtliche darin angelegten Datenblätter werden angezeigt. Wir wählen das Blatt mit den Art-Daten aus. In dem angezeigten Fenster stellen wir nun lediglich die `Species column` und die `Layer column` entsprechend der Anzeige in der 1. Zeile ein. Mit `Next` werden die Daten mit der `species-list` verglichen und angezeigt. Übereinstimmende Namen haben in der 1 (unbezeichneten Spalte) nun eine schwarze Nummer, Namen, die so nicht in der `species list` stehen, eine blaue Null. Nach Doppelklick auf diese Namen erscheint das Arten-Auswahlfenster, in dem wir mit `ok` einen passenden Namen auswählen können. Sind alle Namen schwarz, muss jetzt noch der `cover scale` eingestellt werden, und die einzulesenden Spalten mit der Leertaste markiert werden. Der letzte Schritt ist die Zuordnung der Cover-Codes, die mit `replace all` zugeordnet werden. Achtung: Wenn das Zeichen für die Leerfelder kommt (z. B. Leerzeichen oder Punkt), muss man `ignore all` wählen! Mit `Complete` wird die Tabelle nun eingelesen.

Diese enthält nun noch keine Kopfdaten. Deshalb beginnt die Prozedur mit

**Import → free format header data table**

nun noch einmal. Diesmal wählen wir das Datenblatt mit den Kopfdaten aus. Nun muss jede Spalte durch Klicken auf das oberste (graue) Feld den Datenfeldern der Datenbank zugeordnet werden. Dabei muss man das erste Feld dem Datenfeld `Releve number` zuordnen und hier das Kästchen `Key field` anklicken. Das Kästchen `Add new fields automatically` sollte man nur aktivieren, wenn man alle Kopfdatenfelder aus der EXCEL-Datei einlesen möchte, für die es keine Entsprechung in der bisherigen Database-Struktur gibt. Ansonsten werden diese Felder einzeln nach drücken von `Complete` abgefragt. Jetzt müssen wir nur noch prüfen, ob alles richtig zugeordnet ist, dann können wir diese Aufnahmen mit dem Befehl `Append another database` an eine existierende Database ranhängen, wobei die Aufnahmenummern dann automatisch umbenannt werden

## 1.2.8 Nachträgliches Bearbeiten von Kopfdaten

Mit Doppelklick auf die entsprechende Kopfdatenzeile können bei einer geöffneten Database die Kopfdaten in einem Formular bearbeitet werden. Um Kopfdaten im angezeigten Tabellenmodus zu bearbeiten, muss man vorher über `Edit → Switch edit mode` den Bearbeitungsmodus aktivieren.

Besonders nützlich ist die Möglichkeit, Kopfdaten auch mit dem für Datenbanken üblichen Befehl `replace` zu bearbeiten. Dies funktioniert für ein Kopfdatenfeld für hintereinander liegende Gruppen von Aufnahmen. Man stellt den Cursor in die Aufnahme, ab der man Änderungen vornehmen will, und klickt dann auf das graue Kopffeld der entsprechenden Spalte. In dem sich öffnenden Fenster gibt man den von-bis-Bereich ein, für den der `replace`-Befehl gelten soll, sowie unter `value` die Werte, die man ersetzen will (`old` und `new`, z. B. den falschen Eintrag NW im `Aspect`-Feld gegen den richtigen Eintrag 315). Will man leere Felder mit ein und demselben Wert füllen, lässt man den `old`-Wert leer. Der Befehl `replace with data from other item` ermöglicht das Umschreiben von Kopfdaten

von einem Feld in ein anderes. Das kann sehr nützlich sein, wenn man 2 databases zusammenführt und die fusionierte Database dann 2 verschieden benannte Kopfdatenfelder mit dem gleichen Inhalt enthält.

### 1.2.9 Georeferenzieren mit Google Maps

Für viele Auswertungen braucht man heutzutage Lagekoordinaten der Aufnahmen. TURBOVEG bietet eine komfortable Möglichkeit, sofern man mit dem Internet verbunden ist und die Lage der Aufnahmefläche genau kennt (z. B. aus einer Lageskizze). Dabei kommen jetzt die beiden Felder DEG\_LAT und DEG\_LON zum Einsatz. Nach dem Befehl

**Edit → Georeference using Google Maps**

startet Google Maps. Suchen Sie Ihr Untersuchungsgebiet. Wenn Sie die genaue Lage der Aufnahme gefunden haben (ggf. mit Hilfe der Satellitenbild-Ansicht), klicken Sie mit der linken Maustaste auf die Stelle. Die Koordinaten werden dadurch im Clipboard gespeichert. Gehen Sie nun in der Kopfdatentabelle in das leere Feld DEG\_LON. Mit der Rechten Maustaste fragt Sie das Programm, ob Sie dort die Länge oder die Breits hineinschreiben wollen. Schreiben Sie die Longitude in das Feld DEG\_LON und die Latitude in das Feld DEG\_LAT.

## 1.3 Datenausgabe

**Select → Built query**

Eine der Stärken von TURBOVEG ist die Möglichkeit, alle erdenklichen Gruppen von Aufnahmen zu selektieren und zu exportieren. Dazu können die Kopfdaten (*header data*, z. B. Autor- oder Projektcodes, oder einfach Aufnahmenummern), und/oder Kombinationen von Arten *Species data* (And/Or) unter Berücksichtigung der Bedeckungsgrade und mit der Möglichkeit, Arten auszuschließen (And not), zum Einsatz kommen. Es können auch Listen von Aufnahmenummern eingelesen werden, und Suchkriterien-Kombinationen können gespeichert und später wieder geladen werden. Mit *Execute* wird der Selektionsvorgang gestartet.

Dieser Befehl entfaltet seine Mächtigkeit natürlich erst bei großen Aufnahmezahlen, und ist auch auf mehrere Datenbanken anwendbar. Bei der Nutzung mehrerer Datenbanken dürfen sich die Aufnahme-Nummern (*relevé number*) nicht überschneiden (*Key field!*). Andernfalls bekommen Sie eine Fehlermeldung und müssen ggf. die Aufnahmenummern der betreffende Database mit dem Befehl *Database → renumber* ändern.

Selektierte Aufnahmen werden in dem *Selected relevés*-Fenster angezeigt und sind in der Kopfdatentabelle gelb markiert. Das Selektieren muss stets erfolgen, bevor Sie Daten von TURBOVEG nach JUICE exportieren möchten. Die selektierten Aufnahmen erscheinen anschließend im Fenster „*Selected relevés*“ und können durch Drücken des Filter-buttons in der Anzeige herausgefiltert werden. Durch Drücken der Leer-Taste können einzelne Aufnahmen wieder deselektiert bzw. noch nicht selektierte Aufnahmen händisch dazugenommen werden. Ist die Auswahl zufriedenstellend, kann man die selektierten Aufnahmen exportieren.

**Export → Standard XML-file**

Standard xml-Dateien sind das universelle Austauschformat zwischen TURBOVEG 2.0 und JUICE. Mit dem Exportieren der selektierten Aufnahmen ist die Arbeit mit TURBOVEG vorerst beendet.

## 2. Arbeiten mit JUICE

### 2.1 Einführung

JUICE ist eine von Lubomír Tichý entwickelte Windows-Anwendung zum Ordnen, Klassifizieren und Analysieren großer pflanzensoziologischer Tabellen und deren Kopfdaten. Das Programm ist für den Austausch mit TURBOVEG (HENNEKENS & SCHAMINÉE 2001) optimiert, das derzeit am weitesten verbreitete Datenbankprogramm zur Speicherung pflanzensoziologischer Daten weltweit. Es ist auch möglich, Daten aus einer Tabellenkalkulation (z. B. EXCEL) oder einer Datenbank (z. B. ACCESS) zu importieren. Kernstück von JUICE sind aber die zahlreichen Analysefunktionen und die Schnittstellen mit dem Programmpaket R.

JUICE gibt es kostenlos und ohne Registrierung zum Download auf <http://www.sci.muni.cz/botany/juice/>. Der Autor gibt keine offizielle Garantie für die Funktionsfähigkeit und keinen Support. Die Nutzung für Publikationen sollte mit Verweis auf TICHÝ (2002) immer angegeben werden. Ein ausführliches, mittlerweile dreibändiges Handbuch steht auf der Website zur Verfügung, es deckt aber immer noch nicht alle Funktionen von JUICE ab.

#### 2.1.1 Computer-Konfiguration und Einstellungen

JUICE läuft auf allen Windows-Computern, auf Mac muss man es unter einer Windows-Simulation laufen lassen. Das Programm ist Englisch im US-Format mit Dezimaltrennzeichen '.'. Beim Öffnen auf deutschsprachigen Rechnern wird das Programm automatisch versuchen, die Komma-Dezimalstelle zum Punkt zu konvertieren und greift dazu auf die Grundeinstellungen des Computers zu. Benutzer müssen deshalb über Administratorrechte verfügen, um diese Änderung zu ermöglichen. Die Auswahl wird nach der Beendigung des Programms wieder rückgängig gemacht. Für den Dauereinsatz ist es empfehlenswert, das in den Grundeinstellungen des Computers zu ändern. Der prüfende Zugriff erfolgt aber trotzdem. Ohne Administratorrechte (z. B. in Uni-Netzwerken) lässt sich JUICE deshalb nicht starten.

Anders als in Datenbankprogrammen wie TURBOVEG speichert JUICE nicht jeden Arbeitsschritt. Sie müssen Ihre Daten also regelmäßig über die Save-Funktion sichern. Insbesondere vor und nach wichtigen Arbeitsschritten sollten Sie die Tabelle stets unter neuem Namen speichern, am besten den Namen am Ende mit einer aufwärts steigenden Nummer versehen, so dass man Zwischenstände leicht wiederfindet.

### 2.2 Daten einlesen in JUICE

#### 2.2.1 Importieren von Vegetationsaufnahmen aus TURBOVEG in JUICE

Der Datentransport von der TURBOVEG-Datenbank nach JUICE erfolgt in Form von Standard xml-Dateien mittlerweile problemlos. Vorher müssen Sie die Aufnahmen, die Sie in JUICE weiterverarbeiten möchten, aus TURBOVEG selektieren (siehe oben Kapitel 1.3). Ergebnis ist die Datei tvexport.xml, welche Sie über

**File → Import → Table → From TURBOVEG XML format file**

in JUICE importieren können. Sie müssen die Datei auf der Festplatte finden und werden beim Importieren lediglich aufgefordert, jene Kopfdaten auszuwählen, welche Sie in die Tabelle übernehmen möchten. Nach Abschluss des Imports sollten Sie als erstes die Tabelle unter einem Namen abspeichern.

### 2.3 Der Grundaufbau von JUICE

#### 2.3.1 Die Benutzeroberfläche von JUICE

Die Benutzeroberfläche (siehe **Abbildung 2**) zeigt oben einen Befehlskopf (1–8), als Hauptteil eine Vegetationstabelle (9–13) und unten eine Statuszeile (14). Der Befehlskopf besteht aus einem pull-down Menü (1), welches sämtliche Befehle, teilweise in mehrfacher Verschachtelung, enthält. Wichtige Befehle sind auch als Mouse-Buttons zu finden (2), hier befindet sich mit dem Kopf-Symbol auch der Umschalt-Button zur Kopfdaten-Ansicht.

Markierungen werden in JUICE mit Farben vorgenommen, dazu stehen 8 Art-Farben (3, 4) und 8 Aufnahme-Farben (6, 7) zur Verfügung, mit denen jeweils Arten bzw. Aufnahmen markiert werden können. An den Farbleisten (4 und 7) kann man die aktuell gültige Farbe auswählen; bereits benutzte Farben werden mit einem weißen Ring bzw. schwarzen Punkt gekennzeichnet. Die aktuell gültige Farbe wird in den Feldern 3 (für Arten) und 6 (für Aufnahmen) angezeigt, viele Befehle gelten immer nur für die aktuelle Farbe. Im Feld 5 gibt es noch eine alternative Art-Farbe, die nur bei gleichzeitigem Drücken der Strg-Taste vergeben werden kann. Bestimmte Befehle, wie das Zusammenfassen zweier Arten zu einer, kann man nur mit Hilfe dieser Farbe ausführen. Außerdem kann man dieses Feld nutzen, um in einem Arbeitsgang zwei Farben (jeweils mit oder ohne Drücken der Strg-Taste) vergeben zu können, was bei großen Tabellen nützlich ist.

Ganz rechts in der Menü-Zeile wird die derzeit eingestellte Separator-Hierarchieebene angegeben (8), diese kann hier verändert oder ausgeschaltet werden.

Die Haupttabelle besteht aus 4 wesentlichen Teilen. Ganz links befindet sich das Feld mit der Artenliste (9). Artmarkierungen und das Anzeigen von Arten mittels Doppelklick müssen in diesem Feld vorgenommen werden. Mitte und rechts ist das Feld mit den Art-Abundanzen (10), in diesem Feld können keine Markierungen vorgenommen werden. Darüber befinden sich die „short header“ (11), das wichtigste Feld für alle Auswertungen. Aufnahme-Markierungen und das Anzeigen von Aufnahmen mittels Doppelklick müssen in diesem Feld (letzteres funktioniert auch bei Klicken in das Abundanzfeld) vorgenommen werden. Standardmäßig werden in den Short Header entweder die Eingangsnummer der Aufnahme oder die TURBOVEG-Aufnahmenummer (relevé number) angezeigt, über Head → Group Number kann auch die Nummer der durch zwei Separatoren eingeschlossenen Aufnahmespalten (im Folgenden Cluster genannt) angezeigt werden; mit Head → Ordinal Number kann man die Aufnahmen neu durchnummerieren. Rechts von den Artnamen befindet sich ein weiteres Feld, welches bei Auswertungen eine große Rolle spielt, die „species data“ (12). In der Regel werden hier links die Schichten der Arten angezeigt (gleiche Nummern-Codes wie in TURBOVEG, siehe Tabelle 1), die rechte Spalte kann vom Nutzer mit beliebigen Informationen zu den Arten gefüllt werden (z. B. Frequenzen, Zeigerwerte bis hin zu plant traits). Drückt man auf den Baum in der Menüleiste 2, wird die rechte Spalte ausgeblendet. Die beiden Angaben links oben (13) geben die Anzahl der Aufnahmen und Arten der Tabelle an.

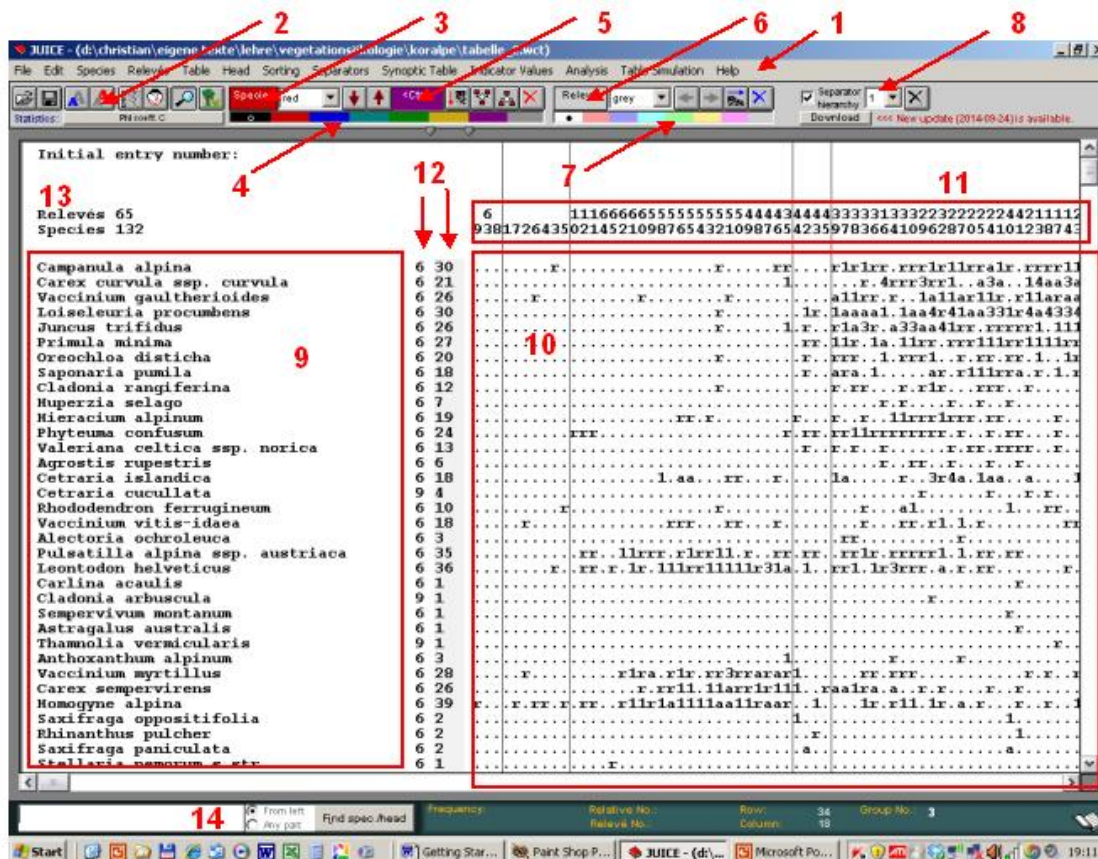


Abbildung 2: Aufbau von JUICE

Unter der Tabelle befindet sich noch eine Statuszeile, die Informationen zum Stand des Cursors anzeigt, interessanter ist aber das Suchfeld (14), welches beim Drücken des Buttons mit der Lupe in der Menüleiste erscheint. Gesucht werden kann nach Arten und nach allem, was sich in den short header befindet, z. B. Aufnahmenummern.

### 2.3.2 Die Dateistruktur von JUICE

Auch eine JUICE-Tabelle besteht aus 3 Dateien: Datei.wct (Artdaten); Datei.exp (Kopfdaten) und Datei.str (Kopfdaten-Codes und Feldname). Bei der Weitergabe müssen immer alle 3 Dateien weitergegeben werden. Beim Versenden über Internet empfiehlt es sich, die 3 Dateien in einer ZIP-Datei zu verpacken.

### 2.3.3 Grundeinstellungen des Programms

File → options

Bevor Sie mit der Tabellenarbeit beginnen, müssen einige Programmeinstellungen in Hinblick auf die Tabelle geprüft und ggf. nachjustiert werden. Sowohl TURBOVEG als auch JUICE erlauben das Arbeiten mit verschiedenen Abundanz-Schätzmethoden (cover abundance scales). Deshalb kann es erforderlich sein, diese vor dem Arbeiten mit JUICE einzustellen.



Um zu prüfen, welche Abundanzsymbole mit welchen Bedeckungsprozenten in der aktuellen Tabelle überhaupt vergeben sind, gibt es den Befehl

**Table → List all covers and their frequency in the table**

Wenn alles in Ordnung ist, brauchen Sie nichts weiter zu unternehmen, erforderliche Korrekturen können Sie dann unter **File → options → display parameters → current scale**

vornehmen. Auch die Verbindung zu externen Programmen sollte stimmen, unter **File → options → external program paths**

können Sie den Pfad zu Twinspan und R einstellen. Achtung: Derzeit korrespondieren nur die R-Versionen 2.10.1 – 2.13 fehlerfrei mit JUICE! Prüfen Sie im Verzeichnis R2.10/library, ob das package vegan installiert ist. Wenn nicht, laden Sie dieses von der entsprechenden CRAN-Seite aus dem Internet herunter oder nutzen Sie die automatisierten downloads von JUICE: Analysis → Ordinations → R-Project → DCA-Analysis → Script Update: nacheinander Associate R-Project with Juice; Installation of R libraries; Update ORDIIJUICE script anwählen.

**File → options → Synoptic tables**

Hier können Sie unter Thresholds einstellen, ab welchem Schwellenwert das Programm Fidelity und Frequency-Werte farblich hervorheben und bei Analysen von Synoptic tables berücksichtigen soll. Für Analysen ist der untere Schwellenwert ausschlaggebend. Empfehlenswert sind untere Schwellenwerte um zwischen 30 und 50 und obere Schwellenwerte zwischen 60 und 80.

JUICE greift stets auf das letzte benutzte Arbeitsverzeichnis zu und merkt sich das auch nach Beendigung des Programms, hier sind keine Einstellungen erforderlich. Letztlich können Sie noch die Anzeige der Tabelle auf Ihrem Bildschirm optimieren. Dazu bietet JUICE links oben den Verkleinerungs-Button (kleines blaues A) und den Vergrößerungs-Button (großes rotes A) an. Kleine Tabellen stellt man eher groß dar, große Tabellen eher klein. Beides können Sie jederzeit nach Bedarf verändern.

## 2.4 Erste Schritte der Tabellenarbeit

### 2.4.1 Cursorbewegungen und Markierungen

Durch klicken in die Farbleiste können Sie die aktuelle Farbe einstellen, die ja Auswirkungen auf jene Befehle hat, die nur für eine Farbe (eben die aktuelle) gelten. In der Tabelle 3 sind die wichtigsten Cursor-Befehle zusammengefasst:

**Tabelle 3:** Befehle für Markierungen, Separator setzen und Verschieben von Einzelaufnahmen

Aufgabe	Aktion
Anzeigen einer Aufnahme	Doppelklick in den Short header oder im Abundanzfeld
Ändern von Artdaten	Doppelklick auf den Artnamen
Markieren einer Art mit der aktuellen Farbe	Rechtsklick auf den Artnamen
Markieren einer Art mit der aktuellen Farbe des Ctrl-Feldes	Strg-Rechtsklick auf den Artnamen
Markieren einer Aufnahme mit der aktuellen Farbe	Rechtsklick auf den short header
Markieren eines Blocks von Aufnahmen	Shift-click auf den short header oder den Artnamen (der Block geht dann hoch bzw. links bis zur nächsten Aufnahme dieser Farbe)
„Aktivieren“ einer Art	Click auf den Artnamen (wird benötigt für solche Befehle, die Arten betreffen)
Setzen eines Separators	Shift-click in die short headers oder in die Artnamen
Löschen eines Separators	Shift-click auf den Separator
Bewegen einer Art	Click in den Artnamen und Ziehen an die neue Stelle
Bewegen einer Aufnahme	Click in die short header und Ziehen an die neue Stelle

### 2.4.2 Vorbereitung der Tabelle

Frisch aus TURBOVEG importierte Daten sind Rohdaten. Bevor Sie damit arbeiten können, müssen Sie die Tabelle auf diverse Ungereimtheiten kontrollieren. Als erste Schritte empfehlen sich `sort species alphabetically` →

`all` und `Species` → `Mark layers`. Bei letzterem Befehl sollte man Schichten, die man nicht haben will (z. B. 0) rot einstellen. Nach diesen beiden Schritten sehen wir, ob manche Arten in ein und derselben Schicht unterschiedliche Namen haben (z. B. *Achillea millefolium*, *Achillea millefolium* ssp. *millefolium* und *Achillea millefolium* agg.) und ob manche Arten keine oder eine falsche Schichteingabe (einer der häufigsten Eingabefehler, z. B. Moose in der Krautschicht) haben. Mit Doppelklick auf den Namen öffnet sich ein Fenster, in dem Sie die Schichtangaben (auch den Namen) bearbeiten können. Verschiedene Namensversionen ein und derselben Art sollte man allerdings nicht ändern, sondern lieber zusammenfassen. Dies geht über den Befehl

**`Species` → `merge (violett) species`**

wobei (violett) hier für die im Ctrl-Fach eingestellte Farbe steht. Rechts neben diesem Fach gibt es auch einen Button für den `merge`-Befehl. Standardmäßig bietet das Programm hier den oben stehenden Namen als den neuen gemeinsamen Namen an (z.B. *Achillea millefolium* agg.), so dass es sich ggf. auszahlt, den gewünschten Namen nach oben zu ordnen. Die Abundanzwerte beider Namen werden alle in der neue Artzeile zusammengefasst. Mit

**`Species` → `Undelete species`**

haben Sie jederzeit einen Überblick, welche Namen aus der Tabelle entfernt wurden. Hier können Sie den Schritt gegebenenfalls auch später noch rückgängig machen (siehe 3.1.5).

### 2.4.3 Editieren von Kopfdaten

Dies ist ein Kapitel von grundsätzlicher Bedeutung, denn die Kopfdaten spielen bei den Analysefunktionen von JUICE eine zentrale Rolle. Hier kann man alle Berechnungen, Sortierungen und Werte, welche die Einzelaufnahmen betreffen, laufend abspeichern und für weitere Analysen, z. B. in Statistikprogrammen, exportieren. Als Austauschmedium hat sich EXCEL bewährt. EXCEL sollte man auch benutzen, wenn man Kopfdaten schnell ändern, löschen oder bloß umbenennen will, denn die hierfür vorhandenen JUICE-Funktionen sind erheblich aufwändiger. Auch zum Auslagern von Kopfdaten eignet sich eine EXCEL-Tabelle, da JUICE ab einer bestimmten Menge an Kopfdaten nicht mehr reibungslos arbeitet. Und letztlich ist das hier beschriebene wichtig für die Zusammenlegung zweier Tabellen, wie im Kapitel 3.1.3 beschrieben.

Zum Exportieren Ihrer Kopfdaten müssen alle Aufnahmen die gleiche Farbe (z. B. weiß) haben. Wählen Sie den Befehl **`Datei` → `Export` → `Export Headers from white releves`**. Markieren Sie alle Header-Felder im folgenden Fenster und drücken Sie **`Continue`**. Sie erhalten eine txt-Datei, für die JUICE den Namen `hlava.txt` vorschlägt. Öffnen Sie EXCEL und öffnen Sie die `hlava.txt`-Datei (semicolon delimited text file). Wie Sie sehen, sind die Kopfdaten transponiert, die Aufnahmenummern verlaufen in der ersten Spalte abwärts. **Achtung:** Überprüfen Sie als erstes die letzte Spalte der Kopf-Tabelle: Wenn nämlich einige Datenfelder ein Semikolon enthalten (oft z. B. in den Remarks), wird die Datei nicht korrekt eingelesen und muss leider mühselig von Hand korrigiert werden! Wenn die Datei ok ist, können Sie jetzt Änderungen, Umstellungen oder Löschungen ganzer Kopfdatenfelder vornehmen. Niemals löschen dürfen Sie das Feld `Relevé number`, denn das ist das Schlüsselfeld für die Rückführung des Datensatzes in die JUICE-Tabelle.

Das wollen wir jetzt also tun: Markieren Sie die Felder der Kopfdaten-Tabelle, die Sie zurückführen möchten und kopieren diese mit `strg-c` in die Zwischenablage. Nun können Sie in JUICE mit dem Befehl **`File` → `Import` → `header data` → `from clipboard`**, die korrigierten Kopfdaten wieder einlesen (Datei vorher speichern und anschließend unter neuem Namen speichern).

## 2.5 Ordnen der Tabelle

Unsere Tabelle hat bisher nur eine Ordnung: Die alphabetische der Arten. Dies hat natürlich keinerlei ökologische Relevanz. Die nächste Aufgabe soll deshalb sein, händisch einige mehr ökologische Ordnungen in die Tabelle zu bringen. Hier kommen erstmalig die `Species data` und die `short header` zum Einsatz.

### 2.5.1 Sortieren nach Schichten und Frequenzen

**`sorting` → `sort species alphabetically` → `within layers`**

Dieser Befehl sortiert die Tabelle nach den Schichten, immerhin schon eine erste ökologische Information. Die Schichten sollten nach dem Befehl `Mark layers` unterschiedliche Farben haben. Als nächstes wollen wir die Arten innerhalb der Schichten nach ihrer Häufigkeit, (Stetigkeit, frequency) sortieren. Dazu müssen wir diese Information in die `Species data` laden:

**`Species` → `Species data` → `Frequency` → `all table`**

Damit steht die absolute Frequenz (n Präsenzen der Art innerhalb der Tabelle) in der Spalte neben der Schichtangabe. Nun stellen wir die Farbe der Baumschicht (z. B. grün) als aktuelle Farbe ein und sortieren die Arten innerhalb dieser Schicht absteigend nach der Frequenz:

**Sorting → Sort species by species data → Green species + numerical + in descending order**

Dieser Befehl sortiert die Baumschicht nach fallender Frequenz, die häufigsten Arten stehen jetzt ganz oben, seltene unten. Wiederholen Sie dies für die Strauch-, Kraut- und Mooschicht. Immer wieder Speichern!

Als nächstes wollen wir Arten mit geringer Frequenz an das Ende der Tabelle verschieben, da diese bei der Grob-Gliederung erstmal keine Rolle spielen. Dazu stellen wir beispielsweise die graue Artfarbe (sofern sie noch nicht benutzt wurde) ein und markieren mit der Block-Funktion alle Arten mit einer Frequenz < 10 (beispielsweise). Da die Block-Funktion nur von oben nach unten anwendbar ist, beginnen wir oben mit der Baumschicht.

Jetzt nutzen wir zum ersten Mal die automatischen Verschiebungsmöglichkeiten von JUICE. Über der Artfarben-Leiste befinden sich zwei Button mit je einem grauen Pfeil (grau sind sie, weil das die aktuelle Farbe ist). Diese beiden Buttons verschieben alle Arten der entsprechenden Farbe bis zur obersten bzw. untersten Art dieser Farbe. Wir klicken auf den Pfeil nach unten und versammeln damit alle grau markierten Arten geringer Stetigkeit am Ende der Tabelle.

Der obere Teil der Tabelle enthält jetzt komprimiert die häufigen Arten, sowie jene Arten mittlerer Stetigkeit, die für eine erste Gliederung der Aufnahmen besonders interessant sind. Speichern nicht vergessen!

## 2.5.2 Ordnen der Aufnahmen nach Kopfdaten-Informationen

Wenden wir uns jetzt der Ordnung der Aufnahmen zu. Ein nächster Schritt könnte sein, die Aufnahmen nach Informationen aus den Kopfdaten vorzusortieren. Dazu kommen die `short headers` erstmalig zum Einsatz. Gut würden sich hier ökologische Zeigerwerte eignen, die wir aber zu diesem Zeitpunkt noch nicht haben. Eine andere Möglichkeit wäre, die Aufnahmen nach Strukturmerkmalen wie Schichtenaufbau oder Vegetationsbedeckung, oder nach Habitaten zu gliedern, eine dritte Möglichkeit, sie geologisch oder geographisch zu stratifizieren, also nach geologischem Untergrund, Höhenstufen, Regionen usw. zu ordnen, sofern solche Informationen in den Kopfdaten vorliegen (mehr dazu in Kap. 3.1.6). Die Werte (nehmen wir mal zu Demonstrationszwecken `altitude` oder `cover`) müssen wir aus den Kopfdaten in die `short header` überführen:

**Head → Store Values to Short header → Header Data**

Hier suchen wir das entsprechende Kopfdatenfeld und laden die Daten mit **Continue** in die `short headers`. Mit **Sorting → Sort Short Headers → All relevés + numerically + ascending**

werden die Aufnahmen sortiert, tiefe Lagen links, Hochlagen rechts, oder lückige Vegetation links, dichte rechts. Jetzt könnte man eine grobe Teilung (z. B. `collin`, `montan`, `alpin`, oder Bedeckung < 30%, Bedeckung 30–70 %, > 70 %) vornehmen, wenn die Unterschiede im Datensatz groß genug sind. JUICE erkennt solche Gruppen aber nicht an Farben, sondern nur an senkrechten Linien, den sog. Separatoren. Deshalb setzten wir mit `shift-click` zwei Separatoren zwischen die sortierten Gruppen. Anschließend geben wir den Aufnahmegruppen noch verschiedene Farben.

## 2.5.3 Erstes Arbeiten mit Stetigkeitstabellen

Sobald man irgendwelche Aufnahmegruppen erzeugt und mit Separatoren abgetrennt hat, interessieren uns die Frequenzen der Arten innerhalb der Gruppen und die qualitativen Unterschiede zwischen den Gruppen. Dies kann man am besten mit Hilfe von Stetigkeitstabellen (`synoptic table`) beurteilen. Diese erzeugt JUICE mit dem Befehl:

**Synoptic table → Percentage frequency**

Die Ansicht wechselt: In der Kopfzeile werden jetzt die Anzahl der Aufnahmen pro Spalte angezeigt und darunter die Nummer der Stetigkeitsspalte 1–3, die hier im Unterschied zu den Spalten der `full table` als Cluster (eine Gruppe von Spalten) bezeichnet werden. Im Gegensatz zu den `Cover values` und den absoluten Frequenz-Daten in den `species data` zeigen die Werte jetzt die relativen (prozentualen) Frequenzen der Arten innerhalb der nun sichtbaren 3 Cluster.

Wir interessieren uns nun dafür, ob es Arten gibt, die dieser Einteilung folgen und für eine der 3 Cluster einen deutlichen Schwerpunkt aufweisen (Differentialarten).

## 2.5.4 Speichern und Laden von Art- und Aufnahme-Farben

Bevor es weitergeht, wollen wir unsere bisher vergebenen Farben speichern. Wir gehen wieder in die `full table` Ansicht durch erneutes Ausführen des Befehls **Synoptic table → Percentage frequency** und speichern mit dem Befehl **Species → Species Colours → Save** die Art-Farben, und mit dem Befehl **Relevés → Relevé Colours → Save** die Aufnahme-Farben. Mit dem Befehl **load** können wir diese Farbe später wieder über die Tabelle legen, auch wenn wir sie verändert haben. Dies ist nützlich, um eine vorgenommene Ordnung (z. B. nach Artengruppen) mit einer älteren (z. B. nach `altitude`, `cover`) zu vergleichen.

## 2.5.5 Suche nach Differentialarten

Danach können wir alle Art-Farben außer die graue durch mehrmaliges Klicken auf das **x** in der Farbleiste auf schwarz setzen, und markieren nun für jede Spalt jene Arten, die in einem Cluster mehr als doppelt so häufig sind wie in den anderen beiden (`constancy ratio`) mit je einer Farbe. Wir haben jetzt den Zusammenhang zwischen unserer anfänglichen

Stratifikation (also Höhenstufen, Bedeckung) und den Arten hergestellt, Arten die in einem Cluster doppelt so häufig sind als in einem oder mehreren Vergleichsclustern, werden als Differentialarten bezeichnet.

Eigentlich wollen wir die Tabelle aber nicht nur ordnen, sondern klassifizieren, d. h. wir wollen nicht unbedingt von den Kopfdaten auf die Artenzusammensetzung schließen, sondern meistens von der Artzusammensetzung auf ökologische Parameter, die wir nicht sofort im Gelände erkennen konnten. Dazu müssen wir Aufnahmen finden, die viele gemeinsame Arten haben, sich also ähnlich sind, und diese dann in eigene Cluster gruppieren. Diese Cluster können dann syntaxonomischen Einheiten (Syntaxa) entsprechen. Der Weg dorthin führt über die Suche nach Arten, die sich im Gelände ökologisch ähnlich verhalten und damit auch in der Tabelle oft zusammen vorkommen (→ Kap. 2.6).

### 2.5.6 Separator-Hierarchie

Vorher noch ein weiterer Einschub: Will man in einer Tabelle kennzeichnen, dass zwei Cluster möglicherweise enger miteinander verwandt sind als mit anderen, kann man mit der **Separator hierarchy** arbeiten. Wir stellen die Hierarchiestufe 2 ein und klicken zweimal zusammen mit der Shift-Taste auf einen Separator, der beim ersten Klick gelöscht wird und beim zweiten als einfach gestrichelte Linie, und damit als Separator der 2. Hierarchieebene, erscheint. Dies hat mehrere Vorteile. Erstens können wir in der Synoptic table-Ansicht die Hierarchieebene frei wählen ohne die Synoptic table-Ansicht zu verlassen (und so in der 1. Hierarchieebene z. B. zwei Cluster, in der 2. Hierarchieebene vier Cluster miteinander vergleichen). Zweitens kann man auf diese Art und Weise (es gibt 6 Hierarchieebenen) Subassoziationen, Assoziationen, Verbände, Ordnungen und Klassen, kurz die syntaxonomischen Ebenen (= Syntaxa), in einer Tabelle bearbeiten und beispielsweise das Differenzialartenkriterium (constancy ratio) in verschiedenen Hierarchieebenen prüfen.

## 2.6 Klassifikation (pflanzensoziologische Tabellenarbeit)

Vegetationsklassifikation arbeitet nie für sich allein, sondern steht in einer langen Tradition von klassifikatorischer Arbeit. Früher, als man noch Neuland beschritt, ergab sich die Klassifikation rein aus dem Verhalten der Arten im Gelände (bzw. in der Tabelle), auch auf die Gefahr hin, dass die gefundene Klassifikation nur eine lokale Gültigkeit hatte. Heute sollte man immer auch Vorinformationen aus der Literatur bei der Klassifikation berücksichtigen (supervised classification), insbesondere in Gebieten, wo es vegetationskundliche Monographien wie beispielsweise die „Pflanzengesellschaften von Österreich (GRABHERR & MUCINA 1993, MUCINA, GRABHERR & ELLMAUER 1993, MUCINA, GRABHERR & WALLNÖFER 1993)“ oder „Die Wälder und Gebüsche Österreichs (WILLNER & GRABHERR 2007a & b)“ gibt. Im Folgenden wollen wir uns kurz mit einigen Möglichkeiten der supervised classification als auch der „unsupervised“ classification beschäftigen. Die Tabellenfarben stellen wir auf Schwarz-Weiß und löschen alle Separatoren. Zu Beginn noch etwas Theorie.

### 2.6.1 Das Konzept der Fidelity („Treue“)

Die „Fidelity“ nimmt in JUICE einen ziemlich zentralen Platz ein. Die Macher von JUICE waren maßgeblich an der Weiterentwicklung dieser Idee beteiligt und haben sie tief in dem Programm verankert. Ausgehend von „Charakterarten“ und der von BRAUN-BLANQUET als Erfahrungswert eingeführten „Gesellschaftstreue“, die angibt, wie eng eine Art empirisch an eine bestimmte Pflanzengesellschaft gebunden ist, haben zuerst BRUELHEIDE (1997, 2000), BRUELHEIDE & CHYTRÝ (2000) und später CHYTRÝ et al. (2002) diese Idee auf objektivere (statistisch-mathematische) Beine gestellt. Treuemaße prüfen die Bindung einer Art an ein bestimmtes Cluster, indem das Vorkommen der Art in diesem Cluster mit deren Vorkommen in der übrigen Tabelle (als Präsenz-Absenz oder auch gewichtet nach Abundanzen) verglichen wird. Damit sind sie also theoretisch geeignet, die Differential- und Charakterarten von Gesellschaften herauszubekommen. Arten mit hohen Fidelity-Werten werden in JUICE deshalb als „diagnostic species“ bezeichnet. Die folgende Zusammenstellung zeigt die Komponenten, die Formeln und die in JUICE verwendete Terminologie im Zusammenhang von Frequency (relative Frequenz, Stetigkeit) und Fidelity („Treue“, hier am Beispiel des  $\Phi$ -Koeffizienten).

#### Ermitteln der *Frequency* (Stetigkeit) einer Art p:

n = Anzahl Aufnahmen im Zielcluster

$$F = \frac{n_p \cdot 100}{n}$$

$n_p$  = absolute Frequenz der Art p im Zielcluster

→ Arten mit hoher Stetigkeit = **Constant species**

→ Arten mit gleich hoher Stetigkeit = **Co-occurring species**

#### Ermitteln der *Fidelity* einer Art p (Beisp. $\Phi$ Koeffizient):

n = Anzahl Aufnahmen im Zielcluster

$$\Phi = \frac{u_{hyp}}{\sqrt{N-1}} = \frac{N \cdot n_p - n \cdot N_p}{\sqrt{n \cdot N_p \cdot (N-n) \cdot (N-N_p)}}$$

$n_p$  = absolute Frequenz der Art p im Zielcluster

N = Anzahl Aufnahmen in der Tabelle

$N_p$  = absolute Frequenz der Art p in der Tabelle

→ Arten mit hoher Fidelity = **Diagnostic species**

→ Arten mit gleich hoher Fidelity = **Interspecific associations**

Statistisch gesehen handelt es sich bei der Fidelity-Berechnung um eine Verteilungsprüfung, für die in der allgemeinen Statistik der  $\chi^2$ -Test und der G-Test sowie der „exakte Test“ nach Fisher geläufig sind. Standardmäßig eingestelltes Prüfmaß ist der  $\Phi$ -(phi)-Koeffizient (CHYTRÝ et al. 2002), es gibt aber noch etliche andere. Ein Nachteil der Fidelity-Werte, dass nämlich ihre Berechnung wie so oft von der Anzahl der Aufnahmen in den einzelnen Clustern der Tabelle abhängt, kann in JUICE abgemildert werden, indem man unter Options → Fidelity measures → Standardization of Relevé Group Size wählt. Damit werden die Werte bei unterschiedlichen n-Zahlen auf ein Einheitsmaß interpoliert. Fishers Exact Test kann man in 3 verschiedenen Konfidenz-Leveln (p-value) im Hintergrund mitlaufen lassen, dann zeigt die Synoptic table nur Fidelity-Werte an, die diesen Test bestehen. Der Nachteil aller Fidelity-Werte ist, dass sich die Berechnung natürlich nur auf die aktuelle Tabelle bezieht und sich sofort ändert, wenn man die Cluster-Einteilung der Tabelle in irgendeiner Weise verändert. Selbst Veränderungen in anderen Clustern beeinflussen die Fidelity-Werte des gerade untersuchten Clusters, da ja die „Treue“ stets über die gesamte Tabelle ermittelt wird. Der Fidelity-Wert ist also stark von der Ordnung der Tabelle abhängig. Auch seine Herleitung aus der Wahrscheinlichkeitsrechnung erschwert das Verständnis und seine Interpretation. Die Bezeichnung „diagnostic species“ klingt verführerisch und verleitet schnell zu Überinterpretationen. Wir empfehlen deshalb auch immer in Synoptic tables die gute alte Stetigkeit (Percentage Frequency) zu verwenden und für Präsentationen von Daten zu verwenden. Stetigkeitswerte sind deutlich robuster und als einfache Prozentwerte für Anfänger leichter zu verstehen, und erfordern auf der anderen Seite mehr „pädagogisch wertvolle“ Auseinandersetzung mit der Tabelle.

Alle Fidelity-Werte entfalten eigentlich erst bei großen Tabellen und über weite geographische Räume ihr „diagnostisches“ Können. Das Fidelity-Konzept hat in der vegetationskundlichen Literatur aber durchaus einen Siegeszug angetreten, und wer sich näher mit Vegetationskunde beschäftigen möchte, kommt daran kaum vorbei. Zum Selbststudium kann man zuerst das JUICE-Handbuch empfehlen und anhand unserer (oder einer simulierten) Tabelle ein bisschen herumprobieren, zum weiteren Verständnis ist das Studium von CHYTRÝ et al. (2002), sowie TICHÝ & CHYTRÝ (2006), WILLNER et al. (2009), LÖTTER et al. (2013) und PEINADO et al. (2013), nebst einiger Erfahrung unerlässlich.

Allerdings sind einige gerade auch für Anfänger äußerst nützliche JUICE-Funktionen an Fidelity-Berechnungen geknüpft. Insbesondere bei der Klassifikation nutzt JUICE im Hintergrund die Fidelity. Einstellen in den Options sollte man den  $\Phi$ -Koeffizient, der als positiver Prozentwert in der Tabelle dargestellt wird. Fangen wir also gleich mal damit an.

## 2.6.2 Klassifikation an Hand von „Leitarten“: Interspecific associations

Vorinformationen kommen nicht nur aus der Literatur, sondern auch bei der eigenen Geländearbeit. Man merkt, dass man in bestimmten ökologischen Situationen immer wieder die gleiche Art oder die gleichen Arten antrifft, oder dass eine Art dominant ist. Eine zu stellende Frage wäre nun: Gibt es um diese auffälligen Arten herum weitere Arten, die häufig mit diesen zusammen vorkommen und deshalb sich in der Tabelle ähnlich „verhalten“? Statt mühselig danach zu suchen, können wir das mit Hilfe des Fidelity-Wertes berechnen lassen.

### Analysis → Interspecific associations

Dieser Befehl zeigt an, in wie weit ( $\Phi$ -Koeffizient in %) sich das Vorkommen einer Art in der Tabelle mit anderen Arten deckt. Wir suchen uns eine Art, die nach Vorinformationen (Geländebeobachtung, Literatur) besonders charakteristisch, interessant, dominant usw. ist oder sein soll. Der Einfachheit halber sortieren wir sie an die erste Tabellenposition und geben ihr eine eigene Farbe. Um sie zu aktivieren, müssen wir den Artnamen einfach mit der Maus anklicken. Wir wählen ALL bei der Frage nach den zu untersuchenden Aufnahmen, denn wir wollen die Fidelity über die gesamte Tabelle berechnen lassen. Es erscheint ein Fenster, in dem nun die mit dieser Art assoziierten Arten nach fallender Fidelity aufgelistet sind. Zusätzlich gibt das Programm in der rechten Spalte noch das Verhältnis von absoluter Frequenz der jeweiligen Art zur gemeinsamen Frequenz mit der markierten Art an. Dies kann man zusätzlich nutzen, um „gute“ Arten auszuwählen, denn seltene Arten haben schnell mal eine hohe Fidelity. Mit etwas Erfahrung sieht man, dass nicht alle Arten mit hohen Werten wirklich ökologisch zu unserer Zielart passen. Solche Arten lassen wir weg, die anderen markieren wir mit `ctrl`-linke Maustaste. Über **Mark selected species in the table** können Sie Ihre Auswahl in der Tabelle farblich markieren (vorher die Farbe einstellen). Interessant ist auch das untere Fenster, das negative Fidelity-Werte anzeigt für Arten, die unsere Zielart „meiden“. Nach Schließen des Fensters sortieren Sie die Arten jetzt nach oben und unsere erste „ökologisch-soziologische Artengruppe“ ist vorerst fertig.

## 2.6.3 Arbeiten mit ökologisch-soziologischen Artengruppen: COCKTAIL

Eine so gefundene Artengruppe mit hohen gemeinsamen Fidelity-Werten oder eine beliebige andere Gruppe von Arten (z. B. aus der Literatur) kann nun nicht nur zur Ordnung der Tabelle genutzt werden, sondern auch weiter optimiert und die dazugehörigen Aufnahmen gesucht werden. Dazu können wir die COCKTAIL-Funktionen nutzen. COCKTAIL (BRUELHEIDE 1996, 2001) ist eine Art Programmiersprache zur Vegetations-Klassifikation, doch soweit kommen wir nicht (einige Möglichkeiten dieses gigantischen Tools sind im Kap. 3.3.2 kurz beleuchtet), wir benutzen hier nur einige sehr basale Funktionen.

### Analysis → COCKTAIL groups

Ausgangspunkt ist also eine farblich markierte Gruppe von Arten. In dem COCKTAIL-Fenster geben Sie als erstes in dem Species-Feld die Farbe der Gruppe ein und schreiben mit **Add species** die Artengruppe in das **Species list**

in **sel. groups**-Feld. Auch in dem Feld **Select relevés as** sollte praktischerweise die gleiche Farbe eingestellt werden. Da ja in der Natur solche Artengruppen selten vollständig vorhanden sind, können wir mit dem Button **Min. No. Sp./Rel.** eine optimale Anzahl von Arten der Gruppe ermittelt, die mindestens in einer Aufnahme vorkommen soll. Wenn wir uns erinnern, dass unser Ausgangspunkt eine einzige Art war, findet das Programm also auf diese Weise auch passende Aufnahmen, in denen diese Ausgangsart gar nicht drin ist. Mit **Select relevés as** werden alle Aufnahmen in der Tabelle markiert, die diese Mindestzahl einhalten. Mit **Manual setting** können auch andere Mindestzahlen ausprobiert werden, denen Sie ggf. eine neue Relevé-Farbe geben. Doch haben wir wirklich schon alle Arten gefunden, die zu unserer Gruppe passen? Mit **Fidelity calculation** können Sie weitere Arten anzeigen, die hohe Fidelity-Werte nicht mehr nur zu unserer Ausgangsart, sondern zur gesamten Artengruppe aufweisen. Sie können daraus weitere ökologisch sinnvolle Arten wählen und in der Tabelle einfärben. Auf diese Weise kann die Artengruppe optimiert werden.

Zurück in der Full table, sortieren wir die eingefärbten Aufnahmen und Arten zusammen, setzen einen Separator und können das so ermittelte Cluster nun auf seine Übereinstimmung mit der Literatur oder seine sonstige Spezifität innerhalb der Tabelle prüfen. Ist das Resultat zufriedenstellend, stellen Sie für den Rest der Aufnahmen weitere Artengruppen zusammen.

## 2.6.4 Händisches Sortieren nach Artengruppen

Auch ohne COCKTAIL können Sie die Tabelle natürlich nach Arten und Aufnahmen sortieren. Nachdem wir eine bestimmte Artengruppe in der Tabelle farblich z. B. gelb markiert haben, wollen wir jetzt noch die Aufnahmen mit diesen Arten zusammenstellen. Der Befehl lautet:

**sorting → Sort relevés → all + by yellow species**

Die Aufnahmen werden jetzt so sortiert, dass erst Aufnahmen mit allen Arten, dann mit einigen Arten und zu Schluss Aufnahmen mit jeweils nur einer Art, fallend sortiert nach Artmächtigkeits-Summe, zusammengestellt werden. Jetzt müssen wir einen sinnvollen Separator setzen und schauen dabei, ob mit sinkender Artmächtigkeit unserer Arten irgendwann andere Arten einsetzen. Den neu entstandenen Aufnahmeblock färben wir nun auch gelb und stellen anschließend wieder weiß als aktuelle Aufnahme-Farbe ein.

Nun lautet die Frage, ob unsere Artengruppe ökologisch „etwas Wert“ ist. Das wäre der Fall, wenn auch andere Arten ein ähnliches Muster zeigen würden, während wieder andere genau diese Gruppe meiden. Das kann man wieder in der **Synoptic Table**-Ansicht → **Percentage Frequency** am besten prüfen. Wir gehen die Tabelle abwärts und markieren alle Arten, die mehr als doppelt so häufig in unserer ersten Spalte als in der zweiten sind, ebenfalls gelb. Mit der gelben Pfeiltaste holen wir die Arten nach oben unter unsere Ausgangsarten. Damit haben wir die erste Spalte erst einmal grob gekennzeichnet.

So geht es weiter, im Folgenden sortieren wir aber nur noch die weißen Aufnahmen, unsere erste gelbe Gruppe wird davon dann nicht berührt. Das Zusammensuchen von Arten, die in einem Cluster besonders häufig sind, kann man auch von JUICE machen lassen, indem man die Arten eines Clusters nach fallender Stetigkeit ordnet. Wir bleiben in der Synoptic table-Ansicht und wählen

**Synoptic table → Sort Species in Synoptic Table**

Da wir nur noch die Arten sortieren wollen, die wir noch nicht markiert und nach oben sortiert haben, müssen wir im Auswahlfenster **Sort species → black, Sort by → Relative frequency** und **Sort → Single column → 2** (falls wir das 2. Cluster sortieren möchten) einstellen. Mit **Sort** starten wir den Befehl. Er sortiert die Arten mit der höchsten prozentualen Stetigkeit nach oben und wir können leicht prüfen, welche Arten dieses Cluster gegenüber den anderen beiden differenziert. Arten farblich markieren, zusammensortieren und mit den verbliebenen weitermachen.

## 2.6.5 Händisches Justieren der Grenze zwischen zwei Artengruppen

Die Vegetationsklassifizierung sucht zuerst nach dem „Typ“ eines Syntaxons, also nach einer besonders typischen weil häufig wiederkehrenden Artenkombination für eine Assoziation. Die zweite, ebenso grundlegende, aber viel schwierigere Frage ist, wann die Artenkombination so anders ist, dass der konkrete Pflanzenbestand nicht mehr dazu gehört? Oder: Wo ist die Grenze zwischen zwei Einheiten zu ziehen? COCKTAIL hat uns vorhin gefragt, wie viele Arten unserer Artengruppe mindestens in einer Aufnahme sein müssen, auch das ist die Frage nach der Abgrenzung. In der Tabelle ist das die Frage: Wo setze ich einen Separator? Jede Einheit hat ihre Differenzialartengruppen, doch ist die Trennung, wie bei allem ökologischen Geschehen, selten scharf: Es gibt Aufnahmen, die Vertreter aus beiden Differenzialartengruppen enthalten. Eine Möglichkeit, das zu optimieren, wäre, die Anzahl der Differenzialarten miteinander zu vergleichen. Beide sind ja eingefärbt, die eine z. B. gelb, die andere violett. Ist gelb die aktuelle Farbe, bietet JUICE in der full table-Ansicht den Befehl:

**Head → Store Values to Short header → Number of yellow species**

Jetzt sehen wir, wie viele unserer Differentialarten in jeder Aufnahme sind. Wir können diesen Wert jetzt sortieren:

**Sorting → Sort Short Heades → All ... Numerically ... In ascending order ... Within Groups**

Dieses **Within Groups** ist hier besonders wichtig, sonst ist unsere Klassifikation dahin. Jetzt stehen die Aufnahmen mit vielen Differenzialarten ganz links, mit wenigen ganz rechts in jedem Cluster. Wiederholen wir den Schritt jetzt mit der violetten Differenzialartengruppe und wählen **In descending order**, so haben wir spiegelbildlich dazu die andere Differenzialartengruppe. Auf diese Weise können wir also abwägen, welche Differenzialartengruppe in einer Aufnahme überwiegt und diese dann ggf. noch in das andere Cluster schieben. Zurück zur Synoptic table-Ansicht können wir feststellen, ob unsere Nachjustierung die Trennung (also die constancy ratio der Differentialarten) verbessert hat.

Diese Methode können Sie auch nach allen automatischen Klassifikationen anwenden. TWINSPAN teilt beispielsweise die Cluster gern in der Mitte, wenn die Unterschiede nur schwach sind. Das muss nicht unbedingt die syntaxonomisch und ökologisch beste Teilung sein.

## **2.6.6 Klassifizieren mit Referenzaufnahmen: Supervized and Semi-Supervised classification**

Eben haben wir von besonders typischen Aufnahmen gesprochen, Aufnahmen, die also von der Artenzusammensetzung und den Standortbedingungen etwa die ökologische Mitte oder eine häufige Artenkombination einer Gesellschaft repräsentieren (nicht zu verwechseln mit dem nomenklatorischen Typus, der einfach nur eine gewählte Aufnahme für den Gesellschaftsnamen ist und mit dem keinerlei ökologische Aussage verknüpft ist). Solche Aufnahmen (z. B. aus der Literatur) können wir nutzen, um die diesen Aufnahmen ähnlichsten Aufnahmen unserer Tabelle zu finden. Dafür bieten JUICE u. a. das K-mean-clustering-Verfahren an. Es ist ein nicht-hierarchisches, agglomeratives Clusteranalyse-Verfahren auf der Basis der Euklidischen Distanz zwischen den Aufnahmen. Die Aufnahmen werden nach Ähnlichkeit gruppiert, die Anzahl der Cluster (K) kann vom Nutzer festgelegt werden. Diese haben keine Cluster-Hierarchie, sind also komplett unabhängig voneinander.

Ausgangspunkt ist also die Auswahl besonders typischer Muster-Aufnahmen. Das Programm sucht uns dann jene Aufnahmen zusammen, die diesen Musteraufnahmen am ähnlichsten sind in Hinblick auf die Euklidische Distanz. Zuerst müssen Sie also, z. B. durch Vergleich mit Literaturquellen Musteraufnahmen ermitteln (oder aus externen Quellen dazuladen). Alle Aufnahmen müssen weiß sein, die Musteraufnahmen (bis 7 Musteraufnahmen sind möglich) markieren Sie mit verschiedenen Farben. Für jedes zu suchende Syntaxon sollte nur eine Aufnahme markiert werden.

**Analysis → Supervized and Semi-Supervised classification → Semi supervised K-Means**

Supervised und Semi-Supervised classification unterscheiden sich lediglich in der Anzahl der zu findenden Cluster, das Verfahren ist identisch. Supervized classification erstellt genau soviel Cluster, wie Sie Musteraufnahmen markiert haben, ordnet also alle Aufnahmen irgendeiner Musteraufnahme zu, auch wenn die Distanz sehr groß ist. Bei der Semi-Supervised classification können Sie eine Cluster-Zahl höher als die Anzahl der Musteraufnahmen eingeben, dann erstellt das Programm ihnen zusätzlich noch diese Anzahl Cluster ohne Musteraufnahmen (der Centroid, also die mittelste Aufnahme, ergibt sich dann aus der Berechnung). Das ist oft besser: Erstens wird die Zuordnung zu den Musteraufnahmen enger, und zweitens erfahren Sie, ob sich in Ihrem Datensatz noch weitere (möglicherweise interessante) Einheiten verstecken, die Ihnen bisher nicht aufgefallen sind.

In dem sich öffnenden Fenster können Sie allerlei Einstellungen vornehmen, z. B. können Abundanzen von Präsenz-Absenz bis hin zu Pseudospecies (siehe Kap. 2.6.7.1) Berücksichtigung finden. Die wichtigste Vorbereitung ist aber der Button **Change Short Header Codes**. Damit weisen Sie ihren Musteraufnahmen jeweils einen Zahlencode > 0, und allen zuzuordnenden (weißen) Aufnahmen eine 0 als Short Header Wert zu. In diesem Fenster können Sie auch zwischen Supervized und Semi-Supervised classification umschalten. Geben Sie unter **No. of clusters** ruhig ein paar mehr ein, als Musteraufnahmen vorhanden sind. Sie können am Ende die neuen Cluster ja dann auf ihren „Wert“ prüfen und ggf. wieder mit einem der anderen Einheiten vereinen. Mit **Run** starten Sie den Rechenvorgang, der auf eine funktionierende R-Verknüpfung mit JUICE angewiesen ist.

Im Ergebnis erhalten Sie eine Tabelle mit der von Ihnen angegebenen Cluster-Anzahl. Das Ergebnis unterscheidet sich von einer TWINSPAN-Analyse u. a. dadurch, dass Sie mit Ihren Musteraufnahmen die Zentren der Cluster selbst festgelegt haben, und dass die Cluster nicht hierarchisch zueinander stehen. Sie sollten jetzt die Arten entsprechend sortieren und in der Synoptic-table-Ansicht die einzelnen cluster anschauen. Nützlich ist dazu der Befehl

**Synoptic table → Analysis of Columns of Synoptic Table**

Wie bei allen rein numerischen Verfahren können Sie das Ergebnis nicht als Endpunkt einer sinnvollen Klassifikation ansehen. Hier muss stets auf ökologische Sinnhaftigkeit gerüft werden und händisch nachgebessert werden. Die Euklidische Distanz ist tückisch, weil sie ein symmetrisches Distanzmaß ist, die also nicht nur gemeinsame Arten, sondern auch gemeinsam fehlende Arten hoch bewertet.

## **2.6.7 Klassifikation mit Hilfe von TWINSPAN**

Nun kommen wir zu einer numerischen Klassifikationsmethode, bei der keine Vorinformationen einfließen und die in der Vegetationskunde eine weite Verbreitung gefunden hat. Vielfach wird nämlich (besonders im angelsächsischen Raum) die

Überlegung, Vegetationsdaten nach „Charakterarten“ oder „ökologisch-soziologischen Artengruppen“ zu bestimmen, als zu subjektiv und wenig nachvollziehbar kritisiert. Die Statistik bietet dagegen über iterative Berechnungsverfahren oder Distanzmaße objektive Klassifikationsmethoden an, bei denen das ökologische Hintergrundwissen über die einzelnen Arten keine Rolle mehr spielt. Dies ist für Anfänger sehr erfreulich, aber zugleich auch sehr gefährlich! Berechnungen beziehen sich immer nur auf die gerade geöffnete Tabelle, und hier können zu geringe Aufnahmezahlen, Pseudoreplikationen, räumliche Besonderheiten, bestimmte Ausbildungsformen der Vegetation oder einfach zufällige Häufungen von Arten Ähnlichkeiten oder Distanzen vortäuschen, die in einem breiteren Kontext nicht haltbar sind. Numerische Klassifikationsmethoden sind also mit der gebührenden Vorsicht zu nutzen und Verallgemeinerungen immer erst nach gründlichem Literaturvergleich vorzunehmen!

In JUICE kann man sehr benutzerfreundlich das von HILL (1979) geschriebene Programm TWINSPAN („two way indicator species analysis“) verwenden. Es ist auch heute noch das am häufigsten verwendete Programm, um vor allem größere Tabellen zumindest vorzusortieren, da es ähnliche Ergebnisse wie eine geübte „händische“ Sortierung liefern kann. Für gute Ergebnisse müssen aber dieselben Voraussetzungen wie für eine Correspondenzanalyse (CA, siehe Kap. 2.8) gegeben sein (Gradient genügend lange, ein ökologischer Hauptfaktor wirkend). Man sollte also vor der Anwendung von TWINSPAN eine im Kapitel 2.8 geschilderte DCA durchführen, um zu überprüfen, ob die Gradientenlänge zumindest 4 SD erreicht. TWINSPAN arbeitet „divisiv“, d. h. es untersucht zunächst den gesamten Datensatz, teilt diesen in zwei Teile (möglichst große Varianz zwischen den beiden Teilen, möglichst geringe Varianz innerhalb der Teile), im nächsten Schritt werden die beiden Teile nach dem gleichen Prinzip wieder geteilt usw. Die Anzahl der Teilungsschritte ist vorwählbar, zumindest die ersten Schritte liefern meist brauchbare Ergebnisse, die weiteren sind meist eher wenig gut, da ja der Gradient innerhalb der Teile immer geringer wird und irgendwann keine unimodale Reaktion der Arten mehr vorhanden ist und deshalb TWINSPAN eigentlich nicht mehr sinnvoll angewendet werden kann.

### 2.6.7.1 Pseudospecies cut levels

TWINSPAN verwendet nur Präsenz-Absenz-Werte der Arten, durch das Prinzip der „pseudospecies cut levels“ können nun aber in gewissem Umfang auch die Deckungswerte der Arten in die Kalkulation eingebunden werden. TWINSPAN bildet intern für Arten, welche die angegebenen Deckungs-Schwellenwerte überschreiten „Pseudospecies“. Geben wir z. B. 3 Pseudospecies cut levels mit 0, 5, 25 ein, vergibt TWINSPAN für die Art *Trifolium repens*, die in einer Aufnahme mit Deckung 4 (=68 % Deckung, Mittelwert der Deckungsklasse 4: 51-75% Deckung) drei Pseudospecies: Trif rep1 (weil die Deckung 0 überschritten wird), Trif rep 2 (weil auch die Deckung 5% überschritten wird) und Trif rep 3 (weil auch der letzte angegebene Deckungsschwellenwert von 25% überschritten wird); Dadurch werden sich z. B. Aufnahmen ähnlicher, die *Trifolium repens* in einer Deckung über 25% enthalten, weil sie mehr gemeinsame Pseudospecies enthalten. Wie viele solcher Pseudospecies cut levels für eine Analyse verwendet werden sollen, kann man nicht vorhersagen; aber wenn man annimmt, dass hohe Deckungswerte für Arten in den Aufnahmen standortsökologische Ursachen haben könnten, sollte man die cut levels darauf abstimmen. Am besten versucht man es mit unterschiedlichen cut-levels, beginnend mit der Grundeinstellung (3 cut levels, 0, 5, 25) und schaut sich an, ob sich die Tabellenstruktur nach Änderung verschiebt und welche Tabelle am besten zu interpretieren ist.

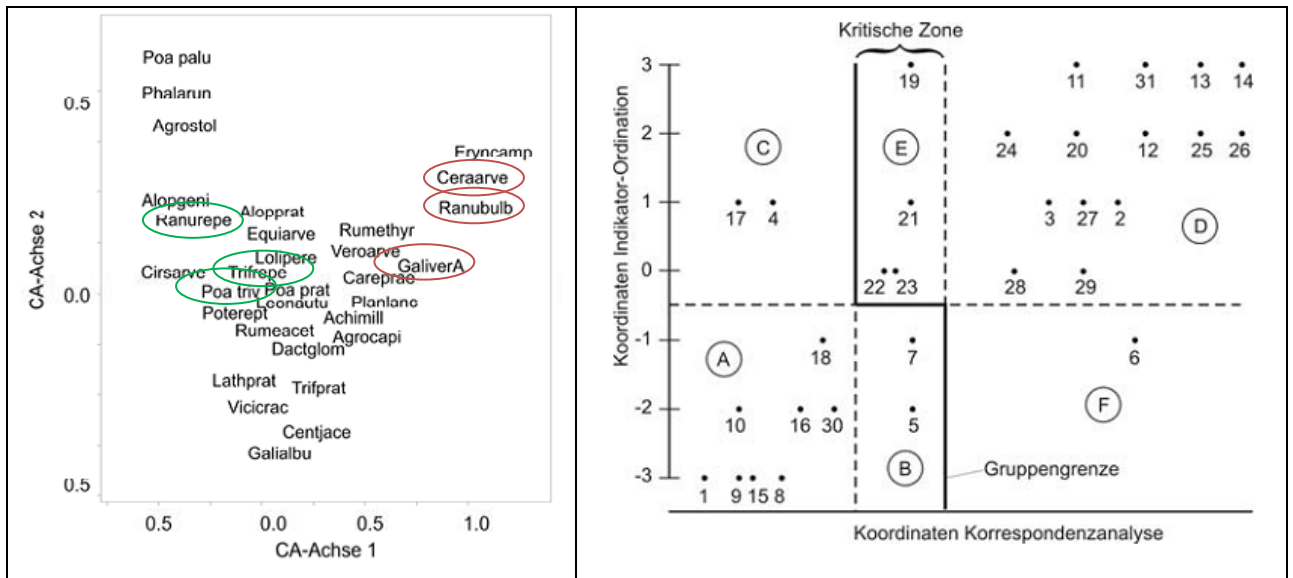
### 2.6.7.2 Ablauf in JUICE

#### Analysis → TWINSPAN → TWINSPAN (Hill 1994)

Nun können die „Pseudospecies cut levels“ gewählt werden und man kann in dieser Version auch nur Tabellenteile bearbeiten (wählen einer colour). Unter **Minimum group size** kann man angeben, ab welcher Aufnahmezahl Tabellenteile nicht weiter geteilt werden sollen. **Maximum level of divisions** legt fest, wie oft TWINSPAN Teilungen durchführen soll, z. B. 3: TWINSPAN führt 3 Teilungen durch, d. h. es entstehen 8 Spalten, weil ja jede neue Teilung im nächsten Schritt noch einmal geteilt wird. Um z. B. auch 5 oder 7 Spalten zu erhalten, kann die von ROLECEK et al. (2009) überarbeitete Version verwendet werden. Hier kann die Anzahl der Spalten (Number of clusters) beliebig eingestellt werden, da in dieser Version TWINSPAN nur jenen Tabellenteil teilt, der die höchste Heterogenität (innere Unähnlichkeit) zeigt. Für deren Bestimmung können verschiedene Ähnlichkeits- bzw. Distanz-Indices gewählt werden (z. B. Jaccard, Sørensen etc.).

TWINSPAN verwendet für die Analyse im Prinzip das gleiche Verfahren wie die in 2.8 beschriebene CA, es ermittelt zuerst das „gewichtete Mittel (weighted average)“ der Aufnahmen und der Arten und sortiert diese nach den Werten. Da nur die erste Achse verwendet wird, kann das „detrending“ unterbleiben (siehe Abbildung 3 links). TWINSPAN soll nun Grenzen zwischen Aufnahmegruppen finden und dies geschieht durch Auswahl sogenannter **Indikatorarten**, d. h. von Arten, die deutlich an der rechten (+) bzw. an der linken Seite (-) der ersten Achse angeordnet sind (die meist seltenen Arten an den äußersten Rändern werden nicht verwendet, „downweighting of rare species“).





**Abbildung 3:** Links – Erste CA-Achse (Twinspan), rot die von TWINSpan ausgewählten +Indikatorarten, grün die – Indikatorarten; Rechts – Aufnahmen aufgetragen nach CA-Werten (x-Achse) und Indikatorkoordinaten (y-Achse). Aus LEYER & WESCHE (2008).

Diese Indikatorarten werden dann für die Aufteilung der Objekte (Aufnahmen) genutzt. Dazu wird ein „Indikatorwert“  $I_j$  für jede Art ermittelt: relative Frequenz einer Art in der negativen Gruppe minus relative Frequenz in der positiven Gruppe (jeweils in % der Gesamtaufnahmen in der Gruppe):

$$I_j = \frac{n_j^+}{n_+} - \frac{n_j^-}{n_-}$$

Die Arten mit den höchsten + (links) bzw. – (rechts) Indikatorwerten werden als Indikatorarten ausgewählt, meist sind es nur wenige (1–5). Nun werden für jede Aufnahme die Indikatorwerte der (meist sehr wenigen!) Indikatorarten addiert, dies ergibt die Indikatorkoordinaten. Aus den Werten der CA für die Aufnahmen und diesen Indikatorkoordinaten erzeugt TWINSpan ein eindimensionales Diagramm (x-Achse: CA-Werte, y-Achse Indikatorkoordinaten, vgl. Abbildung 3 rechts), dies wird als „refined ordination“ bezeichnet. Trotz (oder eigentlich wegen) des gewaltigen Informationsverlustes, es werden ja nur wenige Arten verwendet, zeigt sich nun eine bessere Gruppenstruktur und es ist leichter, Grenzen zwischen den Gruppen zu bilden.

Die meisten Aufnahmen haben ähnliche CA- und Indikatorkoordinaten-Werte, d. h. sie liegen im Diagramm unten links bzw. oben rechts (Gruppen A und D in Abbildung 3 rechts). Aufnahmen nahe der Grenze werden als „borderline positives bzw. borderline negatives“ bezeichnet (B, E in Abbildung 3 rechts).

Einige Aufnahmen unterscheiden sich aber in ihren Werten für die CA- und Indikatorkoordinaten, sie liegen also links oben und rechts unten, sie werden als „misclassified negatives bzw. positives“ bezeichnet (C, F in Abbildung 3 rechts).

Die Grenze zwischen den Gruppen wird iterativ bestimmt, sodass möglichst wenige „misclassified Aufnahmen“ entstehen. Diese „misclassifieds“ sollten in der Tabelle überprüft werden.

Leider erzeugt das unter JUICE laufende TWINSpan keinen vollständigen Ergebnis-File um die eigenvalues der einzelnen Teilungsschritte, die gewählten Indikatorarten sowie die misclassifieds auslesen zu können; Dafür muss TWINSpan mit dem Datei-Manager gestartet werden, oder noch besser, man verwendet WINTWINS, eine an Windows angepasste Version von TWINSpan (freeware, <http://www.canodraw.com/wintwins.htm>). Zuerst müssen wir aber einen file erzeugen, den TWINSpan lesen kann. Das können wir mit JUICE erledigen: **File** → **Export** → **Table** → **to Cornell Condensed File (CC!)** → **name.cc!** in den JUICE-Ordner speichern mit einem möglichst einfachen Namen (er muss beim Start des Programmes im Windows-Explorer genau angegeben werden).

Dann wintwins.exe starten (Icon erscheint nach Installation am Desktop), File angeben, der analysiert werden soll, dann kann der Ergebnisfile (name.twp) gelesen und analysiert werden, z. B. kann man den Eigenvalue der einzelnen Teilungsschritte auslesen (sollte über 0,5 sein) sowie die einzelnen + oder – Indikatorarten oder die „misclassifieds“.

Bedenken Sie, dass der „diagnostische Wert“ einer Art für automatische Klassifikationsprogramme immer gleich ist. So werden zufällige Verteilungen lokal häufiger, euryöker Arten genauso bewertet werden wie das Vorkommen einer stenöken Art mit hohem Zeigerwert. Händisches Nachbearbeiten von automatischen Klassifikationen sollte man also immer erwägen (siehe Kapitel 2.6.5).

## 2.6.8 Abschließende Tabellenordnung

Die meisten Tabellen werden exportiert, designt und präsentiert als Stetigkeitstabellen. Dies verbirgt zwar die Einzelaufnahmen, fasst dafür aber im Sinne wissenschaftlicher Abstraktion wichtige Ergebnisse der Tabellenarbeit

zusammen. Am Ende sollte eine Tabelle so gestaltet sein, dass diagnostisch bedeutende Arten oben stehen und weniger wichtige am Ende. Die Differenzialarten der Einheiten stehen oben von links nach rechts, dann folgen ggf. Differenzialarten höherer Einheiten, die sich also über mehrere Cluster der kleinsten Hierarchie erstrecken. Diese Arten werden meist nach fallender Stetigkeit in dem jeweiligen Cluster geordnet. Dann folgen die häufigsten Arten der Tabelle, die sich also mehr oder weniger über die gesamte Tabelle verteilen. Dies sind oft Differentialarten höherer Syntaxa. Dieser Abschnitt sollte dann auch nach Schichten geordnet sein. Am Ende der Tabelle kommen sonstige Arten, die sich nicht in die anderen Blöcke einordnen lassen, überwiegend Arten mit geringer Stetigkeit (siehe Schema der Tabelle 4). Arten, die in der Tabelle nur einmal vorkommen, könnte man noch als gesonderten Block ans Ende stellen, um die Länge der Tabelle zu verkürzen und sie von unwichtigen Informationen zu entlasten. Diese werden bei einer Veröffentlichung einer Tabelle oft in einem Textblock unter die Tabelle gestellt (den kann JUICE auch automatisch erstellen).

Jetzt sind wir weitestgehend mit der Tabellenarbeit fertig. Speichern wir noch die Art-Farben und die Aufnahme-Farben unserer fertigen Tabelle. Mit dem Befehl **Relevés → Relevé Colours → Load** kann man sich einen Vergleich mit zwischengespeicherten Gliederungen ansehen. Wir könnten die Tabelle jetzt also fertig layouten. Mit dem Befehl **File → Export → Synoptic Table**

können Sie die Tabelle ins csv-Format überführen und in EXCEL laden. EXCEL ist gut geeignet, um ein schönes Layout einer Tabelle zu erzeugen.

Der zweitwichtigste Teil einer Tabelle sind die Kopfinformationen. Hier sollte mindestens der Name des Clusters (oder eine Abkürzung dafür) und die Anzahl der Aufnahmen stehen, weitere Angaben richten sich nach dem Ziel der Arbeit, z. B. geographische oder bibliographische Herkunft der Aufnahmen, mittlere Artenzahlen, mittleren Flächengrößen usw.

**Tabelle 4:** Schematische Ordnung einer JUICE synoptic table. (A = Artengruppen; Gem. A. = Arten, die mehr oder weniger häufig in allen Clustern vorkommen und für kein Cluster das Differenzialarten-Kriterium gegenüber allen anderen Cluster erfüllen; Sonst. = sonstige, oft geringste Arten; tl = Tree layer usw.)

	Einheit 1	Einheit 2	Einheit 3	Einheit 4
Diff.A. 1				
Diff.A. 1				
Diff.A. 1				
Diff.A. 1				
Diff.A. 3+4				
Gem. A. tl				
Gem. A. sl				
Gem. A. hl				
Gem. A. ml				
Sonst. tl				
Sonst. sl				
Sonst. hl				
Sonst. ml				

## 2.7 Arbeiten mit ökologischen Zeigerwerten

### 2.7.1 Einführung

Viele Pflanzenarten zeigen enge Anpassungen an bestimmte Umweltbedingungen. Da Pflanzen während ihrer Lebenszeit ihren Wuchsort nicht ändern können und normalerweise keine Thermoregulation besitzen, kann ihr Vorkommen für Rückschlüsse auf viele ökologischen Eigenschaften ihres Standortes genutzt werden (passives Biomonitoring, Bio-Indikation). Ellenberg (ELLENBERG & al. 1991) hat für fünf Standortfaktoren: Licht L (light) Reaktion R (soil reaction), Feuchtigkeit F (moisture M), Salzgehalt (S, nicht in JUICE verfügbar) und Nährstoffgehalt N (nutrients) sowie die zwei Klimafaktoren Temperatur T (temperature) und Kontinentalität K (continentality C) fast allen Gefäßpflanzen Mitteleuropas Zeigerwerte zugeordnet, meist empirisch aus Erfahrungswissen, teilweise durch direkte Messungen (F, N, R) an Wuchsorten, teilweise durch Arealanalysen (K, T). Diese Zeigerwerte beziehen sich auf das **ökologische Optimum** der Arten, d. h. sie berücksichtigen die Konkurrenz auf den Standorten. In zahlreichen Anwendungen hat sich die Brauchbarkeit dieser Zeigerwertskala für die Abschätzung von Standortbedingungen gezeigt, so z. B. der Zusammenhang zwischen durchschnittlichem Grundwasserspiegel und den F-Werten oder die Biomasseproduktion mit den N-Werten (SCHAFFERS & SÝKORA 2000).

Es handelt sich um Ordinalwerte in einer 9-stufigen Skala (außer F 12-stufig) von 1 bis 9. Besonders gute Rückschlüsse auf Standortbedingungen erlauben synoptische Werte von Vegetationsaufnahmen, die ja alle vorkommenden Gefäßpflanzenarten und oft auch Moose und Flechten berücksichtigen. Dabei werden die einzelnen Zeigerwerte der vorkommenden Pflanzen in einer Vegetationsaufnahme summiert und dann durch die Anzahl der Arten dividiert (arithmetisches Mittel). Da es sich um eine Ordinalskala handelt, ist die Mittelwertbildung statistisch eigentlich falsch und es sollte der Medianwert verwendet werden. In der Praxis hat sich aber gezeigt, dass die Mittelwerte ausreichend robust sind, um interpretierbare Ergebnisse zu liefern.

## 2.7.2 Initialisierung

Als erster Schritt müssen die Namen der Tabelle mit den Ellenberg-Werten verbunden werden (Initiation). Die dazu erforderliche Datei „Ellenb.txt“ wird bei der Installation von JUICE in den JUICE Ordner geschrieben.

**Indicator values → Initiation → Initiation of indicator values**

Jetzt den Pfad suchen, meist C:\Programme (x86)\Juice\Ellenb.txt

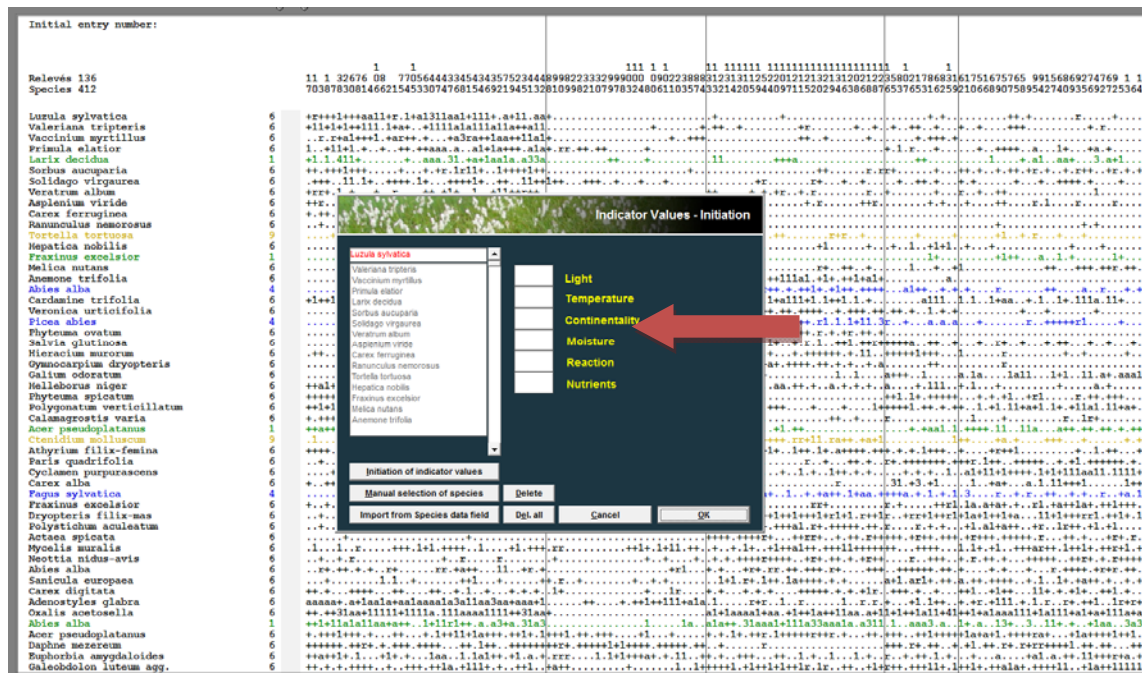


Abbildung 4: Fenster zur Initialisierung der Zeigerwerte

Nun erscheinen für die im linken Block in der ersten Zeile fettgeschriebene Art die Zeigerwerte für L, T, C, M, R und N wenn sie für die Art in der Ellenb.txt Liste eingetragen sind. Um zu prüfen, ob alle Arten gefunden wurden, benutzt man die Pfeiltasten am rechten Rand des Blocks. Arten, die nicht gefunden wurden, werden rot dargestellt (Abbildung 4). Taucht eine solche Art auf, drückt man den Button **Manual selection of species** und macht einen Doppelklick auf den roten Namen. Im rechten Fenster werden nun Varianten dieses Namens geboten (alphabetische Ordnung beachten!), die man wählen kann. Es ist auch möglich, die Zeigerwerte manuell in die Felder einzutragen. Sind allen Namen zur Zufriedenheit Werten zugeordnet, drücken Sie zuerst auf **Close** und anschließend auf **OK**. Abschließend speichern wir die Datei unter neuem Namen ab.

## 2.7.3 Berechnung von Aufnahme-Werten

Nun kann JUICE auf die Ellenberg-Werte zugreifen. Für eine Information der Werte für einzelne Arten kann man sich einen Zeigerwert in die species data Spalte eintragen lassen:

**Species → Species data → Indicator value → z. B. Temperature**

Diese Funktion kann man verwenden, um z. B. Feuchtigkeit zeigende Pflanzen herauszufinden. Mehr Bedeutung hat aber die Berechnung der Mittelwerte für einzelne Aufnahmen. Mit **Indicator values → Calculation for Relevés** → bekommt der Mauszeiger ein Fragezeichen. Sie können nun auf eine beliebige Aufnahme klicken und es erscheint das Fenster wie in **Abbildung 5**.

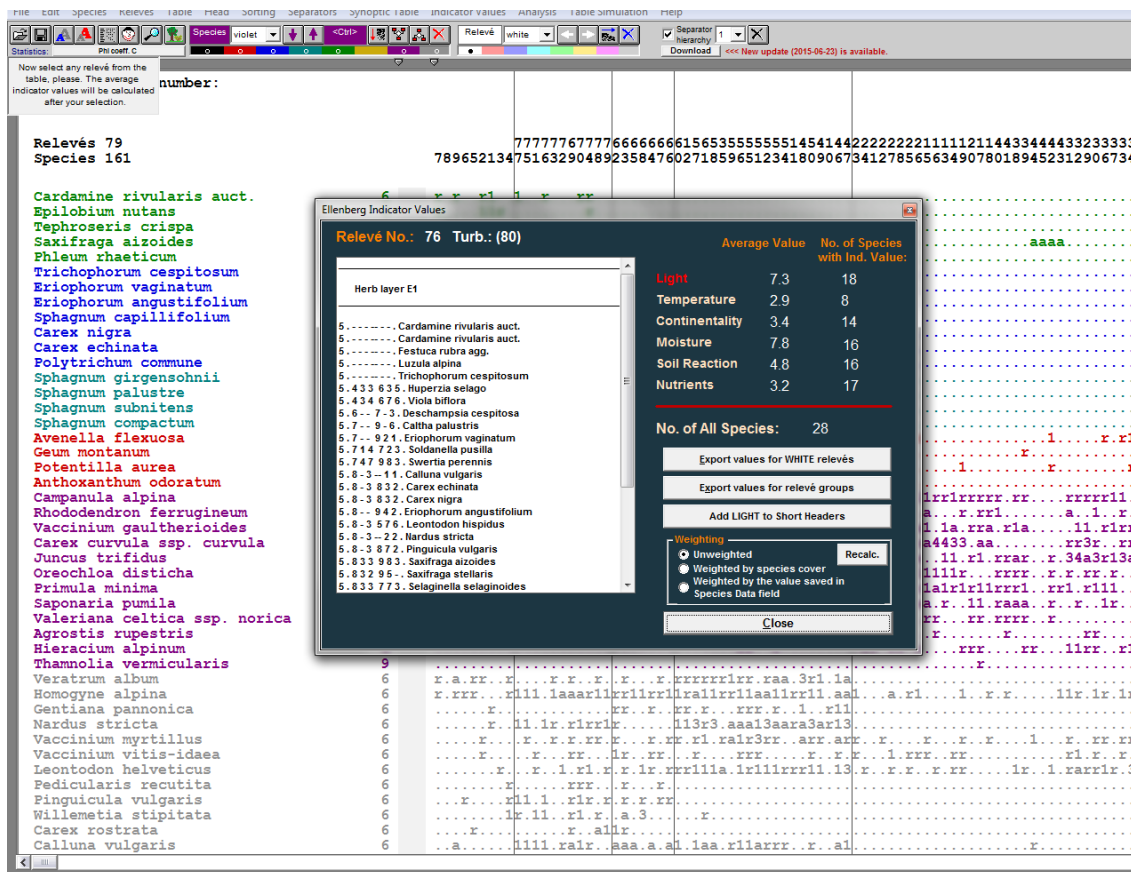


Abbildung 5: Berechnung von mittleren ökologischen Zeigerwerten einer Aufnahme

Der rot markierte Wert wird berechnet. Mit der linken Maustaste können andere Zeigerwerte ausgewählt werden. Im unteren Block kann noch gewählt werden, ob die Arten nach ihren Deckungswerten oder anderen im „species data“ Feld eingetragenen Werten gewichtet werden sollen. Im mittleren Feld kann angegeben werden, was mit den berechneten Werten weiter geschehen soll: **Export values for relevé groups** erzeugt den file `ecol.txt`. Wenn schon Tabellenbereiche mit Separatoren abgetrennt wurden, wird ein in EXCEL importierbarer file erzeugt, der nach den in JUICE erzeugten Spalten getrennt die Werte der einzelnen Aufnahmen dargestellt sind, z. B. die Mittelwerte der einzelnen Zeigerwerte, die Anzahl der Arten, die für die Berechnung zur Verfügung standen und die Standardabweichung. Ist diese sehr hoch, hat man offenbar Arten mit sehr unterschiedlichen Ansprüchen in der Aufnahme und sollte bei der Interpretation vorsichtig sein.

Import von `ecol.txt` in EXCEL: **Datei** → **Öffnen** → Filter auf „alle Dateien“ oder „txt“ stellen → `ecol.txt` wählen → im Textkonvertierungsassistenten „Getrennt“ wählen → **Weiter** → Trennzeichen „Semikolon“ wählen → **weiter** → unbedingt „weitere“ anklicken (sonst versucht EXCEL alle Zahlen mit Dezimalstellen in Datum umzuwandeln) → Dezimaltrennzeichen auf „.“, 1000er Trennzeichen auf „ „ stellen → **Weiter** → **Fertigstellen**. Nun können z. B. am Ende der Gruppen Leerzeilen eingefügt werden und die durchschnittlichen Zeigerwerte der einzelnen Aufnahmen der Gruppen noch einmal gemittelt und miteinander verglichen werden (z. B. in einem Diagramm).

Wir können die Werte aber auch in JUICE weiterverarbeiten. Dazu wählen wir: **Indicator values** → **calculation for Relevés** → **Add „Zeigerwert“ to Short Headers**: Der durchschnittliche Zeigerwert wird in die **Short Headers** übernommen. Um diesen Wert in den Kopfdaten abzuspeichern benutzen wir wieder **head** → **Add short Headers to Header data** → Namen für den Wert angeben, z. B. `Lweight` (für Light, gewichtet nach Deckungswerten). Mit **Head** → **Short Header Averages** können Sie sich die Durchschnittswerte der einzelnen Cluster anzeigen lassen und so Vergleiche der ökologischen Situation durchführen. Wenn die Tabelle keine Cluster enthält oder Sie nach bestimmten Gradienten suchen, können Sie mit **Sorting** → **Sort Relevés by Header data** alle Aufnahmen z. B. nach aufsteigenden Lichtwerten sortieren.

## 2.7.4 Anwendungsbeispiele für Zeigerwerte

Zeigerwerte liefern keine absoluten Aussagen über die Indikatoren, sondern nur Hinweise. Sie eignen sich deshalb vorwiegend für die vergleichende Analyse. Anwendungsbeispiele wären beispielsweise:

- Vorsortieren einer Roh Tabelle: man errechnet die durchschnittlichen Zeigerwerte der Aufnahmen. Unterscheiden sich die mittleren Zeigerwerte der einzelnen Aufnahmen wenig, ist dieser Faktor offenbar nicht als Gradient

wirksam (z. B. L in Wiesenaufnahmen; K beim Vergleich von Aufnahmen aus einer Region). Hat man einen durchschnittlichen Zeigerwert gefunden, der größere Unterschiede zwischen den einzelnen Aufnahmen zeigt, liegt offenbar ein als Gradient wirkender Standortfaktor vor. Dieser kann nach der oben beschriebenen Methode zur Vorsortierung einer Rohtabelle eingesetzt werden.

- b) Landschaftsveränderungen (BÖCKER & al. 1983): liegen hinreichend lokalisierbare historische Vegetationsaufnahmen vor, kann durch Vergleich der durchschnittlichen Zeigerwerte aktueller Aufnahmen auf Änderungen der Standortökologie geschlossen werden (z. B. N-Zunahme nach Düngung; F-Abnahme nach Meliorationen etc.)
- c) Vergleiche von räumlich getrennte Aufnahmegruppen: Gedüngt – ungedüngt; Genutzt – Brache; Nordhang – Südhang; Außendeichs – Innendeichs; Stadt – Agrarlandschaft; Aufnahmen mit und ohne Neophyten; usw.
- d) Klimawandel: Zunahme von wärmeliebenden Arten in Zeitreihen wird als Hinweis auf Global warming angesehen („thermophilization“, ERSCHBAMER et al. 2011)

## 2.7.5 Probleme bei der Anwendung von Zeigerwerten

**Waldbäume:** viele heimische Waldbäume sind in der Jugend auf Beschattung angewiesen (z. B. *Abies alba*, *Fagus sylvatica*), im Alter aber +/- voll dem Licht ausgesetzt. Deshalb sollte die Baumschicht bei der Berechnung der durchschnittlichen Lichtwerte nicht mit einbezogen werden.

**Arealgrenzen:** viele Arten zeigen an der Grenze ihres Areals andere ökologische Schwerpunkte als im Zentrum ihres natürlichen Verbreitungsgebietes;

**Kryptogamen:** Erdmoose und -Flechten zeigen die Situation direkt an der Bodenoberfläche an, diese kann von der Situation wurzelnder Pflanzen abweichen. Dies kann nützlich, aber auch ungünstig sein, je nach Fragestellung.

**Statistik:** Es wurde in der Einleitung darauf hingewiesen, dass es sich bei den Zeigerwerten um Ordinalwerte handelt, d. h. eigentlich keine Mittelwerte gebildet werden dürften. In der Praxis wird dieser Fehler aber häufig in Kauf genommen, weil die Werte sehr plausibel sind. ZELENÝ & SCHAFFERS (2012) bieten ein mit JUICE nutzbares R-script zur Signifikanzkorrektur an. Eine weitere Möglichkeit wäre es, Histogramme der einzelnen Aufnahmen oder Aufnahmegruppen zu erstellen und diese zu vergleichen.

**Sprachgebrauch:** Schreiben Sie nie: „Die Südhänge sind wärmer und trockener.“, sondern immer: „In den Aufnahmen der Südhänge sind mehr Wärme- und Trocknis-Zeiger.“

## 2.8 Multivariate Statistik: Detrended correspondence analysis mit JUICE und R

### 2.8.1 Einführung

Bei der Sortierung von Vegetationstabellen (Rohtabelle, nach der Eingabe in TURBOVEG, import in JUICE) können zwei grundsätzliche Ziele verfolgt werden:

1. **Klassifikation:** einander floristisch ähnliche Aufnahmen werden gruppiert und eventuell hierarchisch gegliedert (z. B. mehrere Assoziationen zu einem Verband). Dies wird im Abschnitt 2.5 und 2.8 behandelt.

2. **Ordination:** Im Vordergrund steht die Suche nach einem ökologischen Gradienten, der mit der floristischen Zusammensetzung der Aufnahmen in Zusammenhang steht (vgl. Abschnitt 2.6), und nach diesem werden die Vegetationsaufnahmen sortiert. Da durch dieses Verfahren am Ende aber ebenfalls die ähnlichsten Aufnahmen nebeneinander stehen, lassen sich die gleichen statistischen Methoden auch für die Klassifikation (mit speziellen Adaptierungen) anwenden.

Wird bei der Suche nach einem ökologischen Gradienten nur die Artenzusammensetzung der Aufnahmen untersucht, spricht man von **indirekter Gradientenanalyse**, d. h. das Sortierprinzip liegt ausschließlich innerhalb des floristischen Datensatzes. Obwohl häufig aus der Not geboren, weil selten verlässliche und über ausreichend lange Perioden gemessene ökologische Standortdaten vorliegen, stellt dies, unter gewissen Umständen, eine sehr elegante Möglichkeit dar, auf stark wirkende Standortparameter schließen zu können.

Dabei werden **multivariate** Ordinationsverfahren angewendet:

- Anordnung von Arten/Aufnahmen in einem vielachsigen Koordinatensystem: Suche nach jenen Achsen („Gradienten“) die maximale Varianz im Datensatz auf möglichst wenige Dimensionen reduzieren.

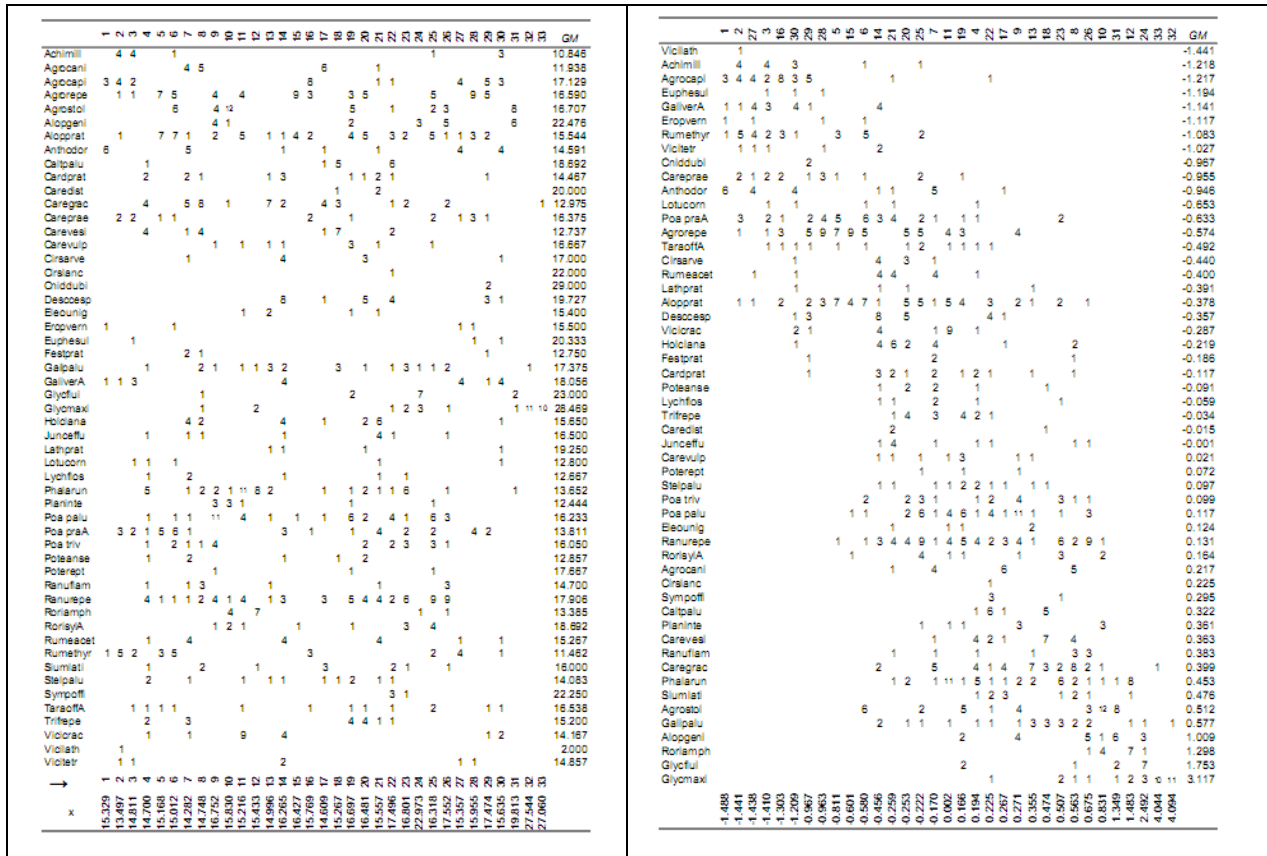
Das Sortierprinzip ist das sogenannte **gewichtete Mittel** (weighted average) für die Arten und die einzelnen Aufnahmen, das Verfahren Correspondence Analysis (Hill 1994), (Abkürzung: CA“, auch „reciprocal averaging“).

Es funktioniert nach folgendem Ablauf (vgl. **Abbildung 6**, nach LEYER & WESCHE 2008):

1. Wählen beliebiger sample scores (Werte für jede Aufnahme, egal wie groß, nur unterschiedlich, z. B. Relevé number)
2. Berechnung der species scores (aus den frei gewählten sample scores) (z. B. Alchem:  
 $[(2 \times 4) + (3 \times 4) + (1 \times 6) + (25 \times 1) + (30 \times 3)] / 13 = 10,846$

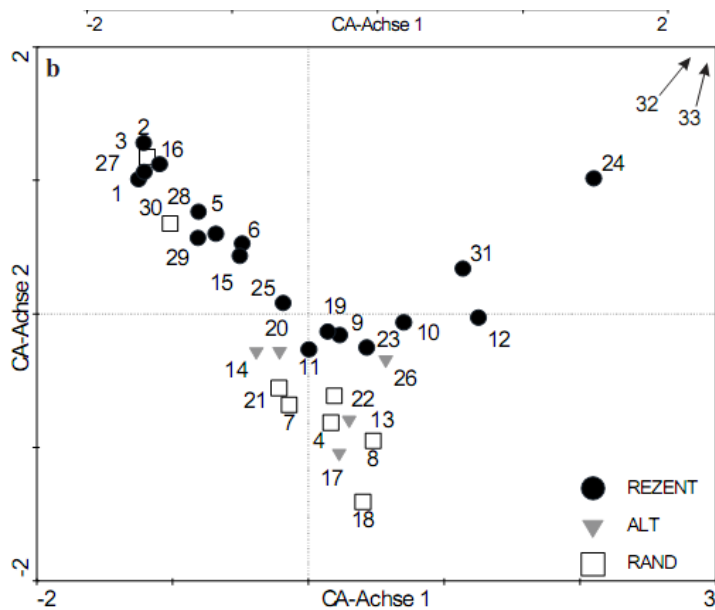
- Berechnung sample scores aus den in 2. errechneten species scores (z. B. Aufn 1:  $[17,129 \times 3 + 14,591 \times 6 + 15,500 \times 1 + 18,056 \times 1 + 11,462 \times 1] / 12 = 15,329$ )
- Wiederholung von 2. und 3. bis sich die species u. sample scores („reciprocal averaging“) stabilisiert haben.

Nun werden die Arten und die Aufnahmen nach den scores sortiert – im Idealfall entsteht eine Tabelle mit Diagonalstruktur, die einen Gradienten abbildet. Welche ökologische Qualität dieser Gradient abbildet, muss nun durch Untersuchung der Arten erklärt werden. In unserem Fall sind die Arten mit niedrigem species-score (z. B. *Galium verum*, *Euphorbia esula*) eher trockenheitstolerant, während die Arten mit hohem Wert (in unteren Teil der Tabelle rechts), wie *Glyceria maxima*, Wasser- oder Sumpfpflanzen sind, d. h. der offenbar am stärksten wirksamste Gradient ist die Feuchtigkeit.



**Abbildung 6:** links – Tabelle mit sample – und species scores; rechts – Tabelle sortiert nach sample- und species-scores. Aus LEYER & WESCHE (2008).

Die errechneten Werte können nun für die Arten und die Aufnahmen auf einer Achse eingetragen werden (1. Achse). Nun wirken auf Pflanzen an einem Standort aber meist mehr als ein ökologischer Faktor – die erste Achse bildet ja nur einen, den am stärksten wirkenden ab. Weitere Achsen können nach einer „Orthogonalisierung“ (damit sie nicht mit den Werten der 1. Achse korrelieren) nach dem gleichen Prinzip wie die erste Achse gewonnen werden (vgl. LEYER & WESCHE 2008). Meist werden 2 Achsen verwendet, aber man kann auch eine 3. Achse ermitteln und aus allen Werten eine dreidimensionale Darstellung bereiten. Werden entweder nur die Aufnahmewerte oder nur die Artwerte eingetragen, spricht man von einem „scatter plot“, siehe **Abbildung 7**, werden Aufnahmen und Arten gemeinsam dargestellt von einem „biplot“.



**Abbildung 7:** Scatter plot von Aufnahmen. Aus LEYER & WESCHE (2008).

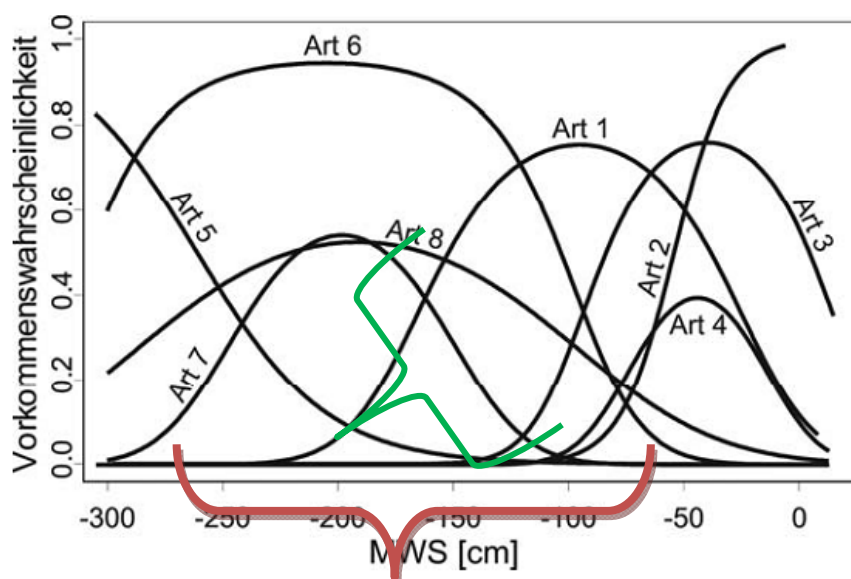
### 2.8.2 Probleme bei der CA

Wie in **Abbildung 7** sichtbar, zeigt die 2. Achse einen „arch-effect“, d. h. die Werte im Zentrum sind niedriger als jene am Rand. Dies wird durch einen groben mathematischen Trick behoben, dem „detrending“: Dabei wird die 1. Achse in gleiche Abschnitte geteilt (Segmente) und die Werte für die 2. Achse um den Mittelwert in den jeweiligen Segmenten reduziert. Dies sollte bei der Interpretation der 2. Achse berücksichtigt werden. Das Verfahren heißt DCA (detrended correspondence analysis).

Seltene Arten: die CA gewichtet seltene Arten zu stark, deshalb sollte immer **downweighting of rare species** gewählt (ist in den gängigen Anwendungsprogrammen anwählbar) oder die seltenen Arten vor der Analyse überhaupt entfernt werden.

### 2.8.3 Voraussetzungen für die Anwendbarkeit der CA bzw. DCA auf einen Datensatz

Eine Voraussetzung für die Anwendbarkeit der Methode ist die **unimodale** Reaktion von Pflanzenarten auf einen ökologischen Faktor, d. h. die Arten besitzen **ein** Optimum des Vorkommens (Stetigkeit, Deckung) entlang eines ökologischen Gradienten und die Aufnahmen müssen sich über und unter das Optimum der auftretenden Arten erstrecken (vgl. **Abbildung 8**)



**Abbildung 8:** Unimodale Reaktion (rot) und lineare Reaktion von Art 7 (grün) auf einen Feuchtegradienten (mittlerer Grundwasserstand MWS). Aus LEYER & WESCHE (2008).

## 2.8.4 DCA aus JUICE mit R

Für eine Anwendung der DCA in JUICE ist es nötig, am Rechner zuerst das R-Programm zu installieren (normalerweise unter <http://cran.at.r-project.org/>). JUICE arbeitet nur mit den älteren Versionen 2.10.1 bis 2.13, diese gibt es noch im Internet, wenn man danach sucht. Achten Sie darauf, dass im Verzeichnis Library das package Vegan installiert ist. Dann muss in JUICE der Pfad zum Programm Rgui.exe eingestellt werden: **File** → **Options** → **External Program Paths** → **R-Projet, change file**, hier das Rgui.exe am Rechner suchen (im Unterverzeichnis bin und den Pfad hier angeben). Dann alle Spaltenfarben auf weiß stellen.

**Analyses** → **Ordinations** → **R-Project** → **DCA Analysis**

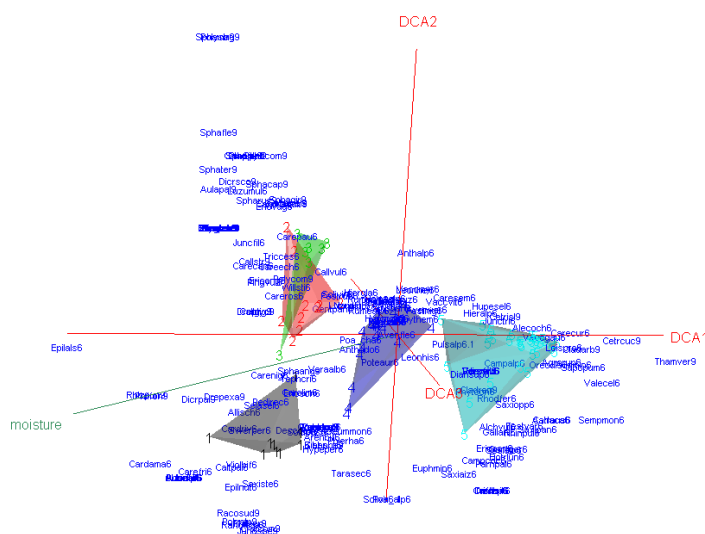
Beim ersten Start im letzten Ordner zuerst **Associate R-Project with Juice**, dann **Installation of R-libraries**, dann **Update ORDIIJUICE script** und **Clean temporary files** drücken. Diese Routine muss später nicht mehr durchgeführt werden. **DCA Analysis** → **2D oder 3D Projektion** einstellen (bei 3D Ansicht sind die Objekte besser sichtbar, aber die 3. Achse hat meist gar keinen Aussagewert) → **Display Parameters**: man kann sowohl die Arten (species) als auch die Aufnahmen (sites) mit Namen (with all labels) oder als Punkte darstellen lassen; **Envelopes**: die Aufnahmegruppen werden in eine farbige Wolke eingehüllt → next → die Deckungswerte können transformiert werden, z. B. mit den Pseudospecies levels, den Button **Downweight rare species** aktivieren → hier kann man einen Wert aus den Kopfdaten (**Use data from Short Headers as**) als Vector oder Surface einblenden lassen, diesen muss man aber vor der Durchführung in die **short header** eintragen lassen) oder man wählt aus den gespeicherten **Header data** einen Wert für „**distinct environmental factor**“ (dieser wird neben die Aufnahme nummer eingehendet) oder einen für den „**Environmental vector**“: dieser wird als Vektor in das Diagramm eingehendet. Auch kann man alle Ellenberg-Werte als Vektoren (ungewichtet) einblenden lassen **Use Ellenberg Indicator Values as vectors** → **Finish**.

Unter **Show summary** kann man sich die Gradientenlänge und die Eigenvalues für alle ausgewerteten DCA-Achsen ansehen. Beurteilen Sie nun, ob die Gradientenlänge ausreicht und wie hoch die Aussagekraft der einzelnen Achsen ist. Durch Drehung des Diagrammes kann man die am besten ablesbare Position auswählen und diese für Verwendung in anderen Programmen als **snapshot** (png-Datei) im R-Ordner (unter bin\3Dsnaphshots) abspeichern.

## 2.8.5 Interpretation einer DCA

Anhand der **Gradientenlänge** (length of gradient) der DCA (Einheit SD, average standard deviation of species turnover) kann beurteilt werden, wie groß der Abstand zwischen den entferntest liegenden Aufnahmewerten auf jeder untersuchten Achse liegen. Eine Gradientenlänge von 8 SD gibt an, dass innerhalb des Datensatzes ein zweifacher Artenwechsel erfolgt – d. h. der untersuchte Gradient ist genügend lang für unimodale Reaktionen der Arten. Die Gradientenlänge sollte für eine sinnvolle Anwendung der DCA mindestens 4 SD (einfacher Artenwechsel) betragen.

Ein weiterer wichtiger Wert ist **Eigenvalue** (Eigenwert) der einzelnen Achsen. Er liegt zwischen 0 und 1, wobei ein Wert über 0,5 anzeigt, dass ein hoher Teil der Variabilität des Datensatzes mit dieser Achse erklärt wird. Die erste Achse hat den höchsten eigenvalue, die weiteren, auch durch das „detrending“ einen niedrigeren und sollen eher vorsichtig interpretiert werden.



**Abbildung 9:** DCA-Diagramm des Korralpen-Datensatzes (DRESCHER et al. 2007) mit drei DCA-Achsen, erstellt in JUICE mit R 2.10 unter Anwendung der JUICE-Schnittstelle. Der Ellenberg Zeigerwert der Aufnahmen ist als Vektor (moisture, grün) eingehendet (Gradientenlänge Achse 1: 8,3 SD, Achse 2: 4,0 SC, Achse 3: 3,0 SD), Eigenvalues Achse 1 0,7732, Achse 2: 0,4338, Achse 3: 0,0601



Wie aus der Unterschrift zur **Abbildung 9** ersichtlich, ist die erste Achse (Gradientenlänge, Eigenvalue) gut zu interpretieren, sie verläuft annähernd parallel zum Vektor für die Feuchtigkeit, d. h. das ist der offenbar wesentlichste wirkende Faktor. Achse 2 bildet die Variabilität des Datensatzes schon viel schlechter ab und die Achse 3 ist nicht interpretierbar und sollte in diesem Beispiel eigentlich auch gar nicht abgebildet werden.

## 2.8.6 Multivariate Statistik: PCA und NMDS mit JUICE und R

Ebenfalls mit der Funktion

**Analyses → Ordinations → R-Project → PCA Analysis ... NMDS Analysis**

können in JUICE mit dem assoziierten R-Paket „Vegan“ mit dem Datensatz entweder eine PCA (=principal component analysis, Hauptkomponentenanalyse) oder eine NMDS (=nichtmetrische multidimensionale Skalierung) errechnet werden, dies vor allem bei für eine DCA zu geringe Gradientenlänge < 4. Leider ist es bei Anwendung dieses Paketes bei der PCA nicht möglich, die „eigenvalue“ auszulesen, sodass man die Qualität der Ergebnisse nicht interpretieren kann. Die NMDS ist ein rangbasiertes Verfahren, welches unabhängig von der Gradientenlänge angewendet werden kann. Für die statistischen Grundlagen vgl. LEYER & WESCHE (2008: 142–152). Die Qualität einer NMDS-Anwendung auf einen Datensatz kann aus dem sogenannten „Stresswert“ abgelesen werden (**Show Summary**), wenn dieser > 20 ist, gelten die Ergebnisse als wenig aussagekräftig.

## 2.9 Verarbeiten von Pflanzen-Eigenschaften in Vegetationstabellen

### 2.9.1 Fragestellung

Das Haupteinsatzgebiet vegetationskundlicher Daten ist zweifellos die Charakterisierung von Landschaften, Schutzgebieten oder Biotopen mit Hilfe von Klassifikations- und Ordinationsmethoden. Oftmals gehen die Fragen aber weiter in Richtung einer Beurteilung und Bewertung bestimmter Zustände von Vegetation. Ist eine Vegetationseinheit schutzwürdig? War eine bestimmte Management- oder Renaturierungsmaßnahme erfolgreich? Wie hoch ist der Anteil an Rote Liste Arten, an Neophyten, an Gräsern, an Naturschutz-Zielarten, an Endemiten usw. in einer Vegetationseinheit? Hier bietet JUICE geniale Möglichkeiten.

### 2.9.2 Datenvorbereitung

JUICE speichert von den Arten nur Namen, Schichtangaben und Abundanzwerte; und nach erfolgreicher Initialisierung auch die ökologischen Zeigewerte von Ellenberg (siehe oben Kapitel 2.7). Alle anderen Informationen müssen wir zusätzlich aus externen Quellen dazu laden. Um die Daten in JUICE in die Species data der Tabelle laden zu können, wird eine Textdatei im fixed length-format benötigt, die wir uns vorher herstellen müssen. Für die Vorbereitung dieser Textdatei bietet sich EXCEL an.

Informationen zu Pflanzeigenschaften bekommt man aus Florenwerken, plant trait Datenbanken, z.B. aus der BIOLFLOR-Datenbank (KLOTZ et al. 2002), aus Roten Listen oder aus verschiedenen anderen **Quellen**. Letztendlich können je nach Fragestellung auch eigene Artengruppen (z. B. Störzeiger, Zielarten) gebildet werden. Die Daten sollten allerdings digital vorliegen, d. h. dass Informationen aus Buchquellen erst mühselig abgetippt oder anders eingelesen werden müssen. Informationen aus elektronischen Quellen ersparen einem zwar das Einlesen, nicht aber die Anpassung der Artnamen und Abkürzungen der Plant traits für die Weiterverarbeitung in JUICE. Die Artnamen müssen 100% mit den Namen in JUICE übereinstimmen. Das kann man mit der Zeit optimieren, indem man immer wieder schaut, bei welchen Arten das Einlesen der Daten nicht funktioniert und deren Schreibweise dann in EXCEL korrigiert. Um die aktuelle Artenliste der Tabelle in EXCEL zu laden, benutzen wir den Befehl **File → Export → Export Species Data**.

Alle Plant traits sollten in Zahlen überführt werden. Kategorische Werte (z. B. Winterannuelle, Sommerannuelle; oder Windverbreitung, Tierverbreitung, Selbstverbreitung) müssen in binomial-Werte (0 und 1, entspricht trait vorhanden oder nicht vorhanden), oder in Ordinal-Werte wie bei den Zeigerwerten (z. B. Blattfläche kein, mittel groß; ein Beispiel zu Rote-Liste-Kategorien folgt in Kap. 2.9.3). Sind alle Werte, die Sie brauchen, vorhanden, dann stellen Sie in EXCEL über **Format → Spaltenbreite optimieren** die optimale Breite der einzelnen Spalten ein und speichern die Arbeitsmappe als Formatierter Text „\*.prn“ ab. Dies erzeugt eine fixed length-Datei. Anschließend im Windows-Explorer die Endung \*.prn durch \*.txt ersetzen (die dabei ausgesprochene Warnung ignorieren) und schon ist die Datei fertig. Die Datei kann ruhig mehrere traits in festen Spalten enthalten, man sollte aber vorher die genauen Spaltenbreiten kennen, also wissen, wie viele Zeichen die Artspalte hat und bei welchem Zeichen innerhalb der Zeile der trait anfängt und aufhört, den Sie gerade laden wollen. JUICE bietet dafür eine Skala an, die aber eher als grobe Richtlinie brauchbar ist.

### 2.9.3 Einlesen von externen Species Data

Jetzt kommt noch das Laden in JUICE. Wir nehmen mal an, wir wollen den Anteil an Rote Liste Arten in unseren Aufnahmen ermitteln. Dazu könnten wir beispielsweise den sogenannten „Gefährdungsinhalt“ (BERG et al. 2004) berechnen, indem wir den RL-Kategorien bestimmte Zahlenwerte zuordnen und diese für die einzelnen Gesellschaften aufsummieren. Zahlenwerte könnten z. B. sein: Ungefährdet = 0, potentiell gefährdet und schwach gefährdet (Kat. 4

und 3) = 1, stark gefährdet (Kat. 2) = 2 und vom Aussterben bedroht (Kat. 0) = 3. Diese Werte könnten dann noch mit der Deckung oder, bei Stetigkeitstabellen, mit der Stetigkeit verrechnet werden.

Mit dem Befehl **Species → Species data → External species data** laden wir die txt-Datei und geben die Breite der Artspalte (z. B. 1 bis 32) und die Breite der Datenspalte RL (z. B. 33–35) an. Dann auf **Continue** drücken und noch einige Fragen beantworten – fertig! Falls es nicht ganz geklappt hat, kann man den Vorgang mit anderen Spaltenbreiten wiederholen. Über **Sort species by species data** können Sie die Arten mit gleichen Einträgen zusammenstellen. Außerdem sieht man dann, welche Arten keinen Eintrag haben (z. B. Moose oder Genus sp.-Angaben) und kann dies ggf. noch korrigieren bzw. diese von weiteren Berechnungen ausschließen (dies ist wichtig bei Mittelwerten, da sonst Arten ohne Einträge als 0 gerechnet werden!).

**Tipp für Leben:** Man kann sich im Laufe des Lebens eine EXCEL-Datenbank zusammenstellen, die möglichst viele functional plant traits bereithält und die möglichst gut mit der Nomenklatur in TURBOVEG und JUICE harmonisiert.

## 2.9.4 Verarbeiten von (externen) Species Data

Das eigentliche Arbeiten beginnt nun, denn wir wollen aus den eingelesenen Werten, ähnlich wie bei den Zeigerwerten, auf unsere Aufnahmen rückschließen und die Ergebnisse in den Kopfdaten abspeichern. In unserem Fall wollen wir also den Gefährdungsinhalt als Summe der Gefährdungswerte berechnen. Alle Arten mit einem positiven Zahleneintrag bekommen jetzt die gleiche Farbe (z. B. grün).

**Head → Store values to Short Header → Sum of GREEN Species Data**

Dieser Befehl schreibt die Summe der Gefährdungswerte für jede Aufnahme in die short headers. Möchte ich beispielsweise die Bedeckungswerte oder den prozentualen Anteil an Rote Liste Arten in den Aufnahmen haben, kommen die nachfolgenden Befehle zum Einsatz:

**Head → Store values to Short Header → Percentage Cover of GREEN Species**

**Head → Store values to Short Header → Percentage Number of GREEN Species**

Mit **Head → Short Header Averages** können Sie sich die Durchschnittswerte der einzelnen Cluster anzeigen lassen. Mit **Head → Add Short Header to Header Data** sichern Sie die Daten dauerhaft in den Kopfdaten und können Sie dann weiteren Auswertungen zuführen.

Die hier exemplarisch gezeigte Möglichkeit des Einlesens externer Daten bietet sehr vielfältige Auswertungen. Bei Statusangaben wie z. B. bei Neophyten oder nach der Frage immergrüner Arten empfiehlt sich der Befehl **Store values to Short Header → (Percentage) Number of GREEN Species**, bei numerischen Daten, wie z. B. Hemerobiewerte oder Strategiewerte, können Sie auch Mittelwerte (**Mean of GREEN Species data**) berechnen lassen.

## 2.10 Biodiversitätsuntersuchungen in Vegetationstabellen

### 2.10.1 Vorbereitung einer Tabelle für Biodiversitätsuntersuchungen

Jetzt wollen wir noch einige Biodiversitätsfragestellungen mit JUICE analysieren. Da gelten etwas andere Gesetze wie in der Pflanzensoziologie, die Tabelle muss für Biodiversitätsfragestellungen besonders präpariert werden. Will man nur einfach ein paar Artenzahlen berechnen, kann man die Methode von Kap. 3.3.3 verwenden, für umfangreichere Fragestellungen empfiehlt es sich, eine eigens zu diesem Zweck vorbereitete Tabelle zu verwenden. Deshalb zuerst die Tabelle unter neuem Namen abspeichern! Für unsere Zwecke wechseln wir zu einer anderen Tabelle Biodiv\_01.wct. Diese enthält die gleichen vier Buchenwald-Vegetationseinheiten, aber etwas mehr Aufnahmen und bessere Flächengrößen-Einträge. Da wir wissen wollen, wie sich ausgewählte Biodiversitätsparameter in unseren 4 Waldgesellschaften unterscheiden, speichern wir unsere Ordnung der Aufnahmen, d.h. unsere 4 Cluster ab. Dazu schreiben wir mit dem Befehl **Head → Group Number** die Cluster-Nummer in die **short headers** und speichern diese Werte in den Kopfdaten als „Cluster“ ab. Damit können wir die aktuelle Ordnung jederzeit wieder herstellen. Jetzt setzen wir den **Head** wieder auf **Original Number** und löschen alle Separatoren und Farben.

#### 2.10.1.1 Vergleichbare Flächengrößen!

Biodiversität ist ja stark von der Flächengröße abhängig. Deshalb soll man bei Biodiversitätsuntersuchungen immer nur Aufnahmen mit exakten Flächenangaben nutzen und nur gleich oder ähnlich große Flächen vergleichen. Vegetationskundliche Daten sind aber oft heterogen, auch was die Wahl der Aufnahmeflächen anbetrifft. Bevor ich also Biodiversitätsfragestellungen angehen kann, muss ich mir die Flächengrößen ansehen.

**Head → Store values to Short Header → Header Data → Relevé area (m<sup>2</sup>)**

Hier gibt es normalerweise stark schwankende Aufnahmeflächengrößen, bisweilen sogar Lücken. Der ausgewählte Datensatz ist allerdings schon „vorgereinigt“. Mit

**Sorting → Sort Short Headers → All + Numerically + In descending order**

können wir die Werte der Größe nach sortieren. Mit **Short Header Averages** schauen wir uns den Durchschnittswert an. Der Durchschnittswert liegt bei rund 400 m<sup>2</sup>. Sinnvoll wäre es nun, Aufnahmen, die einen bestimmten Wert unter- oder überschreiten, zu löschen, wobei wir eine Schwankungsbreite beispielsweise von 300 m<sup>2</sup> bis 500 m<sup>2</sup> als Toleranz ansehen können. Die restlichen Aufnahmen färben wir mit einer Farbe (z. B. grau) und löschen sie mit **Relevés → Delete GRAY relevés**. Mit **Head → Store values to Short Header → Header Data → Cluster** holen wir uns die gerade gespeicherten Cluster-Nummern wieder in die **Short header** und sortieren diese aufsteigend, so dass wir wieder die alte Ordnung haben. Der Befehl **Separators → Make Separators → Within Short Headers** trennt die Cluster wieder automatisch durch Separatoren.

### 2.10.1.2 Beseitigen der Schichtangaben

Als nächstes müssen wir die Artenspalte für Biodiversitätsuntersuchungen vorbereiten. Das JUICE-Handbuch behauptet, dass bei Zähl-Befehlen wie **Number of RED Species** Arten, die in mehreren Schichten vorkommen, als eine gezählt werden. Das stimmt leider nicht. Deshalb müssen wir die Schichtangaben aus dem Datensatz entfernen. Zum Glück bietet JUICE hier den Befehl:

**Species → Merge all same Species Names → Within the Dataset**

Es öffnet sich ein Fenster, in dem die fusionierten Arten angezeigt werden. Jetzt könnten wir noch einige Arten von der Fusion ausschließen, mit **Continue** wird der Befehl dann ausgeführt. Mit dem Befehl **Species → Highlight (S), (D) and (r) species** können wir die fusionierten Arten in der Tabelle sichtbar machen oder nicht.

Nun ist jede Art in dem Datensatz nur einmal vertreten. Will ich allerdings Biodiversitäts-Fragestellungen innerhalb der Vegetationsschichten bearbeiten, z. B. Phytodiversität der Baumschicht, muss dieser Schritt entfallen und ich muss dann genau aufpassen, dass sich die Berechnungen nur auf diese Schicht beziehen.

### 2.10.1.3 Gleiche Aufnahmemengen pro Cluster!

Vom Charakter des Samplings her haben wir es in unserer Tabelle mit einer Species-Sampling-Relationship zu tun. Mit so einem Design können wir nicht nur Aussagen zur Alpha-Diversität, sondern auch zur Beta- und Gamma-Diversität tätigen. Voraussetzung ist, dass wir das Prinzip, immer nur gleich große Flächen vergleichen zu wollen, auch bei den Clustern beherrschen. Wir dürfen also nur Cluster mit (ca.) der gleichen Zahl von Aufnahmen vergleichen. Da die Tabelle mit zahlreichen Aufnahmen (=Stichproben) bestückt ist, können wir es uns leisten, durch eine Neuauswahl (resampling) die Aufnahmezahl der Cluster anzugleichen, vorausgesetzt, wir lassen dabei den Zufall entscheiden.

Wir laden dazu die in den Kopfdaten abgespeicherte Cluster-Nummer wieder in die Short Header, sortieren sie und setzen wieder Cluster Separatoren (dies geht automatisch mit dem Befehl **Separators → Make Separators → Within Short Headers**). In der **Synoptic Table**-Ansicht sehen wir die Zahl der Aufnahmen pro Cluster. Das kleinste Cluster hat immer noch fast 50 Aufnahmen; wir beschließen daher, alle Cluster auf 40 Aufnahmen runter zu kürzen. Dazu gehen wir zurück auf **Full Table**.

**Relevés → Random Relevé Selection → Select Relevés with Manual Switches**

Jetzt stellen wir für die **Separator Group 1** (1. Cluster) unter **No. of Selected Relevés 40** ein und wählen eine Farbe. Dies wiederholen wir jetzt für die 3 anderen Gruppen jeweils mit einer anderen Farbe. Dann wählen wir **Cancel**, stellen Weiß als aktuelle Farbe ein und verschieben die weißen Aufnahmen nach rechts. Die Separatoren können sich dabei verschoben haben, mit **Separators → Make Separators → Within Colours** stellen wir die Separator-Ordnung wieder her. In der **Synoptic Table**-Ansicht müssten die ersten 4 Cluster jetzt 40 Aufnahmen enthalten. Wir speichern das Ganze ab und löschen die weißen Aufnahmen mit **Relevés → Delete WHITE relevés**.

Für noch elegantere Lösungen des Resampling siehe Kapitel 3.1.6.

## 2.10.2 Jetzt kann es losgehen: Alpha- und Gamma-Diversität

Jetzt können wir sicher sein, dass jedes Taxon nur einmal in der Tabelle steht, alle Flächengrößen ähnlich sind und unsere 4 Cluster die gleiche Anzahl von Aufnahmen haben. Nun können wir mit der Analyse beginnen! Es gibt 2 Typen von Ergebnissen: Zum einen jene, die die Einzelaufnahmen betreffen; diese speichern wir via **short header** in den Kopfdaten ab. Zum anderen gibt es auch Ergebnisse, welche nur die vier Cluster betreffen oder die sich nicht in die Short Headers laden lassen, diese speichern wir in einer geöffneten EXCEL-Tabelle ab.

Beginnen wir mit der Artenzahl pro Aufnahme. Diese schreiben wir mit **Head → Store values to Short Header → Number of BLACK Species** in die Short Header und speichern sie anschließend in den Kopfdaten. Mit dem schon vertrauten Befehl **Head → Short Header Averages** können wir uns sowohl die mittleren Artenzahlen als auch die Standardabweichungen dieser Mittelwerte anzeigen lassen. JUICE speichert die Werte beim Anzeigen gleichzeitig in die Zwischenablage (clipboard), so dass wir nur nach jeder Anzeige in unsere EXCEL-Tabelle wechseln und mit **Strg-V** die Werte hineinkopieren brauchen. Damit wir nicht den Überblick verlieren, beschriften wir gleich die EXCEL-Tabelle und ordnen sie sinnvoll.

Ein ähnlicher Befehl **Head → Store values to Short Header → Diversity Indices / Evenness** schreibt einige Indices in die Short Header. Wir versuchen es mal mit allen 4 und schreiben sie in die Kopfdaten. **Short Header Averages** stellt die Durchschnittswerte für die EXCEL-Datei zur Verfügung, die wir ebenfalls in unsere EXCEL-Tabelle kopieren.

Als nächstes interessiert uns der Species pool (Gamma-Diversität) der einzelnen Cluster. Mit **Analysis → Diversity → Components of Diversity** gibt JUICE nochmal die durchschnittliche Alpha-Diversität pro Cluster aus, gleichzeitig aber auch den Species pool pro Cluster und die Gesamtartenzahl der Tabelle. Beides übertragen wir in die EXCEL-Tabelle. Der gleiche Befehl öffnet auch ein Rarefaction-Feld. Hier stellt JUICE in einem iterativen Verfahren kumulative Arten-Akkumulationskurve (Species Accumulation Curves) her. Diese beleuchten grafisch das Verhältnis von Alpha- zu Gamma-Diversität, die Kurven haben für jede Pflanzengesellschaft einen charakteristischen Verlauf. Je flacher der Kurvenverlauf ist, desto geringer ist die Evenness und/oder desto heterogener ist die Pflanzengesellschaft. Wir stellen einmal die Cluster-Nummer und die Farbe des Clusters mit der geringsten Gamma-Diversität ein und klicken auf Rarefaction. In EXCEL speichern wir die Daten ab (ggf. müssen wir Stop drücken, um nach EXCEL wechseln zu können) und wiederholen den Vorgang anschließend nochmal für das Cluster mit der höchsten Gamma-Diversität. Jetzt können wir beide Kurvenverläufe vergleichen und anschließend die Daten der zweiten Rarefaction auch in EXCEL speichern. Die Daten kann man verwenden, um veröffentlichungswürdige Grafiken der Kurven herzustellen.

### 2.10.3 Beta-Diversität

Nicht ganz so leicht zu verstehen ist die Beta-Diversität, die den Unterschied in der Artenzusammensetzung zweier Probestellen (Diversität = Verschiedenheit) bemisst. Komponenten dieses Vergleichs sind gemeinsame Arten und exklusive Arten, die jeweils nur in einer der beiden Flächen vorkommen. Aus diesem Verhältnis ergibt sich Ähnlichkeit und Unterschied. Ein typisches Maß für Beta-Diversität ist der Species turn over, dem wir schon im Kapitel 2.8.5 in Form der „Gradientenlänge“ begegnet sind.

**Analysis → Diversity → Total Inertia/ED/Beta Diversity within relevé groups**

Hier stellen wir bei **Cluster 1** ein und bei **Number of Relevés 40** und drücken **Calculate**. Wenn wir Glück haben, sind die Werte im Clipboard gespeichert, das funktioniert aber nicht immer. Ansonsten notieren wir uns die Werte für Whittakers Beta und Sørensen dissimilarity und führen das für alle 4 Cluster durch.

Am Ende können wir noch fragen, wie unähnlich in Hinblick auf die Artenzusammensetzung unsere vier Cluster sind. Dafür gibt es die Analysis of Similarity, (ANOSIM), welche ähnlich wie eine ANOVA die Beta-Diversität innerhalb der Cluster mit jener zwischen den Clustern vergleicht. Dabei wird ein Prozentwert der relativen Unähnlichkeit errechnet, 100 % bedeutet also, dass die Cluster eine komplett verschiedene Artenzusammensetzung haben. Die Signifikanz wird durch einen Mann-Whitney-Test ermittelt.

**Analysis → Distance Measures → Distance between relevé groups**

Wir stellen den Sørensen Index ein, bei **Structure of the dataset of distances** muss die rechte Möglichkeit eingestellt sein, und wir wählen **All distances**, solange die Available Combinations noch unter 10000 sind. Mit **Calculation** starten wir den Rechengang, es wird ein Percentage distance-Wert und ein p-level ausgegeben. Die Grafik visualisiert den Unterschied zwischen den Aufnahmen: Ergibt sich eine Linie entlang der Winkelhalbierenden, so sind die Cluster recht ähnlich und die Artenzusammensetzung geht eher kontinuierlich ineinander über (die Grenze wird dann mit Hilfe von Differenzial- bzw. Indikatorarten gezogen). Je mehr die Linie sigmoidal verläuft, desto höher ist der Unterschied und desto klarer lässt sich eine Grenze zwischen den beiden Einheiten ziehen.

## 3. Weitere Tipps und Tricks zum Arbeiten mit JUICE

### 3.1 Datenvorbereitung

#### 3.1.1 Import data

Mit TURBOVEG 3 ändert sich das Austauschformat erneut, jetzt werden ähnlich wie in R fix strukturierte Textdateien zum Datenaustausch verwendet. Zum Einlesen von TV3-Dateien in JUICE dient dann der Befehl

**File → Import → Table → from Simple Text File (Database Records)**

#### 3.1.2 Import von Tabellen aus EXCEL nach JUICE

Eine in EXCEL (oder in einem anderen, mit EXCEL kompatiblen Format, z.B. einfache TXT-Dateien) gespeicherte Vegetationstabelle können Sie leicht in JUICE einlesen, wenn Sie den genauen Aufbau der Tabelle beachten. Dies müssen Sie vor dem Import entsprechend in EXCEL vorbereiten. Prinzipiell müssen die Daten so vorbereitet werden wie zum Einlesen in TURBOVEG (siehe oben Kap. 1.2.7), mit dem Unterschied, dass über der Arttabelle eine Leerzeile eingefügt

werden muss, in deren Feld links oben eine Überschrift (Name der Tabelle) eingefügt werden muss, und die Schichtenangabe muss als Zahl (1-9) erfolgen, nicht als Buchstabencode (s. Tabelle 1).

JUICE macht keinen Abgleich der Artenliste (nimmt also Schreibfehler mit in die Tabelle) und ist auch unempfindlich, was die Feldformate anbetrifft. Probleme gibt es lediglich mit Datumsformaten, die sollten wie in TURBOVEG auf JJJJMMTT umgestellt werden Ansonsten kann alles Mögliche importiert werden, z. B. auch Tabellen mit zoologischen Daten.

Nun zum eigentlichen Procedere. Zuerst markieren Sie in EXCEL die gesamte Arten-Tabelle und drücken Strg-C. Damit wird die Tabelle in die Zwischenablage (clipboard) gespeichert. Die JUICE-Befehle lauten nun:

**Import → Table → from Spreadsheet file → from Clipboard as Spreadsheet.**

Bei der Frage nach den Separatoren "Semicolon" eingeben, die Frage nach der 2. Spalte als Schichtangabe bestätigen. Am Ende des Dialogs wird die Anzahl von Reihen und Spalten angezeigt, hier können Sie bereits sehen, ob JUICE die Tabelle richtig erkannt hat. Mit Next gibt es eine Tabellenansicht, erneut Next führt zum Abgleich der Cover-Codes, welche JUICE wie auch TURBOVEG intern in Bedeckungs-Prozente umrechnet. Wenn alles stimmt, kann der Vorgang mit Finish abgeschlossen werden.

Ist der Import geglückt, geht man in die transponierten Kopfdaten, markiert diese und speichert jetzt die header-Tabelle im clipboard durch Drücken von Strg-C. Nach

**File → Import → Header Data → from Clipboard**

müssten nun die Kopfdaten richtig eingelesen sein. Sofort stichprobenartig mit der EXCEL-Tabelle vergleichen und speichern!

### 3.1.3 Zwei Tabellen zusammenführen

JUICE bietet den schönen Befehl **File → Append.** zum Zusammenfügen zweier JUICE-Tabellen. Leider nimmt dieser Befehl die Kopfdaten der anzuhängenden Tabelle nicht mit. Um Kopfdaten der gesamten Tabelle zu bekommen, gehen wir wie folgt vor: Öffnen Sie die Tabelle, die Sie anhängen wollen und löschen Sie alle Aufnahme-Farben. Wählen Sie den Befehl **Datei → Export → Export Headers from white relevés.** Markieren Sie alle Header-Felder im folgenden Fenster und drücken **Continue.** Sie erhalten eine txt-Datei, für die JUICE den Namen hlava.txt vorschlägt. Benennen Sie sie in hlava2.txt um vor dem Speichern. Öffnen Sie nun die Tabelle, an die Sie die andere anhängen wollen und wiederholen Sie den Vorgang. Sichern Sie jetzt die Datei unter dem Namen hlava.txt.

Öffnen Sie EXCEL und laden Sie die hlava.txt-Datei in ein und hlava2.txt in ein anderes Tableau (semicolon delimited). Wie Sie sehen, sind die Kopfdaten transponiert, die Aufnahmeummern verlaufen in der ersten Spalte abwärts. Überprüfen Sie die letzte Spalte der beiden Kopf Tabellen: Wenn nämlich einige Datenfelder ein Semikolon enthalten (oft z. B. die Remarks), wird die Datei nicht korrekt eingelesen und muss von Hand korrigiert werden! Wenn beide Dateien ok sind, kopieren Sie die Daten der hlava2 unter das EXCEL-Tableau von hlava (auf die Feldnamen achten, die Reihenfolge muss nicht übereinstimmen!). Wenn die hlava2-Tabelle Felder enthält, die in der ersten Tabelle fehlen, fügen Sie eine neue Spalte unter diesem Namen ein. Kopfdaten, die Sie mit Sicherheit nicht brauchen, können auch weggelassen oder gelöscht werden.

Nachdem beide Dateien kombiniert sind, kopieren Sie die gesamte Kopfdaten-Tabelle mit **strg-c** in die Zwischenablage. Wechseln Sie in JUICE und hängen Sie die zweite Datei mit dem Befehl **Append** an. Die Kopfdaten der 2. Tabelle fehlen. Jetzt wählen Sie **File → Import → header data → from clipboard.** Überprüfen Sie das Ergebnis sorgfältig!! Vergleichen Sie auch nochmal das Kapitel 2.4.3.

### 3.1.4 Hunting duplicates

Um Duplikate zu finden, verwenden wir den Befehl **Relevés → Hunting Duplicates.** Dieser Vorgang dauert relativ lange Zeit, wenn die Tabelle viele Aufnahmen enthält. Im Ergebnis werden Duplikate werden als "DUPLIC" in den **Short Headers** angezeigt, ohne die Reihenfolge zu verändern, d. h. ohne die Doubletten-Paare anzuzeigen. Um diese zu finden, färben Sie mit der Funktion **head → short header selection (alphabetically, Parameters = DUPLIC),** die Doubletten in eine Farbe ein, sortieren sie in ein Cluster zusammen und lassen über dieses Cluster TWINSPAN mit einer hohen Teilungsrate laufen. Nun sollten sich die Duplikate nebeneinander einfinden.

### 3.1.5 Species merging

Ein wertvolles Feature in den Zeiten der Big Data ist die Vereinheitlichung der Nomenklatur. Dies macht man aber gewöhnlich am Anfang der Arbeit, während man am Ende gerne wüsste, welche Arten man denn nun "gemergt" hat. JUICE bietet eine Funktion, diese Arten in einer Datei aufzulisten.

**Analysis → Expert system → Create Expert file (P1 and P2)**

Dieser Befehl erzeugt eine Datei mit der Dateierweiterung \*.esy, die man abspeichern kann (zum Anschauen die Dateierweiterung notfalls in \*.txt umwandeln). "Copy to Clipboard" speichert die Ergebnisse aber auch ins Clipboard zur

Übertragung in Word. Section 1 zeigt alle Species aggregations an, die Sectionen 2 bis 4 spielen nur eine Rolle im Zusammenhang mit der Nutzung der Cocktail-Gruppen und Cocktail-Formeln.

„Gemergte“ Arten kann man jederzeit zurückholen. Mit **Species → Undelete Species** wird eine Tabelle aller Namen angezeigt, die nach dem Merge-Vorgang von der Tabelle verschwunden sind (hidden species). Mit **Undelete** können die Arten einzeln zurückgeholt werden.

### 3.1.6 Daten-Stratifikation und Resampling

Oftmals sind vegetationskundliche Daten inhomogen verteilt. Bestimmte Gegenden, wo Vegetationskundler sitzen, oder artenreiche oder besonders interessante Habitats, sind mit zu vielen Aufnahmen belegt, abgelegene Gegenden oder langweilige Habitats mit zu wenigen. Dies kann bei der Auswertung zu falschen Ergebnissen führen und muss deshalb vorher durch eine sinnvolle Reduzierung der Daten ausgeglichen werden. Das Ordnen von Stichproben in verschiedene Sachverhalte (Straten) wird Stratifikation genannt, die mengenmäßige Vereinheitlichung dieser Stichproben innerhalb der Straten durch Auswahl einer Teilmenge der Stichproben nennt man resampling. Wichtige Abhandlungen zu dieser Methode sind KNOLLOVA et al. (2005), HAVEMAN & JANSSEN (2008) sowie LENGYEL et al. (2011). JUICE bietet hierfür elegante Möglichkeiten.

#### 3.1.6.1 Geographical stratification

Bei dieser machtvollen Art der Stratifizierung werden die Daten innerhalb eines fiktiven Rasters homogen verteilt. Sie funktioniert nur mit vollständigen Angaben der Koordinaten im Digitalformat, die in den Feldern `deg_lat` und `deg_lon` vorliegen müssen. Mit Hilfe dieser Daten können wir ein Raster über unsere Daten legen, und das Programm weist als erstes jede Aufnahme einem Rasterfeld zu. Wir löschen zuerst alle Farben und Separatoren.

**Head → Resampling → Geographical Resampling**

Hier wird das Geographical grid spacing abgefragt, also ein virtuelles Raster über die Daten gelegt. Je nach Dichte des Datensatzes kann man größere oder kleinere Werte angeben, so dass man ein brauchbares Raster mit einer guten Anzahl Aufnahmen enthält. Dies kann man ausprobieren. Wichtig ist für unsere Zwecke, das Feld **Save geogr. grid code to short headers** anzuklicken. Das Feld **Within groups defined by separators** muss für unsere Zwecke ausgeklickt werden, die Felder in **Preferably retain relevés...** sollten ebenfalls lieber ausgeklickt werden; sie stellen hohe Ansprüche an die Datenqualität. Nach **Continue** schreibt JUICE die Grid-Codes in die **short header**. Nach **Sort short header** und **Make separators between short header** sehen wir nun die Verteilung unserer Aufnahmen auf die einzelnen geographischen Raster.

Jetzt geht das eigentliche resampling los. Mit **Head → Resampling → Random and systematic Resampling** können wir mit **Equal number randomly from each group** und **Max. number of sel. relevés in each relevé group** einige Eckdaten setzen, eine entsprechende Farbe einsetzen und mit **select** die Aufnahmen selektieren.

#### 3.1.6.2 Stratifikation nach 2 Kriterien

Außer geographischer Stratifikation kann man auch nach allerlei anderen Kriterien stratifizieren, z. B. Extremwerte ausschließen. Wir löschen also die Aufnahme-Farbe in der Tabelle und laden nun einen zweiten Wert in die Kopfdaten, z. B. slope, denn wir wollen allzu steile Aufnahmeflächen von der Untersuchung ausschließen, bevor wir die endgültige geographische Stratifikation vornehmen. Nachdem der slope-Wert in die **short headers** eingeladen ist, benutzen wir den Befehl **Sorting → Sort short header** und klicken das Feld **within groups** an. Nun stehen alle stark geneigten Flächen jeder Rasterzelle ganz rechts. Wieder bei **Head → Resampling → Random and systematic Resampling** nutzen wir nun das Feld **Equal number from last relevé of each group** und geben bei **Max. number 1** ein. Nun markiert er in jeder Zelle die Aufnahme mit dem größten slope-Wert. Das funktioniert auch mit dem kleinsten Wert, so dass man auch kleine Extremwerte ausschließen kann. Die Funktion ist auch als einfache Selektionsmöglichkeit zu verwenden. Wenn ich z.B. in einer klassifizierten Tabelle die jeweiligen Aufnahmen mit viel RL-Arten oder Neophyten vergleichen will mit denen, die davon wenig haben, kann ich analog vorgehen.

### 3.1.7 Interpolation fehlender Messdaten anhand der Artenzusammensetzung

Der Zusammenhang zwischen Zeigerwerten und realen Messdaten wurde ausgiebig nachgewiesen. Das kann man nutzen, um fehlende Messwerte zu interpolieren. Zuerst bringt man die Messreihe in die Short Header und färbt die Aufnahmen mit fehlenden Werten z. B. rot an. Dann nutzt man den Befehl

**Head → Imputation of Environmental Variable**

Bei **Data transformation** hat sich **Power 0.5** bewährt, als Distanzmaß der Bray-Curtis-index. Die **Maximal accepted distance** hängt natürlich von der Homogenität des Datensatzes ab, je kleiner die Distanz,

desto genauer wird die Interpolation, aber desto mehr Werte können gar nicht interpoliert werden (angezeigt als unknown), weil entsprechend ähnliche Aufnahmen fehlen. Beginnen Sie mit 0.5 und setzen Sie die Distanz schrittweise hoch, bis die meisten Werte berechnet werden.

Das Ergebnis wird angezeigt und im Clipboard gespeichert, der Transfer in die **Header Data** der Tabelle erfolgt über EXCEL (siehe Kapitel 2.4.3).

## 3.2 Klassifikation

### 3.2.1 Durchschnittliche Bedeckungswerte für einzelne Aufnahme-Cluster

Beim Suchen guter Differenzialarten wird empfohlen, mehrere Prüfwerte nebeneinander zu stellen: die fidelity, die constancy ratio (Stetigkeit doppelt so hoch als die benachbarter Columnen) und die mittlere Bedeckung (average cover). Während JUICE die ersten beiden Werte in den synoptic table einfach pro Columne anzeigt, muss man den dritten Wert für jede Columne einzeln berechnen.

Die entsprechende Columne wird mit einer eigenen Farbe markiert (z. B. rot). Nun braucht es nur noch den Befehl **Species** → **Species data** → **cover values** → **Red relevés** → **Average non zero cover**. Die Werte können nun über **File** → **Export** → **Export species data** in eine Datei und das Clipboard geschrieben werden und in EXCEL abgespeichert werden. Geht man Columne für Columne vor, kann man sich so eine Stetigkeitstabelle mit den entsprechenden Average cover values zusammenstellen.

### 3.2.2 Güte einer Klassifikation mit der Funktion Silhouette beurteilen

Silhouette ermöglicht eine Visualisierung verschiedener Distanzmaße innerhalb eines Clusters. Damit kann die Homogenität eines Clusters in Hinblick auf das gewählte Distanzmaß abgeschätzt werden. Gleich lange Balken zeigen sehr homogene Cluster, bei unterschiedlich langen Balken sinkt die Homogenität. Negativ ist auch die Nähe zum Gesamt-Zentroid des Datensatzes (der 0.0-Wert auf der Skala) zu werten. Die **Abbildung 10** gibt Beispiele für unterschiedlich homogene Cluster. Die Funktion ist R-basiert und läuft nur mit funktionierender Verknüpfung von JUICE und R2.10.1.

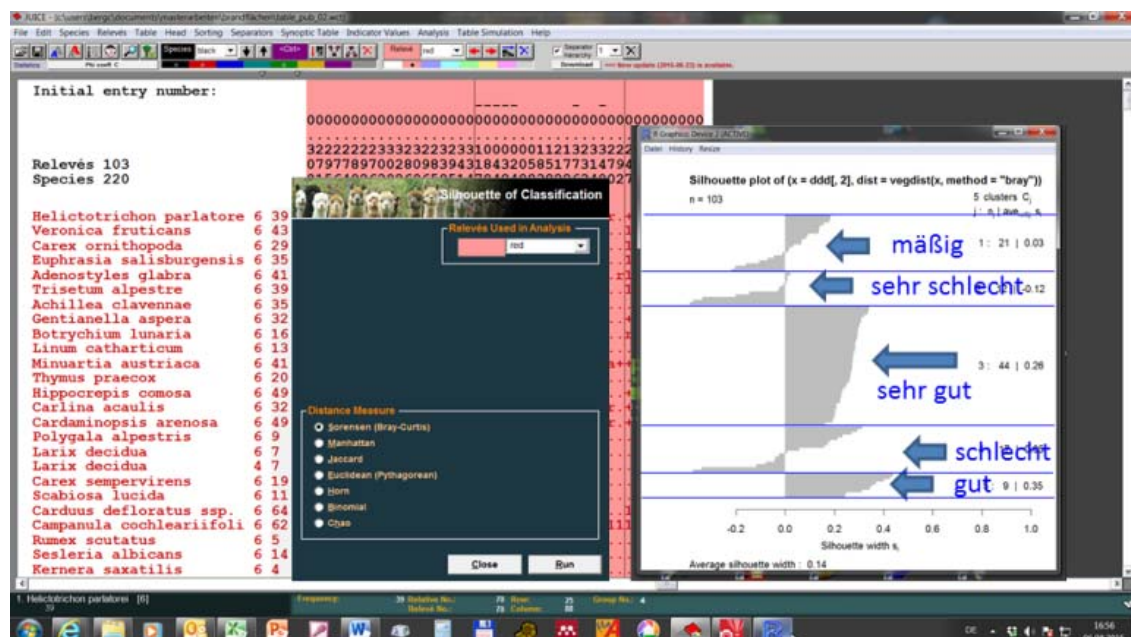


Abbildung 10: Ergebnis einer Silhouette-Analyse

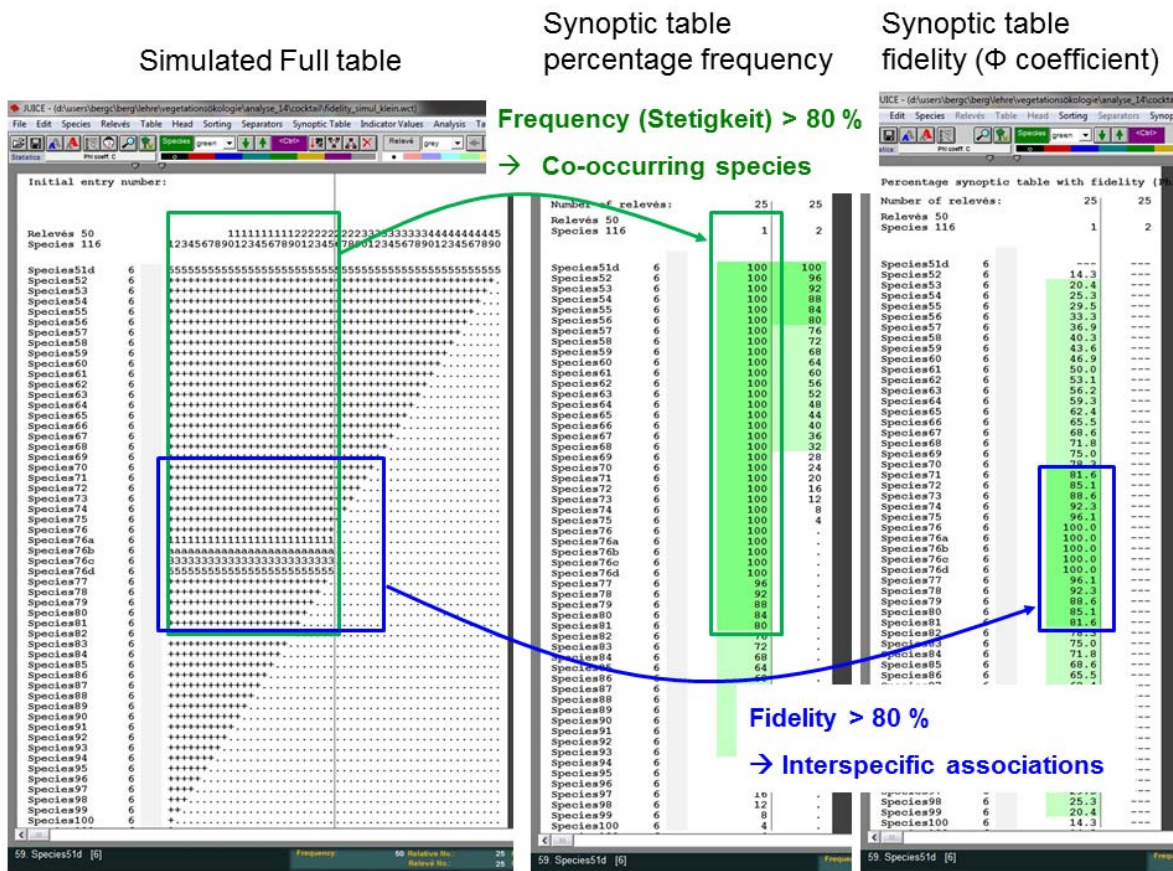
## 3.3 Analyse

### 3.3.1 Co-occurring Species

**Analysis** → **Co-occurring Species**

Etwas außerhalb des Themas Tabellengliederung stehen die co-occurring species. Über diesen Befehl kann man sich eine Liste von Arten anzeigen lassen, die in der Tabelle mit einer markierten Art zusammen vorkommen. Dieser Befehl nutzt nicht die Fidelity, sondern die Frequenz bzw. Stetigkeit, also ohne zu prüfen, in wie weit die co-occurring species auch in Aufnahmen ohne die Zielart vorkommen. Auch hier kann man die Arten im Auswahlfeld markieren und über **Mark selected species in the table** auch in der Tabelle farblich markieren.

Die co-occurring species stellen jene Arten dar, die mit einer ausgewählten Art in einer Tabelle häufig vergesellschaftet sind. Sie geben also die reale Vergesellschaftung einer Art wieder. Die Interspecific associations stellen nur Arten dar, die in der Tabelle annähernd gleiche Vorkommen wie die markierte Art haben. Der grundsätzliche Unterschied zwischen Co-occurring Species und Interspecific associations ist in der **Abbildung 11** an Hand einer simulierten Tabelle mit der Art Species76 und ihrer Frequency- und Fidelity-Werte dargestellt.



**Abbildung 11:** Schematische Darstellung der Unterschiede zwischen Co-occurring Species und Interspecific associations

### 3.3.2 Cocktail

#### 3.3.2.1 Einführung

Die Cocktail-Methode ist eine Art "Programmiersprache" zum formalisierten Charakterisieren von Vegetationseinheiten, und dient damit auch zum Auffinden von bereits bekannten Syntaxa in großen Datensätzen (supervised classification). Die Methode basiert auf Präsenz/Absenz von bestimmten Artengruppen einschließlich deren Artmächtigkeiten (cover values). Sie ist ausführlich im JUCE-Handbuch erklärt und soll hier nur kurz beleuchtet werden, um das Potenzial zu zeigen. Gute Beispiele aus der Literatur liefern u. a. KOČI et al. (2003), LOSOSOVÁ (2004) und ŠILC & ČARNI (2007).

#### 3.3.2.2 Zusammenstellen von Artengruppen

Wenden wir uns zuerst den Artengruppen zu. Diese kann man auch außerhalb von Cocktail zum Klassifizieren und schnellen Ordnen von Tabellen, und zum Suchen von Aufnahmen mit einer bestimmten Artenszusammensetzung verwenden. Deshalb wollen wir hier auch anschauen, wie man eigene Artengruppen erstellt. Es gibt funktionale und soziologische Artengruppen.

#### 3.3.2.3 Functional Groups

Functional Groups sind Gruppen von Arten, die irgendeine meist morphologische Eigenschaft miteinander teilen und denen dadurch eine ähnliche Funktion im Ökosystem zukommt. Ökologische Artengruppen, wie etwa Feuchte- oder Trockenzeiger, oder biogeographische Gruppen, wie ozeanische oder pontische Arten, sind hier dazugerechnet, obwohl sie nicht im strengen Sinne "funktionale" Gruppen sind. Diese müssen vor der Gruppenbildung in der Tabelle mit einer einheitlichen Farbe markiert werden. Für die Zuordnung braucht man eine Tabelle mit Pflanzennahmen und den jeweiligen Eigenschaften; diese kann man dann als "external species data" in die JUCE-Tabelle einlesen (siehe Kap.



2.9.3). Als Beispiel wollen wir mal eine Gruppe aller großen Bäume erstellen, d. h. Baumarten in der Baumschicht. Wir markieren die Arten der Baumschicht (z. B. mit `Mark Layers`) mit einer bestimmten Farbe und gehen jetzt in

#### **Analysis → Cocktail groups**

Zuerst kann man mit `Add species` die markierten Arten in das Listen-Feld laden. Nun kann man der Gruppe im obersten Feld einen Namen geben. Dieser muss bei funktionalen Gruppen entweder mit `#TC` oder mit `#SC` beginnen. `TC` (Total cover) summiert die Deckungswerte der Gruppenmitglieder für die einzelne Vegetationsaufnahme (interpoliert auf max. 100 %) auf, `SC` (single cover) zeigt den jeweils höchsten Bedeckungswert eines Gruppenmitglieds an. Der Name, z.B. "`#TC trees`" bedeutet also eine Gruppe aller Bäume, die in der Tabelle als Bedeckungswert die interpolierte Summe aller Bedeckungen anzeigt. Mit "`Add #TC group`" wird die Gruppe, sozusagen als neue "Art", in die Tabelle geschrieben.

Soziologische Gruppen werden ähnlich erstellt, auch hier werden die Arten vorher markiert, z. B. aus der Beschreibung einer Assoziation. Soziologische Artengruppen werden mit `###` am Anfang des Namens benannt und nicht als Deckungswert, sondern als Stern `*` (Präsenz/Absenz) in der Tabelle dargestellt.

Etwas mehr Schritte braucht es für eigene soziologische Artengruppen. Hier beginnt man meistens mit einer auffälligen Art. Diese wird in der Tabelle markiert, dann wählen wir

#### **Analysis → Interspecific association → ... → for all relevés**

JUICE gibt jetzt in 2 Fenstern positiv und negativ assoziierte Arten (berechnet nach dem jeweils eingestellten Fidelity-Wert) aus. Aus den Arten mit positivem Wert kann man nun einige ökologisch sinnvolle Arten auswählen, die man mit der `shift`-Taste und der Maus markieren kann und mit `Mark selected species in the table` wieder in einer eigenen Farbe in der Tabelle markieren kann. Mit diesen Arten geht man nun in die Cocktail-Groups.

#### **Analysis → Cocktail groups**

Arten einladen und Namen vergeben sind die ersten Schritte. Unter `Min. No. Spec. Rel.` berechnet das Programm eine mathematisch optimierte Mindestanzahl unserer Arten, die mindestens in einer Aufnahme enthalten sein sollten (meist so um 50 % der Arten). Mit "`Manual setting`" kann man diese Zahl auch absenken oder erhöhen. Dann stellt man bei "`Select relevés as`" eine neue Farbe ein und drückt diesen Knopf. Die Aufnahmen, die die eingestellte Anzahl Arten enthalten, werden gekennzeichnet. Man kann auch mit einer niedrigen Zahl anfangen und dabei prüfen, wie sich die Zahl selektierter Aufnahmen verändert. Nachdem wir vorher die interspezifische Assoziationen ausgehend von der ersten Art mit allen anderen Arten berechnet haben, können wir jetzt prüfen, wie es um die Fidelity aller Arten unserer Artengruppe innerhalb dieser Aufnahmen bestellt ist, und ob noch andere Arten eine Affinität zu unserer Artengruppe bzw. zu dem gewählten cluster haben. So kann man nach und nach die Artengruppe und die dazugehörigen Aufnahmen optimieren. Ähnlich kann man vorgegebene Artengruppen, z. B. aus den "Pflanzengesellschaften von Österreich", an Hand neuer Daten erweitern oder einengen.

Haben wir die Artenauswahl optimiert und auch eine Mindestzahl von Arten pro Aufnahme eingestellt, können wir mit `Add ### sociological group` die Gruppe oben in die Tabelle einfügen. Aufnahmen, die die eingestellte Anzahl von Arten dieser Gruppe enthalten, werden mit einem Stern `*` gekennzeichnet.

#### **3.3.2.4 Group aggregation**

Dies ist die eigentliche Power von Cocktail, zum Lernen sei auf die Literatur verwiesen. Hier soll nur kurz angerissen werden, wie es geht. Unter `Group aggregation` kann man Formeln für die Selektion einzelner Vegetationsaufnahmen erstellen und damit die Klassifikation erleichtern und vor allem Standardisieren. Hier ein sehr einfaches Beispiel:

```
<### Bromus erectus >AND #TC Cirsium pannonicum GR 10> NOT <#SC Solidago canadensis GR 5>
```

Das bedeutet: Suche Aufnahmen, die eine bestimmte Anzahl Arten aus der `Bromus erectus`-Gruppe enthalten, außerdem sollen die Arten der `Cirsium pannonicum`-Gruppe zusammen eine Artmächtigkeit über 10 % haben, während Aufnahmen ausgeschlossen werden sollen, bei denen eine einzige Art der `Solidago canadensis`-Gruppe eine Artmächtigkeit über 5 % erreicht.

#### **3.3.2.5 Expert System**

Diese Formeln kann man in eine Datei `*.esy` übertragen, und für alle Zeiten und Datensätze abrufbar halten. Die Datei enthält außerdem im Teil 1 die aggregierten Arten, im Teil 2 die aufgestellten Arten-Gruppen und ihre Mitglieder, im Teil 3 eben diese Formeln und im Teil 4 die Nummern der Aufnahmen, für die keine der angegebenen Formeln zutrifft.

#### **3.3.3 Berechnen von Artenzahlen**

Arbeitet man mit mehrschichtigen Vegetationstypen (Gebüsche, Wälder), dann kann JUICE die genauen Artenzahlen nicht berechnen. JUICE zählt die gleiche Art in der Kraut- und der Strauchschicht als 2 Arten. Deshalb braucht es einiger

Tricks, um Artenzahlen berechnen zu können, siehe Kap. 2.10, Will man aber nur ein paar Artenzahlen berechnen und hinterher mit der Tabelle und getrennten Schichtangaben weiterarbeiten, kann man folgendermaßen vorgehen: 1. Relevé color vereinheitlichen (z. B. white) und Tabelle speichern 2. Tabelle unter neuem Namen speichern 3. Alle Arten mit gleichem Namen fusionieren, 4. Artenzahlen berechnen 5. Berechnete Artenzahlen in die alte, unfusionierte Tabelle übertragen.

Für den Schritt 3 gibt es den Befehl `Species → Merge all same species names → within the dataset`. Danach sollte jeder Name nur in einer Schicht (fusionierte Namen haben die Schicht 0) stehen.

Jetzt berechnet man die Artenzahlen, z. B. `Head → Store values to short header → number of black species`. Ist das der einzige Artenzahlen-Wert, den wir brauchen, kann man die Datei jetzt speichern und dann den Befehl `Edit → Copy white relevé numbers, short headers and group numbers to clipboard`. Und speichert diese Werte in einer EXCEL-Tabelle zwischen. Jetzt öffnet man die alte Tabelle mit den unfusionierten Namen. In EXCEL kopiert man die beiden Spalten Relevé Number und Short header und wählt in JUICE den Befehl: `Edit → Paste clipboard to white short headers`. Dieser Umweg über EXCEL ist eine der Eigenheiten von JUICE, auf direktem Wege funktioniert es eigenartiger Weise nicht fehlerfrei.

Will man mehrere Artdaten berechnen, sollte man in der fusionierten Datei die berechneten shortheaders jeweils in den header data abspeichern und anschließend bei der Ausgangsdatei die gesamten Kopfdaten, zuerst über Export (Datei hlava.txt) und anschließend über Import austauschen. Auch hier muss man den Umweg über EXCEL gehen, in das man die hlava.txt-Datei einlädt, kann von EXCEL nach JUICE aber das Clipboard nutzen (`File → Import → Header data → from clipboard`).

### 3.3.4 Berechnung von Biomasse von Offenland-Gesellschaften

Eine nette Sache ist es, einen Näherungswert für die Biomasse einer Offenland-Gesellschaft (Grünland) zu berechnen (AXMANOVÁ et al. 2012). Man hat damit keinen Masse-Wert im Sinne von Kilo und Gramm, sondern einen Äquivalentwert, mit dem sich die Biomasse verschiedener Offenlandgesellschaften vergleichen lässt. Dazu benötigt wird für die jeweilige Artengarnitur eine Tabelle mit der Wuchshöhe (average plant height). Diese kann man aus diversen Functional trait Datenbanken herausziehen. Mit dem einladen als external species data kann man die Wuchshöhe in die species data schreiben und anschließend mit dem Befehl: `Sum of black species data Weighted by Cover` den Äquivalentwert berechnen.

## 4. References

- AXMANOVÁ, I., TICHÝ, L., FAJMONOVÁ, Z., HÁJKOVÁ, P., HETTENBERGEROVÁ, E., LI, C.-F. F., MERUNKOVÁ, K., NEJEZCHLEBOVÁ, M., OTÝPKOVÁ, Z., VYMAZALOVÁ, M. & ZELENÝ, D. (2012). Estimation of herbaceous biomass from species composition and cover. *Applied Vegetation Science* 15, 580–589.
- BERG, C., TIMMERMANN, T., & DENGLER, J. (2004): Naturschutzfachliche Wertstufe. – In: Berg, C., Dengler, J., Abdank, A. & M. Isermann (2004) [Hrsg.]: Die Pflanzengesellschaften Mecklenburg-Vorpommerns und ihre Gefährdung Textband. S. 67–72. Weissdorn, Jena.
- BÖCKER, R., KOWARIK, I., & BORNKAMM, R. (1983). Untersuchungen zur Anwendung der Zeigerwerte nach Ellenberg. *Verh. Gesellschaft für Ökologie (Festschrift Ellenberg)* XI, 35–56.
- BRUELHEIDE H. (1997). Using formal logic to classify vegetation. *Folia Geobot. Phytotax.*, 32: 44–46.
- BRUELHEIDE H. (2000). A new measure of fidelity and its application to defining species groups. *Journal Vegetation Science* 11: 167–178.
- BRUELHEIDE, H., & CHYTRÝ, M. (2000). Towards unification of national vegetation classifications: A comparison of two methods for analysis of large data sets. *Journal of Vegetation Science* 11, 295–306.
- CHYTRÝ M., TICHÝ L., HOLT J. & BOTTA-DUKÁT Z. 2002. Determination of diagnostic species with statistical fidelity measures. *Journal of Vegetation Science* 13: 79–90.
- DRESCHER, A., THEISS, M., HAFELLNER, J. & BERG, C. (2007). Die Vegetationsverhältnisse des Großen Kars der Koralpe (Kärnten, Österreich). *Mitteilungen Des Naturwissenschaftlichen Vereins Für Steiermark*, 136, 187–238.
- ELLENBERG, H., WEBER, H. E., DÜLL, R., WIRTH, V., WERNER, W. & PAULIBEN, D. (1991): Zeigerwerte von Pflanzen in Mitteleuropa. 2. Auflage. *Scripta Geobot.* 18: 1–258.
- ERSCHBAMER, B., UNTERLUGGAUER, P., WINKLER, E. & MALLAUN, M. (2011): Changes in plant species diversity revealed by long-term monitoring on mountain summits in the Dolomites (northern Italy). *Preslia* 83, 387–401.
- GRABHERR, G., & MUCINA, L. (1993). Die Pflanzengesellschaften Österreichs. Teil II: Natürliche waldfreie Vegetation. 523 pp.

- HAVEMAN, R. & JANSSEN, J. A. M. (2008). The analysis of long-term changes in plant communities using large databases: The effect of stratified resampling. *Journal of Vegetation Science* 19, 355–362.
- HENNEKENS, S. M. & Schaminee, J. H. J. (2001). TURBOVEG, a comprehensive database management system for vegetation data. *Journal of Vegetation Science* 12: 589–591.
- HILL, M. O. (1994). Decorana and Twinspan, for ordination and classification of multivariate species data: a new edition, together with supporting programs, in Fortran 77 (p. 58). Huntingdon, England.
- KLOTZ, S., KÜHN, I., DURKA, W. (2002). BIOLFLOR – Eine Datenbank mit biologisch-ökologischen Merkmalen zur Flora von Deutschland. *Schriftenreihe für Vegetationskunde* 38, 1–334.
- KNOLLOVA, I., CHYTRÝ, M., TICHÝ, L., & HÁJEK, O. (2005). Stratified resampling of phytosociological databases: some strategies for obtaining more representative data sets for classification studies. *Journal of Vegetation Science* 16, 479–486.
- KOČÍ, M., CHYTRÝ, M., & TICHÝ, L. (2003). Formalized reproduction of an expert-based phytosociological classification: A case study of subalpine tall-forb vegetation. *Journal of Vegetation Science* 14, 601–610.
- LENGYEL, A., CHYTRÝ, M., & TICHÝ, L. (2011). Heterogeneity-constrained random resampling of phytosociological databases. *Journal of Vegetation Science* 22, 175–183.
- LEYER, I., & WESCHE, K. 2008. *Multivariate Statistik in der Ökologie. Eine Einführung.* 221 S., Springer, Berlin, Heidelberg.
- LÖTTER, M. C., MUCINA, L., & WITKOWSKI, E. T. F. (2013). The classification conundrum: species fidelity as leading criterion in search of a rigorous method to classify a complex forest data set. *Community Ecology* 14, 121–132.
- LOSOSOVÁ, Z. (2004). Weed vegetation in southern Moravia (Czech Republic): a formalized phytosociological. *Preslia* 75, 65–85.
- MUCINA, L., GRABHERR, G., & ELLMAUER, T. (1993). *Die Pflanzengesellschaften Österreichs. Teil I: Anthropogene Vegetation.* Stuttgart, New York: Gustav Fischer. 578 pp.
- MUCINA, L., GRABHERR, G., & WALLNÖFER, S. (1993). *Die Pflanzengesellschaften Österreichs. Teil III: Wälder und Gebüsche.* Stuttgart, New York: Gustav Fischer. 353pp.
- PEINADO, M., DÍAZ, G., OCAÑA-PEINADO, F. M., AGUIRRE, J. L., MACÍAS, M. Á., DELGADILLO, J., & APARICIO, A. (2013). Statistical measures of fidelity applied to diagnostic species in plant sociology. *Modern Applied Science* 7, 106–120.
- ROLEČEK, J., TICHÝ, L., ZELENÝ, D., & CHYTRY, M. (2009). Modified TWINSpan classification in which the hierarchy respects cluster heterogeneity. *Journal of Vegetation Science* 20: 596–602.
- SCHAFFERS, A. P., & SÝKORA, K. V. (2000). Reliability of Ellenberg indicator values for moisture, nitrogen and reaction: a comparison with field measurements. *Journal of Vegetation Science* 11 : 225–244.
- ŠILC, U., & ČARNI, A. (2007). Formalized classification of the weed vegetation of arable land in Slovenia. *Preslia* 79, 283–302.
- TICHÝ L. 2002. JUICE, software for vegetation classification. *Journal of Vegetation Science* 13: 451–453.
- WILLNER, W., & GRABHERR, G. (2007a). *Die Wälder und Gebüsche Österreichs. Ein Bestimmungswerk mit Tabellen. 1 Textband.* München: Elsevier.
- WILLNER, W., & GRABHERR, G. (2007b). *Die Wälder und Gebüsche Österreichs. Ein Bestimmungswerk mit Tabellen. 2 Tabellenband.* München: Elsevier.
- WILLNER, W., TICHÝ, L., & CHYTRÝ, M. (2009). Effects of different fidelity measures and contexts on the determination of diagnostic species. *Journal of Vegetation Science* 20, 130–137.
- ZELENÝ, D. & SCHAFFERS, A. P. (2012). Too good to be true: pitfalls of using mean Ellenberg indicator values in vegetation analyses. *Journal of Vegetation Science* 23, 419–431.