



# UFZ-Bericht

RSCHUNGSZENTRUM LEIPZIG-HALLE GMBH

Nr. 8/2003

Dissertation

---

**Ordinationsverfahren zur Strukturanalyse  
ökosystemarer Feldinformation und  
Lebensraumeignungsmodelle für ausgewählte  
Arten der Elbauen**

---

Marcus Rink

RS

03

919 MA

ISSN 0948-9452

Helmholtz-Zentrum für Umweltforschung (UFZ)



00409490

**Ordinationsverfahren**  
**zur Strukturanalyse ökosystemarer Feldinformation und**  
**Lebensraumeignungsmodelle**  
**für ausgewählte Arten der Elbauen**

Von der Gemeinsamen Naturwissenschaftlichen Fakultät  
der Technischen Universität Carolo-Wilhelmina  
zu Braunschweig  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften  
(Dr. rer. nat.)

**genehmigte**  
**D i s s e r t a t i o n**

*Archiv*

von **Marcus Rink**  
aus **Wiesbaden**



UFZ-Umweltforschungszentrum  
Leipzig-Halle GmbH  
Zentralbibliothek  
Permoserstraße 15  
04318 Leipzig

03-0916

**1. Referent: Professor Dr. Otto Richter**  
**2. Referent: Privatdozent Dr. habil. Klaus Henle**  
**eingereicht am: 01. August 2002**  
**mündliche Prüfung (Disputation) am: 21. Januar 2003**  
**Druckjahr: 2003**

## Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Gemeinsamen Naturwissenschaftlichen Fakultät, vertreten durch den Mentor der Arbeit, Prof. Dr. O. Richter, in folgenden Beiträgen vorab veröffentlicht:

### Publikationen

- Fuchs, E. & M. Rink (1998): GIS-Anwendung bei der Indikation ökologischer Veränderungen in Flußauen: Versuchsplanung und Modellierung. Strobl/Dollinger [Hrsg.] (1998): Angewandte Geografische Informationsverarbeitung, S. 81-87, Heidelberg.
- Fuchs, E., K. Henle, W. Peter, M. Rink, S. Stab (1999): Versuchsplanung und Zusammenführung von Ergebnissen im RIVA-Projekt. Fachtagung Elbe "Dynamik und Interaktion von Fluß und Aue 4. bis 7. Mai 1999, Wittenberge / Forschungsverbund Elbe-Ökologie, S. 231-232, Universität Karlsruhe.
- Rink, M. et al. (2000). Zur Erstellung einer fachlich-statistisch abgestimmten Datenerhebungsstrategie am Beispiel eines synökologisch orientierten Forschungsprojektes in den Elbauen. Hydrologie und Wasserbewirtschaftung 44 (4): 184-190.
- Stab, S., U. Amarell & M. Rink (2000): Indikation ökologischer Veränderungen in Auen. Tagungsband Statusseminar Elbe-Ökologie Berlin, S. 109-112, Hrsg. BfG Koblenz.
- Henle, K. et al. (2000). Methodische Grundlagen zur Entwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen. ATV-DVWK-Schriftenreihe. Bonn-Hennef, Verlag für Abwasser, Abfall und Gewässerschutz 21: 173-201.

### Tagungsbeiträge

#### international:

- Rink, M. (1997): Study Design and Statistical Approaches in RIVA. (Vortrag) Workshop: Study Design in FAEWE and RIVA, Oktober 1997, Leipzig.
- Rink, M. (1998): Indicator Systems for Floodplain Management. (Vortrag) Scientific Series of Lectures at the Oklahoma State University - Department of Botany, Oktober 1998, Stillwater Oklahoma U.S.A.
- Rink, M. (2000): Strategy of Multivariate Data-Analysis in RIVA. (Vortrag) Multivariate Analysis of Ecological Data, Januar 2000, University of South Bohemia, Budweis – Tschechische Republik.

#### national:

- Rink, M. (1997): Untersuchungs- und Auswertemethodik in RIVA. (Vortrag) Projektübergreifendes Arbeitstreffen Auenökologie/-hydrologie im Rahmen des Förderungsprogramms "Elbe-Ökologie", September 1997, Berlin.
- Rink, M. (1997): Synthese der Daten und statistische Bearbeitung in RIVA. (Vortrag) 1. Statusseminar RIVA, November 1997, Koblenz.
- Rink, M. (1998): Zur statistischen Versuchsplanung im RIVA-Projekt – Analyseverfahren und Stichprobenplan. (Vortrag) Arbeitstreffen Leitbildentwicklung und Bioindikation im Forschungsverbund Elbe-Ökologie, Februar 1998, Berlin.
- Rink, M. (1999): Datenanalysestrategie und Fortgang der statistischen Bearbeitung im RIVA-Projekt. (Vortrag) 2. Statusseminar RIVA, Januar 1999, Höxter.

- Rink, M. (1999): Leitlinien zum Probendesign für indikatorische Fragestellungen. (Vortrag) Arbeitstreffen Biologische Fragestellungen und Bioindikation im Forschungsverbund Elbe-Ökologie, März 1999, Leipzig.
- Rink, M., S. Stab & U. Amarell (1999): Indikationssysteme für ökologische Veränderungen in Auen. (Vortrag) Statusseminar des Forschungsverbundes Elbe-Ökologie, November 1999, Berlin.
- Rink, M. (1999): Indikationssystem für ökologische Veränderungen in Auen – Study Design: Teilaspekt Stichprobenplan "Stratifizierte Zufallsverteilung. (Poster) Statusseminar des Forschungsverbundes Elbe-Ökologie, November 1999, Berlin.
- Rink, M. (2000): Multivariate Statistik und ihre Anwendung in RIVA. (Vortrag) Arbeitstreffen Biologische Fragestellungen und Bioindikation im Forschungsverbund Elbe-Ökologie, März 2000, Leipzig.
- Stab, S. & M. Rink (2001): Planung und Durchführung von Felduntersuchungen zur Entwicklung von Indikationssystemen. (Vortrag) Fachtagung Indikation in Auen, April 2001, Wörlitz.
- Rink, M. (2001): Datenauswertung. (Vortrag) Fachtagung Indikation in Auen, April 2001, Wörlitz.
- Fuchs, E., M. Rink & S. Rosenzweig (2001): Wie arbeitet und was zeigt das Indikationsmodell? (Vortrag) Fachtagung Indikation in Auen, April 2001, Wörlitz.
- M. Rink (2001): Statistische Methoden und deren Ergebnisse im RIVA-Projekt. (Vortrag) Fachtagung Indikation in Auen, April 2001, Wörlitz.
- Fuchs, E., V. Hüsing & M. Rink (2001): Modellbildung im RIVA-Projekt. (Vortrag) Fachtagung Indikation in Auen, April 2001, Wörlitz.

## Vorwort

Das Bundesministerium für Bildung und Forschung (bmb+f) fördert unter dem Ziel der Erarbeitung von Konzepten für eine dauerhaft umweltverträgliche Entwicklung der "Flusslandschaft Elbe" die Erforschung ökologischer Zusammenhänge in Stromlandschaften (BMBF 1995). Im Förderschwerpunkt "Ökologische Konzeptionen für Fluss- und Seenlandschaften – Teilbereich Elbe-Ökologie" wurde im Themenbereich II "Ökologie der Auen" das Verbundvorhaben mit dem Titel "Übertragung und Weiterentwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen" (RIVA) bewilligt. Es wird unter der Federführung des Umweltforschungszentrums Leipzig-Halle GmbH (UFZ) ausgeführt.

Das Verbundprojekt wird durch fünf Teilprojekte realisiert: Koordination (Teilprojekt I), Hydrogeologie und Boden (Teilprojekt II), Vegetation (Teilprojekt III), Fauna (Mollusken in Teilprojekt IV.1, Laufkäfer in Teilprojekt IV.2), Synthese und Modellierung der Daten (Teilprojekt V.1) sowie Robustheit des tierökologischen Indikationssystems (Teilprojekt V.2).

Es hat zum Ziel: (1) abiotische Komponenten von Indikationssystemen zu entwickeln und abzusichern, (2) biotische Komponenten von Indikationssystemen unter Anwendung multivariater statistischer Verfahren weiterzuentwickeln, (3) Ansätze zur Verknüpfung abiotischer und biotischer Indikationssysteme herzustellen, (4) die Möglichkeiten und Grenzen der Übertragbarkeit von Indikationssystemen exemplarisch zu analysieren, (5) die Robustheit des zoologischen Indikationssystems gegenüber einem reduzierten Erfassungsaufwand zu testen und (6) möglicherweise methodische Anleitungen und Datengrundlagen zu entwickeln sowie (7) diese entsprechenden Entscheidungsträgern in der planerischen Praxis zur Verfügung zu stellen.

Die Bundesanstalt für Gewässerkunde (BfG) bearbeitet das Teilprojekt V.1 "Prüfung der Robustheit und modellhafte Gesamtbetrachtung des Ökosystems Elbaue in den Untersuchungsgebieten". Ziel des Teilprojektes V.1 ist es, durch die Synthese der Daten eine "ökologische Gesamtschau" der Aue in den Untersuchungsgebieten herzustellen. Es sollen relevante wissenschaftlich begründete und fachlich fundierte Verknüpfungen, innerhalb wie auch zwischen den abiotischen und biotischen Ökosystemkomponenten, durch den Einsatz multivariater statistischer Verfahren hergeleitet sowie anhand dieser Zusammenhänge eine ökologische Modellierung für die Untersuchungsgebiete aufgebaut werden (Fuchs et al. 1998: Erster Zwischenbericht, Teilbeitrag V.1 im BMBF Verbundvorhaben: Prüfung der Robustheit und modellhafte Gesamtbetrachtung des Ökosystems Elbaue in den Untersuchungsgebieten. Koblenz, Bundesanstalt für Gewässerkunde: 47, unveröffentlicht).

In diesem Anforderungsrahmen entstand eine Promotionsstelle, die aus BfG- und UfZ-eigenen Mitteln anteilmäßig finanziert wurde. Die Arbeit trägt maßgeblich zur Versuchsplanung des Gesamtvorhabens bei, indem sie die kooperative fachlich-statistische Abstimmung des Stichprobenplans zur Sicherstellung einer integren Datenerhebung übernehmen sowie eine geeignete statistische Datenanalysestrategie durch Methodenvergleiche entwickeln soll. Somit wird die Strukturanalyse der Erhebungsdaten hinsichtlich ökosystemarer Muster und beschreibender Wirkungszusammenhänge ermöglicht. Die zur Entwicklung eines integrierten ökologischen Prognosesystems benötigten Informationen, d.h. welche Parameter einen relevanten Erklärungsanteil an der Artenverteilung im Raum besitzen, werden durch statistische Analysen zur Verfügung gestellt. Die Ableitung ökologischer Gruppen und die Selektion typischer stellvertretender Arten für diese sowie die statistischen Modelle selbst bilden dann die Grundlage für ein GIS-basiertes Prognosesystem, für dessen abschließende Erstellung sich die BfG verantwortlich zeigt.

## Inhalt

<b>1</b>	<b>EINLEITUNG: PROBLEMSTELLUNG UND ZIELSETZUNG DER ARBEIT</b>	<b>1</b>
<b>2</b>	<b>WARUM ORDINATIONSVERFAHREN? DIE NATUR UND EINIGE KONZEPTIONELLE GRUNDLAGEN DER MATHEMATISCH-STATISTISCHEN ÖKOSYSTEMFORSCHUNG</b>	<b>6</b>
2.1	VON DER UMWELTREALITÄT ZUM MODELL: MÖGLICHKEITEN UND GRENZEN DER MODELLIERUNG NATURNAHER ÖKOSYSTEME	6
2.2	BIOZÖNOSE, MUSTER UND GRADIENT: KONTINUUMTHEORIE KONTRA DISKONTINUUMTHEORIE	8
2.3	NISCHENKONZEPT UND ÖKOLOGISCHE ANTWORTMODELLE	13
2.4	GLIEDERUNGSANSÄTZE: SOZIOLOGISCH ODER ÖKOLOGISCH?	19
2.5	STELLVERTRETENDE ARTEN FÜR ÖKOLOGISCHE GRUPPEN	20
2.6	BETRACHTUNGSRÄUME: KLEINRÄUMIGE HETEROGENITÄT UND SKALENÜBERGÄNGE	22
<b>3</b>	<b>METHODEN ZUR ERHEBUNG, GEWINNUNG, VORHALTUNG, STRUKTURANALYSE UND MODELLIERUNG ÖKOSYSTEMARER FELDDINFORMATION</b>	<b>25</b>
3.1	DATENERHEBUNG	25
3.1.1	ALLGEMEINE VORÜBERLEGUNGEN	25
3.1.2	STICHPROBENPLAN UND STICHPROBENUMFANG	26
3.1.3	STRATIFIZIERUNGSKRITERIEN UND STRATIFIZIERUNG VON UNTERSUCHUNGSGEBIETEN	34
3.1.4	RÄUMLICHE UND ZEITLICHE ABSTIMMUNG DER DATENERHEBUNG – GRÖÖE UND AUFBAU VON STICHPROBENFLÄCHEN	36
3.1.5	KARTENGRUNDLAGEN UND EINSATZ EINES GEOGRAFISCHEN INFORMATIONSSYSTEMS	40
3.1.6	SATELLITENNAVIGATIONSSYSTEME UND IHRE ANWENDUNG BEI DER ERSTELLUNG VON STICHPROBENPLÄNEN	41
3.1.7	ABIOTISCHER MINIMUMDATENSATZ UND ALLGEMEINE DATENSATZEIGENSCHAFTEN	42
3.2	ABLEITUNG VON SEKUNDÄRDATEN: DATENGEWINNUNG MIT DER ARC/INFO®-APPLIKATION "DIST"	46
3.3	DATENAUFBEREITUNG FÜR DIE STATISTISCHE ANALYSE	49
3.4	DATENHALTUNG: DATENBANKANFORDERUNGEN	53
3.5	GRADIENTENANALYTISCHE VERFAHREN: GRUNDLAGEN, THEORIE, EIGENSCHAFTEN UND HANDHABUNG	54
3.5.1	MULTIVARIATE STATISTIK UND GRADIENTENANALYTISCHE VERFAHREN: EINFÜHRUNG	55
3.5.2	ASSOZIATIONS- UND ABHÄNGIGKEITSMASSE: EIGENSCHAFTEN UND IHRE ZUORDNUNG ZU ORDINATIONSVERFAHREN	62
3.5.3	VARIABLENREDUKTION: KORRELATIONSANALYSE, VORWÄRTSSELEKTION UND ANDERE VERFAHREN	66

3.5.4	HAUPTKOMPONENTENANALYSE	69
3.5.5	KORRESPONDENZANALYSE	75
3.5.6	KORRESPONDENZANALYSE MIT DETRENDINGPROZESS	78
3.5.7	REDUNDANZANALYSE	81
3.5.8	KANONISCHE KORRESPONDENZANALYSE	83
3.5.9	RÄUMLICHE AUTOKORRELATION: MORAN'S I UND PARTIELLE ORDINATION	85
3.6	ORDINATIONS DIAGNOSTIK UND GRAPHISCHE AUFBEREITUNG VON ORDINATIONSERGEBNISSEN	89
3.6.1	ORDINATIONS DIAGNOSTIK	89
3.6.2	GESICHTER DES ORDINATIONS DIAGRAMMS	99
3.6.3	AUSWERTEREGELN FÜR ORDINATIONS DIAGRAMME	105
3.6.4	DIAGRAMME ZUR WEITERFÜHRENDEN DATENEXPLORATION	109
3.7	GENERALISIERTE LINEARE MODELLE	114
3.7.1	LOGISTISCHE REGRESSION	114
3.7.2	DIAGRAMME ZUR VISUALISIERUNG VON ANTWORTKURVEN GEGENÜBER DEN KOMPLEXGRADIENTEN ODER EINZELNEN INFORMATIONSEBENEN	120
3.7.3	ANTWORTOBERFLÄCHEN	121
4	VERGLEICHENDE BEWERTUNG DER METHODEN	123
4.1	ORDINATION: GIBT ES DAS EINZIG WAHRE VERFAHREN?	123
4.2	DIREKTE ORDINATION UND LOGISTISCHE REGRESSION: ZWEI SICH ERGÄNZENDE VERFAHREN?	133
5	GEWINNUNG UND AUSWERTUNG VON ÖKO SYSTEMARER FEL DINFORMATION ZUR ERSTELLUNG EINES PROGNOSE SYSTEMS IN EINEM AUEÖKOLOGISCHEN VERBUNDPROJEKT	137
5.1	UNTERSUCHUNGS GEBIETE	137
5.1.1	ALLGEMEINE VORÜBERLEGUNGEN	137
5.1.2	AUSWAHL DER UNTERSUCHUNGS GEBIETE	137
5.1.3	ZU DEN PROJEKTUNTERSUCHUNGS GEBIETEN	140
5.2	DATENERHEBUNGS STRATEGIE	144
5.2.1	STICHPROBENUMFANG, STICHPROBENPLAN UND STRATIFIZIERUNG	144
5.2.2	ARBEITSSCHRITTE BEI DER ERSTELLUNG UND ETABLIERUNG DES STICHPROBENPLANS "STRATIFIZIERTE ZUFALLSVERTEILUNG" IM RIVA- VERBUNDPROJEKT	148
5.2.3	ZEITLICHE ABSTIMMUNG DER DATENERHEBUNG	151
5.3	DATENSÄTZE UND DATENBANKSTRUKTUR	152
5.3.1	BIOTIK: DIE INDIKATORGRUPPEN	152
5.3.2	ABIOTIK: STANDORTPARAMETER	156
5.3.3	AUFBAU DES GEOGRAFISCHEN INFORMATIONSSYSTEMS	158
5.3.4	ROHDATENAUFBEREITUNG UND GIS-VISUALISIERUNG DER ABIOTISCHEN VARIABILITÄT	159
5.3.5	DATENBANKSTRUKTUR UND "SHUTTLES"	159
5.4	MULTIVARIATE DATENANALYSE: MUSTERERKENNUNG UND MODELLBILDUNG	160
5.4.1	MODELLBILDUNG: VARIABLENSELEKTION MIT ZEITLICHER ROBUSTHEITSPRÜFUNG	160
5.4.2	MUSTERERKENNUNG: ORDINATION	170

5.4.3	WEITERFÜHRENDE AUSWERTUNGEN ZU DEN ORDINATIONSERGEBNISSEN, ABLEITUNG ÖKOLOGISCHER GRUPPEN UND SCHIRMARTENSELEKTION	190
5.4.4	ZEITLICHE UND RÄUMLICHE ÜBERTRAGBARKEIT DER ORDINATIONSMODELLE	206
5.5	GENERALISIERTE LINEARE MODELLE: LOGISTISCHE REGRESSION	215
5.6	GIS-IMPLEMENTIERUNG DER STATISTISCHEN MODELLE ZUR PROGNOSE	224
6	INTERPRETATION UND DISKUSSION DER ERGEBNISSE	227
7	ZUSAMMENFASSUNG	242
8	QUELLENVERZEICHNIS	247

ANLAGEN

NOTATIONEN

DANK



## 1 EINLEITUNG: PROBLEMSTELLUNG UND ZIELSETZUNG DER ARBEIT

*"RIVERS FUNCTION AS ARTERIES AND VEINS ON THE WORLD CONTINENTS, THEY ARE AMONG THE MOST VULNERABLE BUT ALSO MOST ABUSED RESOURCES ON EARTH. IN THE COURSE OF HUMAN HISTORY INTENSIVE USE HAS BEEN MADE OF RIVER RESOURCES, AND CONSEQUENTLY THE QUALITY OF THE RIVER ECOSYSTEMS SLOWLY DEGRADED (NIENHUIS ET AL. 1998, S. 18)"*

Das einleitende Zitat zeigt ganz allgemein, welche Probleme sich aus der Historie der menschlichen Besiedlung und Nutzungsformen wie beispielsweise Schifffahrt, Hochwasserschutz, Landwirtschaft aber auch Energiegewinnung für viele Flussökosysteme ergeben haben. Der Fluss und seine Auen gehören inzwischen weltweit zu den gefährdetsten Biozönoten (Gerken 1988; Nienhuis et al. 1998; Rink et al. 2000; UNESCO 1998; Walters & Korman 1999 u.a.).

Flüsse durchströmen die Landschaft als zentrales, ihren Naturraum auf vielfältige Weise prägendes Element. Noch bis in die erste Hälfte des 19. Jahrhunderts, bevor massive Flussregulierungen vorgenommen wurden, waren die grandiosen Auelandschaften der mitteleuropäischen Flüsse ein bestimmendes Landschaftselement, geprägt durch die Hochwasserdynamik, die Geschiebekraft und ständige Nährstoffzufuhr des Flusses, welches in seiner räumlich-zeitlichen Strukturvielfalt vielen Lebewesen Heimat bot. Mit 12.000 Tier- und Pflanzenarten als regelmäßige Auenbewohner stellen natürliche Auen der Unterlaufregion teils bis heute die artenreichsten Ökosysteme Europas dar (Foeckler & Bohle 1991).

Anthropogene Eingriffe in die Systeme führen zu nachhaltigen Veränderungen der Lebensbedingungen und gefährden den Artenfortbestand. Zu den Maßnahmen der Flussregulierung zählen Hochwasserdämme, welche die Aue in einen rezenten und einen fossilen Teil trennen. Die fossile Aue nimmt damit nur noch über den Grundwasserstrom an der Flusssdynamik teil. Durch die Abschottung von Hochwasserwellen ist eine landwirtschaftlich intensive Nutzung ermöglicht, die meist einhergeht mit Eutrophierung, Restbiotopverinselung und Artenverarmung. Die rezente Aue unterliegt zwar weiterhin der prägenden Hochwasserdynamik, doch die Einengung der Retentionsfläche führt zu einer Anhebung der mittleren Überflutungshöhe und damit zu einer gesteigerten Hochwassergefahr (Dister 1983; Foeckler & Bohle 1991). Umgekehrt stellt sich durch die Verengung der Abflussbreite und Buhnenbau eine erhöhte Fließgeschwindigkeit ein, die häufig mit übermäßiger Sohlenerosion vergesellschaftet ist. Es kommt zur Eintiefung des Flusses mit Absenkung des Grundwasserspiegels. Die Folge kann eine erhöhte Versteppungsgefahr von Auenelementen in der Niedrigwasserperiode sein (Foeckler & Bohle 1991; Hügin 1981).

Es herrscht bereits lange breiter Konsens unter Ökologen, dass die meisten anthropogenen Nutzungseingriffe sich negativ auf die natürliche Ökosystemstruktur, die Funktion und die Lebensgemeinschaften in Flussökosystemen ausgewirkt haben und auswirken. Bis vor kurzem ist aber die Stimme ökologisch gesinnter Visionäre im Umweltmanagement ungehört geblieben. Erst die kostenintensiven Fehlentscheidungen der vergangenen Jahre durch das Ignorieren natürlicher Folgeprozesse bei Managementmaßnahmen haben den historischen Gesinnungswandel vom rein technologischen zu einem ökologisch geprägten Handeln veranlasst. Dieser grundlegende Wandel und zunehmende Umweltprobleme bringen heutzutage Entscheidungsträger, Nutzer, die breite Öffentlichkeit und die Vertreter verschiedener wissenschaftlicher Disziplinen näher zusammen.

Das gemeinsame Interesse ermöglicht ein nachhaltiges Management dieser gefährdeten Räume, indem durch integrative Kommunikation eine genauere Problemdefinition und eine präzise Übersicht über benötigte interdisziplinäre Information zusammengestellt werden kann.

Die neu geschaffene Schnittstelle zwischen Wissenschaft und Politik verlangt nach Instrumenten zu einer nachhaltigen Entscheidungsfindung, um die Klärung der folgenden übergeordneten Fragekomplexe zu ermöglichen:

- (1) Welche bedingenden und erhaltenden Standortfaktoren prägen die faunistischen und floristischen Lebensgemeinschaften der Aue; welche Lebensraumbindung besteht und wann wird die physiologische Toleranz einer Art überschritten und ihre Existenz gefährdet?
- (2) Welche Arten können typische stellvertretende Arten einer ökologischen Typisierung der Lebensgemeinschaften sein und sind modelltechnisch so weit in ihrem räumlich-zeitlichen Verteilungsmuster beschreibbar, um als "Barometer" bei der Entscheidungsfindung in Managementfragen zu dienen?

Die Operationalisierung von wissenschaftlichen Erkenntnissen für die praktische Anwendung in raum- und umweltrelevanten Planungen ist eine aktuelle Forderung der Entscheidungsträger an die Wissenschaft. Statistische Ökosystemanalyse und Modellierung ist eine potenziell viel versprechende Technik, um tiefere Einblicke in die Funktionsweise und Dynamik von Biozönosen zu bekommen. Biozönotische Modelle unterliegen dabei aber in besonderem Maße einer kritischen Betrachtung, da sie unvollständiges und kaum quantifiziertes Wissen über Wechselwirkungen in ökologischen Systemen adäquat abbilden sollen (Asshoff 1999). Jeder Modellierer begibt sich auf den steinigten Weg des Gratwandels zwischen einer geforderten, aber oft unerreichbaren Exaktheit und der Banalität.

Die manifestierte enorme Komplexität von Auenökosystemen lässt auch nach vielen Jahren der Ökosystemforschung noch beträchtliche Wissenslücken bestehen, und bedeutende Anteile existierenden Fachwissens sind nicht jedem direkt zugänglich, wie Asshoff (1999, S. 1) sehr eindrucksvoll beschreibt: "Ein wesentlicher Teil des Fachwissens zu komplexen Fragestellungen im Bereich der Biozönotik besteht aus heuristischen 'Faustregeln', als diffuser Erfahrungsschatz in den Köpfen von Experten, der mit vielen Ausnahmen, Unstetigkeiten und Hintergrundwissen verbunden ist. Dieses subjektive Wissen kann aber in der Fachwelt durchaus verbreitet und als "common sense" anerkannt sein, so dass nicht von weniger Akzeptanz dieses Wissens gegenüber dem aus exakten Mess- und Forschungsergebnissen ableitbarem Wissen ausgegangen werden muss."

Dieses fundamentale Problem der Ökosystemforschung ist Anlass genug, um zu versuchen das über ökologische und biozönotische Zusammenhänge bestehende Wissen zu objektivieren, zu verifizieren, es durch weiterführende Strukturanalysen zu ergänzen und die Erkenntnisse zusammenzuführen. Der ökologische Wissenstransfer in die praktische Anwendbarkeit mit Modelltechniken erscheint dabei zukunftsweisend.

Jedes resultierende Modell besitzt Eigenschaften mit spezifischen Möglichkeiten und Grenzen, d.h. es ist für die Vielzahl biologisch-ökologischer Fragestellungen nicht gleichermaßen gut geeignet und immer eine Approximation und Abstrahierung der natürlichen Realität. Drastisch fassen dies Walters und Korman (1999, S. 414) zusammen: "Every model is constructed as a collection of nasty black boxes and statistical descriptions, no matter how the modeller may try to disguise these descriptions by explaining them in terms of physical and biological mechanisms."

Trotz des bestehenden Pessimismus kann ein Modell dennoch ein wichtiges Entscheidungsfundament sein, auch wenn es generell zu hinterfragen ist. Der Weg zum Modell muss nachvollziehbar sein und es sollte nicht in erster Linie nach seiner mathematischen Komplexität und seinem Detaillierungsgrad beurteilt werden, sondern danach, wie klar und

zuverlässig es Zusammenhänge, die für eine praxisgerechte Aussage und das bessere Verständnis der natürlichen Umwelt von Bedeutung sind, beschreiben und abbilden kann.

Intuitive Sprünge, Frustration und die Furcht vor schwerwiegenden Fehlern liegen eng beieinander und lähmen uns nicht allzu selten, den Schritt nach vorn zu wagen. Erst die Erkenntnis, dass ein ökologisches System nicht in allen seinen Zufälligkeiten erfasst und modelliert werden kann, bringt die Ökologie auf den richtigen Weg (vgl. auch Wissel 1989).

Es gilt zu lernen, dass nicht jedes Detail, welches dem Fachwissenschaftler von Bedeutung erscheint, auch zwingend notwendig ist, um auf entsprechender Skalenebene dem Schutz gefährdeter Räume zuträglich zu sein. Daher sollte man den Rat von Walters & Korman (1999, S. 412) speziell bei der Erarbeitung von Modellen als Prognoseinstrument für die Entscheidungsträger beherzigen:

- (1) "Discount claims about the possibility of modeling everything that might be important"
- (2) "Define the model structure and variables in reference to policy issues rather than scientific interests and disciplinary traditions"

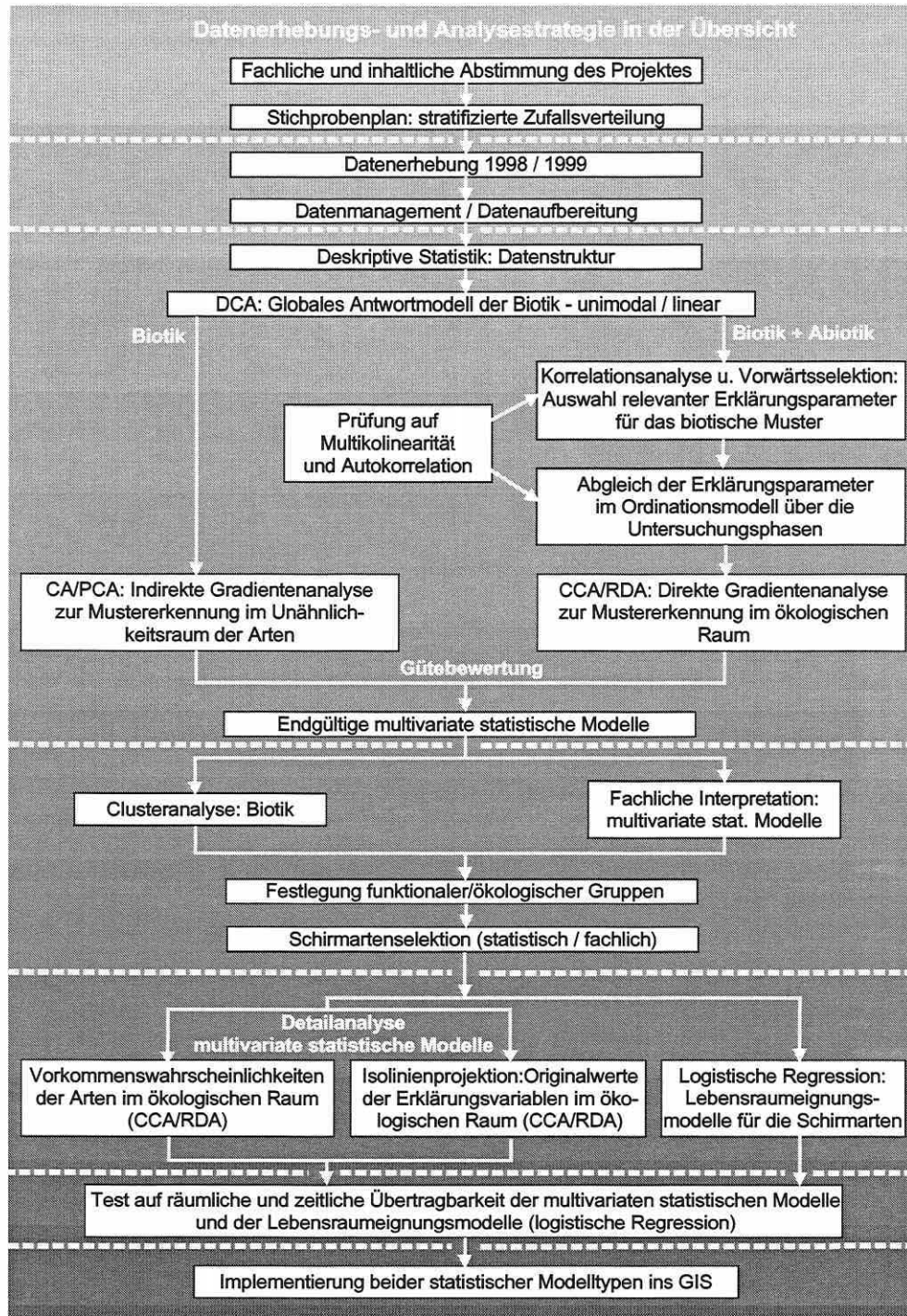
Natürlich muss ein Modell wissenschaftlichen Ansprüchen genügen, andererseits muss es aber auch in der Naturschutzpraxis verwendbar sein.

Im Gegensatz zu den hauptsächlich wegen der Binnenschifffahrt stark kanalisierten Strömen der alten Bundesländer ist die mittlere Elbe noch vergleichsweise unberührt und nicht zu einer modernen Wasserstraße ausgebaut (Kahlenborn & Kraemer 2000; Ständige Arbeitsgruppe der Biosphärenreservate in Deutschland 1995). Die besondere Bedeutung des Elbabschnitts um Dessau wird nicht zuletzt durch das seit 1979 existierende Biosphärenreservat unterstrichen, welches mehrfach erweitert 1997 zur "Flusslandschaft Elbe" erklärt wurde und sich mit etwa 370.000 Hektar über die Elbauen der Länder Brandenburg, Niedersachsen, Mecklenburg-Vorpommern, Sachsen-Anhalt und Schleswig-Holstein erstreckt (UNESCO 1998).

Der Nutzungsanspruch durch die Binnenschifffahrt kollidiert nun besonders deutlich mit den natürlichen Ressourcen. Die zum Teil drastischen Folgen des Wasserstraßenausbaus müssen aber Fehler der Vergangenheit bleiben. Um dieser vielschichtigen Aufgabe gerecht werden zu können, soll das Verbundprojekt mit dem Titel "Übertragung und Weiterentwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen" (RIVA) unter der Federführung des Umweltforschungszentrums Leipzig-Halle (UFZ) einen Teilbeitrag zu Optimierung der zukünftigen naturschutzfachlich-wasserwirtschaftlichen Raum- und Entwicklungsplanung leisten.

Für den Aufbau eines ökologischen Prognosesystems durch das an der BfG angesiedelte Teilprojekt, ist es wichtig, die Absicht des Modellierungsansatzes zu definieren: Mit der Modellierung soll die Veränderung der biotischen Ökosystemkomponenten, speziell der biotischen Deskriptoren, als Reaktion auf Änderungen abiotischer Umweltfaktoren (z.B. Flusswasserstand, Grundwasserstand) prognostiziert werden. Dabei soll eine Prognose des Artenverlustes, der Änderung der Artmächtigkeit, der räumlichen Verschiebung der Artenzusammensetzung und auf höherer Skalenebene eine Prognose der Verschiebung von Biotopen und Habitaten im Raum erreicht werden. Diese Zielsetzung fokussiert auf die Beantwortung der Frage, wie stabil und belastbar ein Auenökosystem ist und wann welche Veränderungen zu erwarten sind.

**Ziel dieser Promotionsarbeit** ist es aufzuzeigen, mit welchen Methoden der mathematisch-statistischen Biozönoseforschung Wissensdefizite aufgearbeitet und bestehendes Wissen zu einer integrativen modellbasierten Lösung zusammengeführt werden können. Die Anforderungen sind vielschichtig und sehr heterogen, so dass der Anspruch erhoben wird, gedankliche Konzepte und die Umsetzung der ganzheitlichen Analysestrategie zu dokumentieren



**Abbildung 1.0-1: Der Lösungsansatz in der Übersicht: Flussdiagramm zum Gesamtkonzept Datenerhebungs- und Datenanalysestrategie.**



und die noch wenig in der Ökologie verbreitete innovative Methodik nachvollziehbar aufzubereiten (vgl. Abbildung 1.0-1, S. 4), um die zukünftige Bearbeitung gleich gerichteter Projekte zu erleichtern.

Im Methodenteil wird der Leser mit der Theorie und Leistungsfähigkeit von Ordinationsverfahren und der logistischen Regression vertraut gemacht. Eine vergleichende Bewertung von Ordinationsverfahren zur Strukturanalyse ökosystemarer Feldinformation folgt, und deren Koppelung mit der logistischen Regression wird diskutiert.

Die praxisrelevante Anwendung des Gesamtkonzeptes am Beispiel des Projektes RIVA bildet den zweiten Fokus. Der Weg von der Datenerhebung im Feld über die Mustererkennung und Modellbildung mit Ordinationsverfahren bis hin zur Erstellung von Lebensraumeignungsmodellen für ausgewählte Arten der Elbauen und deren Möglichkeit zur Implementierung in ein GIS-basiertes Prognosesystem wird erörtert, um die konkrete Ergebnisfindung zu veranschaulichen. Daher ist ein Ergebnis dieser Arbeit, neben den statistischen Modellen selbst, auch der Weg zu diesen sowie die Bereitstellung einer methodischen Leitlinie für zukünftige Anwender.

Die Prüfung der räumlichen und zeitlichen Robustheit und Übertragbarkeit der Modelle ist ein hervorzuhebender Bearbeitungspunkt, da die Modellvalidierung generell eine zentrale Bedeutung hat und ihr speziell für den praxisrelevanten Nutzen des Prognosesystems besondere Relevanz zukommt.

**Die Arbeit gliedert sich neben der Einleitung und der Zusammenfassung in fünf Kapitel:**

Ausgehend von der Frage *"Warum Ordinationsverfahren?"* werden eingangs spezifische Erscheinungen in der Natur den gedanklichen Grundlagen der statistischen Modellierung gegenübergestellt, um Schwierigkeiten einerseits, aber auch die Entscheidung für die Analysestrategie andererseits zu erörtern und in den fachlichen Kontext der Ökosystemforschung zu stellen (**Kapitel 2**).

Die übergreifende Projektfragestellung, ein robustes Indikations-/Prognosesystem für das Grünland als ein Element der rezenten Aue zu entwickeln, nie aus den Augen verlierend, wird in einem Methodenteil die Kette der Verfahren in der Bearbeitungshierarchie ausführlich besprochen (**Kapitel 3**). Anschließend erfolgt ein methodenkritischer Vergleich von gradientenanalytischen Verfahren (Ordinationsverfahren) zueinander und mit generalisierten linearen Modellen (**Kapitel 4**).

Am Beispiel des Verbundprojektes RIVA wird nach einer Einführung in die Untersuchungsgebiete, die Anwendung ausgewählter Verfahren und Module zur Erhebung, Vorhaltung, Aufbereitung, Analyse und Modellierung realer Felddaten mit den jeweiligen Ergebnissen dargestellt (**Kapitel 5**).

Die Diskussion der Ergebnisse aus allen Bearbeitungen beinhaltet **Kapitel 6**. Die Arbeit schließt, indem gesammelte Erfahrungen sowie neue und alte unbeantwortete Fragen zusammengetragen werden. Der Ausblick wird vielleicht einen Interessierten animieren, an dieser Stelle Ideen aufzugreifen, diese weiterzuführen und zu optimieren.

## 2 WARUM ORDINATIONSVERFAHREN? DIE NATUR UND EINIGE KONZEPTIONELLE GRUNDLAGEN DER MATHEMATISCH-STATISTISCHEN ÖKOSYSTEMFORSCHUNG

Das Kapitel möchte den Leser hinführen zu den gedanklichen Konzepten, die sich hinter dem Begriff "Ordination" verbergen. In mehreren Abschnitten wird herausgearbeitet, warum die Gruppe multivariater statistischer Techniken als Instrumentarium für die Bearbeitung der Aufgabenstellung geeignet erscheint.

### 2.1 Von der Umweltrealität zum Modell: Möglichkeiten und Grenzen der Modellierung naturnaher Ökosysteme

*"THE SEARCH FOR PATTERNS AT ALL LEVELS IS ONE OF THE MOST CRITICAL PARTS OF THE SCIENCE OF ECOLOGY TODAY. (...) THE STUDY OF PATTERNS IS THE MOST IMPORTANT SOURCE OF CLUES, AND THE EXISTENCE OF SO MANY PATTERNS IS THE BASIS FOR OUR CONVICTION THAT ECOLOGY CAN BECOME A PREDICTIVE AS WELL AS A DESCRIPTIVE SCIENCE (WEINER 1995, S. 154/155)."*

Der grundlegende Streit, inwieweit sich das System Natur in seinen Zusammenhängen mathematisch beschreiben lässt, reicht weit in die Geschichte zurück, wie allein die fortwährende Diskussion über das "Lotka-Volterra-Modell" zur interspezifischen Konkurrenz zweier Arten beweist (Begon 1998; Richter 1985; Weiner 1995). Auch das intuitive Wissen, dass ein Modell den Kern eines Systems beschreibt, kann zu seiner Akzeptanz führen, und speziell dieses Modell regt allein durch seine interessante Dynamik zu Gedankenspielen an.

Allen Vorbehalten und Zweifeln zum Trotz hat sich die ökologische Modellierung in den vergangenen Jahren zu einem zentralen Werkzeug der Umweltsystem-Analyse und des integrativen Umweltmanagements entwickelt.

Modelle sollen handhabbare und in der Praxis einsetzbare Abstraktionen der **komplexen** Wirkungsgefüge innerhalb ökologischer Raumeinheiten sein (Müller 1997; Wissel 1989). Weiterhin sollen sie die heute auch für den Fachwissenschaftler kaum noch überschaubaren vielschichtigen dynamischen Zusammenhänge der **komplexen** Mensch-Umwelt-Interaktionen darstellen helfen und stets Fragen des Typus "Was wäre, wenn?" beantworten können (Wenkel 1997). Als methodische Eckpfeiler, um die beiden Ziele "sozioökonomische Entwicklung" und "Naturschutz" zu parallelisieren, kommt ihnen ein hoher Stellenwert zu.

Nach den notwendigen Eingangsgrößen eines Modells befragt, fordern die meisten Wissenschaftler einen ausreichenden Detaillierungsgrad des Modells bezogen auf ihr Fachgebiet. Schnell zeigt sich, dass die Berücksichtigung all dieser Belange einen nicht mehr zu leistenden Aufwand bedeutet und die **Komplexität** des Modells ins Unendliche wächst (Hauhs et al. 1996). Es müssen subjektive Entscheidungen getroffen werden, welche Parameter in das Modell einfließen sollen und welche ignoriert werden können. Grundsätzlich finden sich unzählige Argumente für und gegen Modelle in der Literatur (vgl. Hauhs & Lange 1996; Hauhs et al. 1996; Leland et al. 2000; Weiner 1995; Wissel 1989).

Der Leser wird sich zwischenzeitlich schon gefragt haben, weshalb der Terminus "komplex" wiederholt in den vorangegangenen Zeilen hervorgehoben ist.

*Komplexität ist ein Schlüsselwort in der Ökosystemmodellierung*

Hier liegt auch ein Widerspruch vor, den es zu ordnen gilt. Modelle sollen komplexe Zusammenhänge überschaubar machen, scheitern aber oft an exakt dieser Komplexität von Öko-

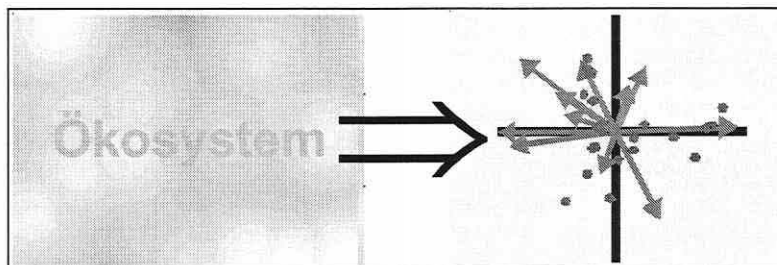
systemen schon in der Phase ihrer Erstellung. Zu Recht werden Gegenstände und Ergebnisse ökologischer Forschung als komplex angesehen, da Gesetzmäßigkeiten und eindeutige Abhängigkeiten oft nicht endgültig aus den teils selbst nicht klaren Organisationsprinzipien eines ökologischen Systems herausgelesen werden können (Hauhs & Lange 1996; Weiner 1995).

Dennoch, der Erfolg einfacher Vorhersagemodelle gibt einem eher reduktionistischen Ansatz bei der Modellerstellung recht. Modelle sollten nicht komplexer als unbedingt nötig sein um ihrem Zweck gerecht zu werden (vgl. Wissel 1989). Viele hervorragende Wissenschaftler entbehren allerdings dem notwendigen Pragmatismus, ein System auf seine essenziellen Eigenschaften zu reduzieren, um dann zielführenden Lösungsansätzen nachzugehen.

Der Vergleich eines Ökosystems mit einem idealen Gas von Hauhs & Lange (1996) mag erstaunen, dennoch möchte ich beipflichten, dass, wenn ein völlig regelloser Zustand auf der mikroskopischen Ebene mit einem einfachen und perfekt beschreibbaren Zustand auf der thermodynamischen Ebene korrespondiert, auch ein vergleichbarer Integrationssprung zwischen der abiotischen Ausstattung eines Raumes und einer intakten selbstorganisierten Biozönose bestehen kann. Das vielleicht einfache Funktionsprinzip ist durch die Konstruktionskomplexität verschleiert. Die Unzahl an irrelevanten Details wirkt als Barriere für unseren menschlichen Intellekt, die nötige Aggregation und Abstraktion zu leisten, um die wesentlichen ordnenden Elemente zu erkennen.

Man darf davon ausgehen, dass in jedem Ökosystem eine teils nicht direkt ersichtliche natürliche Ordnung durch Umweltparameter existiert (vgl. Pielou 1984). Ziel ist es, solche einen Gradienten ausbildenden Variablen aufzudecken und ihre Relevanz in der Strukturbildung abzuschätzen.

Hier liegt die Aufgabe von Ordinationsverfahren: Unterstützung bei der Ergründung einfacher Steuerungsprinzipien und objektive Repräsentation von Zusammenhängen in wenigen Dimensionen (Orlói 1978). Die Suche nach Mustern in Biozönosen und den wenigen Komplex-Gradienten, die sie bedingen, ist der Weg der Komplexität ein ordnendes Element abzurufen (Abbildung 2.1-1) und eine Modellierung denkbar zu machen.



**Abbildung 2.1-1: Vom diffusen Ökosystem zu einer ersten Ordnung durch das Ordinationsdiagramm.**

Das Referenzsystem bilden dabei die Beobachtungsobjekte, Populationen und Individuen, also diskrete biologische Variablen (Hauhs & Lange 1996).

Der reduktionistische Ansatz bei der Modellierung von Ökosystemen und die Vermeidung von Überparametrisierung führen zu einfachen Modellen, die in der Lage sind, mithilfe der Szenariotechnik die Bandbreite möglicher Entwicklungen abzustecken und eine verlässliche Aussage zu treffen, welches dieser Szenarios die gravierendsten Folgewirkungen haben wird. Die Reaktion eines Ökosystems und seiner Biozönosen auf Umweltveränderungen sind zwar nur in Grenzen vorhersehbar, doch auch ohne eine exakte Prognose aller Eventualitäten und Entwicklungspfade ist die Einordnung und Rangfolge verschiedener Eingriffsvarianten möglich. Der relative Aussagecharakter solch einer Vorgehensweise ist stabil genug den Vorwürfen von Modellierungsgegnern standzuhalten, welche die Aussagenunschärfe von Modellen monieren.



Was Ökosystemmodelle wirklich zu leisten vermögen und wo genau ihre Grenzen liegen, können erfahrene Experten verschiedener Disziplinen und Modellierer aber selbst nicht abschließend beantworten: "The workshop did not succeed in drawing a clear line between the possible and the impossible (Hauhs et al. 1996, S. 5)."

Nutzer, Manager und Wissenschaftler haben aber schließlich die Notwendigkeit erkannt, einfache Modelle zu entwickeln, und bescheinigen gerade diesen eine Existenzberechtigung, weil sie effektiv zur Lösung moderner Umweltrisiken beizutragen vermögen (Hauhs & Lange 1996; Hauhs et al. 1996; Weiner 1995; Wissel 1989 u.v.a.).

## 2.2 Biozönose, Muster und Gradient: Kontinuumtheorie kontra Diskontinuumtheorie

*"ALLE BEFUNDE ÜBER RÄUMLICHE VERÄNDERUNG (...) UNTER NATÜRLICHEN BEDINGUNGEN BESTÄRKEN UNS IN DER ANSCHAUUNG, DASS IN DER NATUR WEDER ORDNUNG NOCH UNORDNUNG, WEDER ORGANISATION NOCH CHAOS HERRSCHT, SONDERN EINE MISCHUNG AUS BEIDEM. DIE OBJEKTIVE REALITÄT KENNT WEDER ABSOLUTE IDENTITÄT NOCH IDENTITÄTSLOSES CHAOS (GLAVAC 1996, S. 73)."*

Ausgehend davon das Biozönosen existieren, was kontrovers diskutiert wird (Begon 1998; Keddy 1993; Palmer & White 1994), erscheint die nachfolgende Erörterung, wie diese sich im Raum verhalten, erst sinnvoll. Um eine Lebensgemeinschaft in ihrer räumlichen und zeitlichen Struktur richtig beschreiben zu können, ist deren grundlegendes Organisationsprinzip zu ergründen. Ein eigenständiger Wissenschaftszweig – die Biozönologie – befasst sich damit, wie Artengruppierungen in der Natur verteilt sind, welche Eigenschaften sie besitzen und wie sie durch interspezifische Wechselbeziehungen und physikalisch-chemische Umweltzwänge beeinflusst werden (Begon 1998). Erschwerend wirkt sich dabei aus, dass vielfältige inter- und intraspezifische Verknüpfungen eine Biozönose nicht nur als Summe ihrer Mitgliedsarten erscheinen lassen, sondern dass sie durch die wechselseitigen Beziehungen neue zusätzliche Eigenschaften (emergente Merkmale) erhält, deren "Gesetzmäßigkeiten" nur selten exakt erfassbar sind. Zum Beispiel kommt es durch Konkurrenz zu einer Verschiebung des physiologischen Optimums von Arten gegenüber ihrem ökologischen Optimum (vgl. Abschnitt 2.3). Lebensgemeinschaft und Ökosystem bilden dennoch keine unabhängige Einheit voneinander, sondern beeinflussen sich gegenseitig. Gleich, welche Raum- und Zeiteinheit zur Beschreibung gewählt wird, physiko-chemische und biologische Prozesse und Faktoren sind immer präsent (Palmer & White 1994). Dies schafft die Grundlage für eine Suche nach Wechselwirkungen zwischen Lebensgemeinschaft und abiotischen Faktoren, die deren Raummuster mitbestimmen.

Speziell die räumliche Verteilung von Lebensgemeinschaften oder Arten bezogen auf physikalische Standortfaktoren leitet zu dem Begriff des Gradienten und der Diskussion einer möglichen Abgrenzung entlang diesem über.

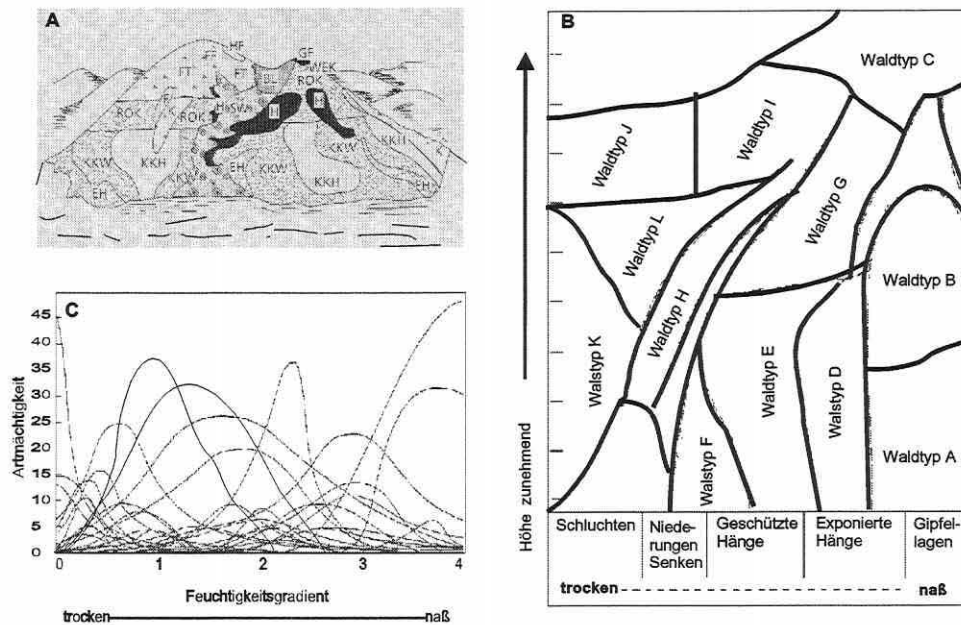
Gradienten, gebildet allein anhand der Variation in der Artenzusammensetzung, bezeichnet man als Gesellschafts-Gradienten oder Coenoclinen. Zwei unabhängige Gradienten innerhalb einer Artenstruktur bilden eine Antwortoberfläche oder Coenoplane (Whittaker 1967; Whittaker 1973; Økland 1990). Weiterhin definiert Whittaker den ökologischen Gradienten als den gerichteten Wechsel der Werte eines Umweltfaktors entlang dessen. Da Umweltfaktoren meist nicht unkorreliert voneinander sind und in ihrer Kombination Einfluss auf das biotische Muster eines Raumes nehmen, prägte er weiterhin den Begriff des Komplex-Gradienten. Er wird durch verschiedene Umweltparameter definiert, die mehr oder weniger parallel zueinander variieren. So verändern sich beispielsweise mit dem Komplex-Gradienten

"Höhe" entlang eines Berghanges auch parallel die Umweltfaktoren Niederschlag, Feuchte, Temperatur und mittlere Windgeschwindigkeit. Auch das Bodenwasser kann als Komplex-Gradient angesehen werden. Es bestimmt neben der Wasserversorgung die Verfügbarkeit an Bodenluft und Nährstoffen (Austin 1985). Aus der Korrelation eines Umweltparameters zur biotischen Struktur darf daher nicht zwangsläufig ein Kausalzusammenhang abgeleitet werden, sondern dieser Schritt ist durch den Fachwissenschaftler näher zu prüfen.

Der Begriff des Komplex-Gradienten kann auch auf eine Eigenschaft der Ordinationsachsen Bezug nehmen. Da Arten generell auf mehr als nur einen Umweltfaktor antworten und erst eine Parameterkombination ihren Standort bedingt, ist die Linearkombination von Umweltparametern als Komplex-Gradient wünschenswert. Genau das Prinzip wird bei der Bildung von Ordinationsachsen in direkten Gradientenanalysen umgesetzt und ist, da die Anzahl solcher Gradienten in räumlich definierten Untersuchungsräumen lokalen Maßstabs grundsätzlich gering ist (Økland 1986a, b), sehr effektiv, so lange die relevanten musterbedingenden Parameter zur Verfügung stehen. Die bestimmenden Zusammenhänge können in wenigen Dimensionen eines Ordinationsraums repräsentiert werden. Sind die den Komplex-Gradienten bildenden Parameter nicht selbst schon Komplex-Gradienten, wird die Interpretation der Art-Gradienten-Zusammenhänge deutlich vereinfacht.

Austin (1985) klassifiziert ökologische Gradienten weiterhin in drei Typen: (1) **indirekte Umweltgradienten**: Der Umweltparameter hat keinen direkten physiologischen Einfluss auf die Arten (z.B. geografische Höhe), (2) **direkte Umweltgradienten**: Der Umweltparameter hat einen direkten physiologischen Einfluss auf die Arten, ohne eine Ressource für deren Wachstum/Populationszuwachs zu sein (z.B. Temperatur) und (3) **Ressourcen-Gradienten**: Der Umweltparameter ist eine essenzielle Quelle für das Wachstum/den Populationszuwachs der Arten mit entsprechender Ressourcen-Konkurrenz (z.B. Nährstoffe oder Wasserversorgung in einzelnen Ökosystemen) (Austin 1985; Økland 1990).

Nach der Einführung des Gradienten als Referenzachse zur Beschreibung der Verteilung von Biozönosen/Arten im Raumgefüge ist die Rückkehr zu den grundlegenden Organisationsprinzipien von Lebensgemeinschaften notwendig. Abbildung 2.2-1 (S. 10) visualisiert die Diskrepanz zwischen zwei speziell in der Vegetationskunde besonders ausgeprägten Paradigmen, die unterschiedlichen erkenntnistheoretischen Positionen folgen: der **Kontinuum- und der Diskontinuumtheorie**. Diese Theorien wurden auf Grundlage von Beobachtungen an höheren Pflanzen erstellt, während das im Abschnitt 2.3 angesprochene Nischenkonzept in der Tierökologie maßgeblich auf Studien von Vögeln und Eidechsen beruht (Austin 1985). Darstellung (a) und (b) in Abbildung 2.2-1 suggerieren eine scharfe Abgrenzung der Arten/Lebensgemeinschaften und unterstützen damit die Diskontinuumtheorie, ähnlich wie das etablierte pflanzensoziologische Klassifizierungssystem und die statistischen Clusterverfahren. Darstellung (c) in Abbildung 2.2-1 hingegen verdeutlicht unmittelbar, dass es keine scharfen Grenzen gibt und die Lebensräume entlang eines Gradienten ineinander übergehen. Whittaker sieht die Kontinuumtheorie bewiesen: "The broad overlap and scattered centres of species populations along a gradient imply that most communities intergrade continuously along environmental gradients, rather than forming distinct, clearly separated zones (in Austin 1985, S. 40)." Austin (1985) diskutiert und bezweifelt dies. Gradienten sind eine abstrakte Dimension im ökologischen Raum und das Kontinuum konnte nie abschließend durch adäquate Mechanismen zur Erklärung bewiesen werden (der Beweis des Diskontinuums steht jedoch ebenfalls aus). Er hält Whittakers Aussage für einen Zirkelschluss. Seine Ansicht lautet: "Vegetation science has no theoretical basis at present, and a satisfactory synthesis of continuum ideas and niche theory is urgently needed, as are mechanistic models of how plant species behave along environmental gradients (Austin 1985, S. 56)."



**Abbildung 2.2-1: Darstellungen für die Raumverteilung von Arten: dreidimensional schematisch (A), zweidimensional schematisch (B) und als Antwortkurven entlang eines Gradienten (C) (Begon 1998, S. 475, verändert).**

Neuere Studien greifen das individualistische Konzept wieder auf (Begon 1998), welches besagt, dass zwar Arten mit ähnlichen Anforderungen an die Umwelt häufig gemeinsam auftreten, aber dennoch keine scharfen Grenzen zwischen Lebensgemeinschaften existieren. Jede Art hat ihre typischen genetischen und physiologischen Ansprüche, Entwicklungszyklen sowie Wechselwirkungen zur abiotischen und biotischen Umwelt, d.h. keine zwei Arten sind identisch in ihrer Verteilung (vgl. Abschnitt 2.5). Begon (1998) schreibt: "Lebensgemeinschaften, die eindeutig und klar voneinander abgegrenzt sind, kommen – wenn überhaupt – äußerst selten vor. Nicht einmal die scharf erscheinende Trennung zwischen terrestrischen und aquatischen Lebensräumen ist ökologisch signifikant." Hierbei spielt Begon auf die Lebenszyklen einer Art an, die durchaus beide Lebensräume integrieren können. Gerade ein extremer Sprung in den Umweltbedingungen entlang eines Gradienten auf engstem Raum ist für die floristische Artenzusammensetzung mit der Ausprägung einer recht deutlichen Grenze zwischen Lebensgemeinschaften verbunden. Faunistische Artengruppen antworten hierauf weniger deutlich, da sie durch ihre Mobilität keine so enge Raumbindung besitzen (siehe Abschnitt 5.4.2 und 5.4.3). Bei vielen statistischen Tests zur Stützung einer der Theorien wurde der Diversität in den Erhebungsdaten keine Aufmerksamkeit geschenkt. Økland (1990, S. 39) weist auf den klaren Einfluss speziell der Beta-Diversität hin: "Segments with high species-turnover are likely to contain many boundaries. Furthermore, the number of boundaries are likely to increase with increasing species richness."

Generell lassen sich Alpha- ( $\alpha$ -), Beta- ( $\beta$ -) und Gamma-Diversität ( $\gamma$ -) unterscheiden. Als Alpha-Diversität bezeichnet man die Diversität in der Artenkomposition auf einer Probe Fläche, als Gamma-Diversität die Heterogenität einer Region oder Landschaft (des Untersuchungsgebietes). Die Beta-Diversität ist ein Maß für die Rate des Wechsels in der Artenzusammensetzung (Species-Turnover) entlang eines Gradienten und kann beispielsweise in Standardabweichungen (SD) der Antwortkurven aller Arten in einer Studie gemessen werden (vgl. Abschnitt 2.3 oder 3.5.6). Sie ist entlang eines gegebenen Gradienten oft sehr verschieden. Ein Vergleich der Beta-Diversität pro Einheit zwischen zum Beispiel einem pH-

und einem Feuchte-Gradienten ist nicht möglich, dennoch kann die gesamte Beta-Diversität (Gradientenlänge, gradient length, Abschnitt 3.5.6) zwischen Gradienten verglichen werden. Kurze Gradienten besitzen generell eine geringe Beta-Diversität (Palmer 1998). Abbildung 2.2-2 verdeutlicht die Zusammenhänge zwischen Alpha- und Beta-Diversität.

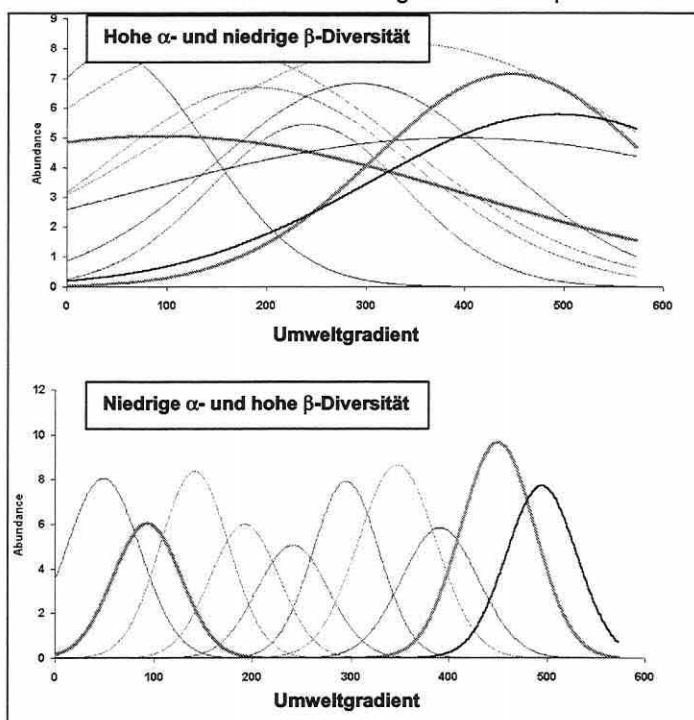


Abbildung 2.2-2: Beurteilung der  $\alpha$ - und  $\beta$ -Diversität anhand der Antwortkurven entlang eines Gradienten (Palmer 1998, coenospa.html, S. 2, verändert: <http://okstate.edu/artsci/botany/ordinate/>).

Trotz bestehender Unklarheiten erweist sich eine Existenz des Diskontinuums im Kontinuum als abschließend am wahrscheinlichsten.

Dies bedarf sicherlich einer Erläuterung.

Ein Kontinuum existiert durch die Überlappung der Lebensraumsprüche von Arten wie in Abbildung 2.2-3 (Teilbild A-C) dargestellt. Dennoch ist dies in der Realität kein lineares Kontinuum (Abbildung 2.2-3, Teilbild A), sondern ein diskontinuierlicher Übergang entlang eines Umweltgradienten. Zwischen räumlich benachbarten Biozönosen finden sich längere oder kürzere Übergangsbereiche, die sogenannten Ökotope. Zwischen den Ökotonen lassen sich Aggregie-

rungen von spezifischen Arten feststellen, deren ökologisches Optimum entlang dieses Bereichs des Gradienten liegt (Glavac 1996; Pfadenhauer 1993).

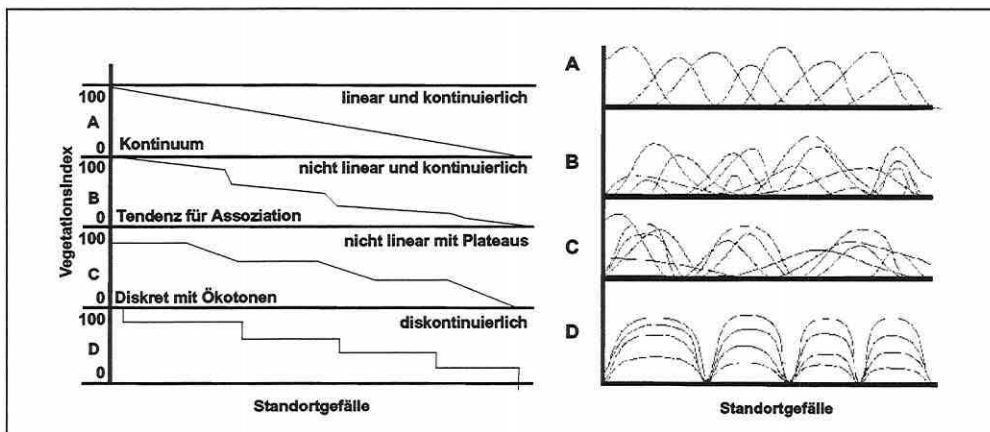


Abbildung 2.2-3: Räumliche Veränderungsmuster von Bestandsstrukturen (links) und Modellvorstellungen entsprechender Antwortkurven (rechts) (Glavac 1996, S. 66, verändert).

Es gibt also ein Kontinuum, welches jedoch deutliche Sprünge zwischen und Häufungen innerhalb der Bestände aufweist (Abbildung 2.2-3, Teilbild B). Ordnet man diese Erkenntnis in die Modelle der Abbildung 2.2-3 ein, so erweist sich Modell B "nicht linear/kontinuierlich" mit Tendenz zur Assoziation als der Realität häufig am nächsten. Der optische Eindruck des zugehörigen Antwortmodells B findet sich in den späteren Analysen der realen Daten wieder (vgl. Kapitel 5). Dieser Umstand erlaubt eine "Abgrenzbarkeit" von Lebensgemeinschaften oder einzelnen Arten bezogen auf beispielsweise einen physikalischen Standortparameter oder einen Parametersatz, der einen Komplex-Gradienten bildet. Letztlich kann auch Begon (1998, S. 478) diese Ansicht teilen: "Die sicherste Aussage, die wir treffen können, ist, dass es keine exakten Grenzen gibt, dass aber manche Biozönosen deutlicher gegeneinander abgesetzt sind als andere."

Speziell die in der vorliegenden Arbeit angewendeten gradientenanalytischen Verfahren (Ordinationen) sind geeignet, diesen Bedingungen gerecht zu werden, da sie im Gegensatz zu Klassifikationsverfahren (Clusterverfahren, s.u.) keine scharfen Gruppengrenzen und Hierarchien erzeugen (vgl. Backhaus 1990, 1996; Legendre & Legendre 1998), sondern die Tendenz einer ökologischen Gruppenbildung unter dem Kontinuumsaspekt und unter Einbeziehung von Standortparametern zulassen (vgl. auch Problem der Kettenbildung unten). Basierend auf der theoretischen Basis ist damit die Visualisierung von Mustern, Trends, Ausreißern und ökologischen Gruppen in den multivariaten Daten möglich.

Da sich Klassifikation und Ordination aber sehr sinnvoll ergänzen können, sollen die Basis-eigenschaften und die prinzipielle Arbeitsweise von Klassifikationsverfahren kurz angesprochen werden.

Klassifikationsverfahren streben die Zusammenführung ähnlicher Objekte in jeweils einer Gruppe unter Anwendung von Ähnlichkeits- bzw. Distanzmaßen an, wobei alle Ausprägungen der Objekte simultan berücksichtigt werden. Die Bildung möglichst homogener Gruppen und die Maximierung der Heterogenität zwischen den Gruppen (Zwischengruppen-Varianz) stellen das Ziel dar. Agglomerative, divisive oder hierarchische Cluster-Algorithmen fusionieren/trennen Objekte in immer größere/kleinere Cluster und bilden im Gegensatz zu nicht-hierarchischen Verfahren als ein typisches Ergebnis das so genannte "Dendrogramm" (hierarchical tree). Dieses zeigt die Zusammenhänge zwischen den Gruppen und in welcher Reihenfolge sie zusammengefasst/aufgespalten wurden (Backhaus 1990, 1996).

Klassifikation und Ordination sind zielführend kombinierbar (Gauch 1982a; Legendre & Legendre 1998), da Ausreißer durch Klassifikationsverfahren leicht entdeckt und aus dem Datensatz entfernt werden können, um die weitere Analyse – z.B. die Ordination – nicht unerwünscht zu beeinflussen (vgl. Abschnitt 3.1.7). Außerdem ist eine Gruppierung der interessierenden Objekte durch Klassifikationsverfahren verlässlicher, da sie frei von dimensions-bezogenen Restriktionen sind. Sie können berechnet werden ohne jeden Bezug zu einem Raum, metrisch oder nicht. Ordinationsdiagramme sind an eine maximal dreidimensionale graphische Darstellung gebunden sind, d.h., die Ausweisung von Gruppen ist im reduzierten Ordinationsraum durch gering-dimensionale Projektionen oft weniger exakt und Zusammenhänge bleiben im Verborgenen bzw. erscheinen obskur. Objekte werden im zweidimensionalen möglicherweise nahe beieinander abgebildet, während sie sich in einer höheren Dimension deutlich voneinander abgrenzen ließen. Zur Kontrolle sollten daher die Klassifikationsergebnisse in die Ordinationsdiagramme übernommen werden. Soll ein n-dimensionales Klassifikationsergebnis im Ordinationsraum visualisiert werden, zeichnet man die Gruppentrennung (Cluster) einfach durch Polygonlinien nach (vgl. Abschnitte 5.4.2 und 5.4.3).

Den Prinzipien von Ordinationsverfahren kommt dabei das Single-Linkage-Verfahren am nächsten, da es raumkontrahierend arbeitet (Legendre & Legendre 1998). Ein Objekt wird immer dem nächsten Nachbarn zugeordnet. Damit verknüpft ist allerdings die Eigenschaft der Kettenbildung, d.h. ein oder wenige Objekte, die mittig zwischen zwei kompakten Gruppen stehen, führen zu einer Vereinigung der Gruppen in einem Cluster. Dies ist natürlich nur



der Fall, wenn solche mittigen Objekte überhaupt auftreten, wodurch wiederum Rückschlüsse auf den Datensatz möglich sind (Legendre & Legendre 1998).

Weiterführende Informationen zu Klassifikationsverfahren finden sich in Backhaus (1990, 1996) und weiterhin in Kent & Coker (1992), Legendre & Legendre (1998), Pielou (1984) und Rink (1997a).

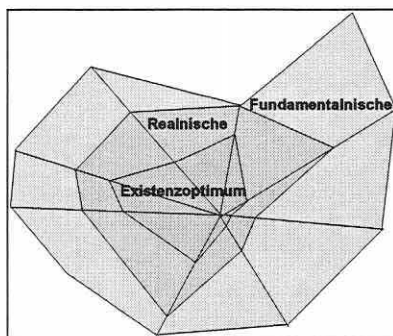
### 2.3 Nischenkonzept und ökologische Antwortmodelle

*"PLANT ECOLOGISTS HAVE EXPRESSED THEIR HYPOTHETICAL PATTERNS IN TERMS OF DISTAL FACTORS (E.G. MEAN ANNUAL RAINFALL) THAT DETERMINE THE BROAD PROPERTIES AND DYNAMICS OF AN ECOSYSTEM. NICHE ECOLOGISTS HAVE BEEN MORE CONCERNED WITH PROXIMAL FACTORS OF FOOD, COMPETITION, AND PREDATION (AUSTIN 1985, S. 57)."*

Diskutiert wird das realistischste Konzept zur Beschreibung der Antwort einer Art entlang eines Gradienten nicht nur in der Vegetationskunde. Das Gedankengut der Tierökologen ist sehr ähnlich (Austin 1985). Die erstmalige Definition des Begriffs "Nische" findet sich bei Grinnell schon 1917 (Asshoff 1999). Mehrere Konkretisierungen und Reformulierungen der Definition folgten. Die klassische Definition im Rahmen des Nischenkonzeptes geht auf Hutchinson im Jahre 1957 zurück (Begon 1998; Økland 1990). Hier liegt die Hypothese zugrunde, dass im wesentlichen ein Umweltfaktorensatz mit Einfluss auf eine Art deren Nische bestimmt. Im Unterschied zur Vegetationskunde stehen in der Zoologie häufig andere Parameter im Vordergrund, wie im Eingangszitat bereits erwähnt. Darüber hinaus ist das raumzeitliche Verhalten solcher Faktoren nicht notwendigerweise das gleiche. Austin (1985) geht davon aus, dass Ressourcen für Pflanzen häufiger monoton gerichtet sind (z.B. Nährstoffgradienten) als solche für Tiere (z.B. Nahrungsangebot oder Brutplätze). Jeder einzelne Faktor bildet eine Dimension der Nische. Es ist aber nicht davon auszugehen, dass die absolute Verbreitung einer Art von nur einem oder zwei Faktoren abhängt. Außerdem existieren keine exakten Begrenzungen (vgl. Abschnitt 2.2), da die Antwort einer Art gegenüber einem Umweltfaktor zwar immer in einem Grenzbereich gegen null geht, aber dies nur bedeutet, dass die Wahrscheinlichkeit des Auffindens dieser Art außerhalb des Bereichs sehr

gering ist. Erst durch die Einbeziehung weiterer Faktoren wird die wirkliche Nische einer Art, in der lebensfähige Populationen aufgebaut werden können, vorstellbar. Das Resultat ist ein mehrdimensionaler Hyperraum, der durch ein Set von Faktoren aufgespannt wird (Abbildung 2.3-1). Dieser imaginäre Raum fasst die Bedürfnisse einer Art zusammen (Begon 1998; Asshoff 1999). Projiziert in den realen Raum ergeben sich die Standorte, welche diese Ansprüche erfüllen.

Der Nischenbegriff, seine Differenzierung in Fundamentalnische und Realnische sowie der Einfluss exogener und endogener Faktoren ist in der ökologischen Literatur ausführlich beschrieben (z.B. Begon 1998; Dierschke 1994). Ergänzt wird die Thematik durch das Gesetz der relativen Standortkonstanz (Dierschke 1994, S. 33) und Modifikationen durch



**Abbildung 2.3-1:** Darstellung der Fundamentalnische, Realnische und des Existenzoptimums einer Art im mehrdimensionalen Hyperraum. Jede Achse bildet eine Dimension (entspricht einem Faktor) der Nische aus.

ökologische Kompensation (vgl. Ellenberg 1952, S. 22; Kuntze 1994, S. 199; Rink 1997c, S. 161; Wilmanns 1989, S. 27).

Je geringer die Konkurrenzkraft einer Art ist, desto weiter muss sie im natürlichen ökologischen Gefüge der Lebensgemeinschaft in Randbereiche ihres Wirklichkeitsfeldes/Potenzbereichs ausweichen (Begon 1998; Økland 1990; Rink 1997c; Walter 1986). Konkrete Ergebnisse von Experimenten solcher zwischenartlicher Konkurrenz und deren Auswirkungen auf Arten, wie beispielsweise der bekannte "Hohenheimer Grundwasserversuch", finden sich in Wilmanns (1989, S. 26), Dierschke (1994, S. 51), Rink (1997c, S. 162) und einige weitere in Begon (1998, S. 186ff). Eine weiterführende Diskussion zur Koexistenz durch Nischendifferenzierung, dem Konkurrenzausschlussprinzip und einer Reihe von Sonderfällen (z.B. unvorhersehbare Lücken durch Feuer) kann man Begon (1998), Mühlenberg (1993) oder Walter (1986) entnehmen.

Die Nische/Antwort einer Art entlang eines Gradienten in der natürlichen Lebensgemeinschaft bezieht sich auf die realisierte Nische. Diese Erkenntnis wird häufig in ökologischen Studien vernachlässigt (Austin et al. 1990), ist aber essenziell bei der Interpretation von Ordinationsergebnissen. Das Auftreten einer Art entlang eines ökologischen Gradienten spricht für deren Konkurrenzvorteil unter spezifischen Bedingungen, und das Überlappen von Antworten/Nischen zeigt eine Ressourcenkonkurrenz an (Austin et al. 1990). Die Breite der Realnische einer Art ist sowohl ein Maß für deren allgemeine Konkurrenzkraft als auch für deren physiologisches Potenzspektrum (Økland 1986c). Ein wichtiger Aspekt – auch im

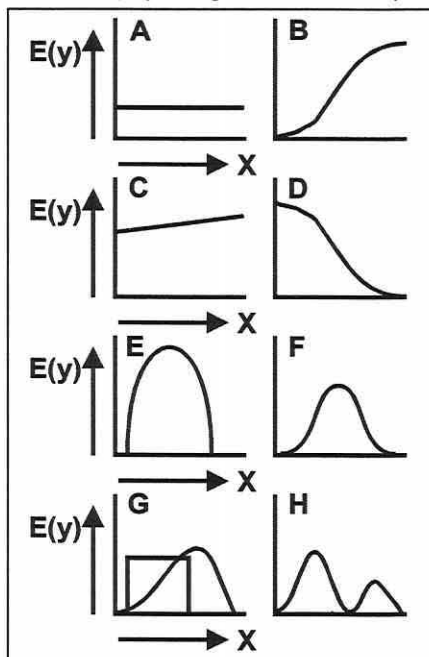


Abbildung 2.3-2: Antwortkurven: Die erwartete Antwort  $E(y)$  ist dem Gradienten einer Umweltvariablen ( $x$ ) gegenübergestellt. Konstante Antwort (A); monoton steigend: Sigmoidkurve (B) oder Gerade (C); monoton fallend: Sigmoidkurve (D); unimodal: Parabel (E), symmetrische Gaußkurve (F), asymmetrische Kurve oder Block-Funktion (G) und schließlich bimodal (H) (Jongman et al. 1987, S. 31, verändert).

Rahmen ökologischer Modellierung – ist, dass das Vorhandensein eines Lebensraums, der die Bedingungen der Realnische für eine Art erfüllt, nicht gleichbedeutend ist mit dem Vorkommen der Art an diesem Ort. Die Bedingungen für die Etablierung und die Erhaltung und Fortentwicklung einer Art sind nicht identisch. Die Etablierung wird auch bedingt durch das arteigene Ausbreitungspotenzial sowie den Isolationsgrad des potenziellen Lebensraums (Begon 1998; Glavac 1996).

Nachdem die theoretischen Grundlagen des Nischenkonzeptes eingeführt wurden, folgt nun der Schritt zum Modellgedanken der Antwortkurve. Die umfassende Diskussion von Verteilungsmustern von Arten und deren Antwort zu Umweltgradienten ist von entscheidender Bedeutung für die Anwendung von Ordinationsverfahren und Generalisierten Linearen Modellen (Kapitel 3). Neben der Suche nach Gradienten, welche die realisierte ökologische Nische einer Art definieren, ist das Modell, welches die Form dieser Antwort zu einem Gradienten beschreibt und abbildet, die Basis für die eigentliche Berechtigung solcher Verfahren bei der Analyse ökosystemarer Wechselwirkungen zwischen Biotik und Abiotik. Es ist nicht erstaunlich, dass trotz vieler Studien bisher kein konsistentes Modell zur Beschreibung von Art-Gradienten-Beziehungen konstatiert werden konnte, da die Natur sich in ihrer Vielfalt wohl kaum in ein einziges Modell "hineinpressen" lässt. Die Anzahl vorgeschlagener Modelle blieb trotzdem bis heute gering (Økland 1986a).



Eine Übersicht der bisher häufiger diskutierten Formen von Antwortkurven gibt Abbildung 2.3-2 (S. 14). Die verschiedenen Modelle sind bis auf die konstante Antwort (Abbildung 2.3-2, Teilbild A) von der oder den Variablen abhängig, die den Gradienten definieren. Die monoton steigenden oder fallenden Kurven können ein Sonderfall der Gaußkurve sein (Abbildung 2.3-2, Teilbilder B und D), wenn das Optimum außerhalb des beprobten Intervalls liegt und somit nur ein Teilbereich (eine Flanke) einer glockenförmigen Antwortkurve erfasst ist. Ähnliches lässt sich für die Gaußkurve als Teilstück einer bimodalen Kurve annehmen (Jongman et al. 1987). Um die realisierte ökologische Nische einer Art in ihrer Form, Position und Breite mit ausreichender Präzision abzubilden, muss ein geeignetes Modell herangezogen werden können (Austin et al. 1990). Die univariate Betrachtung einzelner Faktoren und die Ableitung eines fallweisen oberen und unteren Grenzwertes für das Auftreten einer Art ist auch direkt aus dem Datensatz möglich und bedarf keinerlei modellhaften Annahme. Sie führt aber im besten Fall nur zu wenig differenzierten Aussagen (Asshoff 1999; Begon 1998). Allein die Berücksichtigung quantitativer Maße – Artmächtigkeit, Individuenzahl, Vitalität etc. – als Gewichtungsfaktoren über das reine Anwesenheits-Abwesenheits-Prinzip hinaus lässt schon detailliertere Aussagen zu.

Letztlich wurde und wird die rein deskriptive symmetrische Gaußkurve als einfachstes und oft auch das am besten die Realität approximierende Modell angesehen (Austin 1987; Austin et al. 1984, 1990; Burrows 1990; Causton 1988; Gauch 1982a; Hill 1974; Jongman et al. 1987; Kent & Coker 1992; Legendre & Legendre 1998; van der Maarel 1969; Økland 1986b, 1990; Orłóci 1978; Palmer 1998; Smilauer 1992; ter Braak 1998; Thioulouse et al. 1995). Whittaker (1973, S. 35) propagierte schon 1951: "The general form for distribution of a species population along an environmental complex-gradient is the bell-shaped, apparently Gaussian or binomial curve."

Ruft man sich die Vorstellungen der Niscentheorie in Erinnerung, so erkennt man leicht die Parallelen zwischen ökologischer Antwort und statistischem Modell (Abbildung 2.3-3). In direkten Ordinationsverfahren wird ein Subraum der n-dimensionalen Realnische einer Art als multidimensionale Antwortoberfläche verstanden, die den Hauptanteil der Variabilität einer Spezies im Raum umfasst. Jede Art formt eine Populationswolke – dicht im Zentrum, dünn in den Randbereichen – im Ordinationsraum aus (Achsen des Nischen-Hyperraums), deren gewichteter Schwerpunkt (species score) ihr beobachtetes Optimum in einer kanonischen Korrespondenzanalyse anzeigt.

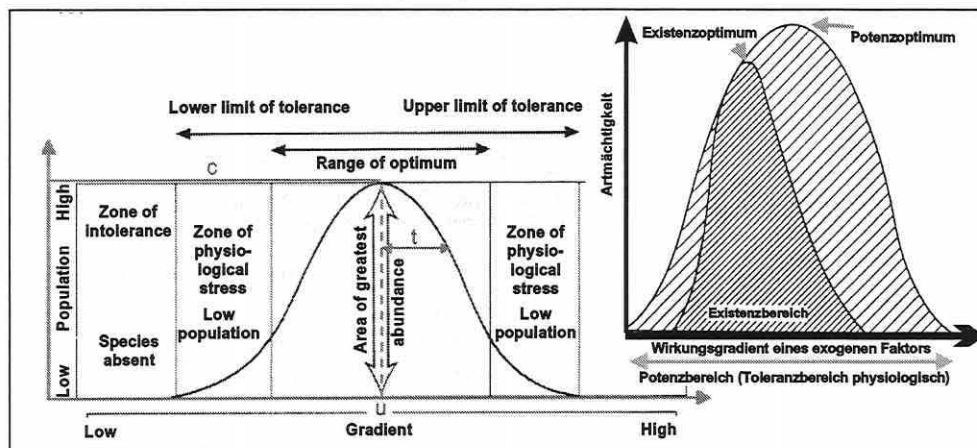


Abbildung 2.3-3: Modellvorstellung und Niscentheorie: Glockenförmige Antwortkurve nach Gauß mit Toleranzzonen einer Art und den drei beschreibenden Parametern c, u, und t sowie die Vorstellung und Begriffe der Niscentheorie (A) (Kent & Coker 1992, S. 14 und Dierschke 1994, S. 52, verändert).

Das Antwortmodell "Gaußkurve" entspricht der Normalverteilungskurve, welche über ein Intervall von annähernd vier Standardabweichungen steigt und wieder gegen null abfällt. Sie wird beschrieben durch die drei Parameter  $c$ : maximale Artmächtigkeit der Art an ihrem Optimum entlang des Gradienten,  $u$ : Optimum bzw. Position entlang des Gradienten und  $t$ : Toleranz der Art (Abbildung 2.3-3). Die häufiger in der "ökologisch-statistischen" Literatur verwendeten Notationen  $u$  und  $t$  (vgl. Jongman et al. 1987; ter Braak 1998) sind Synonyme für die klassischen Verteilungsparameter Standardabweichung ( $\sigma$ ) und Erwartungswert ( $\mu$ ).

Sie bestimmen die exakte Form der Wahrscheinlichkeitsdichte-Verteilung. Die Standardabweichung bezieht sich auf die Wendepunkte innerhalb der Kurve, während der Erwartungswert die Position des Optimums der Kurve gegenüber der Abszisse lokalisiert. Die Höhe der Kurve ist die Wahrscheinlichkeitsdichte korrespondierend zum Wert der Variablen  $x$ . Dieses Wissen wird für die Interpretation der Prognoseergebnisse in Abschnitt 5.6 herangezogen.

Dass die "Gaußkurve" in ihrer vollendeten Symmetrie nur eine Annäherung an die Wirklichkeit sein kann zeigen kritische Studien, welche am häufigsten schiefe unimodale Verteilungen, aber auch bimodale und plateauförmige nachgewiesen haben (Austin 1985, 1987; Austin et al. 1984, 1990). Austin et al. (1984) und Økland (1986b-c) können einige Einflussfaktoren belegen, welche die Form einer Antwortkurve modifizieren. Eine Annahme ist, dass dies besonders bei Arten auftritt, deren bestimmende Parameter nicht erfasst wurden (Austin et al. 1984; Økland 1986b) oder die Skalierung des Umweltparameters nicht mit der Reaktion der Biotik auf diesen Parameter harmonisiert (z.B. ist die Änderung um eine Einheit bei geringer Nährstoffkonzentration sicherlich bedeutender als bei einer bereits sehr hohen). Die Skalierung eines Gradienten beeinflusst deutlich die Form der Antwortkurven (vgl. Økland 1986b).

Speziell schiefe Verteilungen werden meist auch dem Umstand zugeschrieben, dass die Beta-Diversität entlang der Abschnitte eines Gradienten differiert (Økland 1986b). Ebenso führen ungeeignete Stichprobendesigns, zu geringe Stichprobenumfänge und Datenrauschen (noise, vgl. Abschnitt 2.6) zu Verzerrungen (Austin et al. 1984).

Bimodale Verteilungen (zweigipfelig) treten gelegentlich durch Irregularitäten im Datensatz auf, d.h. wenn ein wichtiger Parameter ignoriert wird. Die Erscheinung kann somit durch den Stichprobenplan und/oder die Datenstruktur selbst hervorgerufen werden. Bedenkt man, welche realisierten Nischen durch Konkurrenz auftreten können (vgl. Walter 1986, S. 73 oder Dierschke 1994, S. 172), so muss man bimodale Antworten innerhalb natürlicher Lebensgemeinschaften aber durchaus auch als real ansehen. Konkrete Beispiele liefert hierzu der Artikel von Austin et al. (1984). Leichte Bimodalitäten entstehen mitunter auch durch strukturloses Datenrauschen (Økland 1986b).

Plateauförmige Kurven dürften bedingt werden durch nicht in die Analyse miteinbezogene, aber für die betrachtete Art wichtige Gradienten (Austin et al. 1984; Økland 1986b) oder durch Sättigungserscheinungen.

Eine weiteres Problem rufen Komplex-Gradienten hervor: Da jede Art eine spezifische Antwort gegenüber einem der Umweltfaktoren in der Linearkombination aufweist, wird auch jede der Antwortkurven in Form und Position differieren. Es entsteht somit eine "mittlere Antwort" zum Komplex-Gradienten. Erst wenn der Wert eines relevanten Standortparameters außerhalb des Potenzbereichs einer Art liegt, agiert dieser als Ausschlussfaktor (Økland 1990).

Symmetrische Antwortkurven trifft man letztlich am wahrscheinlichsten an, wenn folgende Bedingungen erfüllt sind:

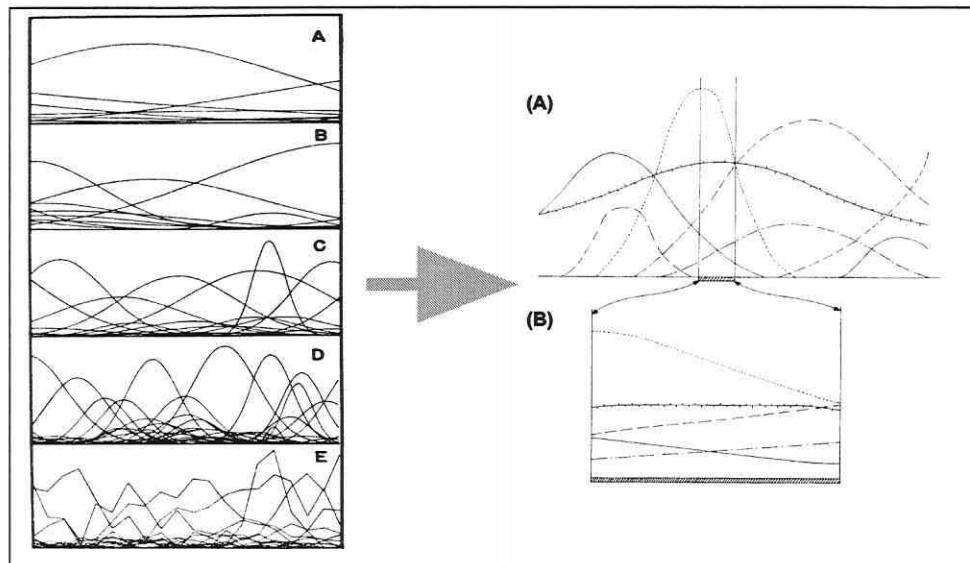
- (1) Es existiert ein relevanter dominanter Gradient ausreichender Länge.
- (2) Die Probeflächen sind adäquat entlang des Gradienten verteilt.
- (3) Der Stichprobenumfang ist groß genug, um eine Antwort zuverlässig zu charakterisieren.
- (4) Das Datenrauschen ist gering.
- (5) Die Skalierung des Gradienten ist passend.

Die letzte Bedingung hat eine besondere Relevanz bei vergleichenden Nischenstudien. Die Nischenbreite wird hier als Maß für die relative Bedeutung der Art in ihrer Lebensgemeinschaft und als Ausdruck der Konkurrenzkraft verstanden. Ist der Gradient in den Einheiten eines abiotischen Parameters skaliert, besteht oft keine Linearität entlang des Gradienten. Die Spanne der Antwortkurven in verschiedenen Abschnitten des Gradienten kann wegen schwankender Beta-Diversitäten nicht vergleichend interpretiert werden (Økland 1986c). Als eine geeignete Skalierung des Gradienten erweist sich die Standardabweichung des Wechsels in der Artenkomposition (Species-Turnover; vgl. Hill's Scaling in Abschnitt 3.5.5 und 3.5.6) (Økland 1986b).

Trotz aller Bedenken gegenüber der glockenförmigen Antwortkurve stellt sie eine meist tragbare Approximation der Art-Gradienten-Zusammenhänge dar. Alle gängigen Ordinationsverfahren arbeiten daher mit einem unimodalen oder einem linearen Antwortmodell.

*Weshalb arbeiten Ordinationsverfahren neben unimodalen Modellen auch mit linearen?*

Es herrscht weitreichender Konsens in der Fachwelt, dass die Zusammenhänge zwischen Biotik und Abiotik überwiegend nicht linearer Natur sind, wie beispielsweise Wildi (1986, S. 52) bestätigt: "In den seltensten Fällen linear ist gar derjenige [Anm.: Zusammenhang] zwischen Vegetation und Standort." Einige Sonderfälle erfordern dennoch den Einsatz eines linearen Modells: Zum einen sollen in manchen Studien Umweltfaktoren einfach linear miteinander kombiniert werden, um Komplex-Gradienten zu bilden, und zum anderen ist oft der beprobte Bereich eines Gradienten zu kurz, um ein unimodales Antwortmodell sinnvoll für die Beschreibung der Antwort anwenden zu können (Abbildung 2.3-4).



**Abbildung 2.3-4: Antwortkurven (links): A und B stellen Anwendungsbeispiele für lineare Modelle, C und D für unimodale dar. E zeigt die identische Situation wie D, jedoch mit dem typischen Grad an Datenrauschen (77%) für ökologische Felddaten (nach Gauch 1982a, S. 105). Der Zusammenhang zwischen unimodaler und linearer Antwort zeigt die Grafik rechts: Wird ein Gradient nur teilweise erfasst (B), so können die in Wirklichkeit unimodalen Antworten der Arten (A) nur linear beschrieben werden (Pielou 1984, S. 38, verändert).**

Dass ein lineares Modell nicht adäquat für die Beschreibung unimodaler Zusammenhänge bei langen Gradienten eingesetzt werden kann, veranschaulicht Abbildung 2.3-5 (S. 18). Die lineare Approximation der Kurve kann die eigentliche Struktur der Antwort nicht erfassen.

Der umgekehrte Fall führt ebenfalls zu Verzerrungen, wie das Beispiel zur Berechnung eines Art-Optimums entlang von Gradienten zeigt (Abbildung 2.3-6 und Tabelle 2.3-1). Das verglichen mit der Parameterschätzung einer Gaußkurve konzeptionell einfachere "Weighted Averaging"-Verfahren soll hier die Zusammenhänge veranschaulichen.

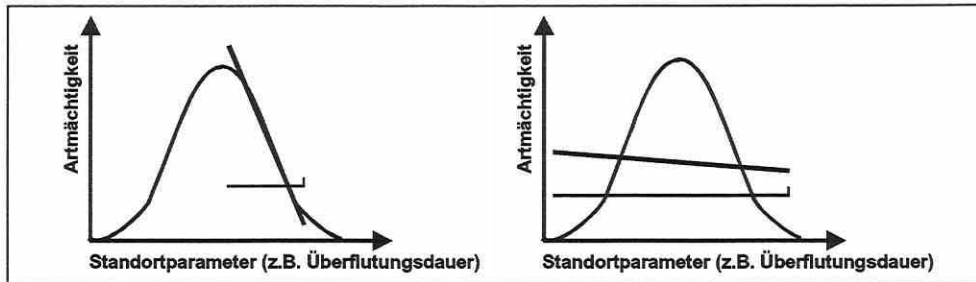


Abbildung 2.3-5: Lineare Approximation einer unimodalen Antwortkurve für einen kurzen (links) und einen langen Gradienten (rechts).

Es schätzt das Optimum effizient, vorausgesetzt, dass die Probestellen homogen über die Spanne der Art entlang des Gradienten verteilt sind und die Antwortkurve entlang des Intervalls vollständig abgebildet wird (Økland 1990; ter Braak & Looman 1987). Wendet man Gleichung 2.3-1 auf Abbildung 2.3-6 bzw. Tabelle 2.3-1 an, so ergibt sich eine deutliche Abweichung (Differenz von 4,8) bei der Berechnung des Optimums für die Beispiele "gesamte Antwortkurve" und "halbe Antwortkurve". Dies legt den Schluss nahe, dass die Wahl des Analyseverfahrens nicht beliebig sein darf, sondern den Datensatzeigenschaften angepasst sein muss.

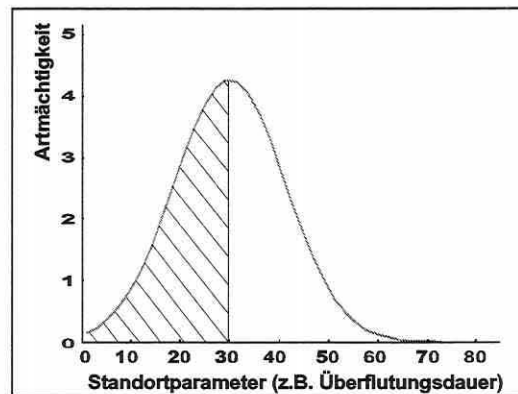


Abbildung 2.3-6: Unimodale Antwortkurve zu einem Gradienten. Die Schraffur stellt den in Tabelle 2.3-1 berechneten Teilbereich dar.

Gleichung 2.3-1: 
$$WA = \sum (WEV \cdot ARTM) / \sum ARTM$$

mit: WA: Weighted Average = Optimum der Art  
ARTM: Gewicht der Art (Artemächtigkeit o.ä.)

WEV: Wert der Erklärungsvariable

Tabelle 2.3-1: Berechnungsgrundlage und Berechnung des "Weighted Average" für die Antwortkurve und deren schraffierten Teilbereich in Abbildung 2.3-6.

Komplette Antwortkurve

Wert Erklärungsvariable (WEV)	Gewicht der Art (ARTM)	Produkt (WEV · ARTM)
0	0,1	0
10	0,5	5
20	2,0	40
30	4,2	126
40	2,0	80
50	0,5	25
60	0,1	6
Summe	9,4	282

$WA = \sum (WEV \cdot ARTM) / \sum ARTM = 282/9,4 = 30$

Schraffierter Teilbereich der Antwortkurve

Wert Erklärungsvariable (WEV)	Gewicht der Art (ARTM)	Produkt (WEV · ARTM)
0	0,1	0
10	0,5	5
20	2,0	40
30	4,2	126
Summe	6,8	171

$WA = \sum (WEV \cdot ARTM) / \sum ARTM = 171/6,8 = 25,2$

## 2.4 Gliederungsansätze: Soziologisch oder ökologisch?

*"JEDE FORM DER KLASSIFIKATION, ALSO DER SYSTEMATISCHEN ABGRENZUNG UND ORDNUNG VON OBJEKTEN, STELLT EINE ABSTRAKTION DAR. EINERSEITS KÖNNEN DIE OBJEKTE SELBST IMMER NUR AUF DER BASIS EINER BESCHRÄNKTEN ZAHL VON EIGENSCHAFTEN DEFINIERT UND ABGEGRENZT – D.H. AUCH WAHRGENOMMEN – WERDEN, ANDERERSEITS IST IHRE SYSTEMATISCHE ORDNUNG DURCH EINE BESTIMMTE SICHTWEISE GEPRÄGT (SCHOLLE 1997, S. 6)."*

Ein Gliederungskonzept, das der Typisierung von biotischen oder abiotischen Einheiten zugrunde liegt, wird nach den Prinzipien der Abgrenzung jener Typen definiert (Scholle 1997). Es ist daher festzulegen, welche Kriterien die Einheiten bilden (soziologische, abiotische etc.). Die sinnvolle Regionalisierung der punktuellen Ergebnisse ist dann davon abhängig, ob die Einheiten sich im Raum wiederfinden. Wird also eine ökologisch abgegrenzte Lebensgemeinschaft ausgewiesen, welche eine bestimmte ökosystemare Nische nachweist, dann ist deren räumliche Übertragung nur auf der Basis einer entsprechend gegliederten Karte realisierbar. Folgt die Flächeninformation einem nicht kompatiblen Gliederungsschema, können die indizierenden Einheiten nicht lokalisiert und zur Prognose herangezogen werden.

Flächenhafte zoologische Kartierungen sind basierend auf der für die meisten Tiere grundlegenden Eigenschaft "Mobilität" und den aufwendigeren Erhebungsmethoden kaum realistisch. Daher ist eher das Prinzip der Kalibrierung in der Zoologie gängig. Unter Kalibrierung versteht man, dass von einem Artenpool mit bekannten Nischenansprüchen auf die abiotische Ausstattung eines Standortes geschlossen wird. Die Prognose verfolgt den umgekehrten Weg, wirft aber in der Zoologie das Problem der Überprüfbarkeit von Ergebnissen auf. Hier sollten zusätzliche Stichproben zur Verfügung stehen (vgl. Abschnitt 3.7.1, 5.4.4 und 5.5), die eine wenigstens punktuelle Modellvalidierung zulassen. Das Auffinden der prognostizierten Arten ist durch die Erhebungsmethoden leider nicht sichergestellt, so dass Fehlinterpretationen auftreten können: Vorkommensprognosen zeigen Präsenz bei beobachteter Abwesenheit. Dieser Fehler 1. Art kann durch nichtgefundene Individuen bei der Datenerhebung verursacht werden (vgl. auch Abschnitt 3.7.1).

Die Validierung von Modellergebnissen in der Fläche bleibt überwiegend der Indikatorgruppe Vegetation vorbehalten. Exemplarisch wird die in der Vegetationskunde gängige soziologische Gliederung nach Braun-Blanquet mit einer ökologisch ausgerichteten verglichen.

Die **soziologische Gliederung** ist monohierarchisch und streng linear aufgebaut. Die daraus resultierenden Hierarchiestufen – wie in Clusteranalysen – bedingen jeweils kennzeichnende Merkmale zur Differenzierung. Es ergibt sich eine zweckmäßige Ordnungsstruktur, die aber inflationär ist. Dem Gedanken einer eindimensionalen Gliederung durch das alle Standortfaktoren integrierende Element Vegetation verhaftet, verlangt die Beschreibung jeder Faktorenkombination einen separaten Vegetationstyp (Scholle 1997).

Die soziologische Gliederung wurde auf der floristischen Ähnlichkeit – teils wurde verwandtschaftliche Ähnlichkeit vermutet – begründet, die letztendlich zwar auch durch die Summe der Standortfaktoren definiert wird, diese aber nicht direkt benennt. Die Ausgliederung ist heute oft nicht mehr fassbar, da beispielsweise Nutzungseinflüsse zu Verschiebungen in der Artenkomposition geführt haben.

Die **ökologisch ausgerichtete Gliederung** ist an das Konzept der floristischen Ähnlichkeit angelehnt, legt jedoch ihr Augenmerk auf die Zusammenfassung von Lebensgemeinschaften mit analogem Verhalten gegenüber der betrachteten Standortfaktorenkombination. Diese Kombination kann ein nach der Fragestellung ausgerichtetes Faktoren-Subset des Hyper-



raums sein, der die vollständige Nische abgrenzt. Im Vordergrund steht die Koinzidenz zwischen Arten/Artenkomposition und Standortfaktoren (Scholle 1997).

Exakt diese Koinzidenzen zwischen Biotik und Abiotik und die daraus resultierenden ökologisch begründeten Einheiten sollen der Fragestellung gemäß über Ordinationsverfahren ermittelt werden und dienen dann als Eingangsgröße einer ökologisch ausgerichteten Gliederung. Ein entscheidendes Argument für die spätere Anwendbarkeit der Resultate ist, dass die soziologische Gliederung durchaus mit einer ökologischen harmonisiert werden kann. Ausgehend von den Grundprinzipien der Gliederungen zeigt sich eine Schnittstelle über den Ansatz der floristischen Ähnlichkeit. Angesichts der häufigen Verwendung der Braun-Blanquet Methode spielt die Kompatibilität eine bedeutende Rolle für die Etablierung von ökologisch basierten Prognosemodellen in der Naturschutzpraxis.

Kurz zusammengefasst bedeutet das:

- (1) Die Indikatoreinheiten müssen sich in dem flächig vorliegenden Gliederungsschema wiederfinden und
- (2) die Gliederungssysteme müssen den Anforderungen der Fragestellung entsprechen, verschiedenen Fachdisziplinen zugänglich und möglichst mit gängigen naturschutzfachlichen Konzepten kompatibel sein.

## 2.5 Stellvertretende Arten für ökologischen Gruppen

*"EIN REGIONALES ZIELARTENSYSTEM WIRD (...) VERSTANDEN ALS EINE MÖGLICHST GERINGE ANZAHL VON ZIELARTEN (ÜBERSCHAUBARKEIT), DIE GLEICHZEITIG STELLVERTRETEND UND REPRÄSENTATIV (GRÖßTMÖGLICHE MITNAHMEWIRKUNG) FÜR MÖGLICHST VIELE ARTEN, LEBENS-GEMEINSCHAFTEN, LANDSCHAFTS- UND LEBENSRAUMTYPEN EINER BEZUGSREGION IST UND DAHER AUCH FÜR EINZELNE LEBENSRAUMTYPEN STETS MEHR-ARTEN-SYSTEME DARSTELLT (VIELGESTALTIGKEIT) (ALTMOOS 1999, S. 44)."*

Das Zielartenkonzept ist eine Naturschutzstrategie und bietet eine pragmatische Möglichkeit, die gesamte Artenanzahl eines Betrachtungsraumes auf eine handhabbare Größenordnung relevanter und erfassbarer Arten zu reduzieren (Foeckler & Bohle 1991). Der Begriff "Zielart" (target species, management indicator species) wurde im deutschsprachigen Raum maßgeblich von Hovestadt et al. (1991) etabliert (Altmooß 1999). Um den Schutzgedanken zuverlässiger zu gewährleisten, sind anstatt des stellvertretenden Schutzes einer Art je ökologischer Gruppe Zielartensysteme entwickelt worden. Mehrere Arten stehen somit stellvertretend für eine Gruppe. Die Berücksichtigung nur sehr weniger Arten führte vielfach zu einer begründeten Skepsis am Gesamtkonzept, da deren Auswahl Unsicherheiten und Willkür nicht gänzlich ausschließt. Die inflationäre Verwendung der Begriffe "Zielart, Leit- und Zeigerart" ist häufig, obgleich die Termini unterschiedliche Sinngehalte besitzen und eine Trennung erfordern. Die Arten können zwar gleichzeitig die verschiedenen Kriterien erfüllen, müssen aber nicht grundsätzlich allen drei Definitionen gerecht werden (Altmooß 1999). Die nachfolgende beschreibende Begriffsbestimmung ist annähernd wörtlich aus Altmooß (1999, S. 36-37) übernommen:

**Zielarten** sind alle Arten, die vorrangig zu erhalten und zu fördern sind. Ihre Auswahl erfolgt nach naturschutzstrategischen Kriterien. Zielarten sollen entweder derart ausgewählt werden, dass beim Schutz ihrer Lebensräume möglichst viele weitere Arten und Naturgüter mitgeschützt werden (Mitnahmewirkung, Schirmeffekt, umbrella effect), oder dass sie unabhängig eventueller Mitnahmewirkungen per se besonders schutzwürdige Arten darstellen. Die

Auswahl erfolgt nicht nach einer wissenschaftlich vorgegebenen Methode, aber baut auf naturwissenschaftlichen Erkenntnissen auf. Eine geeignete Auswahl wird durch fachliche Begründung, Logik, Nachvollziehbarkeit und Reproduzierbarkeit angestrebt.

**Leitarten** im Sinne von Charakterarten sind Arten, die in spezifischen Verbreitungsarealen höhere Stetigkeiten und in der Regel hohe Populationsdichten erreichen. Sie sind fast ausschließlich an eine begrenzte Lebensstätte oder Lebensgemeinschaft gebunden (Schubert & Wagner 1988).

**Zeigerarten** (Indikatorarten) sind Arten, die durch ihr Vorkommen oder bestimmte Lebensäußerungen auf Umweltzustände schließen lassen, die sonst nicht sichtbar oder nicht direkt messbar sind. Ihre Auswahl ist das Ergebnis naturwissenschaftlicher Kausalforschung. Die zahlreichen Kriterienkataloge als methodischer Rahmen (Altmöös 1999; Hovestadt et al. 1991; Mühlenberg 1993; Vogel et al. 1996) sollen hier nicht im Einzelnen diskutiert werden. Eine Übersicht der Kriterien gibt Tabelle 2.5-1. Die Ausschlusskriterien sollten für eine Zielart alle erfüllt sein und von den Einzelkriterien zumindest eines. Die Bedeutung einer Spezies als Zielart lässt sich durch eine Rangfolge definieren, die auf der Anzahl erfüllter Einzelkriterien (naturschutzstrategische Bedeutung), dem Gefährdungsgrad (Dringlichkeit) und der Lebensraumkomplexität (Schirm-, Mitnahmeeffekt) beruht.

**Tabelle 2.5-1: Kriterien und Prüfmethode zur Ableitung von Zielarten. Ausschlusskriterien: Erfüllung erforderlich; Einzelkriterien: Zusatzkriterien, von denen wenigstens eines erfüllt sein sollte (nach Altmöös 1999, verändert)**

Zielart-Kriterium		Kriterienprüfung			
Kürzel	Bezeichnung	Verbreitung	Autökologie	Rote Liste	Ermen
<b>Ausschlusskriterien</b>					
Meth	methodisch gut erfassbar		x		x
Heim	potenziell heimisch / regionsangepasst	x	x	x	
Überleben	wirkliche Überlebenschance	x	x	x	x
Komplex	Schirmeffekt / komplexe Raumansprüche		x		x
Attr	attraktiv für Öffentlichkeitsarbeit				x
<b>Einzelkriterien</b>					
Über	überregionaler Verbreitungsschwerpunkt im Betrachtungsraum	x	x		
Rand	Randbereich der Verbreitung mit Bedeutung für weitere Ausbreitung	x			
Relikt	Endemit / Reliktvorkommen: erhöhte Verantwortung zur Erhaltung	x		x	
Gef	überregional gefährdet und/oder selten	x		x	x
Aus	geringes Ausbreitungs-, Etablierungspotential: Lebensraumzerstörung besonders gravierend		x		x
Stör	besonders störungsempfindlich: Qualitätszeiger		x		x
Ind	Zeigerarten mit ökonomischen Wert in der Forschung		x		
Schlüssel	Arten mit Schlüsselfunktionen in Ökosystemen		x		x

Erwähnenswert erscheint, dass die Gefährdung einer Art in der Praxis oft als Primärkriterium für die Auswahl herangezogen wird. Eine Begrenzung auf gefährdete Arten (Rote Listen), ist nicht wünschenswert, da naturschutzstrategisch auch andere Arten bedeutsam sind (Altmöös 1999). Außerdem ist der wirkliche Gefährdungsgrad einer Art oft nicht abschließend feststellbar, wodurch unter anderem "Rote Listen" einen Mangel an Zuverlässigkeit aufweisen (Mühlenberg 1993). Schirmeffekte sind skalenbezogen, d.h. sie können lokal einen Lebensraum indizieren (z.B. Wassermolluskenart: Temporärgewässer einer rezenten



Aue mit typischer Lebensgemeinschaft) oder aber eine Strukturlandschaft mit verschiedenen Qualitäten unspezifisch charakterisieren (z.B. Großvogelart: unterschiedliche Lebensräume in einer Gesamtlandschaft).

Abschließend wird wiederum die Brücke zu Ordinationsverfahren und Lebensraumeignungsmodellen geschlagen. Die Charakterisierung eines Untersuchungsraumes und die Erstellung von Prognosemodellen zum Praxiseinsatz kann nicht für die gesamte Fülle des erhobenen Materials durchgeführt werden. Ein reduktionistischer Ansatz ist hier notwendig und zielführend zugleich, da allgemein verbreitete Arten ebenso wie nur vereinzelt auftretende Arten entweder nicht differenziert zur Typisierung eines Lebensraums beitragen oder nur unzuverlässig eine Antwortkurve ihrer ökologischen Realnische beschreiben (siehe auch Foeckler & Bohle 1991). Das eine Reduktion des modelltechnischen Prognoseaufwands zwingend ist, zeigt zudem die Planungspraxis. "Angesichts der knappen Haushaltslage und des politischen Sparklimas (Amler et al. 1999, S. 20)" folgt, dass die Realität bei Erfassungsmethoden und -intensitäten in Gutachten anderen als nur den wissenschaftlich geforderten Bedingungen entspricht (siehe Abbildung 2.5-1). Daher werden für die ökologisch gegliederten Artengruppen typische Vertreter bei gleichzeitig guter statistischer Repräsentativität im Datenmaterial selektiert (Abschnitt 5.4.3). Die Definition einer Zielart- oder Zeigerart ist hierbei nicht exakt anwendbar, so dass in dieser Arbeit die verschiedenen Ansätze kombiniert werden und letztlich von typischen Arten oder Schirmarten bzw. Schirmartensystemen gesprochen wird.

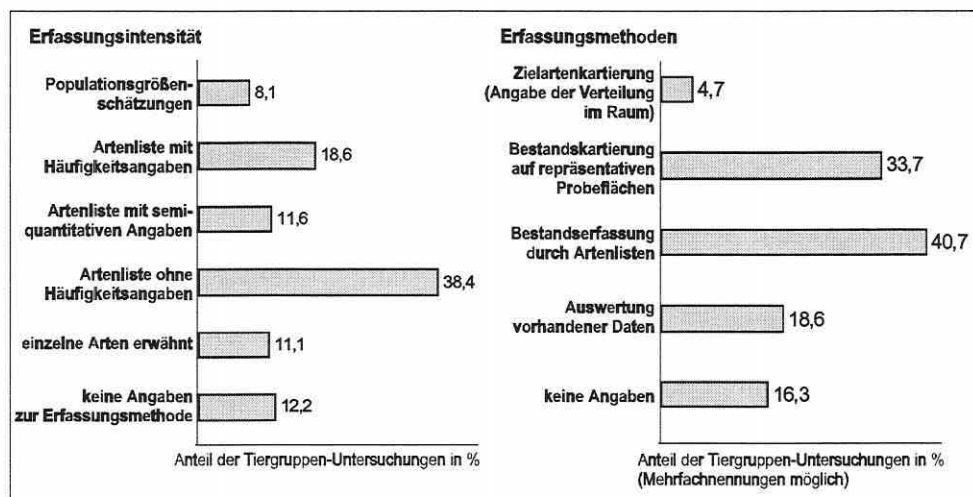


Abbildung 2.5-1: Erfassungsmethoden (links) und Genauigkeit der Bestandsgrößenhebung in 42 zoologischen Gutachten (erfasst von Amler et al. 1999, S. 22/23).

## 2.6 Betrachtungsräume: Kleinräumige Heterogenität und Skalenübergänge

*"THE COMMUNITY SCALE IS RECOMMENDED BECAUSE: (...) FIRST THIS IS THE SCALE AT WHICH POPULATIONS AND INDIVIDUALS OF A PLANT SPECIES CAN BE IDENTIFIED AND GROUPED TOGETHER TO CHARACTERISE THE VEGETATION OF AN AREA OF A FEW SQUARE METRES TO SEVERAL SQUARE KILOMETRES (KENT & COKER 1992, S. 13)."*

Die Textstelle umfasst zwei weitere Gesichtspunkte: "Second, the community scale is important because it is at this scale that humans can make best sense of the nature and varia-

tion of the vegetation cover on the earth. Third, it is at this scale that human activity in changing vegetation cover takes place and thus conservation and environment management practices and policies may be applied (Kent & Coker 1992, S. 13)." Eine geeignete Betrachtungsskala für viele Studien ist somit begründbar (vgl. Goldschmidt 2000).

Die Problemfelder: Kleinräumige Heterogenität und Skalenübergänge müssen im Rahmen einer solchen Arbeit erwähnt werden, können aber dennoch im Projekt "RIVA" nicht befriedigend gelöst werden. Sie sind daher nur bei der Interpretation der Ergebnisse und speziell deren Generalisierbarkeit gebührend zu berücksichtigen.

Der "Streit" um die optimale Stichprobenflächengröße für Felderhebungen ist sowohl in der Vegetationskunde als auch in der zoologischen Forschung ungelöst. Allein die zahlreichen Argumente um das Minimumareal einer vegetationskundlichen Aufnahme fläche füllen bis 1969 schon 170 Publikationen (Tüxen 1970), so dass auf eine Ergänzung dieser Diskussion verzichtet wird. Generell wird der Standpunkt vertreten, dass eine Erhebungsfläche groß genug sein muss, um die Lebensgemeinschaft/den Artenpool adäquat zu erfassen, aber gleichzeitig klein genug, um schnell und effizient bearbeitet werden zu können.

Die Komplikation des Diskussionsfeldes durch Effekte der **kleinräumigen Heterogenität** soll dennoch näher beleuchtet werden, da diese in ökologischen Studien kaum erwähnt wird.

Vergrößert man die Flächen einer Erhebung, wird die Variabilität der Umweltfaktoren innerhalb dieser ebenfalls größer. In der traditionellen Vorgehensweise zur Bestimmung von Wechselwirkungen zwischen Biotik und abiotischen Standortfaktoren wird den beobachteten

Artmächtigkeiten einer Grundfläche ein konkreter Messwert für den jeweiligen Umweltparameter zugewiesen (Abbildung 2.6-1). An sich ist dies inkorrekt, da die Erhebungsfläche keinen Punkt entlang des Gradienten eines Umweltfaktors darstellt, sondern eher eine kleine Region mit einer Spanne veränderlicher abiotischer Konditionen. Es können daher Arten oder Artmächtigkeiten auftreten, die nicht mit dem gemessenen Wert des Umweltfaktors direkt assoziiert sind. Das Problem des effektiven Parameters stellt sich aber nicht nur in der Fläche, sondern auch in der Tiefe. Hier integrieren Wurzelsysteme von Pflanzen über eine Tiefenstufe mit wechselnden Umweltbedingungen.

Es lässt sich schließen, dass die kleinräumige abiotische Variation innerhalb einer Erhebungsfläche – völlig homogene Flächen existieren in der Natur nicht – die beobachtete Nische einer Art durch die Scheinzunahme der Artenkomposition vergrößert (Palmer 1994; Palmer & Dixon 1990). Die Nischenbreite wird somit generell überschätzt, wenn die kleinräumige Heterogenität eines Raumes

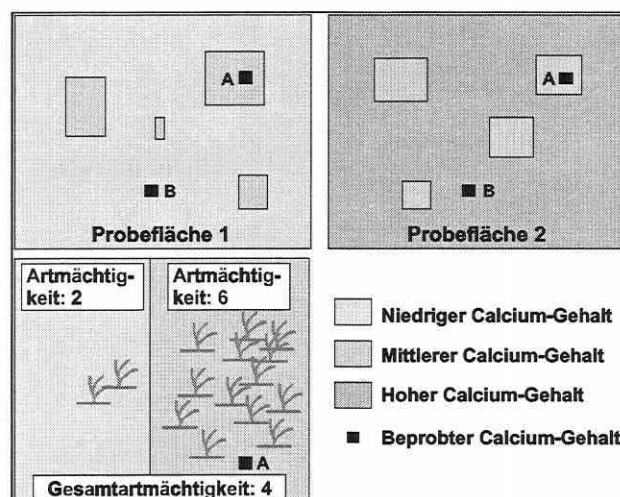


Abbildung 2.6-1: Zusammenhang zwischen kleinräumiger Heterogenität innerhalb einer Probefläche und dem Erhebungswert eines Umweltparameters: Der Erhebungswert (A) ist identisch für beide Probeflächen, während Erhebungswert (B) differiert zwischen den Probeflächen. In beiden Fällen wird ein unterschiedlich großer Fehler bei der Schätzung der Artenkomposition durch die wechselnden Ca-Konzentrationen innerhalb der Probeflächen begangen. Die Intensität dieses Effekts bedingt sich aus dem Fall für Erhebungswert (A) bzw. (B) je Probefläche. Die Graphik unten vermittelt den Fehler am konkreten Beispiel: Dem Messwert (A) würde eine Artmächtigkeit von sechs entsprechen, durch die Betrachtung der gesamten Probefläche erhält er jedoch eine Gesamtartmächtigkeit von nur vier zugeordnet.

nahme der Artenkomposition vergrößert (Palmer 1994; Palmer & Dixon 1990). Die Nischenbreite wird somit generell überschätzt, wenn die kleinräumige Heterogenität eines Raumes

nicht berücksichtigt werden kann. Diese Zusammenhänge korrespondieren mit den Grundprinzipien der Diversitätsforschung auf einer niedrigeren Skalenebene: "Environmental heterogeneity is usually one of the leading explanations for high species richness (Palmer & Maurer 1997, S. 239)."

**Skalenübergänge** stellen ein Problem bei der abiotischen Parameterauswahl und der Generalisierung der Ergebnisse dar. Innerhalb der Hierarchie der Skalenbereiche sind die relevanten Parameter für Prozesse und für einen Wirkzusammenhang zwischen Biotik und Abiotik nicht identisch und von Skala zu Skala übertragbar. Daher müssen die Daten der faunistischen, botanischen und abiotischen Erhebungseinheiten kompatibel zueinander sein. Verdeutlicht werden die Skalenübergänge anhand eines Beispiels aus der Hydrologie.

Die Grundwasserströmungs- und Transportvorgänge obliegen verschiedensten Parametern bzw. werden durch sie beschrieben. Sie sind in der Regional- (km), der Feld- (300 m), der Standort- (m), der Poren- (mm) und der Molekularskala ( $\mu\text{m}$ ) vollkommen unterschiedlich (vgl. Kobus 1995, S. 136). Der Bezug zwischen Umweltfaktor und Lebensgemeinschaft ist deshalb sicherlich nicht auf allen Ebenen gegeben.

Steinhardt (1998, S. 9) stellt weiterhin einen guten Überblick zu den Maßstabsniveaus geowissenschaftlicher Forschung bereit.

Skalenabhängige Erklärungsvariablen stellen für Ordinationsverfahren nicht das Problem dar, weil fehlende Erklärungspotentiale zwischen Umweltparametern und Biotik leicht über die Vergleiche von indirekten und direkten Ordinationsergebnissen entdeckt werden können. Ein Prognosesystem wird durch diesen Sachverhalt jedoch auf eine Skalenebene festgelegt und in seiner Übertragbarkeit bzw. Gültigkeit limitiert.

In diesem Kapitel war es sicherlich nicht möglich, alle Aspekte zu behandeln, welche ökologische Theorie und Ordinationsmethoden verbinden, dennoch zeigt es Hintergründe für die Auswahl der Verfahren zur Lösung der Aufgabenstellung in der Arbeit auf.

### 3 METHODEN ZUR ERHEBUNG, GEWINNUNG, VORHALTUNG, STRUKTURANALYSE UND MODELLIERUNG ÖKOSYSTEMARER FELDMINFORMATION

#### 3.1 Datenerhebung

##### 3.1.1 Allgemeine Vorüberlegungen

*"EVEN IN PUBLISHED WORK, SAMPLING DESIGNS ARE OFTEN INSUFFICIENTLY WELL EXPLAINED OR SOMETIMES IGNORED ALTOGETHER, YET CHOOSING THE RIGHT APPROACH IS A VITAL PART OF ANY (...) SURVEY WORK (KENT AND COKER 1992, S. 52)."*

Methodisch exakt abgestimmte und erhobene Daten sind das entscheidende Glied, um interpretierbare sinnvolle Ergebnisse aus einer Studie zu gewinnen. Die Zusammenführung ökosystemarer Feldinformationen zur Aufdeckung relevanter strukturbildender Wechselwirkungen zwischen Biotik und abiotischen Parametern eines Ökosystems erfordert weiterhin eine integrierende Schnittstelle zwischen den Datengrundlagen und den Auswertekonzepten. Für den Einsatz multivariater Datenanalysemethoden ist ein statistisch bedingter Formalismus bei der Datenerhebung einzuhalten, der mit den fachwissenschaftlichen Belangen zu harmonisieren ist, um die Daten verschiedener Fachdisziplinen und die Verfahren der mathematisch-statistischen Ökosystemforschung miteinander kompatibel werden zu lassen. Willkürliche und Ad-hoc-Datenerhebungen erfüllen typischerweise die Annahmen von multivariaten Analyseverfahren nicht (Stockwell & Peters 1999). Die Analysemethoden über ein exploratives Stadium hinaus anzuwenden, ist daher nicht zu empfehlen. Um dies zu vermeiden, wird eine koordinierende Rahmenvorschrift – der Stichprobenplan – benötigt (Jongman et al. 1987; Kent & Coker 1992; Köhler et al. 1996; Mühlenberg 1993; Økland 1990; Rink 1997c, 1998; Rink et al. 2000).

Die Erstellung des Stichprobenplans (sampling design) ist integriert in die Versuchsplanung (study design) und basiert auf einer Vielzahl hierarchisch gegliederter Arbeitsschritte. Zu berücksichtigen sind unter anderem folgende Aspekte:

- (1) Festlegung der Projektziele
- (2) Auswahl von biotischen und abiotischen Beobachtungsobjekten bzw. Parametern
- (3) Auswahl von Untersuchungsgebieten und -skala
- (4) Prüfung bestehender Datenlage und Schutzstatus der Flächen
- (5) Prüfung logistischer Vorgaben
- (6) Auswahl der statistischen Datenanalysestrategie

Die Erfahrung zeigt, dass in großen Forschungsprojekten mit ökologischem Ansatz zwar häufig Vertreter verschiedener Fachdisziplinen in ein und demselben Areal arbeiten, aber fachinternen Fragestellungen nachgehen und ein interdisziplinär abgestimmter Stichprobenplan nicht existiert (Siebeck 1995). Die Daten bleiben in der "Sprache" der Fachdisziplin und können nur unzureichend zur Beantwortung der projektübergreifenden Fragestellung beitragen (Scholle 1997). Häufig zeigt sich bei den Bearbeitern ökologischer Projekte diese Einsicht erst nach der eigentlichen Versuchsausführung (Legendre & Legendre 1998). Gerade der Anspruch, eine projektübergreifende Fragestellung zu formulieren, der alle beteiligten Fachwissenschaftler zuarbeiten, und die resultierenden Daten zusammenzuführen sowie geeigneten statistischen Methoden zugänglich zu machen, erfordert die Ausarbeitung und konsequente Umsetzung einer gemeinsamen Datenerhebungsstrategie.

Die präzise Formulierung der Projektziele steuert alle Überlegungen der Versuchsplanung, daher stellt sich anfangs immer die Frage:

*Wie lautet das projektübergreifende Ziel?*

Die Betonung dieser Aussage erscheint banal, stellt sich aber sehr schnell als durchaus essentiell heraus, da hierdurch schon zu Projektbeginn Missverständnisse zwischen Arbeitsgruppen vermieden werden und die gemeinsam erarbeitete Leitvorstellung im Projektansatz manifestiert und im Stichprobenplan umgesetzt werden kann (s.a. Kenkel et al. 1989).

Ebenfalls zum Beginn eines Projektes sollte geprüft werden, ob die Einbeziehung externer Projektdaten für die spätere Verifizierung der Auswertergebnisse im Versuchsplan berücksichtigt werden kann. Dazu ist eventuell ein Abgleich bei der Auswahl der Beobachtungsobjekte oder des Untersuchungsraumes bzw. des Stichprobenplans anzustreben. Eine genaue Analyse der Datenlage, z.B. des Stichprobenumfangs, des Datenniveaus und der vorliegenden räumlichen und zeitlichen Stichprobenerhebung ist Grundvoraussetzung, um eine vergleichende Auswertung der Fremddaten mit den im Projekt angewendeten statistischen Verfahren durchführen zu können. Nur wenn die Daten den Anforderungen an die Datenqualität genügen und die Parameter nach den gleichen standardisierten Verfahren erhoben wurden, ist eine nachvollziehbare gleichwertige Zusatzinformation zu erhalten, die Ergebnisse der angestrebten Untersuchung stützen oder modifizieren kann (Rink 1997b).

Zusammengefasst bedeutet dies, dass die Absicht, eine klar fokussierte teilprojektübergreifende Problemstellung zu lösen, eine Vorgehensweise erzwingt, bei der die Fachwissenschaftler von ihrer gewohnten Art, Felderhebungen durchzuführen, abweichen müssen. Nur wenn eine Integrität zwischen den Daten und den Auswertekonzepten existiert, ist eine statistische Auswertung über alle Informationsebenen möglich (Pickett et al. 1994; Richter & Söndgerath 1990; Rink 1998; Rink et al. 2000). Das Resultat ist ansonsten, dass "unzählige" fachwissenschaftlich korrekt erhobene biotische und abiotische Datensätze vorliegen, diese aber nicht auf eine gemeinsame Informationsebene gebracht werden können, und so statistisch bestenfalls spärliche Aussagen zulässig sind. Im Extremfall lässt sich die ursprünglich gestellte projektübergreifende Fragestellung nicht beantworten.

### 3.1.2 Stichprobenplan und Stichprobenumfang

*"THE DECISION-MAKING PROCESS IN SAMPLING MUST BE VIEWED AS A FLEXIBLE EXERCISE, DICTATED NOT BY GENERALIZED RECOMMENDATIONS BUT BY SPECIFIC OBJECTIVES: THERE IS NO PANACEA IN ECOLOGICAL SAMPLING (KENKEL, 1989, S. 195)."*

Eine zentrale Aufgabe des Stichprobenplans als Element des Versuchsplans ist es, die günstigste Anordnung des Versuchs zu finden, um die allgemein gültige Aussagekraft der Daten auf ihr Maximum zu erhöhen. Die Meinungen verschiedener Autoren, welche Methode für welchen Kontext am effektivsten ist, gehen dabei weit auseinander. Konsens herrscht darin, dass jede statistische Analysemethode und deren Ergebnissicherheit grundsätzlich von der sorgfältig erarbeiteten Gesamtstrategie der Untersuchung abhängig ist (Gauch 1982a; Jongman et al. 1987; Kenkel et al. 1989; Kent & Coker 1992; Økland 1990; Richter & Söndgerath 1990; Rink et al. 2000).

Der Stichprobenumfang einer Untersuchung ist eng mit den logistischen Vorgaben, der Projektfragestellung, der Auswertemethode und der Auswahl des Stichprobenplans verknüpft (s.a. Kenkel et al. 1989; Rink 1998). Der statistisch geforderte minimale Stichprobenumfang und Formalismus kollidiert oft mit den rein fachlichen Ansichten der Probeflächenverortung und dem finanzierbaren Beprobungsumfang. Die notwendige Kompromisslösung spiegelt



sich letztlich in der Art des Stichprobenplans wieder. Da die Vielfalt an Basisstrukturen für Stichprobenpläne addiert mit unzähligen Modifikationen Lehrbücher zu füllen vermag, muss eine kontextbezogene Auswahl für die genauere Erörterung getroffen werden.

Die Zielsetzung der Datenerhebung steht in enger Beziehung zu drei grundlegenden Dichotomien:

- (1) Parameterschätzung versus Mustererkennung
- (2) univariat versus multivariat
- (3) diskrete versus kontinuierliche Grundgesamtheit

Die Parameterschätzung verlangt nach einer Präzision, die sich auf die Minimierung der Variabilität über den geschätzten Parameter konzentriert. Mustererkennung fokussiert darauf, die Variabilität zwischen den Beprobungseinheiten zu maximieren und Strukturen in den Daten zu erhellen. Univariate und multivariate Studien unterscheiden sich in der Anzahl der simultan betrachteten Objekte, und diskrete bzw. kontinuierliche Grundgesamtheiten (Population; sampling universe) nehmen entscheidenden Einfluss auf die Gestaltung des Stichprobenplans durch den möglichen Stichprobenumfang (Kenkel et al. 1989).

Diskrete Erhebungseinheiten sind klar begrenzte natürliche Räume finiter Anzahl (Inseln, Teiche etc.), implizierend das die Menge der möglichen Proben ebenfalls endlich ist. Der Stichprobenrahmen ist daher definierbar, während in einer kontinuierlichen Grundgesamtheit durch das Fehlen solcher sichtbar vorgegebener Einheiten dieser Rahmen nicht existiert. Speziell in diesem Falle ist der Stichprobenrahmen per definitionem festzulegen und dann anhand eines Stichprobenplans aus der infiniten Anzahl möglicher, in Form und Größe nicht natürlich bestimmbarer Beprobungseinheiten Stichproben zu selektieren. Folgende Dimensionen sind denkbar: null (z.B. Schnittpunkte eines Rasters), eine (z.B. Abschnitt eines Transeks), zwei (z.B. Probefläche) oder drei (z.B. Bohrkern) (Kenkel et al. 1989).

Bei kontinuierlichen Grundgesamtheiten müssen im Gegensatz zu diskreten, zusätzlich zu der Art der Stichprobenentnahme aus dem Stichprobenareal Vorgaben zur räumlichen Abgrenzung (2-/3-dimensionaler Fall) und dem Arrangement der Beprobungseinheiten getroffen werden.

Bevor der Schwerpunkt auf Stichprobenpläne für multivariate Mustererkennung in kontinuierlichen Grundgesamtheiten gelegt wird (vgl. Abschnitt 5.2.1), sollen einige grundlegende experimentelle Methoden und Standardversuchsanordnungen kurz aufgelistet werden, da viele Ideenkonzepte in experimentellen Designs und Stichprobenplänen zur Mustererkennung aneinander angelehnt sind (z.B. Zufälligkeit, Blockbildung/Stratifizierung, Test gegen Gradienten). "Traditionelle" Versuchsanordnungen in experimentellen ökologischen Fragestellungen beruhen meist auf einer Vorher-Nachher-Situation oder auf einem Vergleich zwischen verschiedenen parallel verlaufenden Einflussnahmen auf ein System gleichen Basiszustandes (Blockbildung). Zufälligkeit auf einem spezifischen Level ist ein Element in allen Versuchsanordnungen. Weiterhin ist die Idee des geschachtelten Designs bedeutsam.

#### **Exkurs "Experimentelle Versuchsanordnungen":**

- (1) **Reine Zufallsstichproben in Mittelwertvergleichen (Completely Randomized Design):** Vergleich des Mittels zweier abhängiger oder unabhängiger Stichproben.
- (2) **Nicht repliziertes BACI-Design (Non-replicated Before After Control Impact):** Es werden zwei Gruppen räumlich getrennter Stichproben selektiert. Eine Gruppe bleibt von dem Eingriff (impact) unbeeinflusst, die andere weist einen "Vor-dem-Eingriff-Zustand" und zeitlich versetzt einen "Nach-dem-Eingriff-Zustand" auf. Der Effekt des Eingriffs kann analysiert werden.
- (3) **Repliziertes BACI-Design (Replicated Before After Control Impact):** Es werden mehrere Blöcke der Ausprägung "Kontrolle" (control) und "Eingriff" (impact, treatment) gebildet, die sich in ihrer Versuchsanordnung wiederholen. Ausgehend von einem analogen Ausgangszustand erfolgt der Eingriff innerhalb der Gruppe "Eingriff" jedes Blocks. Die Auswirkungen werden dann zu einem Zeitpunkt  $T_1$  und  $T_2$  ermittelt. Die Antwort beschreibt z.B. die Interaktion zwischen Eingriff und Zeit.
- (4) **Randomisiertes Block Design (Randomized Complete Blocks):** Sollen verschiedene Behandlungen (treatments) entlang eines Gradienten betrachtet werden, können randomisierte Blöcke mit den Be-



handlungstypen rechtwinklig zur Verlaufsrichtung des Gradienten ausgerichtet werden. Es ist darauf zu achten, dass die Blöcke nicht parallel zum Gradienten verlaufen, da die Behandlungen ansonsten nicht mit einem konkreten Referenzzustand in Beziehung gesetzt werden können.

- (5) Lateinisches Quadrat: Es teilt sich die Eigenschaft mit dem randomisierten Block-Design, dass die Behandlungstypen zu homogenen Blöcken zusammengefasst werden. Es unterscheidet sich aber in der Eigenschaft, die Blöcke entlang zweier Gradienten (Zeile/Spalte) auszurichten, um deren Variabilität zu kontrollieren.
- (6) Hierarchisches (geschachteltes) Design: In dem Versuchsaufbau werden zwei unabhängige Faktoren getestet, die ineinander geschachtelt (genested) sind, um deren Effekt auf die Beobachtungsobjekte zu analysieren.
- (7) Split-Plot-Design: Das Design eignet sich für die Studie eines Blockfaktors (main plot factor), der sich über mehrere Blöcke wiederholt, und eines untergeordneten Faktors mit Einfluss auf die Subplots. Die Subplots werden dafür zufällig innerhalb der Blöcke in Reihenfolge, Lage etc. angeordnet.

Zum grundlegenden Verständnis der im Folgenden beschriebenen Stichprobenpläne sei auf die Studie dieser Versuchsanordnungen verwiesen, die sehr gut nachvollziehbar beispielsweise in Leps (1999) sowie Scheiner & Gurevitch (1993) beschrieben sind. Das Split-Plot-Design lässt sich prinzipiell so an den Stichprobenplan "stratifizierte Zufallsverteilung" (s.u.) anpassen, dass Monte Carlo Tests (Abschnitt 3.6.1) eine Prüfung der Trennkraft von Modellparametern zwischen den Straten (Blöcken) und zwischen den Probeflächen (Subplots) innerhalb der Straten erlauben (vgl. Abschnitt 5.4.2).

### Stichprobenpläne zur Mustererkennung

Bei der Vorstellung von Stichprobenplänen werden zur leichteren Nachvollziehbarkeit grundlegende Kategorien für die Unterscheidung gewählt:

übergeordnet  
linienhaft  
flächig

untergeordnet  
subjektiv  
systematisch  
zufällig

#### **Linienhafte Stichprobenpläne**

Die Erhebung von Daten entlang einer Geländelinie (Transekt) ist sehr beliebt. Sie bietet den Vorteil, dass kontinuierliche Veränderungen oft deutlich erfasst werden, vorausgesetzt, der unterliegende Gradient ist gerichtet und die Richtung bekannt (z.B. Höhengradient). Daraus ergibt sich gleichzeitig, dass die zufällige Wahl des Start- und Endpunkts nicht sinnvoll ist, da das Transekt a priori mit einem bekannten Umweltgradienten parallelisiert werden muss. Das Transekt beschreibt dann die maximale Variation entlang der kürzesten Distanz des unterliegenden Faktors (Kent & Coker 1992).

Ist das Transekt definiert, müssen die Probenahmestellen entlang diesem verortet werden. Dies kann rein subjektiv geschehen durch kontinuierlich benachbarte (geschlossenes Transekt, belt transect), kontinuierlich unterbrochene (vorgegebene äquidistante Entfernungen, open transect) oder zufällige Anordnung (Økland 1990). In allen Fällen kann der Startpunkt der Verortung randomisiert werden, dennoch sind die Stichproben bei systematischen Abfolgen nicht unabhängig voneinander, das Auffinden bisher ungeahnter faktoraler Trends inklusive deren Richtung ist nicht möglich und durch meist intensive räumliche Autokorrelation ist mit einer Verfälschung der Ergebnisse zu rechnen (Wildi 1986). Räumliche Autokorrelation ist allerdings keine exklusive Eigenschaft von Transekten (vgl. Abschnitt 3.5.9).

#### **Flächenhafte Stichprobenverfahren**

Man unterscheidet bei den flächenhaften Stichprobenplänen ebenfalls subjektive und systematische Rasterverfahren von den Verfahren, welche verstärkt Zufallskomponenten integrieren (Abbildung 3.1-2, S. 31). In beiden Fällen kann eine Stratifizierung des Untersuchungsgebietes erfolgen, das heißt eine der Stichprobenverteilung vorausgehende Unterteilung der Gesamtfläche in quasihomogene, sich aber bezogen auf einige Auswahlkriterien voneinander unterscheidende Teilflächen (Straten).

### Subjektive "Stichprobenpläne"

An sich kann bei einer subjektiven Verortung von Probeflächen (selective or preferential sampling) nicht von einem Stichprobenplan gesprochen werden. Die Subjektivität zieht sich durch alle nachfolgenden Analysen. Die statistische Forderung nach Zufälligkeit und unabhängigen Stichproben ist nicht gegeben und folglich sind jegliche Hypothesentests nicht gültig. Die Art der Datenerhebung ist ungeeignet für die Exploration von Wechselwirkungen und Zusammenhängen zwischen biotischen Beobachtungsobjekten und abiotischen Umweltfaktoren (Økland 1990).

Trotz dieser Tatsache und des Wunschs nach Objektivität tendieren Ökologen und speziell Phytosoziologen traditionell dazu, Probeflächen in "homogenen" oder "repräsentativen" Einheiten zu lokalisieren. Die Methode ist das häufigste Stichprobendesign in ökologischen Studien (Økland 1990). Das Verfahren ist auf den ersten Blick attraktiv, da seltene Arten und Lebensgemeinschaften gut erfasst werden und es wenig zeitintensiv ist, Probeflächen zu lokalisieren und zu verorten. Seltene Arten werden aber verglichen mit ihrem Flächenanteil in einem Untersuchungsgebiet als Folge häufig überrepräsentiert, und die Gefahr eines Zirkelschlusses ist hoch: Zielgerichtet gewählte Indikatoren als Zeiger für Umweltbedingungen, erfasst auf entsprechend arrangierten Probeflächen, dokumentieren dann den angenommenen Gradienten durch ihre Verteilung offensichtlich hervorragend.

Allgemein gilt, dass Probeflächen sowohl statistisch als auch ökologisch gefordert nicht rein subjektiv verortet werden sollten. Die statistische Forderung nach Zufälligkeit wird in extremer Weise von Daget & Godron (1982 in Wildi 1986) vertreten. Sie halten alle nicht zufallsbasierten Datenerhebungen für statistisch unbrauchbar. Andere Autoren nehmen aus ökologischer Sicht Stellung: Um das Verhalten einer Art hinsichtlich eines Umweltgradienten zu erfassen, sind gerade die Übergangsbereiche zwischen "homogenen" Einheiten bedeutsam. Außerdem treten viele Arten überhaupt nur in solchen Übergangsbereichen auf (z.B. Palmer 1995).

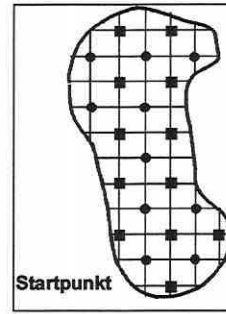
### Systematische Stichprobenpläne

Eine sicherlich sehr präzise Methode ist – je nach Maschenweite – die Erstellung eines regelmäßigen probabilistischen Stichprobennetzes (Raster, Grid; grid sampling), da der Untersuchungsraum eine homogene Abdeckung bei der Probenahme erfährt (Kent & Coker 1992; Mühlenberg 1993). Dabei sollten Ursprung und Richtung des Rasters zufällig gewählt werden. Die Rasterweite ist dabei das entscheidende Element, um objektive und zuverlässige Daten zu erhalten. Økland (1990, S. 77) kommt zu folgender Aussage: "The demand for independence of sample plots is well satisfied with grid sampling provided the mesh width is much smaller or much larger than recurrent structures of vegetation."

Der Bezug zur Vegetation ist nur ein Aspekt und muss verallgemeinert werden. Es ist darauf zu achten, dass keinerlei Koinzidenz zwischen biotischen Raummustern oder kontrollierenden Umweltfaktoren besteht (Kent & Coker 1992). Gibt es wiederkehrende Strukturen, z.B. alle 10m eine Erhöhung im Gelände, und die systematischen Aufnahmen liegen immer in den Senken dazwischen, ergibt sich ein falsches Bild über das gesamte Untersuchungsgebiet (Mühlenberg 1993). Die resultierenden Probeflächen aus systematischen Erhebungen können generell nicht als unabhängig voneinander bezeichnet werden, da sie alle auf den gleichen Ursprung bezogen sind. Streng genommen hat solch ein Design nur einen echten Freiheitsgrad.

Ein weitmaschiges Probenahmenetz entbehrt, je nach Untersuchungsmaßstab, meist der wünschenswerten Genauigkeit. Der positive Effekt homogener Raumabdeckung (Repräsentativität) engmaschiger Raster kann verloren gehen, wenn durch die akkurate Einhaltung des Stichprobenplans viele Aufnahmen auf ungeeignete Flächen fallen (z.B. Wege), so für teils große Lücken im Raster sorgen und die gewünschte Information entfällt. Sollen auch kleinräumige Strukturen mit hoher innerer Variabilität (z.B. Flutrinnen) erfasst werden, muss

das Netz sich an der geringsten physikalischen Größe wichtiger Ökosystemkompartimente orientieren (Kalton 1983). Ist die Maschenweite sehr klein, wird der notwendige Stichprobenumfang in einem großen Untersuchungsgebiet inflationär. In der von Økland (1990) vorgeschlagenen Stichprobenplanvariante "offenes Grid" wird beginnend an einem zufälligen Startpunkt alle  $x$  Knotenpunkte eine Fläche etabliert (Abbildung 3.1-1: runde Signaturen). Sie ist letztlich nichts anderes als ein weitmaschiges Raster und erscheint erst dann sinnvoll, wenn die Probeflächen reihenweise versetzt angeordnet werden, so dass zwei ineinander geschachtelte Raster entstehen (Abbildung 3.1-1: quadratische und runde Signaturen).



**Abbildung 3.1-1: Systematische Beprobung eines offenen Grids (Økland 1990, S. 73, verändert).**

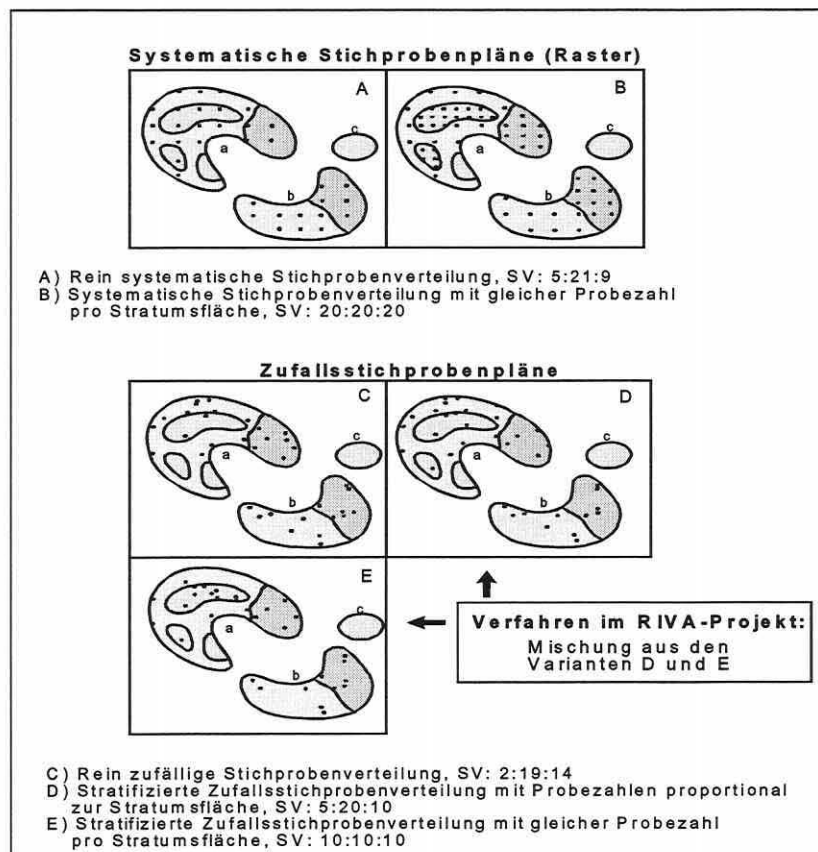
Alternativ kann ein Raster für ausgewiesene Teilflächen (Straten, Blöcke) des Untersuchungsraumes angepasste Maschenweiten erhalten (Wildi 1986). Die strukturelle Variabilität der Teilflächen und der verfügbare Gesamtstichprobenumfang können ein homogenes oder ein nach Flächenanteilen gewichtetes Stichprobenverhältnis bedingen (Abbildung 3.1-2, S. 31, Teilbild A und B). Die Gewichtung der Rasterweite nach Flächenanteil oder Variabilität einer Teilfläche schließt eine gewisse Subjektivität ein.

Der in der Praxis oft viel zu hohe Zeit- und Kostenaufwand einer Rasterlösung mit hohem Stichprobenumfang bei kleiner Maschenweite führt dazu, dass Forschungsvorhaben häufig aus rein logistischen Gründen darauf verzichten müssen (Whittaker 1973). Projekte sind fast immer nur realisierbar, wenn der Stichprobenumfang klein und somit ökonomisch vertretbar gehalten werden kann. Das Verfahren wird aber trotzdem von einigen Autoren für die Studie räumlicher Verteilungsmuster empfohlen (Legendre & Fortin 1989; Økland & Eilertsen 1994).

#### Zufallsstichprobenpläne

Aus statistischer Sicht ist theoretisch die reine Zufallserhebung das Verfahren der Wahl. Jeder Punkt des Untersuchungsgebietes hat eine gleich große Chance, in die Auswahl der Probenahmestellen aufgenommen zu werden. Es gibt daher keinerlei Restriktionen bei der Anwendung inferentieller parametrischer Statistik. Testverfahren basierend auf Wahrscheinlichkeiten fordern Zufälligkeit, d.h. unabhängige Beobachtungen, als essenzielle Voraussetzung (Jongman et al. 1987; Kent & Coker 1992; Økland 1990 u.a.). Eine häufige Situation ist die Exploration und das Testen von Zusammenhängen zwischen Biotik und Umweltfaktoren unter Anwendung von Korrelation bzw. Regression (Kent & Coker 1992). Unter den Zufallsstichprobenplänen hat die reine Zufallsverteilung eine Sonderstellung, weil sie einen ähnlich hohen Stichprobenumfang erfordern kann wie eine Rasterlösung mit enger Maschenweite. Das herdenartige Auftreten vieler Zufallspunkte und die inhomogene Flächenüberdeckung ist typisch für eine echte probabilistische Stichprobenverteilung (Økland 1990; Pielou 1984; Wildi 1986). Abbildung 3.1-2 (S. 31, Teilbild C) veranschaulicht eine reine Zufallsverteilung. Hierbei wird nur das Untersuchungsgebiet begrenzt und die nicht untersuchungsrelevanten Teilflächen ausgegrenzt. Dann wird der gesamte zur Verfügung stehende Stichprobenumfang zufallsverteilt. Als hypothetisches Resultat ergibt sich auf den imaginären Teilflächen im Beispiel ein Stichprobenverhältnis von 2:19:14, d.h. das vom physikalischen Flächenanteil kleinste Stratum wird nur mit zwei Probeflächen belegt. Erst ab einem bestimmten Stichprobenumfang kann aber das Verhalten von Beobachtungsobjekten statistisch gesichert beschrieben werden. Die Folge ist, dass bei einem eingeschränkt leistbaren Stichprobenumfang seltene Standortfaktorenkombinationen und Arten unterrepräsentiert sind oder überhaupt nicht erfasst werden. Extrema entlang von Umweltgradienten beherbergen häufig solch seltene Lebensräume. Einzelmessungen oder ein statistisch zu geringer Stichpro-

benumfang pro Beobachtungsobjekt oder Teileinheit haben einen nur geringen bis keinen Aussagewert. Umgekehrt ist der Grad an Redundanz für dominante Ausprägungen und mittlere Standortbedingungen (große Flächenanteile einnehmend) hoch. Wiederholungs- oder Kontrollmessungen in den gleichen Versuchseinheiten dienen zwar zur Verifizierung von gefundenen Zusammenhängen, sie sollten jedoch in einem angemessenen Verhältnis zum Gesamtstichprobenumfang stehen. Die genannten Umstände nehmen einen nicht unerheblichen Einfluss auf Studien, die Mustererkennung, Nischenbestimmung zu Gradienten und/oder Coenoclinen-Identifikation zum Ziel haben (Økland 1986c, 1990).



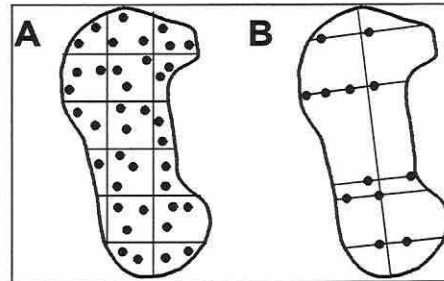
**Abbildung 3.1-2: Verschiedene Stichprobenplanverfahren mit Stichprobenverhältnis (SV) (Wildi 1986, S. 39 u. 41, verändert).**

Zusammenfassend überwiegen die negativen Eigenschaften die positiven einer reinen Zufallserhebung. Da eine absolute Zufälligkeit in der Probeflächenverortung für viele Verfahren (z.B. Klassifikation, Ordination u.a.) nicht zwingend notwendig ist, kommt Økland (1990, S. 75) neben anderen Autoren (z.B. Austin 1987; Gauch 1982a) zu folgender Feststellung: "Relaxation of the demands for strict randomness is advocated by most authors."

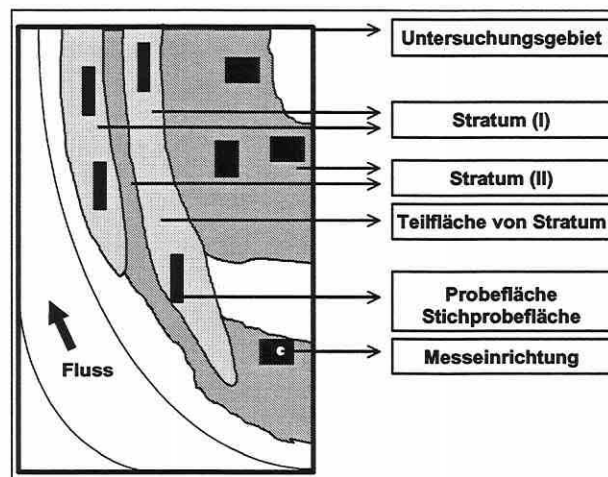
Das für synökologische Fragestellungen und einen eingeschränkten Stichprobenumfang fast generell der reinen Zufallsverteilung vorzuziehende Verfahren ist die stratifizierte Zufallserhebung (stratified or restricted random sampling, s.a. Mühlenberg 1993). Die statistischen Eigenschaften sind der reinen Zufallsverteilung sehr ähnlich, selbst wenn eine gewisse Subjektivität durch die Festlegung der Stratifizierungskriterien in die Erstellung des Designs eingebracht wird. Die Platzierung der Probeflächen erfolgt dennoch innerhalb dieses Rahmens

objektiv zufällig. Aus der Vielzahl an Umsetzungsmöglichkeiten des Basisprinzips sollen einige Beispiele vorgestellt werden.

- (1) Ein Untersuchungsraum kann durch ein grobmaschiges Raster überdeckt und innerhalb dessen Zellen eine homogene Anzahl an Probeflächen – anteilig proportional zum Gesamtstichprobenumfang – zufällig verteilt werden (Abbildung 3.1-3, Teilbild A). Bei dieser Anwendung werden offensichtlich abgrenzbare kleinräumige Einheiten von Interesse nicht speziell mit einer festgesetzten Anzahl an Probeflächen bestückt (keine Gewichtung nach A-priori-Kriterien), sondern nur die Raumabdeckung homogener gestaltet. Zirkelschlüsse zwischen Stratifizierung und beobachteten Objekten sind nahezu ausgeschlossen, weil die Rasterung sich nicht an der räumlichen Verteilung von Biotik oder Abiotik orientiert. Durch entsprechende Grobmaschigkeit und die randomisierte Anordnung der Probeflächen in den Zellen wird eine zufällige Koinzidenz zu Raummustern vermieden.
- (2) Entlang der maximalen Ausdehnung des Untersuchungsraumes oder zufällig wird eine Grundlinie etabliert (baseline approach), entlang derer eine Anzahl quer-verlaufender Linien randomisiert gewählt werden. Je nach Länge der Transversen wird eine gewichtete Anzahl von Zufallskoordinaten auf den Geraden selektiert, welche die Lage der Probeflächen bestimmen (Abbildung 3.1-3, Teilbild B).
- (3) Eine dritte Variante ist die Stratifizierung des Untersuchungsgebietes nach festgelegten Kriterien (Vorwissen) in quasihomogene Teilflächen (Straten) mit homogener oder gewichteter Zufallsverteilung des Umfangs an Stichprobenflächen in den Einheiten nach Flächenanteil oder Homogenität (Abbildung 3.1-2, S. 31, Teilbild D und E). Die Begrifflichkeiten im Rahmen von stratifizierten Zufallsstichprobenplänen beschreibt Abbildung 3.1-4.



**Abbildung 3.1-3: Verschiedene Schemata zur Etablierung von Probeflächen in stratifizierten Zufallsstichprobenplänen. (A) Stratifizierung anhand eines Grids innerhalb dessen Zellen eine homogene Anzahl von Probeflächen zufallsverteilt wurde; (B) Stratifizierung mit einer Basislinie und Transversen (Økland 1990, S. 73, verändert).**



**Abbildung 3.1-4: Begrifflichkeiten im Rahmen von stratifizierten Stichprobenplänen.**

Wird eine gewichtete Zufallsverteilung zur Stratenbildung herangezogen, kann diese auf der Diversität einer Fläche basieren. Je größer diese und der Wechsel der Artenkomposition (Beta-Diversität) ist, desto höher sollte dann auch die Beprobungsintensität sein (Kent & Coker 1992).

Die Zufallsverteilung der Probeflächen innerhalb der Straten kann durch ein Koordinatennetz erfolgen, welches im Geographischen Informationssystem (GIS) mit der Stratifizierung überlagert wird. Aus den straten-spezifischen



Teildatensätzen von Koordinatenpunkten wird dann eine bestimmte Anzahl Zufallskordinaten selektiert und als Basiseckpunkte der Probeflächen im Feld genutzt. Die Verortung kann mithilfe von einem Echtzeit-Satellitennavigationssystem (RTK-GPS; siehe Abschnitt 3.1.6) in freiem Gelände (Rink et al. 2000) oder in Kombination aus RTK-GPS und Laserfeldstecher im überdeckten Gelände erfolgen (vgl. Reiter & Fussenegger 1997). Für eine detaillierte Beschreibung zur Umsetzung dieses Stichprobenplantyps wird auf Abschnitt 5.2.2 verwiesen.

Eine andere etwas weniger objektive Variante beschreibt Kent & Coker (1992) als "random walk procedures": Beginnend an einem beliebigen Punkt des Stratum wird eine Zufallszahl zwischen null und 360 ermittelt. Die resultierende Zahl bestimmt die Kompassrichtung gegen Nord, in der eine ebenfalls zufällige Anzahl an Schritten oder eine jeweils zufällige Distanz zurückgelegt wird. Diese Prozedur wird einmal oder mehrfach pro zu verortender Probefläche wiederholt. Da alle so bestimmten Probeflächen von der jeweils vorangegangenen nicht völlig unabhängig sind, ist auch hier ein höherer Subjektivitätsanteil integriert (vgl. systematische Stichprobenpläne). "Random walks" sind nicht streng zufällig.

Die Anpassung des Stichprobenplans durch Stratifizierung an Gegebenheiten des Geländes gewährleistet, dass trotz Einhaltung des geforderten Zufallsprinzips (Wildi 1986) relevante kleinflächige oder mit einer großen Restinhomogenität ausgestattete Areale ausreichend genau erfasst werden. Außerdem ist in der Regel der Untersuchungsaufwand bei gleicher Ergebnissicherheit gegenüber einer echten Zufallsverteilung reduziert. Die Stratifizierungskriterien zur Abgrenzung der Zielgesamtheiten (Straten) aus der Grundgesamtheit können sehr unterschiedlich sein und werden in Abschnitt 3.1.3 behandelt.

Zwischen den Rasterverfahren und den exemplarisch ausgewählten Stichprobenplänen gibt es eine Vielzahl anderer Stichprobenplanvarianten, die je nach Aufgabenstellung gleichwertig sein mögen oder im Einzelfall vorzuziehen sind. Unter dem Gesichtspunkt, grundlegende Verfahren vorzustellen, erscheint eine Erläuterung weiterer Prinzipien als nicht relevant.

Die in dieser Arbeit überwiegend zur Datenauswertung benutzte Software CANOCO® bietet die Möglichkeit Monte Carlo Tests (vgl. Abschnitt 3.6.1) zur Signifikanzbestimmung für folgende Designs durchzuführen: reiner Zufall, Zeitreihen, Transekte, Raster und Split-Plot-Designs (Stratifizierte Stichprobenpläne).

Da die Datenerhebung, Bestimmungs- und Laborarbeit den kostenintensivsten Teil einer ökologischen Studie ausmachen und quantitative Information immer Datenauswertung mit numerischen Verfahren impliziert, ist der Stichprobenplan unter den gegebenen Rahmenbedingungen zu optimieren. Es ist wichtig, im Voraus die Datenerhebung und die Auswertungsmethoden aufeinander abzustimmen, denn: "It follows that mathematical processing is at the heart of a research; the quality of the results cannot exceed the quality of the numerical analysis conducted on the data (Legendre & Legendre 1998, S. 7)."

### **Stichprobenumfang**

Das Problem des statistisch notwendigen Stichprobenumfangs zur Beantwortung einer spezifischen Fragestellung mit Ordinationsverfahren kann an dieser Stelle diskutiert, aber für eine kontinuierliche Grundgesamtheit nicht beantwortet werden. Es finden sich mehrere statistische Ansätze in der Literatur (z.B. Köhler et al. 1996), dennoch bleibt deren Basis zweifelhaft. Die Aussage: "(...) objective criteria for determination of the minimum number of sample plots necessary for a particular study do not exist (Økland 1990, S. 83)" scheint bis heute gültig.

Dennoch bleibt die Notwendigkeit bestehen, jede Art oder jeden Lebensgemeinschaftstyp durch eine "ausreichende" Anzahl von Stichproben zu beschreiben.

Vorschläge und Daumenregeln in der Literatur differieren beachtenswert: Die Spanne reicht von drei Stichproben pro Art/Typ bis zu 50 (Økland 1990). Bedenkt man, dass eine Absicht



von ökologischen Studien oft die Visualisierung der Realnische einer Art gegenüber einem Gradienten ist, so liegt das absolute Minimum bei drei Stichproben, um eine unimodale Antwortkurve erstellen zu können. Selbstverständlich kann solch ein Konstrukt nicht als zuverlässig bezeichnet werden. Erst fünf bis zehn Vorkommensinformationen bei geringem Datenausschlag und eindeutigem Optimum, also klarer Beschreibung der Antwortkurve einer Art, mögen eine Aussage zulassen.

Es wurde bereits erwähnt, dass die Projektfragestellung, der Stichprobenplan und die Probenflächengröße (siehe Abschnitt 3.1.4) einen relevanten Einfluss auf den Stichprobenumfang haben. Daneben spielt die Heterogenität eines Untersuchungsraumes bzw. einzelner Teilflächen die entscheidende Rolle zur Bestimmung eines angemessenen Stichprobenumfangs. Die Zusammenhänge zwischen räumlicher Heterogenität, Stratenabfolge und Stichprobenumfang werden in Rink (2000) an Beispielen dargestellt und diskutiert.

Letztlich besteht die Krux, dass die Heterogenität in einer Fläche nicht ohne umfangreiche Voruntersuchungen klar einzuschätzen ist. Es bleibt also immer ein Restrisiko hinsichtlich des Verhältnisses Variabilität der Fläche zur Stichprobenanzahl. Grundsätzlich nimmt dieses Restrisiko für die Erfassung der Varianz in einer "quasihomogenen" Fläche mit der Erhöhung des Stichprobenumfangs ab.

Die Folge eines zu knapp bemessenen Stichprobenumfangs ist eventuell eine Fehlinterpretation des Zusammenhangs zwischen Biotik und Abiotik. Natürlich muss man auch bei einem hohen Stichprobenumfang einräumen, dass die Gesamtheterogenität eines Raumes niemals erfasst werden kann, selbst wenn man die Stichprobenzahl beliebig groß machen würde. Der Umkehrschluss soll aber deutlich machen, dass es eine Mindestgrenze des Stichprobenumfangs pro Stratum gibt, die von vielerlei Kriterien definiert wird. Unterschreitet man die minimale Grenze des Stichprobenumfangs, so sind die resultierenden Schlussfolgerungen nicht mehr fachlich vertretbar und statistisch sinnvoll abzusichern. Einsparungen bei der Stichprobenzahl in einem Projekt können sich so im schlimmsten Fall zu einem "ökonomischen Bumerang" entwickeln.

Bei der Anwendung von Klassifikations- und Ordinationsverfahren, kann der empirische Stichprobenumfang zumindest ansatzweise nachträglich begutachtet werden. Verändert sich das Klassifikations- und Ordinationsergebnis durch das Hinzu- oder Herausnehmen von Stichproben nicht spürbar, kann aus der Stabilität der Ergebnisse auf einen adäquaten Stichprobenumfang in der Studie zurückgeschlossen werden.

### 3.1.3 Stratifizierungskriterien und Stratifizierung von Untersuchungsgebieten

**"JE ENGER DIE RÄUMLICHE AUFTEILUNG [ANM.: STRATIFIZIERUNG] VORGENOMMEN WIRD, JE HOMOGENER DIE AUSGESCHIEDENEN RAUMEINHEITEN SIND, DESTO GRÖßER IST DIE REPRÄSENTATIVITÄT DER IN IHNEN ERHOBENEN PROBEN. AUF DIESE WEISE WIRD DIE VARIABILITÄT DER (PFLANZEN-) BESTANDSSTRUKTUREN EINGEENGT, SO DASS AUCH ANHAND EINER GERINGEREN STICHPROBENZAHL ZUTREFFENDE, MIT DER REALITÄT WEITGEHEND ÜBEREINSTIMMENDE AUSSAGEN ÜBER DAS UNTERSUCHUNGSGEBIET ERLANGT WERDEN KÖNNEN (GLAVAC 1996, S. 79)."**

Der in Abschnitt 3.1.2 vorgestellte Stichprobenplan "stratifizierte Zufallsverteilung" erfordert in der Regel die Festlegung von Kriterien zur Abgrenzung der Straten. Der aktuelle Abschnitt wird Pro- und Contraargumente für deren Auswahl liefern. Vegetationskundliche Untersuchungen gehen bei der Stratifizierung einer Untersuchungsfläche, speziell in der mitteleuropäischen kleinräumig gegliederten Kulturlandschaft, oft von einer reinen Einteilung der Pflanz-

zendecke nach physiognomischen und/oder systematisch-strukturellen Gesichtspunkten aus (Glavac 1996; Kent & Coker 1992). Andere Autoren beziehen sich in ihren Untersuchungen auf Nutzungs- bzw. Managementaspekte oder abiotische konstante Merkmale (z.B. hydrologische Einheiten, Höhenlage) eines Raumes (u.a. Kent & Coker 1992; Mühlenberg 1993; Wildi 1986). Generell orientieren sich die Kriterien an dem vorhandenen oder in Feldbegehungen erhältlichen Wissen über einen Raum und an der Fragestellung der Untersuchung. Die Auswahl muss klar nachvollziehbar, repräsentativ und dabei doch so global sein, dass die Objektivität des Versuchs gewahrt bleibt. Ansonsten besteht die Gefahr eines Zirkelschlusses.

Überlegungen, zur Stratifizierung Vegetationsstrukturen heranzuziehen, widersprechen dem Ziel, ökologische Einheiten anhand des biotischen Inventars eines Gebietes auszuweisen (Zirkelschluss) oder deren Verhalten zu einem Gradienten zu erhellen. Durch die subjektive Einteilung wird oft ein Zusammenhang belegt, der real in dieser Deutlichkeit nicht besteht. Ordination und Klassifikation geben somit nur die Struktur bereits vorab konstruierter Klassen wieder. Außerdem steht solch ein Selektionskriterium nicht im Einklang mit dem Prinzip, floristische und faunistische Indikatorgruppen parallel in einem Projekt zu bearbeiten, da es a priori Raumgrenzen zieht, die sich nur an einer spezifischen biotischen Indikatorgruppe (z.B. Flora/Vegetation) orientieren und nicht zwangsläufig relevant sind für die faunistischen Indikatorgruppen. Unabhängiger und objektiver ist daher eine Stratifizierung der Untersuchungsräume anhand abiotischer Merkmale.

Da in der Regel die hydrologischen und bodenkundlichen Verhältnisse nicht konkret in ihrer räumlichen Verteilung bekannt sind, müssen sie durch indirekte Indizien abgeschätzt oder andere durch Feldbegehungen und sonstige Datengrundlagen (topographische Karten, Geländehöhen aus GPS-Messungen) leicht erfassbare Abgrenzungskriterien gefunden werden. Diese Kriterien sollen sich zwischen Untersuchungsgebieten übertragen lassen. Neben den Straten sollten die Untersuchungsräume selbst in ihrer Ausprägung eine gewisse Vergleichbarkeit zulassen, um die gewonnenen Bilder potenzieller Zusammenhänge zwischen Biotik und Abiotik über die Untersuchungsgebiete hinweg verifizieren oder falsifizieren zu können. Unterscheiden sich die Straten grundsätzlich in ihrem Erscheinungsbild bzw. ihren Standortbedingungen ("Ackerland / Buhnenfeld" als Extrembeispiel), so ist diese Übertragung nicht möglich, da sehr wahrscheinlich völlig andere Arten/Lebensgemeinschaften auftreten.

Resultierend aus diesen Überlegungen eignen sich morphologische Merkmale, geographische Geländehöhe und Lageparameter als primäre und z.B. Feuchteverhältnisse als multikausal abgeleitete grobe sekundäre Kriterien zur Ausgrenzung der Zielgesamtheiten/Straten. Die Nutzung als übergeordneter Standortfaktor dient meist einer "Homogenisierung" der Untersuchungsgebiete zueinander. Sie sollte zwischen und in den Untersuchungsgebieten ähnlich sein, um Ergebnisse für eine zugrundeliegende Nutzungsintensität als allgemeingültig ansehen zu können.

Der Zweck einer Stratifizierung und der Zufallsverteilung von Stichprobenflächen in jedem Stratum wird im Überblick nochmals dargestellt:

- (1) Ausschluss aller nicht untersuchungsrelevanten Flächen durch Vorwissen.
- (2) Ausreichende Erfassung relevanter, aber physikalisch kleiner Teilbereiche eines Untersuchungsgebietes bei gleichzeitiger Wahrung eines statistisch notwendigen Minimumstichprobenumfangs pro Stratum.
- (3) Sicherung des Auffindens und Wiedererkennens von Lebensgemeinschaften in ihrer spezifischen Zusammensetzung (Biodeskriptoren) unter ähnlichen Lebensbedingungen in angemessener Zahl (gleichmäßigere Raumüberdeckung als in einer reinen Zufallsverteilung).
- (4) Erfassung ökologischer Randbereiche bzw. vieler "untypischer" Flächen innerhalb eines Stratums.

- (5) Ausschalten eines systematischen Fehlers (Raster entspricht biotischem oder abiotischem Rastermuster; Subjektivität durch z.B. Entfernung zur Unterkunft).
- (6) Ausschalten von Zirkelschlüssen (subjektive Auswahl der Probeflächen nach Attraktivität, hoher Stetigkeit oder Homogenität der Biodeskriptoren etc.).
- (7) Erhöhung der Stichprobenrepräsentativität durch die Abgrenzung von Zielgesamtheiten (Straten, Blöcke) aus der Grundgesamtheit des Untersuchungsraumes.
- (8) In der Regel Reduktion des Untersuchungsaufwandes bei gleicher Ergebnissicherheit gegenüber einem Rasterverfahren oder einer echten Zufallsverteilung (Herdenbildung der Stichproben nur innerhalb einer quasihomogenen Teilfläche möglich, leistbarer Stichprobenumfang wird bestmöglich eingesetzt).
- (9) Einhaltung des Formalismus bei der Stichprobenauswahl, den numerische Verfahren der Datenanalyse voraussetzen.
- (10) Zufallsverteilung erlaubt allgemeingültige Rückschlüsse auf die Grundgesamtheit (Hypothesentests).

### 3.1.4 Räumliche und zeitliche Abstimmung der Datenerhebung – Größe und Aufbau von Stichprobenflächen

*„WESENTLICH IST, DASS DIE ZUORDNUNG DER VEGETATIONS DATEN (ANM. AUTOR: GILT FÜR ALLE BIOTISCHEN DATENSÄTZE) ZU DEN STANDORT DATEN ÖRTLICH UND ZEITLICH STIMMT. DIES IST NACH UNSERER AUFFASSUNG NUR DANN GEGEBEN, WENN ZU JEDER VEGETATIONS AUFNÄHME GLEICHZEITIG EINE STANDORT ANALYSE VORLIEGT (WILDI: 1986, S. 31).“*

Der Stichprobenplan koordiniert nicht nur die Raumverteilung der Probeflächen, sondern umfasst ebenfalls die Vorschrift zur räumlichen Anordnung der Teiluntersuchungsflächen und Messeinrichtungen beteiligter Fachdisziplinen, den Zeitplan zur Durchführung der Datenerhebung und die Gestalt einer Probefläche, welche in diesem Abschnitt beleuchtet werden.

#### **Räumliche und zeitliche Abstimmung der Datenerhebung**

Die grundlegende Notwendigkeit einer konsequenten räumlich benachbarten Erhebung aller biotischen und abiotischen Daten auf sämtlichen Stichprobenflächen wird bedingt durch die Suche nach einem numerischen Nachweis für Wechselwirkungen zwischen verborgenen oder offensichtlichen Umweltfaktoren, welche die Bestandsstruktur eines Lebensgemeinschaft-/Standortsystems ausmachen (Glavac 1996; Jongman et al. 1987; Rink et al. 2000; Wildi, 1986 u.a.). Für die statistische Datenanalyse auf mehreren Informationsebenen ist es weiterhin bedeutsam, gemeinsame Erhebungsphasen zu definieren, um die Ausprägung der Erklärungsvariablen konsistent zur aktuellen Artenkomposition in Bezug setzen zu können (Rink et al. 2000; Wildi 1986). Die Artmächtigkeit der Arten sollte zu ihrem saisonalen Optimum erhoben werden, um die korrekte Antwort zu den Umweltgradienten schätzen zu können (z.B. Frühjahrsgrophyten). Die saisonale Variation reduziert die Genauigkeit der Artmächtigkeitsschätzungen. Da die traditionellen Ansprüche verschiedener Fachdisziplinen hinsichtlich der Datenerhebung offensichtlich zeitlich ebenso unterschiedlich sind wie räumlich, ist eine Abstimmung der Einzelerhebungen oft schwierig und erfordert eine gewisse Kompromissbereitschaft, die letztlich fachlich und statistisch vertretbar sein muss (vgl. Abschnitt 5.2.3).

Die Laufzeit von Projekten beträgt meist drei Jahre, so dass die eigentlichen Beobachtungszeiträume auf maximal zwei Jahre beschränkt bleiben. Der zeitliche Anspruch für die Versuchsplanung zum Projektbeginn und die Datenaufbereitung, Synthese und Auswertung am

Projektende darf nicht unterschätzt werden. Je nach Projektziel sind zwei Jahre Feldarbeit eine sehr knapp bemessene Zeit, da die langfristige natürliche Dynamik von Ökosystemen einige Probleme für die Untersuchung aufwirft. Zwei wichtige Gesichtspunkte sollen an dieser Stelle erwähnt werden: (1) Untersuchungen in sehr dynamischen Bereichen – z.B. der rezenten Aue – können hinsichtlich der hydrologischen Situation nur ein Detailbild erfassen. Die Abfluss- und Überflutungsdynamik des Flusses ist häufig mehrjährigen Zyklen unterworfen (vgl. Abschnitt 5.1). Daher kann von einem reinen "Glücksfall" ausgegangen werden, wenn ein nasses und ein trockenes Jahr erfasst wird, die einen Vergleich der Ergebnisse in der Studie erlauben. Die langfristig angesiedelten Lebensgemeinschaften müssen aber in ihrer Existenzfähigkeit oder Wettbewerbsfähigkeit diesen Extremsituationen gewachsen sein. (2) Aus Langzeitstudien ist bekannt, dass es generell einen Artenwechsel von Jahr zu Jahr gibt, d.h., nur ein Teil aller Arten ist ständig in einem Untersuchungsraum präsent, während die übrigen nur hin- und wieder auftreten.

Beide Gesichtspunkte erschweren es, Artengemeinschaften abzugrenzen und aus ihrem Auftreten (Ist-Zustand, zweijährige Periode) klare Beziehungen zum Standortsystem geschert abzuleiten, da die Reaktionszeiten je nach taxonomischer Gruppe und Art sehr unterschiedlich sein können.

#### **Größe und Aufbau von Stichprobenflächen**

Eng mit der Größe und dem Aufbau von Stichprobenflächen verbunden ist die Größe des Untersuchungsgebietes und wiederum daraus sowie aus der Verortung der Flächen resultierend die Güte der Daten hinsichtlich der Bearbeitungsintensität. Die Bearbeitungsintensität und -genauigkeit sollte teils weiten Wegen zu den Stichprobenflächen möglichst gleich bleiben. Dies gilt nicht nur räumlich für das Untersuchungsgebiet, sondern auch qualitativ für den Bearbeiter bzw. seine Sichtweise (z.B. Schätzung der Artmächtigkeiten bei pflanzensoziologischen Aufnahmen).

Die Ausdehnung des Untersuchungsgebietes nimmt über die Untersuchungsskala (vgl. Abschnitt 2.6) Einfluss auf die Probeflächengröße, da allein dadurch oft andere Beobachtungsobjekte und Erklärungsparameter Relevanz erlangen (Jørgensen & Müller 2000). Beispielsweise sind 1-m<sup>2</sup>-Probeflächen häufig in Grünlandstudien, während 1000-m<sup>2</sup>-Flächen eher in Walduntersuchungen angewendet werden (Økland 1990; Reed et al. 1993). Ein anderer wichtiger Effekt, vergesellschaftet mit der Größe des Untersuchungsgebietes, ist die Autokorrelation (vgl. Abschnitt 3.5.9), die sich zwangsläufig graduell in Lebensgemeinschaften und Umweltparametern findet (Reed et al. 1993).

Die weitere Annahme, dass Korrelationen zwischen Umweltparametern und Beobachtungsobjekten mit der Untersuchungsgebietsgröße ansteigen, da so ein größerer Anteil an Variation erfasst wird, konnte nicht bestätigt werden (Reed et al. 1993). Eine einfache Begründung hierfür könnte sein, dass mit zunehmender Komplexität der Daten auch die Zahl sich gegenseitig beeinflussender Faktoren steigt. Ebenso kann das zufällige Auftreten eines Extremwertes für einen Gradienten die Bedeutung anderer Parameter maskieren. Somit ist die Konzentration der Stichproben auf ein mesoskaliges Gebiet zwar mit einer eventuell mangelnden Generalisierbarkeit der Zusammenhänge auf regionaler bzw. höherer Ebene verbunden, aber die Daten sind eindeutig genug, um lokal bedeutsame Prozesszusammenhänge aufzudecken und zu beschreiben, die ansonsten von weiter gestreuten Faktoren überlagert werden (Reed et al. 1993). Es ist daher zu bedenken "it is important to assess whether the process that determine community structure and function are similar at different scales, or whether, how rapidly, and under what circumstances the dominant processes change with scale of observation (Reed et al. 1993, S. 329)."

Unterschiedliche Gradienten variieren grundsätzlich auf verschiedenen Skalen zueinander (Gauch 1982a; Jørgensen & Müller 2000). Zusammenhänge zwischen Probeflächengröße und Wechselwirkungen zwischen Biotik und Umweltparametern finden sich genauso auf der

Skala "Probefläche" wieder. Ein Ansteigen der Probeflächengröße führt zwangsläufig zu einem Anstieg der abiotischen Heterogenität, welche die Anzahl der Mikro-Nischen erhöht (vgl. Abschnitt 2.6). Die meisten ökologischen Parameter zeigen eine Zunahme der Varianz, wenn die Probeflächengröße bei gleich bleibender Probeflächenanzahl erhöht wird (Økland 1990). Als Resultat ergibt sich, dass die Anzahl signifikanter Erklärungsparameter und Gradienten mit der Probeflächengröße (grain/plot size) steigt. Die Studie von Reed et al. (1993) belegt dies sowohl für die Betrachtung der Einzelparameter als auch für eine Auswertung mit der Kanonischen Korrespondenzanalyse (CCA). Mit der Zunahme der Probeflächengröße werden mehr signifikante Ordinationsachsen extrahiert (Tabelle 3.1-1).

**Tabelle 3.1-1: Signifikanz von CCA-Achsen bestimmt durch Monte Carlo Tests für verschiedene Stichprobenflächengrößen und Datensätze (\* =  $p < 0,05$ ) (nach Reed et al. 1993, S. 335).**

Probeflächengröße (m <sup>2</sup> )	Achsennummer				
	1	2	3	4	5
<b>Datensatz: 196 Probeflächen</b>					
0,0156	0,01*	0,40			
0,0625	0,02*	0,41			
0,25	0,01*	0,55			
1,0	0,01*	0,18			
4,0	0,01*	0,02*	0,03*	0,06	
16,0	0,01*	0,01*	0,01*	0,01*	0,09
<b>Datensatz: 100 Probeflächen</b>					
0,0156	0,95				
0,0625	0,33				
0,25	0,73				
1,0	0,45				
4,0	0,63				
16,0	0,01*	0,14			
64,0	0,01*	0,01*	0,14		
256,0	0,01*	0,01*	0,01*	0,05*	0,39

Nahezu alle Variablen waren bei einer spezifischen Probeflächengröße in dieser Untersuchung signifikant mit der Artenkomposition korreliert. Dennoch gab es keine Probeflächengröße für die alle Variablen signifikante Zusammenhänge zeigten, d.h., es existiert kein allgemeingültiges Gesamtmuster. Die Schlussfolgerung muss lauten: Die signifikant mit der Artenkomposition korrelierenden Umweltfaktoren wechseln mit der Probeflächengröße (Reed et al. 1993). Es liegt ein Skalenproblem vor (vgl. Abschnitt 2.6).

Die nachlassende Signifikanz der Zusammenhänge impliziert mehrere Erklärungsansätze. (1) Bei sehr kleinen Probeflächen könnte durch den physischen Kontakt inmobiler Arten Ressourcenkonkurrenz moderate Korrelationen negativ beeinflussen. Bei großen können Konkurrenzeffekte aufgehoben werden und das Muster der Lebensgemeinschaften reflektiert eher grobgerasterte Variation der physikalischen und chemischen Umwelt. (2) Die Erscheinung ist zurückzuführen auf einen Versuchsfehler (sampling error), da bei sehr kleinen Probeflächen nur wenige Individuen innerhalb der Fläche zu liegen kommen und damit der Zufallseffekt dramatisch ansteigt (vgl. auch Abbildung 3.1-5, S. 39). Erst durch die Vermeidung/Reduzierung dieses Einflusses können Zusammenhänge objektiviert und feinere Muster aufgedeckt werden. (3) Die kleinräumige Heterogenität in den Umweltvariablen verdeckt Zusammenhänge (vgl. Abbildung 2.6-1, S. 23).

In der Studie wurden 16 Meter voneinander entfernt punktuell Bodenproben entnommen. Die punktuelle Entnahme ist für kleinräumige Heterogenität anfällig, d.h., Abweichungen zu den realen mittleren Werten einer Fläche können die Folge sein. Eine Einzelprobe kann an sich nicht die Gesamtfläche befriedigend repräsentieren (vgl. Abschnitt 2.6). Für größere Probeflächen wird so ein Teil der Variation nicht erfasst. Dies scheint dem Ergebnis zu widersprechen, da die Zusammenhänge für Proben flächigerer Ausdehnung eindeutiger wurden. Dazu



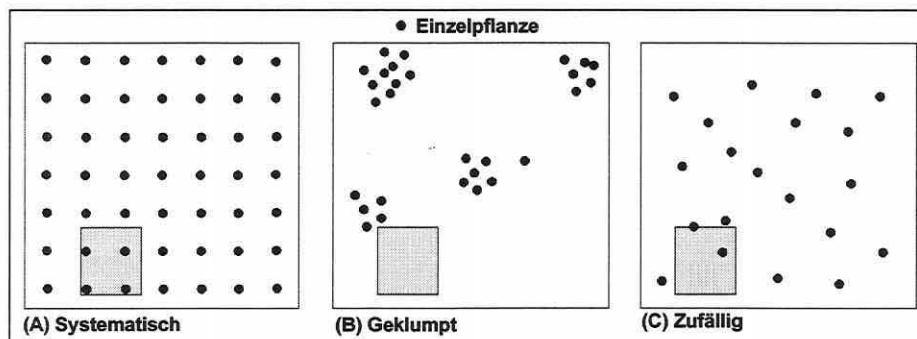
muss man wissen, dass hier ein Mittel aus mehreren Bodenproben berechnet und den kleinen Probeflächen nur ein Wert zugeordnet wurde (Reed et al. 1993).

Die Untersuchung proklamiert folgende Ergebnisse für krautige Lebensgemeinschaften (zoologische Beobachtungsobjekte wurden nicht betrachtet). Die Morphologie bestimmt die Strukturen auf einer 5-20-cm-Skala, Konkurrenz auf einer 0,25-1,5-m-Skala und Umweltfaktoren haben einen dominanten Einfluss auf Skalen  $>1,5$ m. Wie kleinräumig Pflanzenarten auf Umweltparameterunterschiede reagieren, hängt zusätzlich von ihrer physischen Größe und ihrem Wurzelsystem ab. Bäume überspannen kleinräumige Unterschiede einfach und integrieren die Umweltbedingungen zu einem Mittel (Økland 1990). Durch das Mobilitätsmerkmal der Fauna können diese Aussagen nicht uneingeschränkt übertragen werden.

Trotz einiger empirischer Richtlinien gibt es bisher keine objektiven Methoden die kleinste noch repräsentative Probeflächengröße genau zu bestimmen (Økland 1990, Tüxen 1977; vgl. Abschnitt 2.6). Daher müssen folgende Fragen zur Probeflächengröße im Voraus berücksichtigt werden:

- (1) Auf welcher Skalenebene will man Variation studieren? Problem: Dies ist schwierig, wenn die Skala der Variation nicht bekannt ist oder Parameter auf verschiedenen Skalenebenen variieren. Als Lösung bietet sich nur eine Pilotstudie an, oder man verwendet geschachtelte (genestete) Plots (s.u.).
- (2) Ist ein Vergleich zu anderen Studien vorgesehen? In diesem Falle ist dieser nur aussagekräftig, wenn die gleiche Probeflächengröße zugrunde gelegt wurde oder wenn Transferregeln (Eichungen) für Skalenübergänge bzw. die verschiedenen Flächengrößen bekannt sind.

Neben der physikalischen Umwelt bildet auch die Biotik selbst Muster aus. Aufgrund der nicht vorhandenen Mobilität der Vegetation sind diese Muster besser beschrieben als für zoologische Taxa. Es treten systematische, zufällige und geklumppte Verteilungen auf (Abbildung 3.1-5).



**Abbildung 3.1-5: Muster in der Vegetationsdecke: systematisch (A), geklumpt (B) und zufällig (C). Das Quadrat verdeutlicht eine zufällige Stichprobennahme mit unterschiedlichen Trefferquoten bei gleicher Basisgröße innerhalb der verschiedenen Muster.**

Nutzt man das kleine in Abbildung 3.1-5 dargestellte Quadrat, um es beliebig in einem der Muster zu platzieren, wird deutlich, dass die Chance eine entsprechende Anzahl Individuen zu erfassen von (A) über (C) nach (B) hin abnimmt. Die meisten Pflanzenarten tendieren zu Muster (C) (Causton 1988; Kent & Coker 1992).

Obgleich die Tendenzen in der Ökologie bekannt sind, ist eine adäquate Lösung in Multitaxa-Studien im Grunde genommen unmöglich, da jede Art ihr eigenes Muster ausbildet und nach einer jeweils angepassten Probefläche in Form und Größe verlangt. In diesem Zusammenhang können geschachtelte Probeflächen hilfreich sein, d.h. eine Beprobung auf multiplen räumlichen Skalen (Kent & Coker 1992). Durch solch eine Anwendung lässt sich der Skalen-



einfluss auf die Artenverteilung abschätzen. Bemerkenswert ist, dass langgezogene rechtwinklige Probeflächen nach Stohlgren et al. (1995) mehr Arten erfassen sollen als runde oder quadratische. Nicht nur die Größe, sondern auch die Form einer Fläche modifiziert so den Artenreichtum (species richness) pro Probefläche bei gleicher Grundfläche, d.h. geschachtelte Probeflächen sollten von Skala zu Skala nicht in der Form wechseln, wie im Whittaker-Design (vgl. Stohlgren et al. 1995). Darüber hinaus sind direkt benachbarte Flächen ausnahmslos autokorreliert und überlappende Flächen höheren Ranges können unvermeidbar nicht als unabhängige Beobachtungen hinsichtlich des Artenreichtums betrachtet werden (Pielou 1977; Stohlgren et al. 1995). Geschachtelte Probeflächen-Designs sollten somit in erster Linie aus

- (1) nicht oder nur gering überlappenden und
- (2) nicht direkt benachbarten Teilflächen aufgebaut sein (vgl. modifiziertes Whittaker-Design in Stohlgren et al. 1995).

Grundsätzlich kann ein Basisdesign flächig komprimiert oder erweitert werden (zoomen) (Stohlgren et al. 1995).

Die Anforderungen der Fachdisziplinen in einem Projekt begründen letztlich die Minimalgröße und die Lage der Probenahmestellen innerhalb einer Stichprobenfläche. Aus dem summierten Platzbedarf aller Erhebungen und Messeinrichtungen der Teilprojekte ergibt sich dann die Minimalgröße für eine Probefläche bei Multitaxa-Untersuchungen.

Eine klare Systematik der Teilflächenverortung innerhalb einer Probefläche erleichtert das Etablieren und Auffinden der Teilflächen für die Felderhebungen (vgl. Abschnitt 5.2.2). Als Richtlinie zur Anordnung der Untersuchungen auf einer Probefläche sollten Schemata im Stichprobenplan enthalten sein (vgl. Abbildung 5.2-1, S. 145, Teilbild 5). Definiert man einen festgelegten Zutrittspfad, kann die restliche Probefläche geschont werden, um Störungen für faunistische Untersuchungen zu minimieren. Jede Stichprobenfläche ist im Gelände eindeutig durch eine Verpflockung zu kennzeichnen und kann durch Koppelung von RTK-GPS Funktionalität und Metallsuchgerät bei Langzeituntersuchungen leicht wieder lokalisiert werden.

### 3.1.5 Kartengrundlagen und Einsatz eines Geografischen Informationssystems

**"EIN GEOGRAFISCHES INFORMATIONSSYSTEM (GIS) IST DANN SINNVOLL EINZUSETZEN, WENN GEOGRAFISCHE DATEN EINES RAUMES GESPEICHERT, ABGEFRAGT, ANALYSIERT, MANIPULIERT, MIT ANDEREN DATEN VERSCHNITTEN UND SCHLIEßLICH GRAPHISCH (ABER AUCH TABELLARISCH) DARGESTELLT WERDEN SOLLEN. SOMIT DIENT EIN GIS ZUR VERARBEITUNG RÄUMLICH VERTEILTER GEOGRAFISCHER INFORMATIONEN (ROSENZWEIG 1995, S. 2)."**

Der Einsatz eines Geografischen Informationssystems (GIS) ist schon zu Beginn eines Projektes sehr vorteilhaft, da es nicht nur als grundlegendes Instrument zur Erstellung eines Stichprobenplans angesehen werden kann (vgl. Abschnitt 5.2.2), sondern auch der Visualisierung des Untersuchungsraumes (Ausgabe analoger Arbeitskarten beliebigen Maßstabs), der Datenhaltung, der Datensynthese, der Darstellung der Ergebnisse verschiedener Arbeitsschritte (vgl. Abschnitte 5.1, 5.2.1 und 5.3.4) und letztlich der Modellierung des Prognosesystems (vgl. Abschnitt 5.6) dient.

Da der Aufbau des GIS als Basis aller weiteren Bearbeitungen anzusehen ist, muss zuerst eine Kartengrundlage geeigneten Maßstabs und akzeptabler Aktualität gewählt werden. Der geeignete Maßstab einer Arbeitskarte ist abhängig von der physikalischen Größe der Unter-

suchungsräume und von den auf die Projektzielsetzung ausgerichteten Feldarbeiten. Die Untersuchungsgebiete im RIVA-Projekt umfassen beispielsweise jeweils eine Fläche von ca. 0,3-1,3km<sup>2</sup>. Der Flächeninhalt einer Probefläche beträgt etwa 450-500m<sup>2</sup>. Diese Ausgangsdaten lassen einen Übersichtsmaßstab von 1:2000 und einen Arbeitskartenmaßstab für die Geländeeintragungen von 1:500 geeignet erscheinen. Die Ausgabe von analogen Karten aus dem GIS für die genannten Maßstabszahlen erfordert daher eine Kartengrundlage, welche die Topographie der Räume genau genug abbildet. Die dargestellte Topographie sollte in Deichvorländern nicht älter als fünf Jahre sein, um morphologische Veränderungen beispielsweise durch Hochwasserereignisse so weit zu erfassen, dass die aktuelle Realität noch widergespiegelt wird (vgl. Abschnitt 5.3.3).

### 3.1.6 Satellitennavigationssysteme und ihre Anwendung bei der Erstellung von Stichprobenplänen

*"MIT DER ENTWICKLUNG DES NAVIGATION SATELLITE TIMING AND RANGING GLOBAL POSITIONING SYSTEM, KURZ: NAVSTAR GPS, IST DER TRAUM VON DER BLACK BOX, DIE EINE GENAUE POSITIONIERUNG AUF JEDEM PUNKT DER ERDE ZU JEDER ZEIT BEI JEDEM WETTER ERLAUBT, WIRKLICHKEIT GEWORDEN (FRÖHLICH & GRIMM 1995, S. 3)."*

Der Einsatz eines Satellitennavigationssystems (global positioning system; GPS) zur Umsetzung des Stichprobenplans im Gelände, dem Einmessen und Wiederauffinden von Messeinrichtungen und der Probeflächen sowie einer Überprüfung des Digitalen Höhenmodells (DGM) ist zeitlich effizient und erfordert nur eins bis zwei Personen.

Um die allgemeine Funktionsweise verständlich zu machen, wird eine kurze Übersicht über den Grundaufbau gegeben. Die drei Segmente des Systems gliedern sich in: (1) das Raumsegment, (2) das Kontrollsegment und (3) das Nutzersegment.

Das **Raumsegment** wird aufgebaut aus der Gesamtheit der verfügbaren Satelliten (z.Z. 24). Sie sind auf sechs Flugbahnen in einer Höhe von etwa 20.000km angeordnet, so dass sich jederzeit wenigstens vier Satelliten bei einer durch die Flughöhe bedingten vorteilhaft kleinen Gesamtzahl empfangen lassen.

Das **Kontrollsegment** umfasst die "Master Control Station" in Colorado Springs und vier weitere Stationen. Die Aufgabe des globalen Netzes besteht darin, die gemessenen Pseudoentfernungen aller sichtbaren Satelliten zu sammeln, diese hinsichtlich spezifischer Fehler zu korrigieren und das Messrauschen zu eliminieren. Die aufbereiteten Daten werden dann zum eigentlichen Datensignal zusammengestellt. Alle acht Stunden erfolgt die Übertragung zu den jeweiligen Satelliten, die das Signal wiederum zum Empfang mit einer GPS-Antenne bereitstellen.

Das **Nutzersegment** besteht in der Basisausführung aus Empfänger und Antenne. Die Satellitenempfangsanlage zeichnet alle Signale der sichtbaren Satelliten auf. Der Signalcode wird praktisch zeitgleich im Satellit und im Empfänger auf der Erde erzeugt. Aus dem Vergleich der beiden Codes, d.h. ihrer Verschiebung zueinander, lässt sich die Laufzeit des Signals berechnen. Der kalkulierte Weg wird als Pseudoentfernung bezeichnet. Insgesamt sind vier Satelliten notwendig, um eine genaue Positionsbestimmung durchzuführen (Fröhlich & Grimm 1995).

Der konkrete mathematische Vorgang der Positionsberechnung und das meist beschrittene Auswertekonzept der doppelten Differenzen (Relativbeobachtung mit zwei Empfängern: RTK-GPS) führt im Rahmen dieser Übersicht zu weit und es sei auf die Fachliteratur verwiesen (u.a. Fröhlich & Grimm 1995, S. 29ff und Seeber 1996, S. 58f).

Die in der Versuchsplanung des RIVA-Projektes angewendeten Systeme werden in Rink (1998) näher beschrieben, ihre Vor- und Nachteile gegenübergestellt und ihre Anwendung im Feld dargelegt.

### 3.1.7 Abiotischer Minimumdatensatz und allgemeine Datensatzeigenschaften

**"DAS ZIEL DER SYNBIOLOGISCHEN FORSCHUNG MUSS NATÜRLICH DAS SEIN, DASS MAN FÜR JEDEN ÖKOLOGISCHEN FAKTOR DIE GRENZWERTE ANGEBEN KANN, ZWISCHEN DENEN DER BETREFFENDE FAKTOR IN EINER GEWISSEN ASSOZIATION VARIERT (DU RIETZ 1921, S.246)."**

Es wird aufgezeigt, welche Anforderungen an die grundlegenden Daten in einem Projekt gestellt werden sollten und wie eine Auswahl an erklärenden Variablen zustande kommen kann.

#### **Abiotischer Minimumdatensatz**

Die Erfüllung des zitierten Ziels synbiologischer Forschung setzt zwar voraus, für jeden Standortfaktor eine arten- oder lebensgemeinschaftsbezogene Spanne anzugeben, in der Praxis ist dies aber nur ansatzweise realisierbar. Rein pragmatisch muss schon aus ökonomischen Gründen eine Auswahl an abiotischen Feld- und Laborparametern getroffen werden, die erhoben werden sollen. Eine allgemeine Reduktion an Informationen muss daher in jeder Ökosystemanalyse – mangels Datenvollständigkeit – akzeptiert werden.

Ein Beispiel aus der Vegetationskunde soll das Problem verdeutlichen: Meist ist die Anzahl pflanzensoziologischer Aufnahmen umgekehrt proportional zu der Intensität der durchgeführten Standortuntersuchungen, d.h. Vegetationsaufnahmen auf breiter biometrischer Grundlage, welche eine induktive Verallgemeinerung der Befunde und eine statistisch gesichertere Aussage rechtfertigen, können aus finanziellen und arbeitstechnischen Gründen nur von verhältnismäßig wenigen Standortuntersuchungen begleitet werden (Glavac 1996). Bestätigt wird dies auch durch die Stichprobensituation im RIVA-Projekt: Theoretisch sind für die Zielsetzung des Projekts 200 oder mehr Stichprobenflächen statistisch wünschenswert, um mit gut gesicherten Ergebnissen rechnen zu können. Demgegenüber sind umfangreiche, über längere Zeit verlaufende mit apparativem Aufwand verbundene abiotische Standortuntersuchungen nur auf maximal 60 Stichprobenflächen leistbar (vgl. Abschnitt 5.2.1).

Daraus resultiert sowohl eine Einschränkung des Untersuchungsumfangs auf wenige oder nur einen Landschaftsbereich (siehe Abschnitt 5.1.2) als auch eine Beschränkung der abiotischen Erhebungen auf einen sorgfältig zusammengestellten Minimumdatensatz. Dieser Mindestdatensatz soll als unterster Rahmen der definitiv zu erhebenden Daten gelten und orientiert sich an den beteiligten Fachdisziplinen, dem Vorwissen von Experten und Erkenntnissen aus der Literatur, einer praxisnahen Erhebbarkeit der Parameter und den Anforderungen der ökologischen Modellierung selbst. Das spätere Prognosesystem erfordert die Regionalisierbarkeit der Punktdaten in die Fläche. Ist die nicht möglich, muss ein Parameter trotz seiner Erklärungsanteile am biotischen Raummuster als Modellparameter verworfen werden.

Der "Gebrauchswert" eines Parameters hängt also mitunter von seiner Beschreib- und Messbarkeit ab, da er ansonsten nur schwer zum Nachweis der Zusammenhänge im Beziehungsgefüge Lebensgemeinschaft/Standorttyp verwendet werden kann (vgl. Abschnitt 5.4.1).

Weiterhin ist die allgemeine Frage nach den entscheidenden Standortfaktoren eine kurze Überlegung wert:

*Welches ist der wichtigste Standortfaktor oder die entscheidende Standortfaktorenkombination? – Was bestimmt das räumliche Vorkommen und die Artenkombination einer Lebensgemeinschaft?*

Die Kriterien, die eine Auswahl der Faktoren bedingen, sind immer geprägt von dem Blickwinkel des Forschers (Landwirtschaft, Naturschutz, Bodenkunde, Hydrologie etc.), der Fragestellung, der oben schon angesprochenen Mess- und Nutzbarkeit sowie anderen subjektiven Einflüssen. Meist werden dabei die Standortfaktoren als wichtiger interpretiert, die eine deutliche Veränderung des zu erklärenden Phänomens bewirken. Letztlich ist dies aber nur der Versuch, ein komplexes kybernetisches System mit allen seinen Beziehungen zwischen den Faktoren zu vereinfachen, um sich dann die Wechselwirkungen des gewählten Faktorenkomplexes besser vorstellen zu können. Der Effekt eines solchen Bestrebens ist, dass ein wichtigerer Faktorenkomplex von einem weniger wichtigen oder interessanten abgetrennt wird, zumal die in der "Black Box" enthaltenen Faktoren (bisher unbekannte Faktoren) überhaupt nicht mit in die Auswahl eingehen können. Prinzipiell ist also das Auffinden des wichtigsten Faktors unter allen möglichen Einflussfaktoren fast unmöglich (vgl. Reed et al. 1993; Richter & Söndgerath 1990; Abschnitt 3.1.4), ebenso ist eine getroffene Auswahl immer relativ. Auch die Gradientenanalyse kann nur von der Annahme ausgehen, dass eine enge Korrelation zwischen Art und Standortfaktor einen engen ökologischen Zusammenhang beschreibt, der nicht kausaler Natur sein muss (Gigon 1983).

Trotzdem dient dieser Ansatz dem Verständnis der Wirkzusammenhänge und es ist möglich, die entscheidenden Faktoren einzugrenzen. Findet man einen Parameter gleicher Größenordnung in zwei unterschiedlichen Lebensgemeinschaften vor oder unterliegt dieser Umweltfaktor weiten regellosen Schwankungen innerhalb der Artenkombination, so kann er nicht entscheidend sein für eine Differenzierung der betrachteten Lebensgemeinschaften. Ein Ergebnis von Literaturstudien ist, dass meist kein Einzelfaktor als Unterscheidungsmerkmal der Lebensgemeinschaften herangezogen werden kann, sondern dass nur eine spezifische Kombination aus Einzelfaktoren die Artenzusammensetzungen voneinander zu trennen vermag (vgl. Abschnitt 2.3). Bach schreibt hierzu: "Die entscheidenden Standortfaktoren resultieren aus dem Zusammenspiel vieler Einzelfaktoren, die zum Teil unabhängig voneinander variieren und sich in der Wirkung ersetzen können. Die Einzelstandorte einer Pflanzengesellschaft sind darum nicht in allen Eigenschaften gleich, sondern nur in ihrer Gesamtwirkung: Sie sind analog (Bach in Glavac 1996, S. 90-91)."

Dieser Exkurs soll zeigen, dass es sinnvoll ist, einen umfangreichen abiotischen Datensatz zu erheben. Es besteht aber weiterhin die Notwendigkeit, Vorwissen in die Parameterauswahl einzubringen, da nicht jedes Forschungsprojekt bestehendes Wissen wieder neu entdecken muss, sondern dieses ergänzen sollte. Als Übersicht des Umfangs direkt und indirekt auf z.B. Pflanzen wirkender Faktoren dient Abbildung 3.1-6 (S. 44, s.a. Glavac 1996, S. 76). Allgemein sind häufig von besonderer Bedeutung für eine Biogeozönose die Bewirtschaftungsart und innerhalb eines Klimagebiets die Faktoren des Wasser- und Nährstoffhaushalts. In der Literatur finden sich zum Teil Empfehlungen, welche Umweltfaktoren für einzelne Taxozönosen von Relevanz zu sein scheinen (z.B. Glavac 1996).

Bei der Auswahl von Erklärungsvariablen sollten somit folgende Kriterien bedacht werden:

- (1) Repräsentativität für die Beobachtungsobjekte und den Untersuchungsraum
- (2) Messtechnischer Aufwand; Erhebung auf allen Probeflächen notwendig
- (3) Quantifizierbarkeit; z.B. Beweidungsdruck, Erosionsprozesse
- (4) Vergleichbarkeit mit anderen Studien (häufig beprobte Variablen)
- (5) Kleinräumige Heterogenität und zeitliche Stabilität des Parameters
- (6) Einflüsse des Parameters auf verschiedenen Skalenebenen
- (7) Bestreben, neue beschreibende Variablen aufzudecken

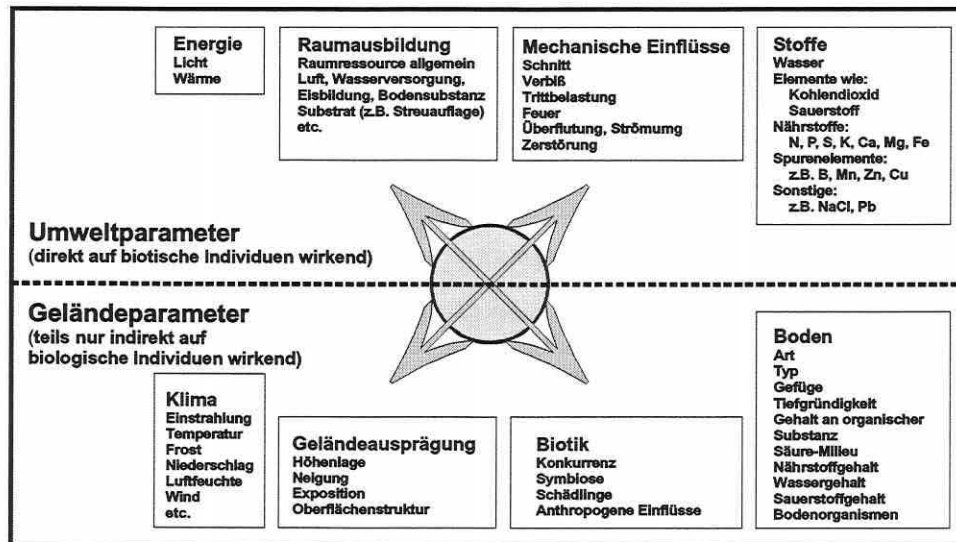


Abbildung 3.1-6: Verzahnung direkt und indirekt auf biologische Individuen wirkender Standortparameter.

### Allgemeine Datensatzeigenschaften

Die Zielsetzung und Erhebungsskala einer Untersuchung definiert automatisch die Anforderungen an die **Datengenauigkeit**. Auf Stichprobenflächengrößen von mehreren 100m<sup>2</sup> ist nicht mit einer völligen abiotischen Homogenität zu rechnen. Ebenso ist die biotische Datenerhebung nicht exakt: Die Güte der pflanzensoziologischen Aufnahmen ist von der Qualifikation und die Artmächtigkeitswerte sind von der subjektiven Einschätzung des Bearbeiters abhängig. Selbst zwei hoch qualifizierte Vegetationskundler kämen sicherlich nicht zu einem absolut identischen Ergebnis. Gerade bei wiederholtem Monitoring der Flächen spielt der Einfluss des Bearbeiters eine Rolle. Bei der erneuten Schätzung von Artmächtigkeiten ein Jahr später können geringe Veränderungen nicht realisiert werden. Frequenzerhebungen liefern oft genauere Ergebnisse als Schätzungen des Deckungsgrades (Økland 1990). Grundsätzlich sollten auch bei der Betrachtung biotischer Felddaten drei Eigenschaften – Datenrauschen (strukturlose Variation, noise), Redundanz und Ausreißer (outlier) – nicht übersehen werden.

Wiederholte Aufnahmen einer Lebensgemeinschaft bei gleichen Standortbedingungen sind selten identisch, d.h., die Daten reflektieren teils interessierende Struktur und teils unerwünschtes **Datenrauschen**. Diese strukturlose, nicht systematisch erklärbare Antwort der Arten wird bedingt durch: zufällige Verteilung und Etablierung von Individuen, lokale externe Störfaktoren (z.B. anthropogene Aktivitäten) und Heterogenität der Standortbedingungen auf einer der Untersuchungsskala untergeordneten räumlichen Skala. Weiterhin können die Annahmen in Auswerteverfahren und Genauigkeiten der Messgeräte bzw. der Erhebungsmethode modifizierend wirken. Als typischen Grad des Datenrauschens gibt Gauch einerseits 10-50% (Gauch 1982a, S. 8) und andererseits einige Seiten später durchschnittlich 77% an (Gauch 1982a, S. 105). Trotz dieses "Widerspruchs" bleibt die Existenz eines erheblichen Datenrauschens bei Felddaten unangetastet (Gauch 1982a; Kent & Coker 1992).

**Redundanz** kann als Gegenstück zum Datenrauschen beurteilt werden. Es ist leicht nachvollziehbar, dass mittlere Standortbedingungen mit entsprechenden Lebensgemeinschaften prozentual häufiger beprobt werden als Extrema. Felddatensätze haben somit einen deutlichen Grad an Redundanz. Dieser auf den ersten Blick keine Zusatzinformation beinhaltende Datensatzanteil erweist sich aber dennoch als wichtig, um Strukturzusammenhänge bestäti-



gen zu können, da mit zunehmender Redundanz ein Zusammenhang besser abgesichert werden kann. Datenrauschen lässt sich erst durch redundante Daten von der strukturierten Variation in den Daten unterscheiden (Gauch 1982a; Kent & Coker 1992).

Die Reaktion aller oder nahezu aller zusammen lebender Arten zu relativ wenigen Komplexgradienten ist der Hauptgrund für die **strukturierte Variation** in einem Ökosystem. Sie resultiert aber ebenso aus Interaktionen zwischen den Arten. Die strukturierte Variation ist die prinzipiell vorhersagbare Variation aus (a) den Standortbedingungen und (b) dem Wissen über An- oder Abwesenheit konkurrierender, prädativer oder symbiontischer Arten (s.a. Økland 1990). Ziel der Mustererkennung ist es, einen möglichst großen Anteil der Struktur in den Daten zu extrahieren, wobei Ordinationsverfahren sehr effektiv sind, da sie hierarchisch arbeiten und den deutlichsten Gradienten der ersten Ordinationsachse zuweisen. Durch dieses Prinzip wird strukturloses Datenrauschen auf nachfolgende Achsen verlagert, die dann auf ihre Signifikanz hin getestet werden können (vgl. Abschnitt 3.6.1).

**Ausreißer** sind in Felddaten die Regel, da gestörte ungewöhnliche Standorte oder Lücken in der Datenerhebung normal sind. Ein einziger Wert bzw. eine Probestfläche, die dem Datenschwarm gänzlich unähnlich ist, wird als Ausreißer im eigentlichen Sinne der Terminologie verstanden. Ausreißer werden in den Ordinationsverfahren verschieden gehandhabt. Sie erscheinen rein zufällig entweder um den Ursprung oder weit in der Peripherie des Diagramms oder verursachen sehr lange Gradienten, wobei sie den eigentlichen Datenkörper in einer unbefriedigenden Weise auf kleinem Raum zusammenpressen. Egal welche Variante zutrifft, es entsteht nicht wirklich ein Fehler, da diesen Ausreißern ohnehin keine Information zugrunde liegt. Dennoch können die Ergebnisse verwirrend und/oder gestaucht sein. Es daher ratsam, solche Ausreißer vorab zu ermitteln und zu entfernen (Gauch 1982a). Sie können sowohl durch die Ordination selbst oder effektiver durch ein Klassifikationsverfahren bestimmt werden.

Bilden sich zwei Datengruppen in der Ausgangsmatrix ab, die sich deutlich in ihrer Artenzusammensetzung oder ihren Standortbedingungen unterscheiden, spricht man von **disjunkten Blöcken**. Die erfolgreiche Analyse des Übergangsbereichs ist unmöglich aufgrund der fehlenden Information. Dennoch schätzen direkte Ordinationsverfahren solche Lücken gut aus der vorhandenen Information, wenn sie nicht zu groß werden (vgl. Abschnitt 3.5.8). Es kann sinnvoll sein, die Gruppen einer Einzelbetrachtung zu unterziehen (vgl. Abschnitt 5.4.3). Speziell Disjunktionen ganzer Datengruppen lassen sich oft durch Zusatzinformation erklären. Als extremes Beispiel kann man sich die Wasserversorgung in einer Oase und in der umliegenden Wüste vor Augen führen.

Besondere Aufmerksamkeit sollte auch der Erhebung abiotischer Daten geschenkt werden:

Die **bodenkundlichen Standortparameter** lassen sich in dynamische (z.B. Bodentemperatur) und langfristig invariante Faktoren (z.B. Korngrößenverteilung) einteilen. Nach ihrer zeitlichen Stabilität richtet sich letztlich der Zeitpunkt einer sinnvollen Erhebung, das notwendige Erhebungsintervall bzw. ihre generelle Anwendbarkeit für ein Prognosesystem.

Die **hydrologischen Messgrößen** sind als dynamische Parameter zu bezeichnen, deren Erhebung nur Sinn macht, wenn ihre Schwankungen in Zeitreihen festgehalten werden. Einzelwerte geben keinen Aufschluss über die realen Verhältnisse. Hügin (1990) verbalisiert einige wichtige Parameter des Grundwassers, und wie generell aus den Ergebnissen synökologischer Untersuchungen hervorgeht (s.a. Hügin 1990), ist für die Pflanzengesellschaften von Flussauen der Wasserfaktor das prägende Element. Hügin & Henrichfreise (1992, S. 10) schreiben: "Das Wasser ist in Flusslandschaften entscheidend für die Abwandlung der Standorte. (...) Die Standortdifferenzierung einschließlich der Nährstoffverfügbarkeit wird überwiegend vom Wasser bestimmt."

Indirekt ist also nicht nur die Wasserversorgung der Pflanze, sondern auch die Nährstoffzufuhr mit diesem Faktor gekoppelt: "Die Mobilität der Nährstoffe in Böden ist im Wesentlichen



eine Funktion ihrer Konzentration in der Bodenlösung (Marschner 1985, S. 292)." Diese Aussage bestätigt sich auch in den Ergebnissen des RIVA-Projekts (vgl. Abschnitte 5.4.1 und 5.4.2).

Weiterhin ist der Literatur zu entnehmen, dass die biotische Ausgestaltung von Flussauen hinsichtlich ihres Grundwasserregimes nur mit der nötigen Genauigkeit analysiert werden kann, wenn die Auflösung der **Geländehöhen** und die der Grundwasserflurabstände in 1cm-Genauigkeit vorliegen (Hügin & Henrichfreise 1992). Diese geringe Fehlerrate in den Geländehöhen erscheint jedoch unrealistisch, da diese im 1cm-Bereich durch vielerlei Faktoren modifiziert wird und das exakte Mikrorelief zum aktuellen Stand der Technik (Befliegung) nicht flächendeckend erfassbar ist. Genauigkeiten im 1cm-Bereich sind bisher nur durch statische Einmessungen der Stichprobenflächen mit einem Nivellier erzielbar und würden alleine auf einer Gesamtfläche von mehreren 100m<sup>2</sup> pro Probefläche nur einen Bruchteil aller Möglichkeiten abdecken ohne inflationär zu werden. Eine flächige Anwendung zur Höhenbeschreibung eines Untersuchungsgebietes ist bei einem engen Raster unvorstellbar aufwendig. Um eine realisierbare Genauigkeit zu erhalten, können RTK-GPS Messungen ( $\pm 5\text{cm}$ ) und/oder die Erstellung eines DGM aus Luftbildern mit einer nachfolgenden photogrammetrischen Auswertung ( $\pm 15\text{cm}$ ) durchgeführt werden (Peter et al. 1999; Tabelle 5.3-5, S. 158). Die genannten Höhengenaugigkeiten wurden in einer Studie für die Untersuchungsgebiete des RIVA-Projekts nach dem Schema "Kontrolle des DGM durch ein GPS-System und Kontrolle dieses durch ein Nivellier überprüft". Das Ergebnis zeigt, dass die GPS-Messungen – im Falle korrekter Bezugspunkte – Lage- und Höhenwerte ergeben, die in fast allen Fällen nicht mehr als  $\pm 5\text{cm}$  von der Wirklichkeit abweichen (vgl. Jurisch 1998; Peter et al. 1999).

Die Lagedifferenzen zwischen DGM und GPS liegen zu 90% im Intervall von  $\pm 5\text{cm}$  und in 0,2% der Fälle ( $n=1176$ ) wird eine Abweichung von 10cm überschritten. In der Höhengenaugigkeit treten reliefbedingt größere Ungenauigkeiten auf. Nur 47% aller Rasterpunkte erfüllen die geforderte Genauigkeit von  $\pm 15\text{cm}$ . 96% aller Werte finden sich im Intervall  $\pm 30\text{cm}$  wieder (Peter et al. 1999). Da das DGM eine der wichtigsten Datengrundlagen darstellt, weil es als Referenz für viele Sekundärdaten (z.B. Grundwasserflurabstände bei der Regionalisierung von Punktdaten) herangezogen wird, sollten Fehlerquellen auf ein Mindestmaß reduziert werden. Vor der Luftbilderstellung ist auf die Auswahl des geeignetsten Höhensystems und entsprechender Festpunkte zu achten. GPS-Vergleichsmessungen zur Kontrolle sollten immer erfolgen (Peter et al. 1999).

### 3.2 Ableitung von Sekundärdaten: Datengewinnung mit der ARC/INFO®-Applikation "DIST"

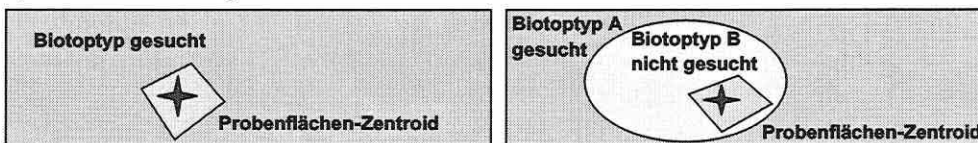
*ES KANN NICHT DAVON AUSGEGANGEN WERDEN, DASS DIE ANTWORT DER BEOBACHTUNGSOBJEKTE UNABHÄNGIG VON DER SKALIERUNG EINES PARAMETERS IST. HÄUFIG IST EINE REAKTION DER ARTEN ERST NACH EINER SKALENTRANSFORMATION DER ROHDATEN NACHWEISBAR (NACH PALMER 1993).*

Die Ableitung von Sekundärdaten aus Rohdaten ist eine gebräuchliche Vorgehensweise. Es erübrigt sich, gängige Ableitungen wie Differenzbildungen, Mittelwerte oder zeitabschnittbezogene Aufteilung von Datenreihen näher zu erläutern. Es soll aber eine ARC/INFO®-Applikation vorgestellt werden, deren Idee und technische Umsetzung im Zuge der vorliegenden Arbeit entstand: Die ARC/INFO®-Applikation "DIST" zur Berechnung von Flächenanteilen und Distanzen (Rink & Schröder 1999, BfG Koblenz, unveröffentlicht).

Die Applikation wurde zur interaktiven automatisierten Berechnung von Distanzen zwischen Punktelelementen und Flächen (Polygonen, wie z.B. Biotoptypen) erstellt. Außerdem kann die



Der erste Fall (Abbildung 3.2-2, links) ist unproblematisch: Bei einer A-priori-Kontrolle der Eingangsdaten wird für den Fall, dass die Probenahme­fläche bzw. ihr Zentroid sich innerhalb eines Polygons des gesuchten Biototyps befindet, die Distanz auf null gesetzt und die entsprechende Flächengröße ermittelt.



**Abbildung 3.2-2: Problemfälle bei der Programmierung von DIST: Probenahme­fläche liegt innerhalb des Biototyps (Distanz = 0; links); Probenahme­fläche liegt innerhalb einer Inselfläche (rechts).**

Der zweite Fall (Abbildung 3.2-2, rechts) ist wesentlich schwieriger zu lösen. Der Umfang und die Fläche des gesuchten Biototyps werden immer durch mindestens zwei Polygone bestimmt. Ein Rückschluss auf die Polygonfläche über einen Vergleich der Längsinformation der Umgrenzungslinie mit dem Umfang des Polygons ist also nicht ohne weiteres realisierbar. Außerdem ist die Distanz zur Umgrenzungslinie des Biototyps B das gesuchte Ziel­element und nicht die Distanz zur äußeren Begrenzung des eigentlich gesuchten Biototyps A. Es wird fälschlicherweise auf die Fläche des Inselfpolygons geschlossen, nicht aber auf die des gesuchten Polygons. Die Fehlinterpretation kann nur umgangen werden, wenn vorher das Inselfpolygon eliminiert wird. Als Folge ist jedoch keine Distanzberechnung zum Inselfpolygon mehr möglich.

Dieser Fall wird in der Ausgabe der Daten vermerkt und statt der Flächengröße des gesuchten Biototyps, die der Inselfläche ausgegeben. Im Bedarfsfall muss dann der Flächeninhalt des gesuchten Biototyps manuell ermittelt werden.

Weitere Schwierigkeiten treten bei der Berechnung der Flächenanteile einzelner Biototypen innerhalb eines festgelegten Radius um das Zentroid der Probefläche auf. Der Befehl "BUFFER" legt zwar entsprechend der gewünschten Distanz Kreisflächen um die Zentroide der Probeflächen, bei Überschneidungen gehen jedoch die Informationen zu den einzelnen Probenahme­flächen (Nummerierung) verloren und die entstandenen Kreisflächen werden zu einer Fläche verschmolzen. Die Folge ist, dass kein Bezug zur ursprünglichen Flächengröße mehr besteht und darüber hinaus auch nicht mehr zur Nummer (Kennung) der Probefläche. Somit müssen alle Zentroide der Probeflächen einzeln gebuffert, zunächst untereinander und danach mit den gesuchten Biototypen verschnitten werden. Anschließend werden die jeweiligen Informationen wieder zusammengefasst. Bei 15 Probenahme­flächen entstehen 50 Zwischenstände und bei 30 Biototypen 1500 temporäre Themen nur für die Berechnung der Flächenanteile in einem Untersuchungsgebiet. Ein bisher nicht gelöstes Problem in ARC/INFO® ist die Flächenberechnung der Bufferflächen. Wenn eine Kreisfläche um ein Zentroid mit einer Distanz von 60m berechnet wird, erhält man als Flächeninformation den Wert 11252,08m². Die eigentliche Kreisfläche beträgt allerdings  $A = r^2 \cdot \pi = 11309,73\text{m}^2$  (A: Fläche; r: Radius). Dieser Unterschied resultiert aus der Tatsache, dass Bufferflächen um das Zentroid als kreisähnliches Vieleck berechnet werden. Eine hinreichende Genauigkeit, um die prozentualen Anteile bestimmter Biototypen in einer Bufferzone zu berechnen, ergibt sich bei Bufferdistanzen von 10 bis 150m. Für die Bearbeitung dreier Untersuchungsgebiete mit jeweils ca. 30 verschiedenen Biototypen resultieren ca. 10.000 Arbeitsschritte. Aus diesem Grund wurde zur Vereinfachung die Applikation geschrieben, welche menügesteuert die Bearbeitungsschritte weitestgehend automatisiert. Für jeden Biototyp pro Untersuchungsgebiet muss die Applikation einmal angewendet werden (90 Arbeitsschritte). Durch die allgemeine Form der Applikation steht sie über die konkrete Fragestellung in RIVA hinaus auch für andere Aufgaben oder Nachbearbeitungen zur Verfügung. Die Abbildung 3.2-3 (S. 49) gibt einen Einblick in die Ausgabetabelle der Applikation.

	A	B	C	D	E	F	G	H	I	J	K
1	PF_Nr	DISTANCE	ASPECTS	AREA_409	ISLAND	AREA_BUF	RELATIV	DISTANCE_M	ASPECTS_MI	AREA_409MI	IS_MIN
2	1	0,00000	Inside	342170,00000	0,00000	26843,198290	86	0,00000	Inside	342170,00000	0,00000
3	2	116,17320		342170,00000	0,00000	0,000000	0	116,17320		342170,00000	0,00000
4	3	0,00000	Inside	342170,00000	0,00000	26641,956160	85	0,00000	Inside	342170,00000	0,00000
5	4	0,00000	Inside	342170,00000	0,00000	29186,594460	93	0,00000	Inside	342170,00000	0,00000
6	5	0,00000	Inside	342170,00000	0,00000	27422,294300	88	0,00000	Inside	342170,00000	0,00000
7	6	0,00000	Inside	342170,00000	0,00000	17199,316680	55	0,00000	Inside	342170,00000	0,00000
8	7	0,00000	Inside	342170,00000	0,00000	30716,627340	98	0,00000	Inside	342170,00000	0,00000
9	8	0,00000	Inside	342170,00000	0,00000	31255,785300	100	0,00000	Inside	342170,00000	0,00000
10	9	0,00000	Inside	342170,00000	0,00000	27527,340500	88	0,00000	Inside	342170,00000	0,00000
11	10	0,00000	Inside	226522,50000	0,00000	28970,770350	93	0,00000	Inside	226522,50000	0,00000
12	11	0,00000	Inside	226522,50000	0,00000	22685,561430	73	0,00000	Inside	226522,50000	0,00000
13	12	9,93417	Surround	0,00000	50757,50000	17071,861380	55	9,93417	Surround	0,00000	50757,50000
14	13	0,00000	Inside	226522,50000	0,00000	25785,636280	82	0,00000	Inside	226522,50000	0,00000
15	14	0,00000	Inside	226522,50000	0,00000	29993,293220	96	0,00000	Inside	226522,50000	0,00000
16	15	0,00000	Inside	226522,50000	0,00000	27727,940980	89	0,00000	Inside	226522,50000	0,00000

Abbildung 3.2-3: Datenausgabe von DIST als dBase-File

**Feldbeschreibungen zu Abbildung 3.2-3:**

Pf\_Nr: Probeflächennummer

Distance: Minimumdistanz zwischen Probeflächenzentrum und Zieleinheit

Aspects: Zusatzinformationen wie INSIDE (innerhalb) und SURROUND (Insellage)

Area\_409: aktuelle Fläche der zugeordneten Zieleinheit

Island: Distanz zum Inselpolygon

Area\_Buf: Flächenanteil der Zieleinheit innerhalb des gewählten Buffers (Summe)

Relativ: Prozentanteil der Fläche einer gewählten Zieleinheit an der Gesamtfläche des Buffers

XXX\_Min: Analoge Angaben für die Zieleinheit, welche die interaktiv gewählte Minimumfläche überschreitet

**3.3 Datenaufbereitung für die statistische Analyse**

*"THE RANGE OF RESULTS OBTAINABLE BY DIFFERENT DATA EDITINGS IS GREATER THAN THAT OBTAINABLE BY EMPLOYING DIFFERENT MULTIVARIATE TECHNIQUES BECAUSE DATA EDITING CHANGES THE DATA SUBMITTED FOR ANALYSIS (GAUCH 1982A, S. 212)."*

Daten lassen sich in verschiedene Gruppen einteilen. Daher sollen in diesem Abschnitt die grundlegenden Begrifflichkeiten geklärt werden, bevor spezifische Kodierungen und Datenaufbereitungen zu erörtern sind.

Die **Rohdaten** bilden die Gesamtheit der zur Verfügung stehenden Daten in einer Analyse. Diese Daten werden aus den Antwortvariablen (Beobachtungsobjekte) und – falls verfügbar – den Erklärungsvariablen (Deskriptoren) gebildet. Zur Unterscheidung müssen die Begriffe "Beobachtungsobjekt" und "Deskriptor" näher definiert werden.

Die **Beobachtungsobjekte** in ökologischen Studien können Raumeinheiten unterschiedlicher Größe und Dimension ebenso wie Individuen/Arten oder sonstige durch die Datenerhebung/den Versuch festgelegte Einheiten umfassen. Sie können oft in einem sehr präzisen quantitativen oder semi-quantitativen Weg erhoben werden. Als Beispiele seien Artmächtigkeiten nach Braun-Blanquet, Prozentdeckungen, die Biomasse von Pflanzenarten oder Zählraten von zoologischen Individuen angeführt. Die einfachste Variante der Datenerhebung stellen Präsenz-/Absenzdaten (1/0) für die Beobachtungsobjekte dar.

Die Bezeichnungen **Deskriptor** oder **Variable** werden synonym für Erklärungsvariablen verwendet. Sie beziehen sich auf Eigenschaften und Charakteristika, welche die Beobachtungsobjekte beschreiben. Hierzu zählen physikalische, chemische, ökologische und biologische Eigenschaften. Deskriptoren können quantitative Variablen (z.B. pflanzenverfügbare Phosphor in mg/100g Boden), semi-quantitative Variablen (z.B. Grad der Nutzungsintensität auf einer 0-3 Skala) oder qualitative Variablen (z.B. Bodentyp) sein.

Es wird offensichtlich, dass die Trennung nicht eindeutig ist, da beispielsweise die Arten/Artenkompositionen zur Beschreibung von Probeflächen und umgekehrt die Probeflächen und deren abiotische Ausgestaltung zur Charakterisierung der Arten herangezogen werden. Die Rolle einer Gruppe von Daten kann somit wechseln zwischen Erklärung und Antwort.

Je nach Analyseverfahren sind die Daten auf ein bestimmtes Datenniveau anzupassen. Die Logistische Regression (vgl. Abschnitt 3.7.1) erfordert binärkodierte Daten für die Objekte. Zur Berechnung von Gradientenanalysen mit der Software CANOCO® müssen kategoriale Variablen, die meist eine Klassifizierung der Beobachtungsobjekte oder Erhebungseinheiten zum Ausdruck bringen, rekodiert werden. Die resultierenden Variablen werden als **Indikatoren oder Dummy-Variablen** bezeichnet. Dabei erhält jede Klasse eine separate Variable zugeordnet, welche die Ausprägung des Faktors beschreibt. Gehört eine Erhebungseinheit dieser Klasse an, so wird der Indikatorvariablen eine eins zugeordnet, während die verbleibenden Indikatorvariablen für diese Erhebungseinheit eine null erhalten.

Durch diese Rekodierung eröffnet sich auch die Möglichkeit einer **Fuzzy-Kodierung** (Studien auf der Basis von Fuzzy-Set-Theorien haben Castella & Speight (1996) sowie Roberts (1986) durchgeführt). Haben beispielsweise zwei Klassen in gleicher Weise zu der momentanen Ausprägung der Erhebungseinheit beigetragen, so ist es wünschenswert, der jeweiligen Indikatorvariablen einen Wert größer null und kleiner eins zuzuweisen. Im Falle von drei Klassen erhalten die Indikatorvariablen z.B. die Werte 0.5, 0.5 und 0. Wichtig hierbei ist, dass die Werte sich zu eins summieren (Leps & Smilauer 1999, unveröffentlicht). Dieser Umstand resultiert aus der Matrixalgebra. Wenn eine Variable die Linearkombination anderer ist, entsteht eine singuläre Matrix, was gleichbedeutend ist mit der Matrixoperation: Division durch null. Bei jeder kategorialen Variablen können daher nur k-1 Klassen berücksichtigt werden, da sich sonst die Werte aller Objekte zu einer Konstanten von eins summieren. Trotz dem Entfernen einer der Kategorien aus der Analyse wird der Informationsgehalt nicht reduziert, da die Summe der Werte aller anderen Klassen den Wert der fehlenden Kategorie eindeutig vorhersagt (Gesamtsumme immer gleich eins).

CANOCO® entdeckt solche überflüssigen Variablen und schließt diese automatisch aus der Analyse aus. Probleme können allerdings in folgenden Situationen auftreten: Werden alle Bodenkationen und zusätzlich der Gesamtkationengehalt bestimmt, dann ist letztere Variable eine Linearkombination. Gleiches findet sich bei Variablen, die sich zu 100 Prozent (unity) summieren, wie Bodenarten. In beiden Fällen wird die überflüssige Variable nicht automatisch entdeckt, wenn die Linearkombination nicht exakt ist. Für das Bodenartenbeispiel heißt dies, wenn durch Ungenauigkeiten im Labor oder Rundungseinflüsse sich die drei Hauptbodenarten nur zu 99,99 Prozent summieren, dann können Ergebnisverzerrungen auftreten (Palmer, mündliche Mitteilung).

Häufig ergibt sich die Situation, dass einige Umweltparameter Erklärungsanteile an dem biotischen Muster aufweisen, welche zwar berücksichtigt werden sollen, aber nicht Gegenstand des Interesses sind. Um die Effekte anderer Variablen in den Vordergrund zu stellen, werden Variablen gezielt als **Kovariablen oder Kovariate** gesetzt. Ein typisches Beispiel für die Vorgehensweise sind das Ausblenden (partialling-out) eines Höhengradienten, wenn der Fokus der Studie auf Unterschieden in der Artenkomposition durch Managementaspekte liegt (vgl. Abschnitt 3.5.9).

Die **Datenaufbereitung** verändert den Informationsgehalt der Rohdaten. Sie umfasst alle Eingriffe an der Primärdatenmatrix (Økland 1990):

- Gewichtung, lineare und nicht-lineare Transformation
- Operationen an den Zeilen und Spalten einer Matrix: z.B. Standardisierung
- Operationen bei denen alle Matrixelemente eingeschlossen sind: z.B. Ordnung



Die Umweltdatenmatrix erfordert in der Regel andere Transformationen als die biotische Datenmatrix, weshalb sie getrennt betrachtet werden.

Es ist anzumerken, dass bei Ordinationsmethoden keine Transformationen an den Daten durchgeführt werden müssen, um statistische Annahmen wie die Normalverteilung zu erfüllen, da Signifikanzen nur anhand von Monte Carlo Tests (randomization tests) bestimmt werden.

**Abiotik:** Die Datenaufbereitung der abiotischen Variablen (Deskriptoren) ist bei quantitativen Daten unumgänglich, um die verschiedenen Maßeinheiten der Umweltparameter (z.B. ppm, mmol/l, mg/100g) vergleichbar zu machen (Jongman et al. 1987; Legendre & Legendre 1998; Sokal & Rohlf 1995; Zar 1996). Mehrere Methoden stehen hierfür zur Verfügung, die entweder nur die Größendifferenz (Zentrieren, centering, Translation) oder sowohl diese als auch die Variabilität (Standardisierung) in eine einheitliche Skala überführen. Die häufig angewandte Methode der Standardisierung führt zu sogenannten "z-Werten". Aus der Transformation folgt die Dimensionslosigkeit der Daten. Die Position der Objekte entlang der so standardisierten Variablen  $z_i$  wird dann konsequenterweise in Standardabweichungen bezogen auf den ökologischen Raum in Ordinationsdiagrammen gemessen. CANOCO® standardisiert alle Daten im Erklärungsvariablen-File (\*.env) automatisch (ter Braak 1998). Sollen einzelne Variablen jedoch in eine andere Skala transformiert werden, um ihnen einen ökologisch adäquateren Informationsgehalt zu geben, muss dies vorab geschehen. Da die Spanne an möglichen Transformationen viel weiter und komplexer ist als für Spezies-Daten, stellt CANOCO® keine Werkzeuge hierfür zur Verfügung. Dies soll aber weder implizieren noch empfehlen, die Umweltparameter nicht zu transformieren. Folgende Aussage bereichert die Diskussion: "As an aside note, the obvious fact is that if you do not transform, this is also a transformation, as your measurement scale is by no means any elite choice among the possible other scales (Smlauer 2000, Ordnews Diskussionsforum/Internet)." Eine Skaleninvarianz von Ordinationsmethoden liegt zwar nicht im Sinne von Ergebnisunterschieden zwischen einer Meter- oder Kilometerskala vor (lineare Transformationen), dennoch können Erklärungsvariablen eine differierende Aussagekraft besitzen, wenn beispielsweise Nährstoffkonzentrationsunterschiede durch eine nicht-lineare Transformation im kritischen Bereich entsprechend betont werden. Hohe bis sehr hohe Konzentrationen, die beide schon im toxischen Bereich liegen, erbringen keinen Informationsgewinn. Wann und wie letztlich zu transformieren ist, bedingt sich aus den Daten, vorliegender Erfahrung und den Ansichten des Bearbeiters (vgl. Abschnitte 5.3.2 und 5.3.3).

**Biotik:** Nur in den linearen Verfahren stellt CANOCO® einige Möglichkeiten für die Transformation der Matrix Y zur Auswahl: Zentrieren, Standardisieren und Standardisieren plus Gewichten anhand der Fehlervarianz. Dies begründet sich auch daraus, dass speziell bei der Hauptkomponentenanalyse diese Matrix entweder biotische oder abiotische Variablen zur Ordination der Objekte enthalten kann.

Wählt man keine Transformation für die biotischen Daten, so bleiben die Eingangsdaten zwar ohne Gewichtsveränderung, doch resultiert dies in einem ökologisch nicht sinnvollen Modell mit dem Ursprung in einer hypothetischen Probefläche, in der die Artmächtigkeiten/Häufigkeiten aller Arten gleich null sind. Zusätzlich stellen sich meist modelltechnische Probleme ein (vgl. Abschnitt 4.1).

Das Zentrieren der biotischen Daten steht meist nicht im Einklang mit dem Zweck multivariater Analysen, da Arten mit einer hohen Varianz ein noch höheres Gewicht erhalten. Hinzu kommt, dass Probeflächen die tendenziell stark von Durchschnittsbild abweichen (deviant sample plots) deutlich betont werden (vgl. Økland 1990). Im Ordinationsdiagramm dominieren oft einige Arten, während die Masse der Arten dicht gedrängt um den Ursprung liegen. Das Zentrieren korrespondiert zu einer PCA/RDA basierend auf der Kovarianzmatrix (Jongman et al. 1987; ter Braak 1998).

Die Standardisierung der Arten bedeutet eine Gleichgewichtung aller Arten, d.h. seltene Arten oder Ausreißer werden von ihrer Gewichtung her, relativ zu Arten höherer Dominanz, aufgewertet. Gleiches gilt für die Standardisierung der Probeflächen: Sie erhalten alle das gleiche Gewicht in der Analyse. Die Bedeutung artenarmer Probeflächen und von Flächen mit geringer Häufigkeit der Arten, wird relativ zu anderen Probeflächen verstärkt. Alle diese Effekte sind aus ökologischer Sicht ebenfalls nicht wirklich wünschenswert (Økland 1990). Dennoch wird die Standardisierung häufig angewendet. Sie korrespondiert zu einer PCA/RDA basierend auf der Korrelationsmatrix (Jongman et al. 1987; ter Braak 1998). Eine weitere Variante, das "Standardisieren und Gewichten anhand der inversen Fehlervarianz", reiht sich nahtlos in die Problematik einer geeigneten Transformation ein. Die Analyse wird in zwei Schritten durchgeführt: Zuerst wird die Eingangsdatenmatrix standardisiert und dann eine neue Gewichtung für die Arten vergeben, die proportional zu deren inversen Fehlervarianz in der Modellanpassung ist, d.h., gut durch das Modell repräsentierte Arten erhalten höhere Gewichte. Gerade bei Arten mit wenigen Vorkommen kann es zu einer extremen Überbewertung kommen, wenn diese zufallsbedingt gut mit dem Modell harmonisieren (Økland 1990).

Alle angesprochenen Transformationen/Skalierungen in linearen Verfahren sind im ökologischen Kontext nur bedingt empfehlenswert (vgl. auch Abschnitt 4.1). Es bleibt dem Bearbeiter überlassen, welche Eigenschaften er als am unbedenklichsten für die zu bearbeitende Fragestellung ansieht.

Sowohl für lineare als auch unimodale Ordinationsverfahren bietet die Software CANOCO® zwei weitere Transformationen an: Die Logarithmus- und die Wurzeltransformation.

Speziell mit der Logarithmus-Transformation ist das Problem negativer Ergebniswerte und von Nullwerten verbunden, da der Logarithmus von eins gleich null wird (nicht definiert) und Werte zwischen null und eins negativ werden. Beinhalten die biotischen Daten solche Werte, muss bevor logarithmiert wird eine Konstante zu den Daten addiert werden um das Problem zu vermeiden (ter Braak 1998).

Die Frage, wann entsprechende Transformation angebracht erscheint, ist ein viel diskutiertes Thema (Legendre & Legendre 1998; ter Braak 1998 u.a.). Allgemein soll der Effekt einiger hoher Werte in den Daten biotischer Felderhebungen auf die Ordination gedämpft werden, da schiefe Verteilungen der Daten nahezu Standard sind (Jongman et al. 1987; ter Braak 1998). Ungünstig erscheint jedoch die Tatsache, dass die Ordinationsergebnisse instabil sind für unterschiedliche Einheiten: z.B. Biomasse in Gramm, Pfund oder Kilogramm (Palmer, mündl. Mitteilung). Bei floristischen Daten sind sehr detaillierte Skalen aufgrund der Schätzunsicherheit in der Deckung von Pflanzenbeständen für gewöhnlich wenig effektiv und es wird die quantifizierte Braun-Blanquet Skala angewendet. Diese Kombination aus Individuenzahl und Deckung resultiert in sieben Klassen, die bereits eine logarithmusähnliche Skalierung implizieren. Häufig werden diese Klassen in mittlere Prozentdeckungen übersetzt. Den Klassen *r* und *+*, die auf der Individuenzahl beruhen, ordnet man in der etablierten Vorgehensweise Werte von 0,1 und 0,5 zu, da dies die Analyseergebnisse kaum beeinflusst (Leps & Smilauer 1999, unveröffentlicht).

Die alternativ in CANOCO® integrierte Transformation für Spezies-Daten ist die Wurzeltransformation. Palmer (mündliche Mitteilung) wendet sie zur Dämpfung der Einflüsse dominanter Arten an, weil sie keine instabile Ergebnisse – bedingt durch unterschiedliche Einheiten – erzeugt. Leps & Smilauer (1999, unveröffentlicht) empfehlen die Transformation speziell für faunistische Häufigkeitsdaten. Sie wird auch in dieser Arbeit durchgehend angewendet.

Weiterhin bietet CANOCO® die Option, seltenen Arten ein geringeres Gewicht zu geben (downweighting of rare species, vgl. Abschnitt 3.5.2). Bei diesem Verfahren werden die Originalwerte  $y_{ij}$  der seltenen Arten nach Gleichung 3.3-1 durch neue Werte  $y'_{ij}$  ersetzt.

**Gleichung 3.3-1:** 
$$y'_{ij} = y_{ij} \cdot [f_i / (f_{i, \max} / 5)]$$

Eine Art gilt als selten, wenn die Zahl der Erscheinungen im Datensatz  $f_i$  geringer als  $f_{i, \max}$  dividiert durch fünf ist, wobei  $f_{i, \max}$  die maximale Häufigkeit einer beliebigen Art im Datensatz ist. Die Reduktion ihres Gewichtes erfolgt somit im Verhältnis zur Häufigkeit der betroffenen Art, vorausgesetzt, sie ist seltener als  $f_{i, \max}$ . Arten häufiger als  $f_{i, \max}$  bleiben unverändert in ihren Werten (Økland 1990; ter Braak 1998). Die Beziehungen von Arten, die weniger als fünfmal im Datensatz erscheinen, zu anderen Arten, sind wahrscheinlich nur unzureichend beschrieben. Sie agieren als Ausreißer im Datensatz und deshalb wird der Divisor fünf in der Gewichtungsprozedur angewendet (ter Braak 1998). Dieses Ad-hoc-Verfahren entbehrt nicht einer gewissen Willkür. Es führt in vielen Fällen zu einer zu strengen Gewichtung der selteneren Arten in der Analyse, d.h. zu viele Arten werden manipuliert, wie die Studien von Eilertsen et al. (1990b) belegen. Außerdem werden dem Verfahren in der Literatur unvorhersehbare Auswirkungen auf die Daten vorgeworfen (vgl. Eilertsen et al. 1990a; Eyre et al. 1986; Lahiti & Väisänen 1987; Oksanen 1983).

In dieser Arbeit wird eine manuelle Vorgehensweise bevorzugt: Alle Arten mit einer Häufigkeit  $n \leq 2$  werden per Hand im Datensatz selektiert und passiv gesetzt, da speziell in unimodalen Verfahren allein schon das Modell wenigstens drei Erscheinungen als Minimum erfordert (s.a. Legendre & Legendre 1998, S. 415). Diese manuelle Gewichtung Passivsetzung oder Zuweisung eines konkreten Gewichts pro einzelner Art und Probestrich stellt die letzte durch CANOCO® angebotene Transformationsmöglichkeit dar.

### 3.4 Datenhaltung: Datenbankanforderungen

**DIE GEWICHTIGE BEDEUTUNG EINER GUT ABGESTIMMTEN DATENBANK FÜR EIN PROJEKT UND DIE ABLEITUNG VON SEKUNDÄRPARAMETERN FÜR STATISTISCHE ANALYSEN SOLLTE NICHT UNTERSCHÄTZT WERDEN.**

Die Datenbank und deren Strukturierung ist eine essenzielle Ressource für eine reibungslose Datenübergabe, -haltung und -bereitstellung. Die Datenbank eines Projektes muss somit auf einem Datenmodell basieren, das die problemlose Zusammenführung der Teilprojekt-Erhebungsdatenbanken in eine Gesamtdatenbankstruktur ermöglicht. Die thematisch verschiedenen Dateninhalte sollen klar getrennt werden und die Daten bei teils sehr unterschiedlichen Erhebungsintervallen und -zeiträumen noch überschaubar und handhabbar bleiben (Rink 1997b).

Weiterhin muss die zentrale Datenbank für alle Projektteilnehmer zugänglich sein, so dass sich ein Onlinezugriff über das Internet anbietet, wie es in RIVA realisiert wurde.

Die allgemeine Struktur der Datenbank baut sich aus den Stammdaten (Referenzdaten: z.B. Probestrichennummer), den Sachdaten (teilprojektbezogene Daten: z.B. hydrologische Parameter) und den Metadaten (eventuell auch Kontext-, Memofelder; z.B. Bearbeiter) auf. Herausgehoben aus der Gesamtstruktur ist eine so genannte Anwender-Schnittstelle (user-interface), in der eine hierarchische Zuordnungsliste (z.B. Roll-up-System) den Nutzer ohne eine direkte Dateneingabe zur gewünschten Stelle der Stammdatenbank oder Sachdatenbank führt. Dieses System vermeidet allgemein Eingabefehler und kann so programmiert werden, dass ein Nutzer der zentralen Datenbank zwischen verschiedenen Optionen (z.B. Stamm- oder Sachdaten anschauen, ändern, grafisch darstellen bzw. flächen-, datumsbezogenes Recherchieren) wählen kann. Ebenfalls ist eine hohe Datensicherheit gegeben, da die Nutzung der Datenbank nur innerhalb einer temporären Datenkopie stattfindet.

Das eigentliche Datenmodell (Felddefinitionen etc.) ist im zugriffsgeschützten "Overhead" abgelegt (Abbildung 3.4-1, S. 54). Es besitzt eine relationale Architektur (Spalte gleich Parameter; Zeile gleich Fall/Datensatz) und die Stammdatenbank enthält die Primärschlüssel

(primary key; 1:1 Verknüpfung), über die auf Sachdaten zugegriffen wird. Innerhalb der Teildatenbanken kann eine Indizierung (Indexfelder, innerhalb deren Duplikate erlaubt sind; z.B. Datumsfelder) erfolgen. Die verschiedenen Verknüpfungen und Folgeverknüpfungen bieten umfangreiche Recherchemöglichkeiten und automatischen Zugriff auf alle relevanten Eintragungen bei Selektions- oder Löschvorgängen. Ein anwendungsoptimiertes Datenmodell macht eine reine Datenbank zum Informationssystem (Rink 1998).

Wichtig in Verbundprojekten ist die Kompatibilität des Datenbanksystems zu anderer Software (z.B. für die Auswertung der Daten), um einen reibungslosen und für alle Anwender komfortablen Datenaustausch zu gewährleisten.

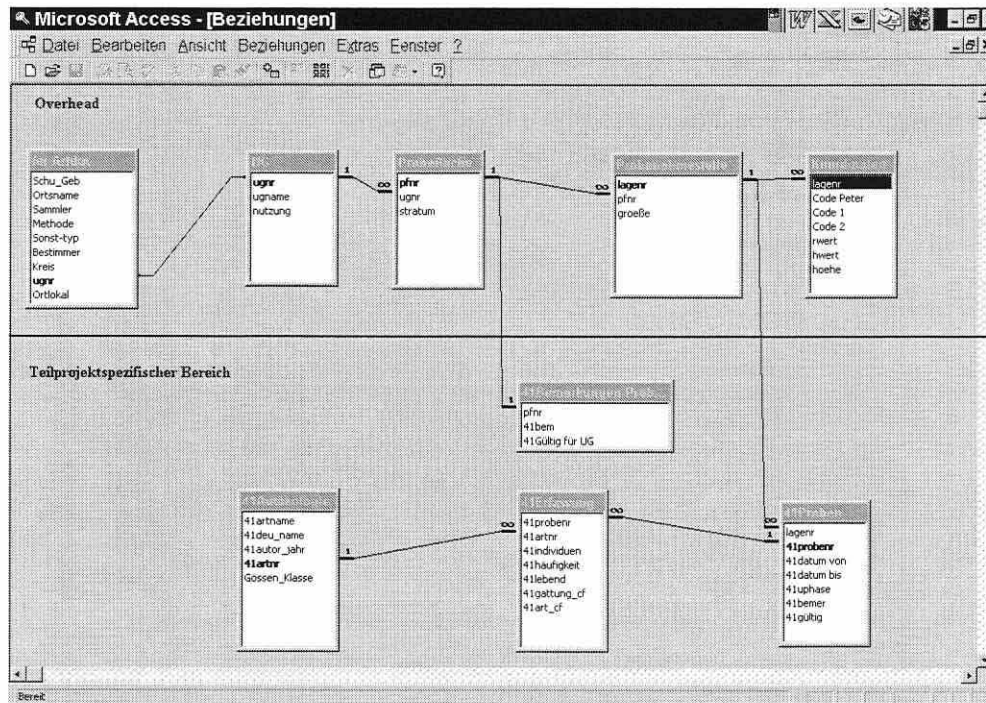


Abbildung 3.4-1: Datenbank mit Unterscheidung zwischen 'Overhead' und teilprojektspezifischem Bereich.

### 3.5 Gradientenanalytische Verfahren: Grundlagen, Theorie, Eigenschaften und Handhabung

Der Abschnitt 3.5 soll den Leser mit der Theorie und Leistungsfähigkeit von Ordinationsverfahren vertraut machen. Neben der allgemeinen Einführung in die gradientenanalytischen Verfahren werden zwei bedeutsame in sie integrierte bzw. mit ihnen eng verbundene Grundlagen Assoziationsmaße und die Variablenreduktion besprochen.

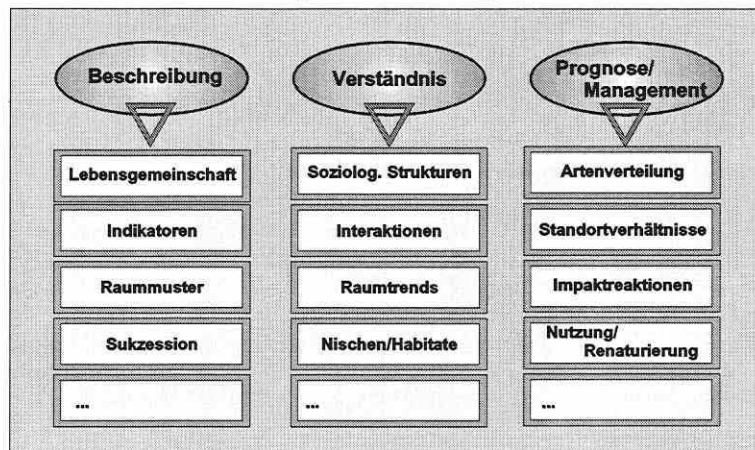
Anschließend werden die Ordinationsverfahren selbst jeweils erst beschreibend und dann soweit notwendig mathematisch dargestellt. Die Abschnitte zu den einzelnen Ordinationen schließen mit der Verknüpfung zwischen der dargestellten Theorie und den Einstellungen zur Berechnung verschiedener Ordinationstypen in der Statistiksoftware CANOCO®, welche zentral in dieser Arbeit verwendet wird.



### 3.5.1 Multivariate Statistik und gradientenanalytische Verfahren: Einführung

**"NUMERICAL APPROACHES CAN NEVER DISPENSE RESEARCHERS FROM ECOLOGICAL REFLECTION ON OBSERVATIONS. DATA ANALYSIS MUST BE SEEN AS AN OBJECTIVE AND NON-EXCLUSIVE APPROACH TO CARRY OUT IN-DEPTH ANALYSIS OF THE DATA (LEGENDRE & LEGENDRE 1998, S. 482)."**

Mit Verfahren der multivariaten Statistik lassen sich verschiedenste Fragestellungen in der Ökologie bearbeiten, da der explorative Charakter der Verfahren eine weit gefächerte Anwendungsspanne erlaubt. Die in Abbildung 3.5-1 exemplarisch dargestellten Blöcke Beschreibung, Verständnis und Prognose/Management erheben keinen Anspruch auf Vollständigkeit und lassen sich nicht klar voneinander trennen, da allein die angeführten Schlagworte schon eine sehr deutliche Verflechtung miteinander zeigen.



**Abbildung 3.5-1: Ökologische Fragestellungen und Einsatzbereiche multivariater Statistik.**

Die meisten Arbeiten in der Biozönoseforschung unter Anwendung multivariater gradientenanalytischer Verfahren sind in den letzten drei Jahrzehnten entstanden (Gauch 1982a). Eine Bibliographie von Birks et al. (1994) enthält bereits 379 Titel ökologischer Studien mithilfe von Ordinationsverfahren. Die Vorteile dieser Verfahren liegen unter anderem darin, dass sie

- (1) den Anwender zu einer präzisen Formulierung der Fragestellung, einer vorausschauenden Datensammlung basierend auf einem Stichprobenplan und der Vorauswahl einer Datenanalysestrategie drängen sollten (Orlóci 1978);
- (2) eine Einheitlichkeit und Konsistenz in der Vorgehensweise bei der Verwendung der Handhabungsregeln und Entscheidungen durch den ganzen Prozess der Datensammlung, der Analyse und der Bewertung/Schlussfolgerung fordern (Orlóci 1978);
- (3) dadurch nach Kooperation und einem engen Austausch zwischen den beteiligten Wissenschaftlern verlangen (Orlóci 1978);
- (4) die Entdeckung, Würdigung und das Verstehen geringfügiger Unterschiede erlauben (Gauch 1982a);
- (5) relativ objektive, organisiert zusammengefasste Ergebnisse bereitstellen, welche sowohl das Verständnis von Zusammenhängen, die Erinnerung als auch eine effektive Kommunikation darüber erleichtern. Rohdaten sind zu umfangreich und unorganisiert, um diesen Zwecken zu dienen (Gauch 1982a);
- (6) logistische Vorteile bei ständig anwachsender Datenanzahl und Komplexität der Studien bieten (Gauch 1982a).



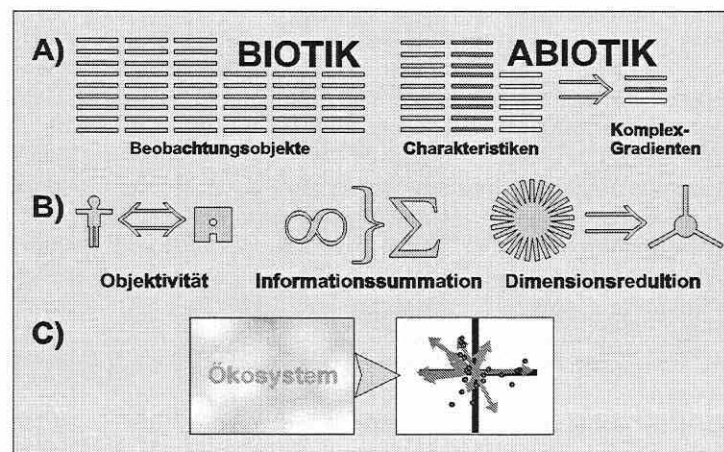
**Was bedeutet multivariate Statistik?** Sie ist der Zweig der Mathematik, der sich mit der Untersuchung zahlreicher Zufallsvariablen gleichzeitig auseinander setzt: "The need for multivariate analysis arises whenever more than one characteristic is measured on a number of individuals, and relationships among the characteristics make it necessary for them to be studied simultaneously (Gauch 1982a, S. 1)". Die Datenauswertung in (syn-)ökologischen Studien weist diese Besonderheit, die Daten als Ganzes zu behandeln, auf, weil jede Probe- fläche durch beispielsweise Artmächtigkeiten von Pflanzenarten und diese wiederum durch eine Vielzahl von Umweltfaktoren und so fort beschrieben wird. Das heißt, eine Vielzahl von Antwortvariablen muss mit mehr als einer Charakteristik gleichzeitig in Beziehung gesetzt werden, um daraus die innere Ordnung der Beobachtungsobjekte in der Natur zum Vor- schein zu bringen und zu erklären (Abbildung 3.5-2, Teilbild A). Bedenkt man allein den Umfang einer Matrix mit 100 Proben und 300 Arten – sie besitzt 30 000 Einträge – erschei- nen Hilfsmittel zur Analyse solcher Datenmengen unverzichtbar. Daher ist der **Begriff Ordi- nation**, der Ordnung bedeutet und von dem lateinischen Wort "ordinatio" kommt gut gewählt (Legendre & Legendre 1998).

**Ziel** ist es, das undeutliche Bild der in jedem Ökosystem vorhandenen natürlichen Ordnung so zu übersetzen (Abbildung 3.5-2, Teilbild C), dass die durch Umweltparameter erklärbaren Strukturen deutlich werden und ihre Relevanz für das biotische Raummuster abgeschätzt werden kann. Eine möglichst objektive Repräsentation von Raummustern und Wirkzusam- menhängen im reduzierten meist zwei- oder dreidimensionalen Ordinationsraum (Orlóci 1978) unterstützt eine der Schlüsselaufgaben in der Ökologie – die Mustererkennung (Weiner 1995).

Die entscheidende **Voraussetzung** für den Erfolg von Ordinationsverfahren ist die grund- sätzlich geringe Anzahl von Komplex-Gradienten mit bedeutendem Einfluss in terrestrischen Ökosystemen, wodurch die angestrebte Dimensionsreduktion auf wenige beschreibende Kernfaktoren erst möglich wird (Abbildung 3.5-2, Teilbild A).

Die Erklärungsvariablen sollen regelhafte Strukturen beschreiben, um das Wissen über die betrachteten Organismen zu erweitern und etwas über den strukturellen Aufbau eines Öko- systems zu lernen. Gewonnene Erkenntnisse können dann hinleiten zu Modellen und Mana- gementkonzepten für die Erhaltung und Förderung von Lebensräumen und deren Bewoh- nern.

Die grundlegenden **Funktionalitäten** der Verfahren lassen sich allgemein kurz zusammen- fassen (Abbildung 3.5-2, Teilbild B). Der Begriff Objektivität steht für die gleich gewichtete



**Abbildung 3.5-2:** Schematisierte Darstellung der Voraussetzungen (A), der grundlegenden Funktionalitäten (B) und dem Ziel von Ordinationsverfahren (C). Erläuterungen im Text.

Datenbehandlung bei der Analyse. Der Ausschluss von bereits a priori bestehenden gerichteten Aufmerksamkeiten des Bearbeiters lässt die Entdeckung von latent verborgenen Datenstrukturen wahrscheinlicher werden. Weitere Eigenschaften sind die Summation redundanter Dateninformation und die Dimensionsreduktion auf relevante Komplex-Gradienten. Das **Resultat** gewährleistet die Handhabbarkeit der vielschichtigen Information in den Daten für das menschliche Vorstellungsvermögen (Abbildung 3.5-2, Teilbild C).

Die Schwelle, bei der ein Anwender eine Effizienz in der Anwendung multivariater Methoden sieht, ist willkürlich und hängt von den Eigenschaften der Datensätze ab. Die Zahl simultan zu betrachtender Variablen liegt zwischen etwa 10-15. Mit weniger, speziell fünf oder darunter, erscheinen diese Techniken nicht mehr als geeignet und Varianzanalysen oder Streudiagramme sind vorzuziehen (Gauch 1982a). Dies gilt jedoch nicht für stark asymmetrische Matrizen mit beispielsweise  $5 \cdot 1000$  Einträgen.

Ein **Manko** komplexer statistischer Verfahren ist, dass "(...) mathematics may obscure, under an often esoteric language, the natural phenomena that scientists are trying to elucidate (Legendre & Legendre 1998, Vorwort)." Das eigentliche Problem liegt aber nicht in den Verfahren begründet, sondern bei den Anwendern. Hochentwickelte Computerprogramme ermöglichen einfachen Zugang zu multivariaten und anderen Analyseverfahren. Sie vermitteln aber nicht den notwendigen statistischen Hintergrund und reißen die Datenanalyse aus dem Kontext eines gut strukturierten Versuchs. Datenerhebungsstrategie, Eigenschaften der Daten, Anforderungen der Analyseverfahren und verschiedene Limitierungen bleiben unberücksichtigt. Unerfreulicherweise gibt es viele Beispiele in der ökologischen Literatur, in denen mangels Kenntnis mathematische Verfahren unbedarft und inkorrekt angewendet werden, was dazu führt, dass das vorangegangene Zitat untermauert wird. Dennoch ist Optimismus geboten, da es offensichtlich mehr Beispiele gibt, die belegen, dass, wenn Ökologen die theoretischen Hintergründe numerischer Verfahren verinnerlichen und diese folglich korrekt anwenden, ein tieferes Verständnis natürlicher Phänomene und Zusammenhänge aus den Ergebnissen zu gewinnen ist (s.a. Legendre & Legendre 1998).

Gerade die **Entwicklung von Modellen**, die in der Lage sind, Veränderungen in Variablen – ausgelöst durch Änderungen in anderen Variablen – vorherzusagen, stehen in engem Zusammenhang mit Methoden der numerischen Ökologie. Die Hypothesen, die den kausalen Modellen unterliegen, werden nicht selten unter Anwendung numerischer Verfahren aufgestellt. Die multivariate Statistik stellt eine Vielzahl von Analyseverfahren bereit, die sich gegenseitig ergänzen können. Jedes Verfahren erlaubt es dann, einen anderen Aspekt an Information in den unterliegenden Daten zu entdecken.

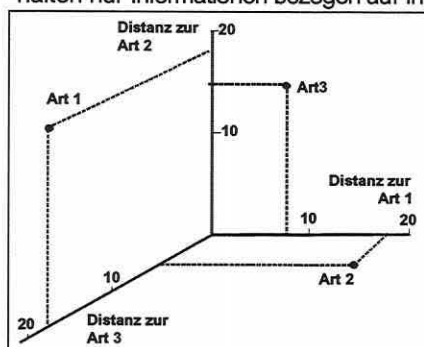
Die ökologische Datenmatrix kann grundsätzlich ausgehend von zwei **Analysetypen** betrachtet werden: den Beziehungen zwischen den Beobachtungsobjekten (Q-Mode-Analyse, beschrieben durch beobachtete Deskriptoren/Erklärungsvariablen) oder den Deskriptoren (R-Mode-Analyse, beschrieben durch beobachtete Objekte; vgl. Abschnitte 3.3 und 3.5.2). Neben weiteren Betrachtungsweisen erscheint die T-Mode-Analyse noch von Relevanz. In diesem Falle werden Variablen über mehrere Objekte bezüglich verschiedenen Zuständen entlang einer Zeitachse verglichen. Hierzu seien die in Abschnitt 3.1.2 angesprochenen BACI-Designs in Erinnerung gerufen. Jeder Analyseweg ist mit spezifischen Assoziationsmaßen verknüpft (Legendre & Legendre 1998). Die Suche nach Mustern in ökologischen Daten basiert immer auf Assoziationsmatrizen (vgl. Abschnitt 3.5.2). Auch in sogenannten assoziationsfreien Methoden, wie Hauptkomponenten-, Korrespondenz- oder Clusteranalyse ist immer ein Assoziationsmaß verborgen (Jongman et al. 1987; Legendre & Legendre 1998).

Das Verständnis der Arbeitsweise von Ordinationsverfahren erfordert, die Einführung spezieller **Raumkonzepte**, d.h. von **fünf konzeptionellen Räumen**. Da die geometrischen Konzepte meist leichter nachvollziehbar sind als deren Äquivalente in der Matrixalgebra wird ihnen besondere Aufmerksamkeit zu Teil.

Der **Artenraum** (species space) wird gebildet durch Achsen, die von einzelnen Arten definiert sind. Die Probeflächen werden in diesem Raum als Punktsignaturen abgebildet. Umgekehrt bildet sich der **Probeflächenraum** aus den Probeflächen als Achsensystem und den Arten als Punktsignaturen innerhalb dessen aus. Das Wesen der Achsen ist also entscheidend für die Bezeichnung des Raumes, der letztlich die identische Information der Ausgangsdatenmatrix repräsentiert (Gauch 1982a).

Zwei weitere Räume leiten sich aus einer jeweiligen Sekundärmatrix für die Arten oder Probeflächen ab: der **Unähnlichkeitsraum für die Arten bzw. Probeflächen**. Die Sekundärmatrix enthält dann ein Unähnlichkeitsmaß (Distanzmaß; vgl. Abschnitt 3.5.2) für jedes Paar Arten oder Probeflächen der Originaldatenmatrix. Bei  $n$  Arten gibt es  $[n(n-1)] / 2$  Vergleichsmöglichkeiten, d.h. gegenüber einer Originaldatenmatrix mit 100 Arten auf 36 Probeflächen mit 3600 Elementen (Zeilen · Spalten) besitzt die Assoziationsmatrix zusammen  $\{[i(i-1)] / 2\} + \{[j(j-1)] / 2\} = 5580$  Elemente ( $i$ : Arten und  $j$ : Probeflächen).

Im Unähnlichkeitsraum für die Arten (Abbildung 3.5-3) definiert folglich die Unähnlichkeit der Arten die Achsen und die Arten die eingeordneten Punktsignaturen. Im Unähnlichkeitsraum für die Probeflächen verhält es sich invers (Gauch 1982a). Die Unähnlichkeitsräume beinhalten nur Informationen bezogen auf ihren Fokus, d.h., der Unähnlichkeitsraum für die Arten trifft



**Abbildung 3.5-3: Unähnlichkeitsraum für die Arten.**

nur eine Aussage hinsichtlich der Beziehungen der Arten zueinander, aber keine Aussage zu den Probeflächen. Die geometrische Darstellung ist sozusagen informationsverarmt gegenüber ihrem Pendant: dem Artenraum bzw. den Originaldaten.

Die bisher beschriebenen vier Räume bilden die Ausgangsbasis für die meisten Ordinationsverfahren aufgrund dreier Eigenschaften (Gauch 1982a): (1) Beziehungen zwischen Arten oder Probeflächen lassen sich aus der Distanz der Punktsignaturen zueinander erkennen. Arten mit ähnlichem Verhalten werden nahe beieinander abgebildet und umgekehrt.

(2) Einige Artenkombinationen sind häufiger als andere, so dass die Verteilung der Probeflächensignaturen nicht zufällig, sondern meist gerichtet oder punktuell verdichtet erscheint (s.a. Orlóci 1978). Ebenso lässt sich bei der Darstellung der Punktsignaturen der Arten durch die Muster in ihrer natürlichen Verbreitung eine Struktur nachvollziehen. Konsequenterweise okkupieren die jeweiligen Punktsignaturen einen Subraum innerhalb des höherdimensionalen Gesamttraums. Diese Eigenschaft macht die Dimensionsreduktion durch Datensummutation in Ordinationsverfahren erst möglich. (3) Die grafische Darstellung der Ergebnisse des Ordinationsalgorithmus in solchen Räumen hat einen heuristischen Wert, da sie den meisten Personen einen einfacheren Zugang zu den Datenstrukturen verschafft, als das die korrespondierende komplexere Matrix-Algebra vermag (Gauch 1982a).

Arten- und Probeflächenraum besitzen die volle Dimensionalität der Originaldatenmatrix, d.h., sie umfassen leicht Hunderte von Dimensionen. Da diese sich weder übersichtlich visualisieren noch überschauen und interpretieren lassen, ist der Wunsch nach einer Projektion der Zusammenhänge in wenige Dimensionen bei maximaler Realitätsnähe nahe liegend. Der Erfolg einer solchen Projektion setzt voraus, dass die Daten eine innere Struktur aufweisen. Im Falle reiner Zufälligkeit gibt es nahezu unendliche Projektionsmöglichkeiten, die alle den gleichen Wert bzw. eigentlich keinen Wert besitzen. Gibt es aber einige Dimensionen, in denen sich die Information in den Daten konzentrieren lässt und zu denen die restlichen Dimensionen korreliert sind, dann ist die entsprechende Projektion entlang der aussagekräftigen Achsen effektiv. Sie repräsentieren folglich einen deutlich höheren Anteil der originalen Datenstruktur und zeigen ein recht akkurates Abbild der Strukturierung des Originalraums

(Gauch 1982a). Die in Abbildung 3.5-3 dargestellte Projektion nutzt Punktsignaturen zur Repräsentation der Arten oder Probenflächen (z.B. CA, CCA; Abschnitte 3.5.5 und 3.5.8). Eine Alternative bilden Vektordarstellungen (z.B. PCA, RDA; Abschnitte 3.5.4 und 3.5.7).

Der fünfte konzeptionelle Raum wird als **ökologischer Raum** bezeichnet, weil er Umweltgradienten als Achsen nutzt. Daraus resultiert eine Redundanz in den Speziesdaten, da ein Gradient gleichzeitig viele Arten in ihrer Verteilung beeinflusst. Anders ausgedrückt lässt sich aus einem hohen oder niedrigen Wert einer Art bereits Information zu dem erwarteten Wert einer anderen Art ableiten. Der ökologische Raum ist daher von Beginn an geringdimensional, wie die meist angestrebte Ausgabegrafik bzw. Lösung einer Ordination. Dieser Zweck von Ordinationsverfahren wird in folgendem Zitat plakativ in Worte gefasst: "Field data must be high dimensional because of the large number of species and samples, whereas the final results must be low dimensional because of the human limitations. Ordination is a means for bridging this gap, for producing effective, low-dimensional summaries from field data by relatively convenient and objective means (Gauch 1982a, S. 118)." Statistisch betrachtet bilden jeweils die beiden ersten extrahierten Achsen eine Ebene durch die Punktwolke der Daten, welche deren Variabilität in der Projektionsebene maximiert und die Residuen minimiert (Doledec & Chessel 1997).

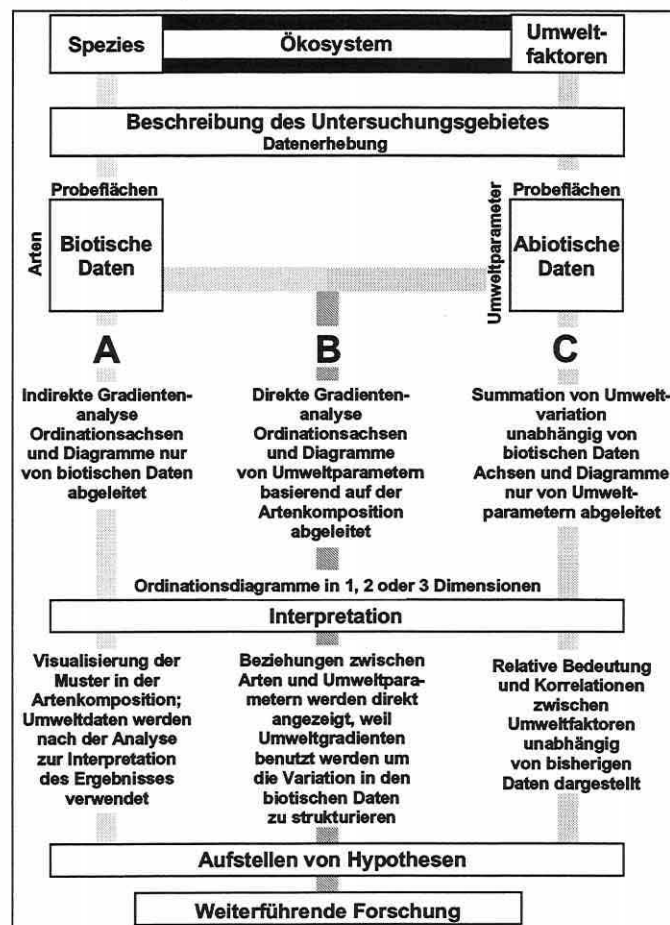


Abbildung 3.5-4: Direkte und indirekte Ordination. Weg A und C ist indirekt, d.h., nur eine Datenmatrix wird geordnet. In B wird der direkte Weg beschriftet und eine bereits geordnete Datenmatrix aus Weg A anhand einer zweiten erklärt (Kent & Coker 1992, S. 163, verändert).

Die Unterscheidung von Gradientenanalysen in **direkte und indirekte Ordination** bildet die grundlegende Dichotomie zwischen den Verfahren (vgl. Abbildung 3.5-4, S. 59). In den indirekten Analyseverfahren werden die Probeflächen bzw. die biotischen Daten unabhängig von abiotischen Erklärungsvariablen nur basierend auf der Artenkomposition oder dem Artvorkommen geordnet (CA, PCA; Abbildung 3.5-4, Weg A). Die unterliegenden dafür verantwortlichen abiotischen Erklärungsvariablen bzw. die von ihnen beschriebenen Gradienten können nur anhand der Trends in der Verteilung geschlussfolgert werden, d.h., der Bearbeiter übernimmt eine Art "Künstlerrolle", basierend auf seinem persönlichen Fachwissen. Es gilt: "ordination axes can be considered as latent variables, or hypothetical environmental variables, constructed in such a way as to optimize the fit of the species data to a particular (linear or unimodal) statistical model of how species abundance varies along gradients (ter Braak & Prentice 1988, S. 272)". Anstatt der Artenkomposition können aber auch Umweltvariablen die Ordnung von Probeflächen definieren, d.h., die Summation der Variabilität in der Umweltinformation erfolgt unabhängig von der Biotik (PCA; Abbildung 3.5-4, Weg C).

Direkte Techniken eröffnen die Chance, Probeflächen und/oder Arten direkt auf einen Satz gemessener abiotischer Erklärungsvariablen zu beziehen, daraus Zusammenhänge zu erkennen und die Analysen generell objektiver werden zu lassen (Abbildung 3.5-4, Weg B). Bevor es die Möglichkeit gab, direkte Gradientenanalysen zu rechnen, war man darauf angewiesen, separate Regressionen für jede einzelne Art durchzuführen, um die Artenkomposition eines Raumes auf Umweltvariablen beziehen zu können (Gauch 1982a; Jongman et al. 1987; Kent & Coker 1992 u.a.). Die Achsen indirekter Verfahren korrespondieren zur Richtung maximaler Variabilität einer Dimension des Datensatzes. In direkten Verfahren korrespondieren sie zu der Richtung größter Variabilität einer Dimension im Datensatz, die durch Umweltvariablen in der Analyse erklärt werden kann.

Direkte Ordinationsverfahren kombinieren Konzepte und Funktionalität von indirekter Ordination und Regression oder genauer gesagt, Redundanzanalyse (RDA) und kanonische Korrespondenzanalyse (CCA) sind eng verbunden mit der multiplen Regression.

Das Verfahren der multiplen Regression modelliert das Verhalten einer Antwortvariable  $y$  (vgl. weiterhin das Notationsverzeichnis) unter Benutzung eines Satzes von Erklärungsvariablen in Matrix  $X$  (Abbildung 3.5-5, Teilbild B). Dabei ist die eindimensionale Ordination der Objekte bezogen auf die Originalantwortvariable  $y$  eine andere als die Ordination der Objekte anhand der angepassten Modellvariablen  $\hat{y}$ . Das Modell erzeugt also eine Korrespondenz zwischen den Ordinationen an  $y$  und  $\hat{y}$ , da die Ordination an  $\hat{y}$  der Einschränkung unterliegt, eine optimale lineare Relation zu den Variablen in Matrix  $X$  aufzuweisen. Die Einschränkung ist optimal im Sinne einer Lösung, die zu den kleinsten Abweichungsquadraten führt und somit das (multiple) Bestimmtheitsmaß  $R^2$  für die lineare multiple Regression maximiert. Diese Eigenschaft teilt sich das Verfahren mit den direkten Ordinationsmethoden.

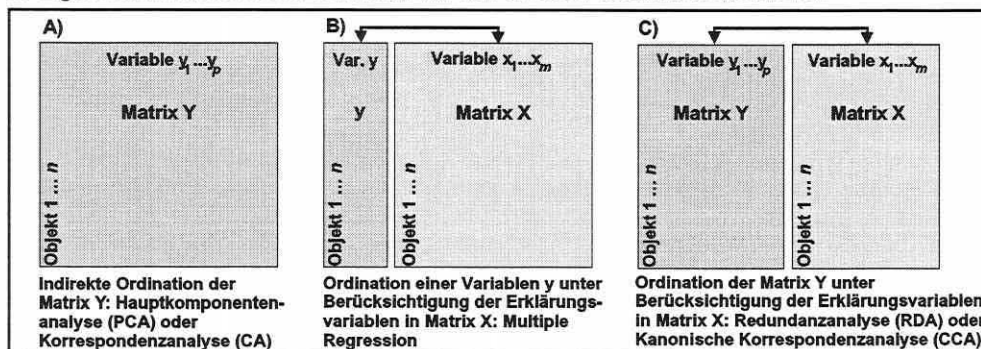


Abbildung 3.5-5: Mathematische Beziehungen zwischen indirekter Ordination (A), multipler Regression (B) und direkter Ordination (C).



Matrizenbezogen bedeutet dies, dass Matrix X in direkten Verfahren unmittelbar in die Berechnungen eingreift und die Ordination von Matrix Y "zwingt", einen maximalen Bezug zu den Linearkombinationen der Erklärungsvariablen in Matrix X aufzuweisen (Abbildung 3.5-5, Teilbild C). Ziel ist es die gesamte Variation von Y zu extrahieren, die mit X in Beziehung steht. Es wird eine Ordination der Matrix Y unter der Einschränkung des Informationsgehaltes in Matrix X erzeugt, d.h. der Datensatz in Y wird durch die Variablen in X erklärt. Die Redundanzanalyse ist zu bevorzugen, wenn die Daten in X lineare Beziehungen zu den Variablen in Y aufweisen, die kanonische Korrespondenzanalyse im Falle unimodaler Zusammenhänge.

Viele Verfahren der multivariaten Statistik erfordern es, eine algebraische Basis zu schaffen, die aus einer orthogonalen Form einer nicht orthogonalen symmetrischen Matrix besteht. Da ökologische Daten fast immer viele nicht voneinander unabhängige Variablen (d.h. linear miteinander korrelierte) umfassen, ist eine grundlegende Idee, die zahlreichen Verfahren obliegt, die Reduktion der Variablenanzahl durch Bildung einiger unkorrelierter Linearkombinationen (vgl. Abschnitt 3.5.1). Jede dieser Linearkombinationen erklärt dann einen exklusiven Teil der beobachteten Variation. Mit wenigen linear kombinierten Variablen soll so ein möglichst großer Anteil der Variabilität erfasst und unterschiedliche Phänomene der Datenstruktur erklärt werden. Die Mathematik dahinter zielt auf die Berechnung von **Eigenwerten**  $\lambda$  und **Eigenvektoren**  $u$  (Jongman et al. 1987; Legendre & Legendre 1998; ter Braak 1998). Der Begriff Eigenwert (eigenvalue, characteristic root, latent root) kommt aus dem Deutschen, wo "Eigen" für zu etwas gehörend bzw. charakteristisch steht. Er ist ein Maß für die Bedeutung einer Ordinationsachse (Linearkombinationen von Erklärungsvariablen, s.u.), den Anteil erklärter Variation entlang der Achse und in diesem Zusammenhang auch ein Maß für die Relevanz des beschriebenen Umweltgradienten. Die exakte Bedeutung wird aber durch die angewandte Ordinationsmethode festgelegt (Palmer 1998). Die orthogonalen Eigenvektoren bilden neue voneinander unabhängige Variablen (Ordinationsachsen).

Ausgehend von einer quadratischen Matrix A (meist eine symmetrische Assoziationsmatrix), deren Elemente außerhalb der Diagonalen den Grad der Assoziation zwischen den Fällen oder Variablen beschreiben, wird eine korrespondierende Diagonalmatrix  $\Lambda$  gesucht, die neben Nullwerten in der Diagonalen die Eigenwerte  $\lambda$  der nun orthogonalen Eigenvektoren beinhaltet, d.h. deren Assoziation zueinander beschreibt. Die Matrix der Eigenwerte wird auch als kanonische Form der Matrix A bezeichnet. Die Eigenwerte und -vektoren der Matrix A können anhand folgender Gleichung ermittelt werden:

**Gleichung 3.5-1:**  $Au_i = \lambda_i u_i$

Es ist notwendig, eine Einheitsmatrix I einzufügen, um die Differenz zwischen zwei quadratischen Matrizen bestimmen zu können.

**Gleichung 3.5-2:**  $[A - \lambda_i I] = 0$

Die Lösung von Gleichung 3.5-2 ergibt die mit der Ausgangsmatrix A assoziierten Eigenwerte. Die Gleichung wird auch als charakteristische Gleichung (determinantal or characteristic equation) bezeichnet.

Da weiterhin ökologische Daten eine räumliche Struktur besitzen, ist es Ziel, die beobachteten Daten zu ordnen und dann die geeignetste Projektion im reduzierten Ordinationsraum zu lokalisieren, damit das resultierende Modell die realitätsnächste Darstellung der ursprünglichen Datenstruktur abbildet (Ordinationsdiagramme, vgl. Abschnitt 3.6.2). Das Maximieren der Variabilität in einer Projektionsebene ist gleichbedeutend mit dem Minimieren der Residuen um die Projektionsebene herum (Doledec & Chessel 1997). Projektion als geometrische Operation und Modellierung der Struktur als numerische Operation haben die gleiche Natur.

### 3.5.2 Assoziations- und Abhängigkeitsmaße: Eigenschaften und ihre Zuordnung zu Ordinationsverfahren

*"THE MOST USUAL APPROACH TO ASSESS THE RESEMBLANCE AMONG OBJECTS OR DESCRIPTORS IS TO FIRST CONDENSE ALL THE INFORMATION AVAILABLE IN THE ECOLOGICAL DATA MATRIX INTO A SQUARE MATRIX OF ASSOCIATION AMONG OBJECTS OR DESCRIPTORS (LEGENDRE & LEGENDRE 1998, S. 251)."*

In der numerischen Ökologie ist die Struktur, die durch Ordinationsverfahren visualisiert wird, die der Assoziationsmatrix (vgl. Abschnitt 3.5.1). Die Ergebnisse der Analyse reflektieren daher nicht notwendigerweise die Gesamtinformation der Originaldatenmatrix, sondern nur die kondensierte Information der Assoziationsmatrix, d.h. die quantifizierte Ähnlichkeit zwischen den Beobachtungsobjekten oder den Variablen (Deskriptoren) der Rohdatenmatrix.

Im aktuellen Abschnitt bezieht sich die Diskussion von Assoziations- und Abhängigkeitsmaßen (s.u.) auf die zwei Basistypen der Datenbetrachtung (vgl. Abschnitt 3.5.1), d.h. Maße berechnet zwischen Beobachtungsobjekten (Q-Mode) oder zwischen Deskriptoren (R-Mode). Die Einteilung der Daten in Objekte und Deskriptoren wird letztlich durch die Zielsetzung in der Studie bedingt und führt häufig zu einer Konfusion der Termini bzw. des verfolgten Analysetyps (vgl. Abschnitt 3.3). Jeder Analysetyp erfordert sein spezifisches Assoziations- oder Abhängigkeitsmaß, und die Zuverlässigkeit von Aussagen in multivariaten Methoden hängt stark von der Kompatibilität zwischen angewendetem Maß und der Datenstruktur ab, wodurch dessen Bedeutung betont wird.

Die Anwendung und Eigenschaften einiger in Ordinationsverfahren integrierter Maße werden daher in diesem Abschnitt besprochen. Einen Überblick über weitere Assoziationsmaße geben Legendre & Legendre (1998), Liebetrau (1983), Økland (1990) und Orłóci (1978).

#### Q-Mode-Analyse

Assoziationsmaße charakterisieren die Intensität der Beziehung zwischen zwei Objekten. Dabei werden sich die Objekte bei Ähnlichkeitsmaßen zunehmend ähnlicher, je größer der Wert des Maßes wird. Bei den Distanz- oder Unähnlichkeitsmaßen verhält es sich genau umgekehrt: Je kleiner der Wert des Distanzmaßes, desto ähnlicher sind sich zwei Objekte (Backhaus 1990). In Ordinationsverfahren kommen Ähnlichkeitsmaße nicht zur Anwendung, da sie mangels metrischer Eigenschaften nicht geeignet sind, um Objekte in einem entsprechend metrischen Raum zu positionieren. Zu Distanzmaßen transformierte Ähnlichkeitsmaße bzw. Distanzmaße mit "vollmetrischen" Eigenschaften sind gefordert (Legendre & Legendre 1998).

Der einfachste Weg zur Berechnung von Distanzen in einem mehrdimensionalen Raum ist die Berechnung euklidischer Distanzen. In einem zwei oder dreidimensionalen Raum ist dieses Maß die räumliche geometrische Distanz zwischen Objekten.

**Euklidische Distanz (ED):** Sie ist wahrscheinlich das gebräuchlichste Distanzmaß überhaupt (vgl. Gauch 1982a; Økland 1990; Orłóci 1978):

Gleichung 3.5-3: 
$$D(x_1, x_2) = \left[ \sum_{j=1}^p (y_{1j} - x_{2j})^r \right]^{\frac{1}{r}}$$

mit:

$D(x_1, x_2)$ :

Distanz zwischen den Fällen  $x_1$  und  $x_2$

$y_{1j}; y_{2j}$ :

Wert der Variablen  $y_j$  ( $j=1, 2, \dots, p$ ) für Fall  $x_1$  bzw.  $x_2$  (meist Probeflächen)

$r = 2$ :

Minkowski-Konstante ( $L_2$ -Norm);  $r = 1$  ( $L_1$ -Norm) besitzt semimetrische Eigenschaften und ist daher weniger geeignet für Ordinationen

Die ED quadriert die Daten (Gleichung 3.5-3) und betont daher die dominanten Arten in ökologischen Studien (van der Maarel 1979). Damit wird das Linearitätskriterium verletzt (Gauch 1982a; Økland 1990). Der Mangel einer oberen Skalenbegrenzung bewirkt, dass die ED zweier Probestflächen ohne gemeinsame Arten geringer ist als die zwischen zwei Probestflächen mit vielen gemeinsamen Arten, aber mit beträchtlichen Unterschieden in den Häufigkeiten einzelner Arten (Økland 1990). Die ED der Flächen  $x_1$  und  $x_2$  ohne gemeinsame Arten ist geringer als die der Flächen  $x_1$  und  $x_3$  mit gemeinsamen Arten. Ein numerisches Beispiel dieses Paradoxons gibt Tabelle 3.5-1.

**Tabelle 3.5-1: Euklidische Distanz für die Probestflächen eines Beispieldatensatzes.**

Orig. Werte	Arten			ED	Fläche		
	$y_1$	$y_2$	$y_3$		$x_1$	$x_2$	$x_3$
Fläche				Fläche			
$x_1$	2	2	0	$x_1$	0	3,464	5,657
$x_2$	0	0	2	$x_2$	3,464	0	8,718
$x_3$	6	6	0	$x_3$	5,657	8,718	0

$$D(x_1, x_2) = [(2-0)^2 + (2-0)^2 + (0-2)^2]^{\frac{1}{2}} = 3,464$$

Im Allgemeinen bedeutet dies, dass das Auftreten von doppelten Nullwerten (vgl. unimodaler Modellgedanke einer Nische in Abschnitt 2.3) zu einer Distanzreduktion führt (DoppelnulLEN-Problem). Dieser Umstand macht die ED eher ungeeignet als biotisches Assoziationsmaß, wenn Probestflächen anhand von Artmächtigkeits-/Häufigkeitsdaten verglichen werden (Legendre & Legendre 1998). Ein essentielles Problem der Hauptkomponentenanalyse (PCA) und Redundanzanalyse (RDA) leitet sich aus diesem Faktum ab (vgl. Abschnitt 4): Die Verfahren ordnen die Objekte im multidimensionalen Ordinationsraum der Deskriptoren basierend auf der L2-Norm der euklidischen Distanz (Euklidian or Pythagorean distance; vgl. ter Braak 1998, S. 134).

**Chi-Quadrat Distanz ( $\chi^2$ , Chi-squared distance):** Sie ist als biotisches Distanzmaß in die CA-Ordinationsfamilie integriert (Biplot-Skalierung; vgl. Abschnitte 3.5.5 und 3.5.8). Die implementierte Multiplikation verstärkt das Gewicht insgesamt dominanter Arten, weshalb die Wurzeltransformation der Rohdaten angeraten wird (Økland 1990, Palmer 1993, 1998; Abschnitt 3.3).

Das Maß basiert auf der Chi-Quadrat-Metrik (Legendre & Legendre 1998; Økland 1990):

**Gleichung 3.5-4:**

$$D(x_1, x_2) = \sqrt{\sum_{j=1}^p \frac{1}{y_{+j}} \left( \frac{y_{1j}}{y_{1+}} - \frac{y_{2j}}{y_{2+}} \right)^2}$$

mit:  $D(x_1, x_2)$ : Distanz zwischen Objekt in Zeile 1 zu Objekt in Zeile 2  
 $y_{+j}$ : Summe der Häufigkeiten in Spalte j  
 $y_{1+}$ : Summe der Häufigkeiten in Zeile 1

Die Abweichungsquadrate der Differenzen zwischen den Profilen der konditionalen Wahrscheinlichkeiten zwischen zwei Reihen oder Spalten einer Häufigkeitstabelle werden berechnet, weshalb die Matrix der konditionalen Wahrscheinlichkeiten herangezogen werden kann:

$$\begin{bmatrix} Y \\ y_{+j} \end{bmatrix} \begin{bmatrix} y_{i+} \\ y_{++} \end{bmatrix} \Rightarrow \left[ \frac{y_{ij}}{y_{i+} y_{+j}} \right] \text{ Matrix der konditionalen Wahrscheinlichkeiten}$$

Dabei wird jeder Term der Abweichungsquadrate gewichtet mit der inversen Häufigkeit der Zeile oder Spalte in der Gesamtmatrix. Jeder Deskriptor enthält einen exklusiven Teil an Information. Zusätzlich kann er einen Teil seiner Information mit einem anderen Deskriptor gemeinsam haben, d.h., durch den Informationsgehalt von einem Deskriptor 1 kann gleichzeitig auf einen bestimmten Anteil des Informationsgehaltes von Deskriptor 2 geschlossen werden. In der Matrix konditionaler Wahrscheinlichkeiten enthält jede Zeile die Wahrscheinlichkeiten, ein Objekt für den entsprechenden Zustand des Deskriptors 2 unter der Annahme, dass das Objekt den zu der Zeile korrespondierenden Zustand des Deskriptors 1 besitzt, zu selektieren. Die Zeilensumme der Wahrscheinlichkeiten beträgt immer eins, weil die Häufigkeitswerte durch die zugehörige Zeilensumme dividiert werden. Obgleich die Chi-Quadrat-Metrik kein oberes Limit besitzt, sind deren Werte meist kleiner eins. Das Chi<sup>2</sup>-Distanzmaß ist eng an die Chi<sup>2</sup>-Metrik angelehnt. Es unterscheidet sich nur darin, dass die Abweichungsquadrate durch die Wahrscheinlichkeiten (relative Häufigkeiten) dividiert werden und nicht durch die absoluten Häufigkeiten:

**Gleichung 3.5-5:** 
$$\chi^2 = D(x_1, x_2) = \sqrt{\sum_{j=1}^p \frac{1}{y_{+j} / y_{++}} \left( \frac{y_{1j}}{y_{1+}} - \frac{y_{2j}}{y_{2+}} \right)^2}$$

mit:  $y_{++}$ : Summe aller Häufigkeiten in der Datenmatrix

Berechnet man die Distanzen zwischen Fall  $x_1$  und  $x_2$  bzw.  $x_1$  und  $x_3$  aus dem Beispiel in Tabelle 3.5-1 erhält man die Lösungen in Tabelle 3.5-2. Die erste Art, welche auf Probefläche  $x_1$  und  $x_3$  nicht auftritt, kürzt sich aus der Gleichung heraus, weshalb das Distanzmaß metrisch ist und das Doppelnullen-Paradoxon ausschließt. Soll umgekehrt die Metrik zwischen den Arten kalkuliert werden, müssen die Wahrscheinlichkeiten zwischen den Spalten berechnet werden. Bevor Gleichung 3.5-4 angewandt wird, ist die Matrix zu transponieren (Legendre & Legendre 1998).

**Tabelle 3.5-2: Chi<sup>2</sup>-Distanz für die Probeflächen eines Beispieldatensatzes.**

Orig. Werte	Arten			ED	Fläche		
Fläche	$y_1$	$y_2$	$y_3$	Fläche	$x_1$	$x_2$	$x_3$
$x_1$	2	2	0	$x_1$	0	10,125	0
$x_2$	0	0	2	$x_2$	10,125	0	10,125
$x_3$	6	6	0	$x_3$	0	10,125	0

$$D(x_1, x_2) = \left[ \frac{(0,5 - 0)^2}{8/18} + \frac{(0,5 - 0)^2}{8/18} + \frac{(0 - 1)^2}{2/18} \right]^{\frac{1}{2}} = 10,125 \quad \text{und}$$

$$D(x_1, x_3) = \left[ \frac{(0,5 - 0,5)^2}{8/18} + \frac{(0,5 - 0,5)^2}{8/18} + \frac{(0 - 0)^2}{2/18} \right]^{\frac{1}{2}} = 0,000$$

**Die generalisierte Mahalanobis Distanz:** Sie ist ein weiteres wichtiges Assoziationsmaß in unimodalen Verfahren, welches die Korrelation zwischen Deskriptoren berücksichtigt und unabhängig von deren Skalen ist. Das Maß berechnet die Distanz zwischen zwei Punkten in einem Raum, dessen Achsen nicht notwendigerweise orthogonal zueinander sind. Obgleich Orlóci (1978, S.48) die Formel zur Berechnung der Distanzen zwischen zwei Punkten modifiziert hat, wird sie sonst nur für den Vergleich zweier Gruppen ( $w_1$ ,  $w_2$ ) mit jeweils  $n_1$  bzw.  $n_2$  Objekten, die durch die gleichen Deskriptoren ( $p$ ) bestimmt werden, herangezogen (Legendre & Legendre 1998). Das Quadrat der generalisierten Mahalanobis Distanz wird gegeben durch Gleichung 3.5-6 (S. 65).

**Gleichung 3.5-6:**

$$D^2(w_1, w_2) = \bar{d}_{12} V^{-1} \bar{d}_{12}'$$

mit:  $\bar{d}_{12}$ : Vektor der Differenzen zwischen Mitteln aller Variablen (p) in zwei Gruppen von Objekten w1, w2: Objektgruppen

Die Matrix V (Gleichung 3.5-7) beinhaltet die zusammengefasste Innergruppen-Varianz (pooled within-group dispersion) der beiden Gruppen, welche geschätzt wird aus den Matrizen der Summen der Abweichungsquadrate und den Kreuzprodukten zwischen den gruppenzentrierten Deskriptoren der jeweiligen Gruppe an Objekten, die schrittweise aufaddiert und dann dividiert werden durch  $(n_1 + n_2 - 2)$ . Dieser Weg wird in der Diskriminanzanalyse und der Korrespondenzanalyse (Skalierung nach Hill in CANOCO®, vgl. Abschnitte 3.5.5 und 3.5.8) angewendet (Backhaus 1990; Sokal & Rohlf 1995; ter Braak 1998; Zar 1996).

**Gleichung 3.5-7:**

$$V = \frac{1}{n_1 + n_2 - 2} [(n_1 - 1)S_1 + (n_2 - 1)S_2]$$

mit: S1, S2: Streuungsmatrizen für jede Gruppe  
V: Matrix, die Kovarianzen zwischen den Deskriptoren berücksichtigt

Die generalisierte Mahalanobis Distanz in der von Orlóci beschriebenen Variante (Hill's Scaling) lässt sich auch auf kurzem Wege berechnen, indem man die Ordinationswerte der Objekte (sample scores) durch  $\sqrt{(1-\lambda)/\lambda}$  und die Ordinationswerte der Deskriptoren (species scores) durch  $\sqrt{\lambda(1-\lambda)}$  dividiert (Jongman et al. 1987; ter Braak 1998). Die Interpretation der resultierenden Achsenskalierung ist so gerichtet, dass Probeflächen, die weiter als vier Standardabweichungen voneinander entfernt im Ordinationsdiagramm liegen, keine oder nahezu keine Arten mehr gemeinsam haben (vgl. Abschnitte 2.3 und 3.5.5).

Bei allen Distanzmaßen existieren drei Mängel:

- (1) Aufgrund von Datenrauschen haben Probeflächen selbst bei theoretisch identischen Standortbedingungen nie dieselbe Artenkomposition. Somit kann die Untergrenze des Unähnlichkeitsmaßes von null nicht erreicht werden, auch wenn die ökologische Distanz null ist. Über sehr kurze ökologische Distanzen haben diese Maße somit immer einen Mangel an Robustheit. Die tatsächliche Untergrenze des Unähnlichkeitsmaßes lässt sich nicht exakt bestimmen (Økland 1990).
- (2) Probeflächen, die mehr als vier Standardabweichungen entlang einer Coenocline entfernt voneinander liegen, haben nur noch wenige bis keine Arten mehr gemein. Sie besitzen damit eine biotische Gleichheit von null (Gauch & Whittaker 1972). Distanzmaße sind deshalb vage hinsichtlich Coenoclinen-Distanzen größer vier Standardabweichungen.
- (3) Alle Maße der ökologischen Distanz zeigen ein typisches nicht lineares Verhalten. Die biotische Separierung nimmt geringfügig zwischen 0 bis 0,5 Standardabweichungen zu, um dann stärker anzusteigen bis zum Wendepunkt um etwa zwei Standardabweichungen. Dann erfolgt eine allmähliche Abnahme/Einnivellierung bis hin zu vier Standardabweichungen (Gauch 1982a; Økland 1990; Whittaker 1967). Dies basiert auf den nicht linearen Beziehungen zwischen dem artendimensionalen und dem ökologischen Raum bzw. dem Faktum, dass die Artenhäufigkeiten entlang der Komplexgradienten nicht monoton zu- oder abnehmen (Gauch 1982a; Økland 1990).

### R-Mode-Analyse

In R-Mode-Analysen werden Abhängigkeitskoeffizienten (Korrelationskoeffizienten, Kovarianzen) angewendet (Legendre & Legendre 1998; ter Braak 1998).



Die Kovarianz zwischen zwei Deskriptoren wird anhand der zentrierten Variablen und der Korrelationskoeffizient nach Pearson anhand der standardisierten Variablen berechnet (vgl. Abschnitt 3.5.4).

Am häufigsten wird der Produkt-Moment-Korrelationskoeffizient nach Pearson angewendet. Da die Antwortkurven von Arten zu einem langen Gradienten meist nicht linear, sondern glockenförmig sind, erscheint der Koeffizient aber als ungeeignet. Er sollte nur bei linearen Zusammenhängen, d.h. bei einer geringen Beta-Diversität, genutzt werden (vgl. Abbildung 2.3-5, S. 18). Besser ist die Anwendung nicht parametrischer oder verteilungsfreier Korrelationskoeffizienten, welche die Beziehungen nur basierend auf der hierarchischen Ordnung (rank order) der Beobachtungen betrachten. Sie sind robuster gegen nicht lineare Variation, nehmen aber dennoch an, dass die Variablen einen monoton ansteigenden oder abfallenden Zusammenhang besitzen gegenüber dem unterliegenden Komplex-Gradienten.

Spearman- oder Kendalls-Rangkorrelationskoeffizient sind gängige Maße. Unempfindlicher gegen die Verteilung der Variablen einerseits, ergeben sie die gleichen Ergebnisse mit allen monotonen Transformationen der Variablen andererseits. Sie haben nur eine geringfügig kleinere Analysekraft als der Produkt-Moment-Korrelationskoeffizient, wenn die Variable normal verteilt ist, dafür aber eine deutlich höhere Kraft, wenn die Verteilung abweicht (Økland 1990).

Der Einsatz von Korrelationskoeffizienten beim Vergleich von Objekten (Q-Mode) ist generell zu vermeiden. Selbst wenn die Deskriptoren standardisiert wurden, ist dennoch der Koeffizient zwischen zwei Objekten eine Funktion basierend auf allen Objekten im Datensatz. Typisch für biotische Felddaten ist, dass viele Arten auf den Probeflächen nur durch geringe Häufigkeiten vertreten sind und dadurch eine große stochastische Variation entsteht. Als Folge sind die Ränge der Arten unzuverlässig bezogen auf das Referenzökosystem. Die nicht parametrischen Koeffizienten zwischen Objekten sind anfällig für diese zufallsbedingte Variation, da ihre Werte auf einer somit großen Anzahl unsicherer Ränge kalkuliert wurden. Die Vorgehensweise ist gleichbedeutend mit der Tatsache, dass dadurch entscheidendes Gewicht auf viele nur dürtig repräsentierte Arten gelegt wird (Legendre & Legendre 1998).

Weiterhin bedeutet eine geringe Korrelation zwischen Deskriptoren – auch Arten können wie bereits bekannt, als solche genutzt werden – nicht unbedingt, dass sie zueinander unkorreliert sind. Kovariieren die Arten nicht in einer linearen Art, dann verfehlt das Modell seine Aussage und die eigentliche Information bleibt verborgen.

Wie deutlich wurde, sind bestimmte Assoziationsmaße der Q-Mode-Analyse in verschiedene multivariate Ordinationsmethoden integriert, wodurch eine genaue Kenntnis ihrer Eigenschaften essentiell ist, um eine zuverlässige Interpretation der Ergebnisse zu gewährleisten. Fasst man die Aussagen dieses Abschnitts zusammen, sei nochmals betont, dass der CA/CCA anstatt der euklidischen Distanz (PCA/RDA) die  $\chi^2$ -Distanz (bzw. die generalisierte Mahalanobis Distanz) unterliegt und sich dadurch das Problem der Doppel-Nullen ausschließt. Allgemein gilt daher, dass die Korrespondenzanalyse allein wegen des unterliegenden Distanzmaßes besser als die Hauptkomponentenanalyse geeignet ist um die Artenverteilung im Raum zu studieren.

### 3.5.3 Variablenreduktion: Korrelationsanalyse, Vorwärtsselektion und andere Verfahren

**"OKHAM'S RAZOR: PLURALITAS NON EST PONENDA SINE NECESSITATE (WILLIAM OKHAM 1290-1349 IN LEGENDRE & LEGENDRE 1998, S. 520)."**

Die Auswahl einer möglichst geringen Anzahl relevanter Deskriptoren zur Erklärung des Musters in den Beobachtungsobjekten ist der grundlegende Schritt jeder Modellbildung (vgl.

Abschnitte 3.7.1, 5.4.1 und 5.5). Insgesamt gibt es mehrere nennenswerte Gründe für die Durchführung einer Variablenreduktion, wobei das Prinzip der Sparsamkeit (principle of parsimony) die zentrale Stellung einnimmt. Unnötige Komplexität eines Modells kann vermieden werden, indem (multi-) kollineare und nicht signifikant zur Modellverbesserung beitragende Parameter ausgeschlossen werden. Signifikanzen sind nur für unabhängige Variablen eindeutig (Legendre & Legendre 1998; Zar 1996).

Mit zunehmender Multikollinearität werden einerseits die Schätzungen der Regressionskoeffizienten in einer multiplen Regression unzuverlässig (Backhaus 1990; Jaccard et al. 1990) und andererseits verändern sich die Modellkoeffizienten erheblich, wenn eine hochkorrelierte Variable hinzukommt oder herausgenommen wird (Legendre & Legendre 1998). Bei perfekter Multikollinearität ist eine Regressionsanalyse nicht mehr durchführbar (Backhaus 1990). Der Effekt der Multikollinearität sollte daher minimiert werden.

Weiterhin ist es notwendig, eine Variablenreduktion durchzuführen, wenn der Stichprobenumfang und die Anzahl der unabhängigen Variablen sich annähern. Wird dies nicht berücksichtigt, können Scheinzusammenhänge durch zufällige Linearkombinationen auftreten. Die Analysekraft eines direkten Ordinationsverfahrens sinkt im Falle einer Modellübersättigung, d.h. die Ordnung der Matrix Y wird nicht länger durch die Variablen in Matrix X eingeschränkt. Die Ergebnisse der direkten Ordination nähern sich der einer indirekten an (Causton 1988; Gauch 1982a; Jongman et al. 1987; Pielou 1977; ter Braak 1998 u.a.).

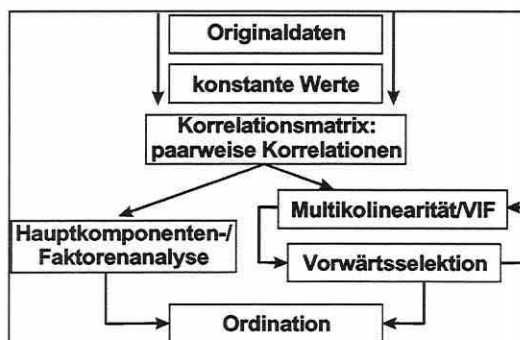


Abbildung 3.5-6: Hierarchische Prüfschritte bei der Variablenreduktion.

Zur Durchführung einer Variablenreduktion stehen mehrere Wege zur Verfügung, die im besten Falle kombiniert angewendet werden sollten, wie in dieser Arbeit geschehen (vgl. Abschnitt 5.4.1 und Abbildung 3.5-6). Grundsätzlich kann die Variablenreduktion nach externen Kriterien (z.B. Kompatibilität zu anderen Studien) oder nach datensatzspezifischen Eigenschaften erfolgen. Letzteres umfasst eine Prüfung auf konstante Werte über die Stichprobenflächen in den Originaldaten (kein Trennvermögen) und die Auswertung von Korrelationsmatrix, multipler Regression, Vorwärtss Selektion

(Forward Selection) und/oder Faktoren- bzw. Hauptkomponentenanalyse. Alle diese Verfahren prüfen auf verschiedene Art die (Multi-) Kollinearitäten und Erklärungsanteile der Erklärungsvariablen und geben dem Bearbeiter Kennwerte an die Hand, welche die Einschätzung von Bedeutung bzw. Redundanz der Erklärungsvariablen und damit die Variablenreduktion erlauben.

Nach dem **Entfernen konstanter Variablen** im Datensatz, ist die **Prüfung der Korrelationsmatrix** auf hohe Werte der Korrelationskoeffizienten ( $> 0,9$ ) der erste Schritt Kollinearitäten zu ermitteln und eine Variablenreduktion durchzuführen. Dabei ist zu bedenken, dass Korrelationen kleiner eins immer noch einen Resterklärungsanteil besitzen und der Ausschluss einer solchen Variablen mit einem geringen Informationsverlust einhergeht. Es werden nur paarweise Abhängigkeiten mit dem Korrelationskoeffizienten gemessen, daher kann trotzdem eine hochgradige Multikollinearität zu Variablenkombinationen bestehen (Backhaus 1990). Zur **Aufdeckung von Multikollinearität** (zweiter Schritt) kann für jede unabhängige Variable die Berechnung des multiplen Bestimmtheitsmaßes  $R_j^2$  gegenüber den übrigen Variablen erfolgen. Ein Wert von  $R_j^2$  gleich eins bedeutet, dass die Variable  $x_j$  sich durch eine Linearkombination bestimmter anderer Variablen in der Betrachtung erzeugen lässt und damit keine Zusatzinformation liefert. Die Toleranz der Variablen ( $1-R_j^2$ ) ist der inverse Wert

und damit ist null gleichbedeutend mit perfekter Multikollinearität. Ein eindeutiger Grenzwert, wann Multikollinearität ernsthafte Unzuverlässigkeiten in einer multiplen Regression verursacht, lässt sich nicht angeben (Backhaus 1996). CANOCO® prüft die Erklärungsvariablen auf Multikollinearität anhand des "Variance Inflation Factors" ( $VIF = 1 / (1 - R_j^2)$ ), der bezogen ist auf die multiple Korrelation zwischen der Variablen  $x_j$  und den anderen Variablen in der Analyse. Ein VIF größer 20 bedeutet eine fast perfekte Korrelation (Toleranz = 0,05 bzw.  $R_j^2 = 0,95$ ) und null eine perfekte Korrelation. Umgekehrt steht ein VIF-Wert von eins für völlig unkorrelierte Variablen (ter Braak 1998). Diese Prüfung ist kombiniert mit dem in CANOCO® implementierten **Verfahren der Vorwärtsselektion** (FS, forward selection, dritter Schritt) und wird mit der FS vorrangig in dieser Arbeit zur Variablenreduktion eingesetzt.

Die Vorwärtsselektion geht von einem Nullmodell ohne Erklärungsvariablen aus und überführt iterativ jeweils die Variable ins Modell, die den höchsten Beitrag zur Verbesserung der Erklärungsanteile des Modells hat. Die Selektion endet, wenn alle verfügbaren Variablen in das Modell aufgenommen wurden, ein vordefinierter Minimumlevel für den Zuwachs von  $R^2$  oder eine vorbestimmte Signifikanzschranke unterschritten wird (vgl. Abschnitt 5.4.1). Die schrittweise Modellbildung reduziert die Gefahr einer Modellbeeinflussung durch (Multi-) Kollinearität, da hoch korrelierte Erklärungsparameter kaum noch zur Modellverbesserung beitragen können, wenn bereits eine oder eine Kombination der korrelierten Variablen ins Modell integriert wurde. Sie bleiben meist in den weiteren Schritten der Vorwärtsselektion unberücksichtigt. Dennoch kommt es gelegentlich zur Aufnahme, wenn geringe Erklärungsanteile einer weitgehend kollinearen Variablen (Variablen erklären sich nie zu 100% gegenseitig) einen signifikanten Beitrag zur Modellverbesserung liefern und diese integriert wird (vgl. auch Palmer 1993). Durch die abschließende Kontrolle des VIF und der Korrelationsmatrix lässt sich das Kollinearitätsproblem allerdings leicht aufdecken, so dass die Variable wieder entfernt werden kann (vgl. Abbildung 3.5-6, S. 67). Ein Problem der Vorwärtsselektion ist, dass sie Variablen streng hierarchisch ins Modell integriert und daher die weitere Auswahl von den bereits im Modell befindlichen Variablen abhängt, da selten absolute Unkorreliertheit besteht (Legendre & Legendre 1998).

CANOCO® stellt neben der automatischen eine halbautomatische Variante der Vorwärtsselektion bereit, die sowohl die Einzeleffekte (marginal effects) als auch die bedingten Effekte (conditional oder partial effects) berechnet. Bevor die erste Variable in das Modell aufgenommen wird, zeigen die Einzeleffekte den jeweils alleinigen Erklärungsgehalt einer Variable an der Gesamtvarianz (total inertia) der zu erklärenden Daten an. Ab dem zweiten Durchlauf erfolgt die Ausgabe einer geordneten Liste aller restlichen Erklärungsvariablen, welche den Zuwachs des Erklärungsanteils in der Regressionsfunktion nach deren Einschluss ins Modell angibt (vgl. Abschnitt 5.4.1). Jede Variable kann vor ihrer Aufnahme ins Modell mit einer Monte Carlo Permutation (vgl. Abschnitt 3.6.1) darauf getestet werden, ob ihr Beitrag zur Modellverbesserung signifikant oder zufällig ist (ter Braak 1998).

#### **Zusammengefasste Vorgehensweise der halbautomatischen Vorwärtsselektion:**

- (1) In der ersten Iteration werden separate kanonische Korrespondenzanalysen (CCAs) für jede Variable gerechnet und deren Erklärungsanteile ausgegeben.
- (2) Der Anwender wählt einen Modellparameter aus, indem er dessen Beitrag zur Erklärung der Antwortvariablen und/oder dessen Signifikanz berücksichtigt. Bei mehreren signifikanten und etwa gleich gut erklärenden Parametern kann er die Eigenschaften der Parameter (z.B. einfache Erhebbarkeit) für eine Auswahl gegenüberstellen oder sich von seiner persönlichen Einschätzung bezüglich der Bedeutung für die Studie leiten lassen.
- (3) Nach dem Auswählen der ersten Modellvariablen werden partielle CCAs mit Berücksichtigung der bereits im Modell befindlichen Parameter kalkuliert.
- (4) Die Iteration beginnt erneut bei Schritt (2).

- (5) Das Modell wird fertiggestellt, wenn das "Signifikanzniveau"\* von  $\alpha = 0,05$  unterschritten wird oder wenn eine sachlogische Anzahl an Variablen ausgewählt wurde.

\*Der Grund dafür, dass der Begriff "signifikant" in Anführungszeichen steht, ist, dass keine wahre statistische Signifikanz vorliegt, sondern durch multiple Vergleiche des Monte Carlo Tests (vgl. Abschnitt 3.6.1) Fehler auftreten können und der Typ 1 Fehler erhöht ist (Palmer 1998; Sokal & Rohlf 1995).

- (6) Das Modell wird fertiggestellt und die Parameter in die aktuelle Studie übernommen.

Anzumerken bleibt, dass kategoriale Variablen, die als Dummies binär kodiert wurden, in der Vorwärtsselektion als unabhängige Deskriptoren behandelt werden, d.h. der Beitrag des gesamten Faktors lässt sich nicht ermitteln, sondern nur der einer jeweiligen Kategorie. Es stellt sich die Frage, ob alle (d.h.  $k-1$ ) Kategorien des Faktors in das Modell integriert werden sollen, auch wenn nur eine Dummy-Variable durch die Vorwärtsselektion vorgeschlagen wird. Die Antwort ist naheliegend: Geht man davon aus, dass es sich um eine kategoriale Variable mit fünf abhängigen Zuständen handelt, dann ja, sonst – im unabhängigen Falle – nein (Palmer 1998). Die beiden letzten Kategorien einer solchen Variablen haben den identischen Erklärungsgehalt (vgl. Abschnitt 3.3) und die häufigere Kategorie sollte gewählt werden.

Ein Problem aller schrittweisen Verfahren ist, dass man nie ausschließen kann, ob ein anderer Variablenteildatensatz existiert, welcher in seiner spezifischen Kombination einen höheren Gesamterklärungsanteil besitzt. Dies bedeutet: "The only way to determine the best set of variables is to perform a separate CCA for every conceivable combination of variables; this is, in most cases, impossible with current technology because it involves an astronomically large number of combinations (Palmer 1998, reduce.html, S. 3: <http://okstate.edu/artsci/botany/ordinate/>)". Die mangelnde Garantie, eine perfekte Lösung zu erhalten, muss nicht zu sehr beunruhigen, wenn das Ziel ist, eine eingeschränkte Anzahl von Variablen zu selektieren, welche die Variation in den biotischen Daten zufriedenstellend erklären.

Betrachtet man mehrere Zeitphasen in einem Untersuchungsgebiet, so kann die optimale Variablenkombination je Phase differieren. Die Einzelmodelle weichen im Normalfall voneinander ab, da Modellparameter sich beispielsweise nur im Frühjahr als relevant erweisen und andere im Herbst Bedeutung erhalten. Durch den Vergleich der Erhebungsphasen und einen Abgleich der Deskriptoren ist unter geringen Verlust an Erklärungsanteilen ein zeitlich stabileres Modell zu entwickeln (vgl. Abschnitt 5.4.1).

Abschließend darf gesagt werden, dass speziell die parallele Anwendung mehrerer Verfahren zur Variablenreduktion empfehlenswert ist, um die Erkenntnisse der verschiedenen Methoden gegenüberzustellen und zu verifizieren. So kann die Anwendung einer Hauptkomponenten- oder Faktorenanalyse auf die Matrix der standardisierten Umweltparameter eine Alternative zur Variablensumme oder -reduktion sein (vgl. Abbildung 3.5-6, S. 67). Die Funktionalität der Verfahren wird in Abschnitt 3.5.4 erläutert.

### 3.5.4 Hauptkomponentenanalyse

**"IN SUMMARY, PCA IS AN ORDINATION TECHNIQUE FOR PROJECTING A MULTIDIMENSIONAL CLOUD OF POINTS INTO A SPACE OF FEWER DIMENSIONS, USING RIGID ROTATION TO DERIVE SUCCESSIVE ORTHOGONAL AXES, WHICH MAXIMIZE THE VARIANCE ACCOUNTED FOR (GAUCH 1982a, S. 143)."**

Die Hauptkomponentenanalyse (PCA: principal component analysis) ist eine algebraische Methode, die schon 1901 von Pearson eingeführt wurde (Kent & Coker 1992). Sie bildet die Grundlage aller anderen nachfolgend beschriebenen Ordinationsverfahren (Abschnitte 3.5.5 bis 3.5.8).

Das indirekte gradientenanalytische Verfahren ist in der Lage, simultan Objekte und Deskriptoren in einer integrierten Analyse zu ordinieren (Gauch 1982a). Eine zweite Matrix mit



Erklärungsvariablen wird dabei nicht berücksichtigt (vgl. Abbildung 3.5-5, S. 60). Die Datenmatrix kann entweder als Deskriptoren biotische Information (Arten) oder Umweltparameter enthalten (vgl. Abschnitt 3.3 und Abbildung 3.5-4, S. 59). Je nach Datensatzeigenschaften und Betrachtungsweise erfordert die PCA eine unterschiedliche Vorbehandlung der Variablen: Die **Hauptvarianten der PCA** können daher nach der angewandten Transformationen unterschieden werden (vgl. auch Abschnitt 3.3):

- (1) **Untransformierte PCA (non-centered PCA)**: Es wird keine Vorbehandlung der Ausgangsmatrix durchgeführt. Sie basiert auf der Streuungsmatrix der Originaldaten. Das Koordinatensystem hat damit seinen Ursprung in einer hypothetischen Probefläche mit der Artmächtigkeit null für alle Arten.
- (2) **Zentrierte PCA (centered PCA)**: Werden die Originaldaten zentriert, basiert die PCA auf der Varianz-Kovarianzmatrix. Somit ist der Ursprung des Achsensystems zum Zentroid der Daten verschoben, d.h. der mittleren Artmächtigkeit aller Arten.
- (3) **Standardisierte PCA (centered and standardized PCA)**: Die Deskriptoren werden nach dem Zentrieren standardisiert zu einer einheitlichen Varianz von eins. Bei dieser Transformation basiert die PCA auf der Korrelationsmatrix. Die Position des Ursprungs ist identisch zur zentrierten PCA. Durch die Standardisierung werden die Deskriptoren geometrisch gesehen zu einer einheitlichen Länge kontrahiert oder expandiert (Gauch 1982a). Diese PCA-Variante ist speziell gefordert, falls die Einheiten der Daten nicht homogen sind, wie generell im Fall von Umweltparametern.

Der PCA unterliegt ein lineares Modell, d.h., es wird angenommen, dass sich die Antworten der Deskriptoren adäquat als eine lineare Funktion entlang der latenten Variablen (PCA-Achse) beschreiben lassen. Als Assoziationsmaß ist generell die euklidische Distanz implementiert (vgl. Abschnitt 3.5.2). Die Vektoren im Ordinationsdiagramm indizieren die Richtung des steilsten Anstiegs der Deskriptorwerte (beispielsweise Artmächtigkeiten oder standardisierte Umweltparameter) relativ zu den PCA-Achsen. Die Ordinationswerte der Probeflächen sind die Koordinaten ihrer Position entlang der Achse (Jongman et al. 1987; Økland 1990).

Nach der kurzen Einführung der Methode wird als nächstes die oft unklare **Differenzierung zwischen Hauptkomponenten- und Faktorenanalyse** umrissen und dann die eigentliche Arbeitsweise des Verfahrens dargelegt. Das primäre Ziel der Hauptkomponentenanalyse ist die Erklärung der maximalen Varianz in den Daten, während die Faktorenanalyse die Erklärung der Kovarianz zwischen den Deskriptoren anstrebt. Ein weiterer charakteristischer Unterschied zwischen den beiden faktorenanalytischen Modellen besteht darin, dass in der Hauptkomponentenanalyse davon ausgegangen wird, dass die Gesamtvarianz der Variablen im Datensatz relevant ist, um die Achsen zu bestimmen (Legendre & Legendre 1998). In der Faktorenanalyse wird nur die Kommunalität (communality) eines Deskriptors verwendet. Die Kommunalität beschreibt den Varianzanteil, den eine Variable mit anderen Variablen teilt und der durch einen gemeinsamen Faktor erklärt werden kann. Der Anteil an Varianz, welcher einer Variablen eigen ist, trägt zur spezifischen Varianz bei und findet keine Berücksichtigung in der Faktorenanalyse (StatSoft 1997).

Die Grundidee des Verfahrens, welches die Deskriptoranzahl auf eine deutlich geringere Anzahl von gemeinsamen Faktoren reduziert (Datenreduktionsmethode), kann am Regressionsmodell dargestellt werden: In einem Streudiagramm wird eine Regressionsgerade so angepasst, dass die "beste" Beschreibung der linearen Beziehung zwischen den Variablen repräsentiert wird. Kann man dann eine Variable definieren, welche die Regressionsgerade in einem solchen Diagramm approximiert, erklärt diese Variable das Wesen der beiden Deskriptoren. Die einzelnen Werte dieses neuen Faktors, der durch die Regressionsgerade repräsentiert wird, können in zukünftigen Analysen verwendet werden, um die beiden Deskriptoren zu ersetzen. In gewisser Hinsicht sind die zwei Variablen auf einen Faktor redu-



ziert worden, der eigentlich eine Linearkombination aus beiden Originalvariablen ist. Das Prinzip lässt sich auf den Fall mehrerer Variablen verallgemeinern.

Man erhält somit als Vorteil wenige voneinander unabhängige linear kombinierte Faktoren aus dem Satz der Deskriptoren (Ausschluss von Kolinearität), die aber gleichzeitig den Nachteil haben, dass man nun mit Variablenkombinationen arbeitet, wodurch ein Rückschluss auf die einzelnen Originalparameter schwierig ist und die Interpretation eines Ordinationsdiagramms (direkte Verfahren) sehr verwirrend sein kann (Palmer 1998).

In CANOCO® ist nur die Hauptkomponentenanalyse integriert, so dass für weitere Details zur Faktorenanalyse auf die Fachliteratur verwiesen wird (z.B. Backhaus 1990, 1996; StatSoft 1997).

Die Hauptkomponentenanalyse kann in anderer Weise als die Faktorenanalyse ebenfalls zur Variablenreduktion eingesetzt werden. Die Hauptkomponenten verlaufen in Richtung der maximalen Varianz je Dimension in den Daten, d.h., die Selektion der jeweils am engsten mit einer PCA-Achse korrelierten Variablen (vgl. Korrelationen/Ladungen) und deren Einsatz als Modellparameter ist – speziell wenn einige Originalparameter beibehalten werden sollen – eine effektive Methode zur Variablenreduktion.

Die **Arbeitsweise der Hauptkomponentenanalyse** wird im Folgenden zuerst graphisch-deskriptiv veranschaulicht, bevor auf mathematisch-technische Einzelheiten des Verfahrens eingegangen wird.

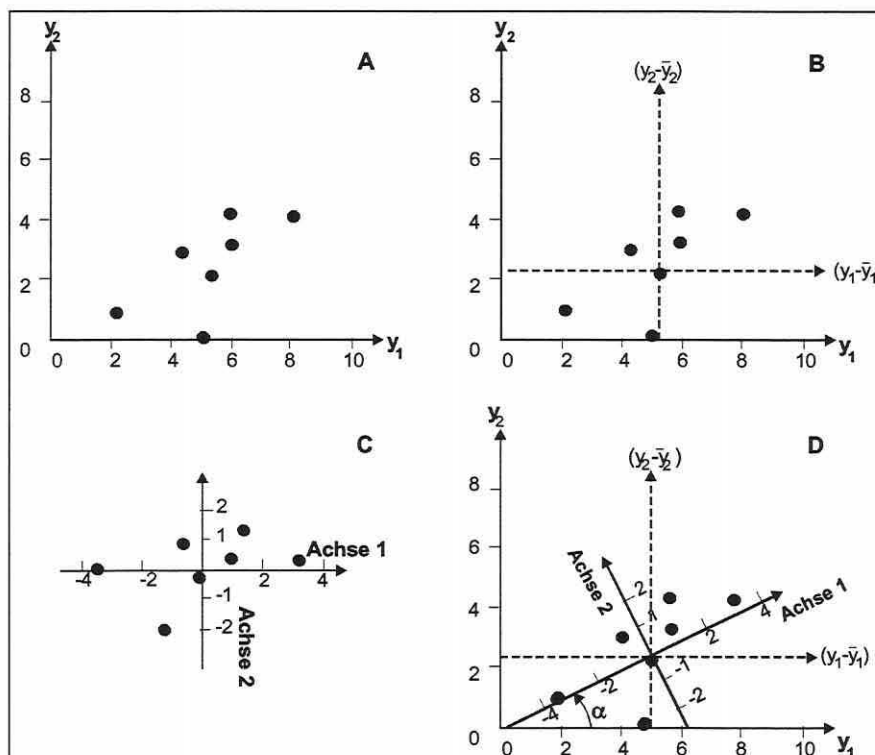


Abbildung 3.5-7: Grafische Repräsentation der Arbeitsschritte einer zentrierten PCA. (A) Objekte im Artenraum der Deskriptoren  $y_1$  und  $y_2$  (Originaldatenmatrix). (B) Darstellung der Objekte nach dem Zentrieren der Originaldaten (gestricheltes Achsensystem ist in den Schwerpunkt der Verteilung verschoben) (Sekundärmatrix: Varianz-Kovarianz Matrix). (C) Repräsentation der Objekte im Hauptkomponentensystem. (D) Überlagerung der Achsensysteme (B) und (C) durch die varianzmaximierende Drehung um den Winkel  $\alpha$ . Die relativen euklidischen Distanzen der Objekte bleiben unverändert. (Legendre & Legendre 1998, S. 393, verändert).

Das Verfahren umfasst mehrere Schritte: Ausgehend von den Originaldaten kann ein mehrdimensionaler Artenraum aufgespannt werden, in dem die Objekte als Punktsignaturen angeordnet werden (Abbildung 3.5-7, Teilbild A, S. 71). Dann erfolgt die gewählte Transformation der Originaldatenmatrix zu einer symmetrischen Streuungsmatrix  $S$  (Kovarianz- oder Korrelationsmatrix) mit entsprechendem Einfluss auf den Ursprung des Achsensystems (Abbildung 3.5-7, Teilbild B). Der eigentliche Rechenschritt extrahiert danach orthogonale varianzmaximierende Achsen (Abbildung 3.5-7, Teilbild C) in der Datenwolke, d.h., die Varianz der "neuen" Variablen (Hauptkomponenten) wird maximal, während die Varianz um die neue Variable minimal wird. Nach dem gleichen Prinzip werden weitere Achsen festgelegt, welche die jeweils verbleibende Variabilität maximieren. Auf diese Weise ergeben sich voneinander unabhängige Hauptkomponenten, die mit zunehmender Zahl jeweils weniger Variabilität im Datensatz erklären. Abschließend dreht man bei einer PCA das Koordinatensystem in die neue Position (Abbildung 3.5-7, Teilbild D), mit dem Resultat, dass die relativen Objektdistanzen erhalten bleiben, aber die erste PCA-Achse nun in Richtung der ersten Hauptkomponente (maximale Variabilität in der Datenwolke) verläuft (Legendre & Legendre 1998; ter Braak 1998). Ist die Datenwolke im multidimensionalen Raum in bestimmten Richtungen konzentriert, dann wird der Hauptanteil an der Gesamtvarianz (total inertia) durch wenige Hauptkomponenten (Dimensionen) erklärt und die Eigenwerte der nachgeschalteten Achsen tendieren gegen null. In einer typischen zönotischen Studie (community study) summieren die ersten drei Eigenwerte etwa 40% der Gesamtvarianz einer Streuungsmatrix. Bei sehr umfangreichen Datensätzen mit einem hohen Anteil an strukturloser Variation reduziert sich der Wert auf etwa 5%. Dennoch kann solch ein Ergebnis ökologisch informativ sein (vgl. Gauch 1982a).

Da **Kovarianz- und Korrelationsmatrix** eine große Rolle spielen für das Verständnis der unterschiedlichen PCA-Varianten, wird deren Sinngehalt und deren Zusammenhang diskutiert. Die Kovarianz  $\sigma_{jk}$  ist eine Erweiterung des Varianzkonzepts auf zwei Deskriptoren. Sie misst die gemeinsame Streuung der beiden Variablen um ihr Mittel. Die Kovarianzmatrix  $\Sigma$  der biotischen Daten in Matrix  $Y$  beinhaltet die Varianzen und Kovarianzen aller Deskriptoren  $p$  (z.B. Arten) und ist im Allgemeinen eine Assoziationsmatrix. Sie kann positive und negative Werte annehmen und hat keine obere und untere Grenze. Die Kovarianz stellt Information über die Orientierung einer Datenwolke im Deskriptorenraum zur Verfügung, trifft aber keine Aussage über die Intensität des Zusammenhangs zwischen zwei Variablen. Dennoch kann die Kovarianz zu- bzw. abnehmen mit veränderter Korrelation zwischen zwei Deskriptoren (vgl. Legendre & Legendre 1998; Sokal & Rohlf 1995, 1996; Zar 1996). Die Relation zwischen der Kovarianzmatrix  $\Sigma$  und der Korrelationsmatrix  $R$  ist gegeben durch:

**Gleichung 3.5-8:** 
$$\Sigma = D(\sigma) R D(\sigma)$$

mit:  $D(\sigma)$ : Matrix der Standardabweichungen von  $Y$   
 $R$ : Korrelationsmatrix

Bei der nachfolgenden mathematischen Ableitung des Verfahrens muss darauf hingewiesen werden, dass es zahlreiche Algorithmen für die Berechnung einer PCA gibt. Zur Vertiefung der mathematischen Eigenschaften faktorenanalytischer Techniken kann auf die umfangreiche Literatur zurückgegriffen werden (z.B. Backhaus 1990, 1996; Causton 1988; Gauch 1982a; Jongman et al. 1987; Legendre & Legendre 1998; Økland 1990; Orłóci 1978; Pielou 1977, ter Braak 1998).

#### **Berechnungsschritte einer Hauptkomponentenanalyse basierend auf der Streuungsmatrix**

**Eigenwerte und Eigenvektoren:** Die PCA ist eine eigenvektorbasierte Methode, die auch als Eigenanalyse bezeichnet werden kann. Wie bereits erwähnt, verläuft die erste PCA-Achse in Richtung der maximalen Variabilität durch das  $p$ -dimensionale Datenellipsoid einer

Streuungsmatrix. Die Extraktion der Hauptachsen erfolgt durch Lösung der Gleichung 3.5-9, die genutzt wird, um die Eigenwerte zu berechnen (vgl. auch Gleichungen 3.5-1 und 3.5-2, S. 61).

**Gleichung 3.5-9:**  $(S - \lambda_k I) u_k = 0$

mit: **S:** Streuungsmatrix       $\lambda_k$ : Eigenwert  
**I:** Identitätsmatrix       $u_k$ : Eigenvektor

Die Eigenvektoren können durch sukzessives Einsetzen ( $\lambda_1 > \lambda_2 > \dots \lambda_p \geq 0$ ) der Eigenwerte in Gleichung 3.5-9 bestimmt werden. Die resultierenden Werte füllen die Spalten in der korrespondierenden Matrix U. Die Elemente der Matrix entsprechen dem Richtungskosinus der Winkel zwischen den Originaldeskriptoren und den Hauptachsen, d.h. anhand der Matrix U kann die Rotation des Achsensystems (Winkel  $\alpha$ ) durch den arc cos von  $u_1$  berechnet werden (vgl. Abbildung 3.5-7, Teilbild D, S. 71).

Die Gesamtvarianz (total inertia) in der Varianz-Kovarianzmatrix S (Multiplikation der zentrierten Matrix Y mit ihrer transponierten Matrix dividiert durch n-1 Objekte) berechnet sich aus der Summe der Hauptdiagonalen-Elemente. Bezogen auf die Matrix S beschreibt der erste Eigenwert gleich  $100 \cdot \lambda_1 / \sum \lambda_k$  der Gesamtvarianz.

**Hauptkomponenten:** Die Elemente der Eigenvektoren, welche die varianzmaximierten Achsen der Streuungsmatrix repräsentieren, sind als Gewichte oder Ladungen der Deskriptoren bezogen auf die Linearkombination der Deskriptoren zu bezeichnen, von denen die Hauptkomponenten berechnet werden. Die Positionen der Objekte im neuen PCA-Achsensystem werden durch die Hauptkomponenten bestimmt. Die Position von Objekt  $x_i$  entlang der 1. Achse ist definiert durch folgende Linearkombination der Deskriptoren:

**Gleichung 3.5-10:**  $f_{i1} = (y_{i1} - \bar{y}_1)u_{11} + \dots + (y_{ip} - \bar{y}_p)u_{p1} = [y - \bar{y}]_i u_1$

mit:

$u_{j1}$ : **Ladungen (Gewichte) der Deskriptoren auf den 1. Eigenvektor/die erste PCA-Achse**

$[y - \bar{y}]$ : **Matrix der zentrierten Beobachtungen**

$(y_{ij} - \bar{y}_j)$ : **Werte der Koordinaten des Objekts i gegenüber den zentrierten Deskriptoren j**

Die Matrix der zentrierten Beobachtungen multipliziert mit der Matrix der Eigenvektoren ergibt die Matrix der transformierten Daten F (matrix of component scores, Gleichung 3.5-11), welche die Koordinaten der Punktsignaturen aller Objekte im Ordinationsraum beinhaltet. Man spricht hier von einer PCA auf der zentrierten Matrix (Varianz-Kovarianzmatrix).

**Gleichung 3.5-11:**  $F = [y - \bar{y}] U$

Wird anstatt an der Matrix der zentrierten Beobachtungen die Multiplikation von U an der Originaldatenmatrix durchgeführt, dann handelt es sich um den Spezialfall nicht zentrierte PCA (non-centered).

**Gleichung 3.5-12**  $F = Y U$

Die PCA ist also eine reine Rotation der Achsensysteme im multidimensionalen Raum, wobei die Originaldeskriptorachsen auch nach der Analyse (Rotation) noch im 90° Winkel zueinander stehen. Die euklidischen Distanzen der Objekte werden in beiden Fällen im rotierten Achsensystem beibehalten (Jongman et al. 1987; Legendre & Legendre 1998).

**Die Matrix der Eigenvektoren:** Die häufige Methode der Normalisierung von Eigenvektoren, wobei die Länge der Zeilen- und Spaltenvektoren gleich eins gesetzt wird (Matrix U), kann in der PCA durch eine zweite Anwendung ersetzt werden. Dabei wird die Eigenvektormatrix U mit der Eigenwertmatrix  $\Lambda^{1/2}$  multipliziert und man erhält die neue Matrix  $U\Lambda^{1/2}$ . Die Länge

jedes Eigenvektors  $u_i$  wird so zur Quadratwurzel seines Eigenwertes ( $\sqrt{\lambda_i}$ ), also seiner Standardabweichung skaliert, da der Eigenwert gleich der Varianz ist. Das Ergebnis dieser Skalierung bedeutet, dass der Kosinus der Winkel zwischen den Deskriptorachsen proportional zu deren Kovarianz ist. Ein Winkel von  $90^\circ$  ist gleichbedeutend mit einer Kovarianz von null bzw. mit der Orthogonalität der Achsen. Wird die beschriebene Skalierung angewendet, ist die euklidische Distanz zwischen den Objekten nicht länger gegeben (Jongman et al. 1987; Legendre & Legendre 1998).

#### Die Hauptkomponentenanalyse basierend auf der Korrelationsmatrix

Die Hauptkomponentenanalyse kann nicht nur basierend auf einer Streuungsmatrix  $S$ , sondern auch auf der Korrelationsmatrix  $R$  ausgeführt werden, da die Korrelationen gleichbedeutend mit den Kovarianzen der standardisierten Deskriptoren sind. Die Werte der Hauptdiagonalen in der Matrix  $R$  sind alle gleich eins, daher ist die Gesamtvarianz der Matrix identisch mit der Summe der Deskriptoren  $p$  in der Analyse. Hauptkomponenten berechnet auf der Korrelationsmatrix unterscheiden sich von den Hauptkomponenten bezogen auf eine Varianz-Kovarianzmatrix. Da die Summe der Eigenwerte für Matrix  $S$  gleich der Summe der Varianzen  $s^2$  und die Summe der Eigenwerte für Matrix  $R$  gleich der Summe der Deskriptoren  $p$  ist, müssen die Eigenvektoren zwangsläufig differieren. Die Distanzen zwischen Objekten sind in den beiden Fällen nicht die gleichen. Im Falle der Korrelationsmatrix sind die Deskriptoren standardisiert und daher Einheiten unabhängig, während die Originaldeskriptoren auf verschiedenen Skalen basieren.

Die Berechnung der Punktsignaturkoordinaten aller Objekte im Ordinationsraum geht zurück auf die Matrix  $U$  der Korrelationsmatrix und der Matrix der standardisierten Beobachtungen (Gleichung 3.5-13, vgl. auch Gleichung 3.5-11, S. 73).

Gleichung 3.5-13:

$$F = \begin{bmatrix} y - \bar{y} \\ s_y \end{bmatrix} U$$

Die PCA ist nach wie vor eine reine Rotation der Achsensysteme. Dennoch sind durch die Standardisierung der Deskriptoren die Objekte anders positioniert (ter Braak 1998).

Weil die Standardabweichungen und Varianzen aller Deskriptoren nach der Standardisierung gleich eins werden, sind die Elemente der Matrix  $U\Lambda^{1/2}$  Korrelationskoeffizienten zwischen den Deskriptoren  $j$  und den Komponenten  $k$  (Legendre & Legendre 1998).

#### Hauptkomponentenanalyse in CANOCO®

Zum Ende dieses Abschnitts soll dem Anwender der Schritt von der Theorie zur Praxis, d.h. der Berechnung verschiedener Hauptkomponentenanalysen mit CANOCO®, erleichtert werden, indem die Einstellungen der Software mit den Ausführungen abgeglichen werden (vgl. hierzu auch ter Braak 1998). Zur weiteren Unterscheidung zwischen Distanz- und Korrelationsbiplot siehe auch Abschnitt 3.6.2.

##### Centering and Standardization:

Die Einteilung der Datenmanipulation der Originaldatenmatrix  $Y$  erfolgt nach "Samples" (= Zeilen) und "Species" (= Spalten). Grundsätzlich sind hier die Arten-Optionen anzuwenden, da in den Spalten der Originalmatrix die Deskriptoren stehen und diese manipuliert werden sollten. Die Anwendung der Optionen auf die Zeilen führt zu fraglichen Berechnungen (vgl. Abschnitt 4).

Non-centered: Es erfolgt keine Standardisierung, d.h., Gleichung 3.5-12 (S. 73) wird basierend auf  $Y$  berechnet. Der durch die Regression der Art  $i$  zur latenten Variable ermittelte Vektor hat den Ursprung in einem hypothetischen Plot mit der Abundanz null für alle Arten -> PCA auf der Originaldatenmatrix.

Center by species: Das Mittel der jeweiligen Spalte wird von den Daten abgezogen, so dass die resultierende Datenmatrix Spaltenmittel des Wertes null besitzt. Gleichung 3.5-11 (S. 73)

wird auf Matrix  $[y - \bar{y}]$  berechnet. Somit ist der Ursprung aller Deskriptorachsen zum Zentroid verschoben, d.h. zur mittleren Artmächtigkeit aller Arten -> PCA auf der Varianz-Kovarianzmatrix.

Center and standardize: Standardisierung der Erklärungsvariablen durch Subtrahieren des Mittels und die Division durch die Standardabweichung. Gleichung 3.5-13 (S. 74) wird auf Matrix  $[y - \bar{y}] / s_y$  berechnet. Dadurch wird ebenfalls der Ursprung in das Zentroid gelegt, d.h., die Position ist identisch zur zentrierten PCA. Zusätzlich wird die Länge aller Deskriptorachsen vereinheitlicht. Der Nachteil kann sein, falls die Deskriptoren Arten sind, dass die Variation aller Arten eins gesetzt wird, egal ob z.B. 1-10000 oder 100-102 Individuen gezählt wurden. Die Summe der Eigenwerte ist durch die Auswahl der Option gleich der Anzahl der Deskriptoren, da diese jeweils die Varianz eins besitzen -> PCA auf der Korrelationsmatrix.

#### Scaling:

Inter-sample distances: Die Skalierung bewirkt die Projektion der PCA-Ergebnisse als Distanzbiplot basierend auf den Matrizen U und F. So können die Distanzen der Objekte zueinander interpretiert werden, da die euklidischen Distanzen des multidimensionalen Raumes bestmöglich bewahrt werden. Siehe auch Eigenschaften des Distanzbiplots im Abschnitt 3.6.2.

Inter-species correlations: Die Skalierung bewirkt die Projektion der PCA-Ergebnisse als Korrelationsbiplot basierend auf den Matrizen  $U\Lambda^{1/2}$  und G. So können die Winkel der Deskriptorachsen zueinander interpretiert werden, da deren Kosinus eine Approximation der Korrelation der Deskriptoren miteinander ist. Siehe auch Eigenschaften des Korrelationsbiplots im Abschnitt 3.6.2.

Divide by standard deviation (PCAcentered oder PCAnon-centered): Die Option hat keinen Effekt, wenn die Deskriptoren bereits standardisiert sind (PCA auf der Korrelationsmatrix; centered and standardized). Die Längen der Deskriptorachsen werden bei dieser Option zueinander relativiert. Da die Deskriptorachsen im Korrelationsbiplot eine Länge von  $s_j$  im voll-dimensionalen Raum besitzen, kann anhand der Länge der Projektion des Deskriptors im reduzierten Raum seine Standardabweichung abgeschätzt werden. Um den Effekt zu dämpfen und die Deskriptoren besser vergleichbar zu machen, kann die Division mit der Standardabweichung durchgeführt werden.

Ist die Skalierung "Inter-species correlations" eingestellt, kann die Länge der Vektorpfeile der Arten als Korrelation (Fit) mit dem Ordinationsdiagramm angesehen werden.

Do not post-transform (PCAcentered oder PCAnon-centered): Die Länge der Deskriptorachsen spiegelt bei dieser Option die Standardabweichung der Arten wider, d.h., Arten mit einer großen Varianz (meist dominante Arten) liegen weit in der Peripherie des Diagramms und dominieren dieses.

### 3.5.5 Korrespondenzanalyse

**"AS A METHOD, CORRESPONDENCE ANALYSIS IS VERY IMPORTANT BECAUSE IT PROVIDES THE BASIS FOR MOST DEVELOPMENTS IN ORDINATION TECHNIQUES SINCE 1970 (KENT & COKER 1992, S. 215)."**

Die Korrespondenzanalyse (CA: correspondence analysis) geht auf Hirschfeld 1935 und Fisher 1940 zurück. Sie wurde in verschiedenen Ländern weiterentwickelt und ist ebenfalls unter den englischen Begriffen: "Contingency Table Analysis", "RQ-technique", "Reciprocal Averaging", "Reciprocal Ordering", "Dual Scaling", "Homogeneity Analysis" oder im Französischen unter den Term "analyse factorielle des correspondances" (COA) bekannt. Das Verfahren führte Hill zum ersten Mal 1973 in die Ökologie ein.



Die Korrespondenzanalyse ist ein indirektes Ordinationsverfahren (vgl. PCA) zur Analyse einfacher Zwei-Wege-Kontingenztafeln, in denen die Zustände der Objekte mit denen der Deskriptoren einer Matrix verglichen werden (s.a. Abbildung 3.5-5, S. 60). Der primäre Zweck besteht in der vereinfachten niedrigdimensionalen Darstellung der Information aus der Datenmatrix (vgl. auch Abschnitt 4.1). Die Daten in den Zellen der Matrix sind in der Ökologie am gebräuchlichsten Spezies-Daten erhoben auf verschiedenen Probestellen. Das Verfahren wird auf dimensional homogene Daten (gleiche Einheiten) angewendet. Es ist konzeptionell angelehnt an das "Weighted-Averaging-Verfahren", welches in Abschnitt 2.3 eingeführt wurde, stellt sich mathematisch aber als Eigenanalyse – wie die PCA – dar (Gauch 1982a). Trotz der Ähnlichkeit zur PCA, nutzt die CA aber ein unimodales Antwortmodell. Je nach Skalierung wird die  $\chi^2$ - oder die generalisierte Mahalanobis Distanz als Assoziationsmaß verwendet (vgl. Abschnitt 3.5.2).

Gleich dem Schema des vorangegangenen Abschnitts wird die **Arbeitsweise der Korrespondenzanalyse** zuerst graphisch-deskriptiv und danach mathematisch abgeleitet.

Graphisch kann die Korrespondenzanalyse ebenso wie die Hauptkomponentenanalyse als Achsenextraktionsmethode angesehen werden, so dass die prinzipiellen Arbeitsschritte auch anhand von Abbildung 3.5-7 (S. 71) nachvollzogen werden können. Auch hier verlaufen die Achsen durch den Datenschwarm in Richtung maximaler Varianz und es wird die Projektion in einem niedrig-dimensionalen Raum durch Informationssummutation angestrebt.

Beschreibt man die Korrespondenzanalyse in geometrischen Ansätzen, sind drei fundamentale Größen nötig, um die Eigenanalysetechnik zu spezifizieren: das Assoziationsmaß, die den Ordinationswerten zugeschriebenen Gewichte und die Position des Ursprungs für die Projektion (Gauch 1982a; Hill 1974; Orłóci 1978).

Als Assoziationsmaß zur Quantifizierung der Beziehungen zwischen den Zeilen- und den Spalteninhalten wird die  $\chi^2$ -Distanz herangezogen, welche das DoppelnulLEN-Problem vermeidet (vgl. Abschnitt 3.5.2). Den Objekten wird ihr Zeilentotal und den Deskriptoren ihr Spaltentotal als Gewicht zugeordnet. Der Ursprung liegt im Schwerpunkt der Daten (Gauch 1982a). In der PCA werden im reduzierten Ordinationsraum die euklidischen Distanzen zwischen den Profilen der gewichteten konditionalen Wahrscheinlichkeiten bewahrt, während die CA äquivalent die  $\chi^2$ -Distanzen zwischen den Zeilen und Spalten der Kontingenztafel beibehält und repräsentiert (Legendre & Legendre 1998).

Zum Verständnis der Arbeitsweise eignet sich die Prinzipskizze in Tabelle 3.5-3:

**Tabelle 3.5-3: CA-Prozedur der Matrixneuordnung um die Diagonale (A, B, ... : Arten und 1, 2, ... : Probestellen).**

Rohe, ungeordnete Datenmatrix						Neu arrangierte Datenmatrix aus der CA					
	A	B	C	D	E		E	A	C	B	D
1	5	0	0	0	10	1	10	5	0	0	0
2	0	10	5	5	0	5	5	10	5	0	0
3	5	5	10	0	0	3	0	5	10	5	0
4	0	5	0	10	0	2	0	0	5	10	5
5	10	0	5	0	5	4	0	0	0	5	10

Ausgehend von einer Matrix, in der die Objekte und Deskriptoren zufällig angeordnet sind (Tabelle 3.5-3, links), erarbeitet der CA-Algorithmus eine spezifische Ordnung. Die Ordinationswerte der ersten Achse maximieren als Folge die Korrelation der Objekte und Deskriptoren, d.h., wenn die Objekte oder Deskriptoren einer Datenmatrix in der Ordnung ihrer Ordinationswerte umgeordnet werden, resultiert dies in einer Matrixstruktur, in der die höchsten Werte der Zeilen- und Spaltenprofile um die Hauptdiagonale konzentriert sind (Tabelle 3.5-3, rechts). Die anfangs nicht ersichtliche Struktur in den Daten – speziell bei umfangreichen Matrizen – wird durch die Umordnung/Ordination deutlich. Letztlich sind ähnliche Objekte in

den Zeilen sowie Arten in den Spalten zusammen und entsprechend unähnliche auseinandergerückt. Dadurch das bestimmte Probenflächen bestimmte Arten enthalten und umgekehrt, erklärt sich das Prinzip der Korrelationsmaximierung. Eine Datenmatrix ist ein begrenzt dimensionales Medium zur Darstellung von Datenstrukturen, da die Objekte und Deskriptoren nur in einer einfachen Sequenz umgeordnet werden können. Beinhaltet ein Datensatz Strukturmerkmale mehrerer Gradienten, kann die neu arrangierte Matrix dies nicht mehr verdeutlichen und es müssen effektivere Darstellungsmöglichkeiten gefunden werden. Die multidimensionalen Datenstrukturen lassen sich leicht in einem Diagramm präsentieren, für das weitere Matrixumordnungen nach dem grundlegenden Prinzip die Ordinationswerte der nächsthöheren Dimensionen liefern. Grundsätzlich gilt: Sind biotische Daten durch einen Gradienten klar strukturiert, wird dieser durch das Matrixrearrangement abgebildet (Gauch 1982a; Jongman et al. 1987).

Zur vertiefenden Studie des Verfahrens können weiterhin die Texte von Causton (1988), Clausen (1998), Gauch (1982), Gower (1987), Hill (1974), Hirschfeld (1935), Kent & Coker (1992), van der Maarel (1969), Mercier (1992), Økland (1990), Orlóci (1978), Palmer (1998), Pielou (1984), ter Braak (1987a, 1998) und Whittaker (1973) empfohlen werden.

#### Berechnungsschritte einer Korrespondenzanalyse

Zur Berechnung der CA wird die Kontingenztafel umgeformt in eine Matrix mit den Beiträgen zur  $\chi^2$ -Statistik nach Pearson. Die berechnete Matrix  $\bar{Q}$ , auf der die Korrespondenzanalyse basiert, enthält die relativen Häufigkeiten der Kontingenztafel.

Gleichung 3.5-14:

$$\bar{Q} = [\bar{q}_{ij}] = \left[ \frac{p_{ij} - p_{i+}p_{+j}}{\sqrt{p_{i+}p_{+j}}} \right]$$

mit:  $p_{ij}$ : Relative Häufigkeit (absolute Häufigkeit  $f_{ij}$  der zugehörigen Zelle dividiert durch das Matrixtotal  $f_{++}$ )  
 $p_{i+}$ : Zeilengewichte ( $f_{i+}/f_{++}$ )  
 $p_{+j}$ : Spaltengewichte ( $f_{+j}/f_{++}$ )

Die Werte  $\bar{q}_{ij}$  differieren von den Werten  $\chi_{ij}$  (standardisierte Residuen der absoluten Häufigkeiten) nur durch eine Konstante. Alle Eigenwerte sind aber bei der Analyse der Matrix  $\bar{Q}$  kleiner oder gleich eins. Die Trägheit der Matrix (total inertia) ist definiert als Wert für die Pearsonsche  $\chi^2$ -Statistik (Gesamtvarianz) dividiert durch die Gesamtsumme aller Einträge ( $\chi_p^2/f_{++}$ ) (ter Braak 1998). Anders ausgedrückt, ist die Quadratsumme aller Werte in Matrix  $\bar{Q}$  das Maß für die Gesamtvarianz der Matrix und identisch zu der Summe aller Eigenwerte, die durch die Eigenanalyse der Matrix  $\bar{Q}$  bestimmt werden können.

Im nachfolgenden Analyseschritt kann eine Singulärwertzerlegung an der Matrix  $\bar{Q}$  (vgl. Press et al. 1986) oder alternativ eine Eigenanalyse an der Kovarianzmatrix  $\bar{Q}'\bar{Q}$  (führt zu U) bzw.  $\bar{Q}\bar{Q}'$  (führt zu  $\hat{U}$ ) ausgeführt werden (Legendre & Legendre 1998). Die Berechnung der Eigenanalyse muss nur einmal durchgeführt werden um die Matrizen U und  $\hat{U}$  zu erhalten (vgl. Gleichung 3.5-15).

Gleichung 3.5-15:

$$\hat{U} = \bar{Q} U \Lambda^{-1/2} \quad \text{und} \quad U = \bar{Q}' \hat{U} \Lambda^{-1/2}$$

mit:  $\hat{U}$ : Eigenvektormatrix von  $\bar{Q}\bar{Q}'$   
 U: Eigenvektormatrix von  $\bar{Q}'\bar{Q}$

#### Korrespondenzanalyse in CANOCO®

Um die besprochenen CA-Varianten zu rechnen, werden die entsprechenden Softwareoptionen kurz angesprochen (vgl. auch ter Braak 1998). Zur weiteren Unterscheidung zwischen Distanz- und Korrelationsbiplot siehe auch Abschnitt 3.6.2.

Focus:

Scaling on Inter-sample distances: Die Skalierung entspricht dem Skalierungstyp 1. Die Probeflächen werden im gewichteten Schwerpunkt der sie definierenden Deskriptoren (Arten) abgebildet und die Arten liegen gestreut um den Ordinationspunkt der Probefläche. Der Eigenwert entspricht der Varianz der Zeilen (Probeflächen) entlang der Achse.

Scaling on Inter-species distances: Die Option ist identisch mit Skalierungstyp 2. Die Arten befinden sich im Schwerpunkt der Probeflächen, in denen sie auftreten. Die Probeflächen selbst verteilen sich um den Ordinationspunkt der Art. Der Eigenwert entspricht der Varianz der Spalten (Arten) entlang der Achse.

Scaling:

Biplot scaling (Scaling +): Als Assoziationsmaß wird die  $\chi^2$ -Distanz gewählt. Die Option führt zu einer mehr quantitativen Interpretationsmöglichkeit durch die Anwendung der Biplot-Regel (vgl. Abschnitt 3.6.3) und ist bei kürzeren Gradienten mit der Tendenz einer häufigeren linearen Antwort zur Achse angebracht.

Hill's scaling (Scaling -): Als Assoziationsmaß wird die generalisierte Mahalanobis Distanz gewählt. Die Option lässt bei langen Gradienten (vgl. Abschnitt 3.5.6) die Anwendung der Distanz-Regel zu (vgl. Abschnitt 3.6.3) und das Zentroid-Prinzip wird insofern erweitert, dass die Ordinationspunkte der Arten als Optimum der Realnische einer Art gegenüber dem Gradienten angesehen werden können. Sie ist geeignet für lange Gradienten (> 3-4 Standardabweichungen Länge) mit deutlich überwiegenden unimodalen Antworten. Probeflächen, die mehr als vier Standardabweichungen voneinander entfernt liegen, haben dann für gewöhnlich keine Arten mehr gemeinsam. Dies ist in Kombination mit "Scaling on Inter-sample distances" die optimale Skalierung, um den Artenwechsel entlang von Achsen zu betrachten (Beta-Diversität).

### 3.5.6 Korrespondenzanalyse mit Detrendingprozess

*"IN A FEW YEARS, THE METHOD BECAME THE STATE-OF-THE-ART METHOD FOR ORDINATION, STATED NOT TO BE FAR FROM THE THEORETICAL OPTIMUM FOR ORDINATION METHODS (ØKLAND 1990, S. 149)."*

Die Korrespondenzanalyse mit Detrendingprozess (DCA; detrended correspondence analysis) wurde von Hill & Gauch 1980 in Abwandlung der CA als Ad-hoc-Methode entwickelt, um den zwei Problemen des Basisverfahrens entgegenzuwirken (Jongman et al. 1987):

- (1) Bogenförmige Verzerrungen ab der zweiten Achse und
- (2) Stauchung der Achsenenden verglichen zum Achsenzentrum.

Die DCA liefert nach Gauch (1982) Ergebnisse, die meist qualitativ gleichwertig oder besser sind als die anderer indirekter Ordinationstechniken, da sie darauf ausgerichtet wurde, die Probleme der CA und der PCA zu vermeiden. Speziell in den 80er Jahren wurde sie weit verbreitet eingesetzt (Kent & Coker 1992; Økland 1990) und man postulierte, dass die DCA nicht weit entfernt ist vom Optimum einer indirekten Ordinationsmethode (Gauch 1982a, b). Dennoch blieb das Verfahren nicht ohne Kritik (Dargie 1986; Knox 1989; Minchin 1987; Oksanen 1988; Wartenberg et al. 1987) wie Abschnitt 4.1 belegen wird.

Der Ursprung der Probleme in CA und PCA ist zu erläutern, um die Arbeitsweise der Korrekturverfahren nachvollziehbar zu machen. Ausgehend vom Nischenkonzept und der Antwortmodelldiskussion ist der Effekt eines Gradienten auf die Distanzbeziehungen zwischen den Probeflächen, berechnet auf den Art-Daten, meist nicht linearer Natur (vgl. Abschnitt 2.3). Ordinationsmethoden versuchen, das nichtlineare Phänomen in einer meist zweidimensionalen graphischen Projektion wiederzugeben. Dies führt zu kurvenförmigen Verzerrungen (Gauch 1982a; Hill & Gauch 1980; Jongman et al. 1987 u.v.a.), deren Form vom Distanzmaß

des Verfahrens abhängt (vgl. Abschnitt 3.5.2). Der Detrending-Prozess – ausgeführt an den Ordinationsachsen der CA, da sie grundsätzlich durch die  $\chi^2$ -Distanz dem unimodalen Modellgedanken näher ist – versucht diese Verzerrungen zu korrigieren.

### Bogenförmige Verzerrungen der PCA und CA

Wendet man eine PCA an, um die sich nicht monoton verändernden Zusammenhänge in Abbildung 3.5-8 zu ordinieren, führt dies dazu, dass die Werte der euklidischen Distanz beim Vergleich der ersten Probe mit allen weiteren anfangs zunehmen, dann ab, um wieder zuzunehmen und so fort.

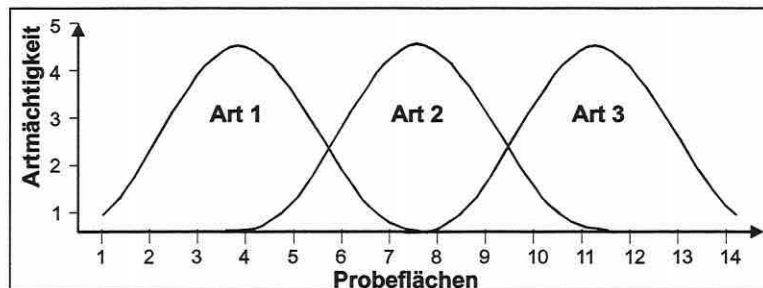


Abbildung 3.5-8: Theoretische Antwort dreier Arten gegenüber einem Probeflächen-Gradienten.

In den weiteren Zeilen der Distanzmatrix ergibt sich ein ähnlich komplexes Muster, welches die PCA in maximal drei Dimensionen darzustellen hat, da sie nicht mehr Achsen extrahieren kann als Deskriptoren in der Originaldatenmatrix vorhanden sind (vgl. Abschnitt 3.5.4).

Im Ordinationsdiagramm sind die Enden des Gradienten entlang der ersten Achse nach innen umgebogen und somit die unterschiedlichsten Flächen näher beieinander abgebildet als ähnliche. Diese durch die Eigenschaften der euklidischen Distanz verursachte Form wird als Hufeiseneffekt (horseshoe) in der Fachliteratur angesprochen.

In der CA ist zumindest das Einfalten der Enden eines Gradienten durch das  $\chi^2$ -Distanzmaß nicht gegeben. Die dennoch mögliche bogenförmige Verzerrung wird auch als Bogeneffekt beschrieben. Neben dem Begriff "arch" finden auch Termini wie "bow" und "Guttman effect" Verwendung in der englischen Literatur.

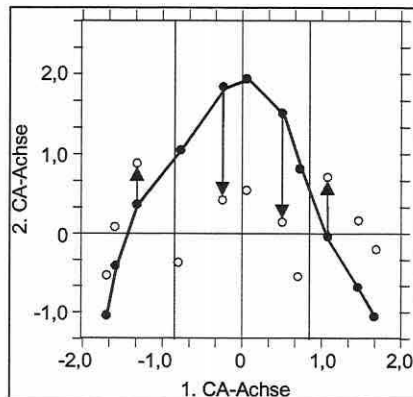
Das mathematische Artefakt entsteht in der Korrespondenzanalyse, wenn die erste Achse einen sehr hohen Erklärungsanteil gegenüber der nächst höheren Dimension hat. Der quadrierte Eigenwert der ersten Achse entspricht dem Polynom zweiten Grades dergleichen (Peet et al. 1988), d.h., da die Achsen in der Ordnung fallender Eigenwerte extrahiert werden, kommt dieser Fehler zustande, wenn die reale zweite Achse einen geringeren Eigenwert aufweist als die quadratische Funktion der ersten (Gauch 1982a; Hill & Gauch 1980). In anderen Worten: Wenn die zweite Achse kürzer als die Hälfte der ersten ist, führt dies zu einem höheren Eigenwert der quadratischen Funktion der vorangegangenen Achse und somit zum Bogeneffekt (Økland 1990). Das Erscheinen einer Achse, die ein Polynom einer vorangegangenen darstellt, ist trotz intensiver Forschung noch immer nicht im Detail verstanden (ter Braak & Prentice 1988).

Wenn die Eigenwerte höherer Achsen allmählich abnehmen, findet sich eine dritte Variante von Achsen ein: Gradienten, welche nur auf dem stochastischen Teil des Modells bzw. dem Datenrauschen beruhen. Diese können gelegentlich höhere Zufallseigenwerte besitzen als reale Achsen, falls die Datenmatrix nur sehr geringe Zusammenhangsstrukturen beinhaltet (Gauch 1982a). Die Schlussfolgerung muss sein, dass es wenig zuverlässig ist, Achsen höher als die vierte noch zu interpretieren (Økland 1990). Daher ist in CANOCO® per Voreinstellung die Extraktion von nicht mehr als vier Achsen für Ordinationsmethoden vorgesehen.

Die **Korrekturverfahren der CA**, welche den Detrending-Prozess ausmachen, sind von zweierlei Ausprägung: (1) "Detrending-by-segments" und (2) "Detrending-by-polynomials" (Hill & Gauch 1980; Jongman et al. 1987; ter Braak 1998).

Das erste Verfahren bewirkt, dass innerhalb einer gebildeten Anzahl von Raumsegmenten entlang der ersten Achse, das Mittel der darin liegenden Ordinationspunkte jeweils null gesetzt wird, d.h., das Mittel wird exakt entlang der zweiten Achse auf die Position der ersten Achse verschoben. Entsprechend werden relativ dazu auch alle anderen Ordinationspunkte entlang der zweiten Achse bewegt (Abbildung 3.5-9).

Die Positionen der Probeflächensignaturen bezogen auf die zweite Detrending-Achse werden dadurch ökologisch bedeutungslos. Das Verfahren verschiebt weiterhin die Signaturen, welche in der Originalordination entlang der ersten Achse nahe beieinander liegen, willkürlich, wenn sie in unterschiedlichen Segmenten zu liegen kommen.



**Abbildung 3.5-9: Prinzip des "Detrending-by-segments" an willkürlich definierten Raumsegmenten. Gezeigt wird ein Bogeneffekt der gefüllten Signaturen, die nach der Korrektur in die Lage der ungefüllten Signaturen überführt werden.**

Die Distanzen sind daher nicht länger verlässlich. Durch verschiedene Segmentierungen entstehen differierende Ordinationsergebnisse entlang der Detrending-Achse (Achse 2). Die Anzahl der Segmente in CANOCO® liegt zwischen 10 und 46, wobei die Voreinstellung 26 ist (Legendre & Legendre 1998, ter Braak 1998). Für die dritte DCA-Achse wird die Korrektur unter Berücksichtigung der ersten und der zweiten Achse durchgeführt. Diese Vorgehensweise gilt auch für jede weitere extrahierte Achse, um Polynomfunktionen bereits berechneter Achsen zu unterdrücken (Økland 1990).

Das zweite Verfahren (Detrending-by-polynomials) ist eine Erweiterung des Orthogonalisierungsschrittes im CA-Algorithmus (Hill & Gauch 1980).

Es bewirkt, dass nicht nur die Achsen zueinander unkorreliert sind, sondern auch Orthogonalität zu den Polynomfunktionen – derer zweiter bis vierter Ordnung – herrscht (Økland 1990). "Detrending-by-polynomials" ist ein geeignetes Verfahren, wenn das angenommene Modell mit den Daten harmoniert. Die Diskontinuitäten des ersten Verfahrens entfallen. Umgekehrt wird aber das Problem der gestauchten Achsenenden (s.u.) nicht gelöst (Legendre & Legendre 1998). Die vertikalen Residuen zwischen dem gewählten Polynom und den Probeflächensignaturen werden berechnet und zur Projektion der Probeflächen entlang der zweiten Achse verwendet. Die Residuen der Regression verhelfen dem DCA-Diagramm zwar zu einer eleganten Formgebung, sind aber ebenfalls bezogen auf ihren Informationsgehalt bedeutungslos. Die Position der Punkte entlang der zweiten Achse ist ökologisch nicht relevant (Legendre & Legendre 1998). Knox (1989) zeigt, dass die "Detrending-by-segments"-Prozedur gegenüber dem "Detrending-by-polynomials" zu bevorzugen ist, da die Lösungen "realistischer" sind. Sehr detailliert hat sich auch Økland (1986a-c) mit dem Verfahren auseinandergesetzt.

Dem zweiten Problem der PCA und der CA, der **Kontraktion der Achsenenden** (edge effect) wird begegnet durch eine nichtlineare Reskalierung der ersten Achse nach dem Detrending (nur Detrending-by-segments). Um die Spannweiten der Antwortkurven anzupassen, wird die Achse in schmale Segmente unterteilt, und Abschnitte mit einer sehr geringen Innergruppen-Varianz werden expandiert, während solche mit einer großen kontrahiert werden.



Als attraktives "Abfallprodukt" des Reskalierens hat die Achsenskalierung die Eigenschaft, dass sie die durchschnittliche Spanne der Antwortkurven (Standardabweichung; SD) der Arten und damit den Artenwechsel in der Analyse wiedergibt (Gauch & Wentworth 1976; ter Braak 1998 u.a.). Da die Kurve über vier SD zu ihrem Optimum ansteigt und wieder gegen null abfällt, haben Probeflächen die weiter als vier SD im Diagramm auseinander liegen, gewöhnlich keine Arten mehr gemeinsam. Ein Wechsel der Artenkomposition um 50% (half-change) erfolgt über eine Abweichung von 1 bis 1,4 SD, je nach Grad des Datenrauschens (Gauch 1982a; Legendre & Legendre 1998). Man bedenke, dass diese Relativangaben von Konstanz sind, die eigentliche Spanne einer SD entlang der Achse aber schwankt, da sie durch die Arten und deren Varianz in der Analyse definiert wird. DCA-Achsenlängen (length of gradient, Abschnitt 3.6.1) für Probeflächen-Ordinationen sind innerhalb von 10 bis 20% bei der Schätzung von der natürlichen Gradientenlänge exakt (Hill & Gauch 1980). Eine Lücke im Gradienten, verursacht durch das Stichprobendesign, kann bis zu einer Länge von drei SD adäquat geschätzt werden und verursacht keine drastische Verzerrung der Information (Gauch 1982a). Die Länge der ersten DCA-Achse zeigt die Spanne des wichtigsten ökologischen Gradienten in der Analyse und wird herangezogen, um abzuschätzen, welches Ordinationsverfahren bzw. welches mathematische Modell für die weitere Analyse geeignet ist. Bei einer Länge kleiner drei SD (2 bis 3 SD nach Økland 1986a) erweist sich meist ein lineares Modell als vorteilhaft, zwischen 3 bis 4 SD können beide Modelltypen ähnlich gute Ergebnisse liefern und bei einer Länge größer vier SD ist ein unimodales Modell angeraten (Jongman et al. 1987; Økland 1986a; ter Braak 1998).

### 3.5.7 Redundanzanalyse

**"RDA IS (...) A MULTIPLE REGRESSION FOR ALL SPECIES SIMULTANEOUSLY (I.E. A MULTIVARIATE REGRESSION) WITH LINEAR CONSTRAINTS ON THE REGRESSION COEFFICIENTS (TER BRAAK 1998, S. 53)."**

Die Redundanzanalyse (RDA) wurde erstmals von Rao in 1964 beschrieben (Jongman et al. 1987; Legendre & Legendre 1998). Sie ist die kanonische Form der Hauptkomponentenanalyse und eine direkte Erweiterung der multiplen Regression zur Modellierung multivariater Antwortvariablen (vgl. Abbildung 3.5-4, Weg C, S. 59).

#### Arbeitsweise der Redundanzanalyse

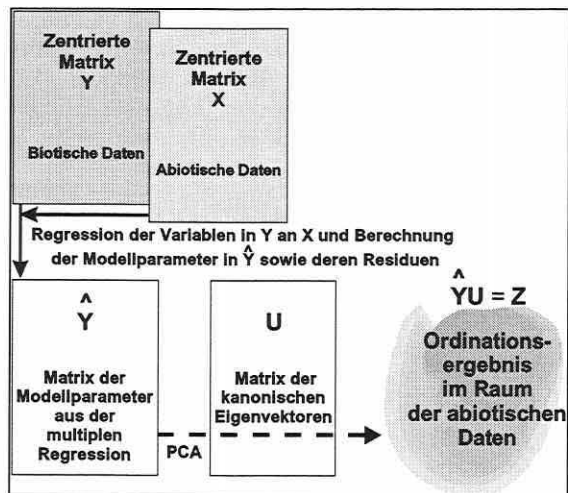
Kanonisch kann in dem Sinne verstanden werden, dass die Ordination der Matrix mit den biotischen Daten anhand der Matrix mit den abiotischen Daten erfolgt, d.h., die Ordination unterliegt der Einschränkung, dass die resultierenden Vektoren der Ordination Linearkombinationen der Variablen in X darstellen (Matrix Z, vgl. auch Abbildung 3.5-5, S. 60).

Abbildung 3.5-10 (S. 82) gibt einen Überblick zur Arbeitsweise der Redundanzanalyse. Matrix  $\hat{Y}$  umfasst die Werte der Variablen in Y nach der Regression an den Erklärungsvariablen in X, deren euklidische Distanzen unter den Objekten das Verfahren bewahrt.

Die RDA-Achsen sind den Hauptkomponenten der PCA zwar ähnlich, verlaufen aber in der Richtung maximaler Variabilität innerhalb der biotischen Datenwolke in Y, welche optimal zu der besterklärenden Linearkombination der Variablen in Matrix X für die jeweilige Dimension korrespondiert. Die Ordination unterliegt somit einer Einschränkung durch die vorgegebenen Variablen in der Matrix X bzw. deren Erklärungsanteile an der biotischen Variation in Y.

Die Gleichung der Eigenanalyse für eine Redundanzanalyse lautet:

**Gleichung 3.5-16:** 
$$(S_{YX}S_{XX}^{-1}S'_{YX}-\lambda_k I) u_k = 0$$



**Abbildung 3.5-10: Prozessschritte der Redundanzanalyse.** Nach der multiplen Regression zur Berechnung der Modellwerte  $\hat{Y}$  wird eine Hauptkomponentenanalyse an Matrix  $\hat{Y}$  ausgeführt um Matrix  $U$  zu erhalten. Als resultierende Ordinationen ergibt sich Matrix  $[Y \ U]$  im Raum der Erklärungsvariablen.

Regression an allen Variablen in Matrix  $X$  berechnet, um die angepassten Modellwerte  $\hat{y}$  (fitted values) zu erhalten. Aufgrund der Zentrierung aller Variablen in  $X$  und  $Y$  enthalten die Vektoren in der Matrix der Regressionskoeffizienten  $B$  keinen y-Achsenabschnitt. Die Vektoren in  $\hat{Y}$  sind ebenfalls zentriert. Ist die Zahl der Objekte  $m$ , gleich der Zahl der Variablen  $n$ , folgt eine quadratische Matrix  $X$ , welche die Variablen in  $Y$  zu 100% erklärt, so dass  $Y = \hat{Y}$  ist (vgl. Abschnitt 3.5.3).

Als nächster Schritt wird an der Matrix  $\hat{Y}$  eine Hauptkomponentenanalyse ausgeführt und das bereits aus der PCA bekannte Eigenwertproblem gelöst (vgl. Abschnitt 3.5.4). Es ergibt sich die normalisierte Matrix  $U$  mit den Eigenvektoren, welche die Beiträge der Deskriptoren in  $\hat{Y}$  zu den verschiedenen RDA-Achsen angeben.

Nach der Ordination sind die Positionskoordinaten der Objekte im Raum der Erklärungsvariablen (Matrix  $X$ ) in den Spalten der Matrix  $Z$  enthalten (vgl. Abbildung 3.5-10). Die durch das Modell angepassten Ordinationskoordinaten der Objekte werden dann als "sample scores" bezeichnet (sonst: site scores) und gehen auf die Linearkombinationen der Erklärungsvariablen zurück. Die Varianzen der Vektoren in Matrix  $Z$  entsprechen den Eigenwerten.

Die Korrelation  $r_k$  zwischen den Ordinationen der Objekte in den Räumen  $Y$  und  $X$  für eine Dimension  $k$  (Achse) wird in CANOCO® als Art-Umwelt-Korrelation (species-environment-correlation) bezeichnet. Sie ist ein Maß dafür, wie eng der Zusammenhang zwischen den beiden Datensätzen bezogen auf eine RDA-Achse ist. Eine hohe Art-Umwelt-Korrelation ist jedoch nicht gleichbedeutend mit einem großen Anteil erklärter Varianz in  $Y$ . Diese ist gegeben durch den Eigenwert der Achse (vgl. auch Abschnitt 3.5.8).

Letztlich ist noch von besonderer Bedeutung zu wissen, welche der Erklärungsvariablen hohe Beiträge zu der jeweiligen Dimension  $k$  leisten. Die Matrix der kanonischen Koeffizienten  $C$  ( $[BU]$ ) enthält die Gewichte (Ladungen) der Variablen in  $X$  zu den Achsen, d.h. die Beiträge, welche die Erklärungsvariablen zur Formierung der Matrix  $\hat{Y}$  leisten (Jongman et al. 1987; Legendre & Legendre 1998; ter Braak 1998).

Auch die RDA greift auf den Korrelations- und Distanzbiplot zurück (vgl. Abschnitt 3.6.2 und Abbildung 3.6-3, S. 100).

Beinhalten die Matrizen  $Y$  und  $X$  identische Daten, so reduziert sich die Gleichung 3.5-16 (S. 81) zur Gleichung 3.5-9 (S. 73), wobei  $S$  auch die Notation  $S_{YY}$  erhalten kann. Die RDA wird zur PCA.

Die möglichen Transformationen der Datenmatrix  $Y$  entsprechen denen im Abschnitt 3.5.4 beschrieben.

Die Matrix  $X$  wird unter der Annahme, dass sie Erklärungsvariablen unterschiedlicher Einheiten beinhaltet, in CANOCO® automatisch standardisiert (ter Braak 1998).

### Berechnungsschritte einer Redundanzanalyse

Die algebraische Ableitung einer Redundanzanalyse kann im Überblick an Abbildung 3.5-10 nachvollzogen werden und wird im Folgenden dargelegt. Als erstes wird für jede Antwortvariable in Matrix  $Y$  eine multiple lineare

### Redundanzanalyse in CANOCO®

Der in CANOCO® verwendete Algorithmus geht auf den "two-way weighted summation algorithm" der PCA zurück, wurde aber insofern modifiziert, dass nach jeder Iteration eine Regressionsanalyse ausgeführt wird (Jongman et al. 1987). Die Optionen zur Berechnung der besprochenen RDA-Varianten, sind identisch zu denen für die PCA und können in Abschnitt 3.5.4 nachgeschlagen werden.

### 3.5.8 Kanonische Korrespondenzanalyse

*"... CCA IS BEST DEFINED AS A METHOD OF DIRECT ORDINATION WITH THE RESULTING ORDINATION BEING THE PRODUCT OF THE VARIABILITY OF THE ENVIRONMENTAL DATA AS WELL AS THE VARIABILITY OF THE SPECIES DATA (KENT & COKER 1992, S. 229)."*

Das direkte Ordinationsverfahren wurde Mitte der 80er Jahre von ter Braak entwickelt (Jongman et al. 1987, ter Braak 1986, 1987a). Das Grundprinzip der kanonischen Korrespondenzanalyse (CCA: canonical correspondence analysis) ist gleich dem der CA (Abschnitt 3.5.5). Die Analyse stellt – wie die RDA zur PCA – die kanonische Form des Basisverfahrens dar und kombiniert die Funktionalität der CA mit der multiplen Regression an der Matrix der Erklärungsvariablen (ter Braak 1998; vgl. Abbildung 3.5-4, S. 59, und Abbildung 3.5-5, S. 60). Jede für eine CA geeignete Datenmatrix kann auch in der CCA als Antwortvariablenmatrix Anwendung finden.

Im Unterschied zur RDA (linear) beruht das Verfahren auf dem unimodalen Modellgedanken (vgl. Abschnitte 2.3 und 3.5.2). Es wird im Ordinationsdiagramm angestrebt, das Optimum der Nische einer Art zum Gradienten zu bestimmen, während die RDA die Richtung des steilsten Anstiegs eines Artvorkommens im Raum der Erklärungsvariablen durch einen Vektor beschreibt.

#### Arbeitsweise der kanonischen Korrespondenzanalyse

Grundsätzlich ist auch die CCA eine Achsenextraktionsmethode (vgl. Abschnitte 3.5.4 und 3.5.5 sowie Abbildung 3.5-7, S. 71), wobei die CCA-Achsen nach dem Prinzip der kleinsten Abweichungsquadrate etabliert werden, d.h. durch die multiple Regression der Antwortvariablenmatrix an der Erklärungsvariablenmatrix. Es wird die Linearkombination an Variablen mit dem höchsten Erklärungsanteil für die Dimensionen in Matrix Y aus Matrix X selektiert (vgl. Abbildung 3.5-10, S. 82 und Abbildung 3.5-5, S. 60). Dies führt zu dem Ergebnis, dass die Achsen in Richtung der maximal durch die Parameter in X zu erklärende Variabilität im Datensatz der Matrix Y verlaufen (Jongman et al. 1987; Kent & Coker 1992). Alle Dimensionen werden unter dem Zwang extrahiert, unkorreliert mit den vorangegangenen zu sein (Jongman et al. 1987). Mit anderen Worten, die CCA ist eine auf den Erklärungsanteil der Matrix X eingeschränkte Korrespondenzanalyse (constrained ordination).

In der CCA ist das Varianzkonzept gekoppelt mit dem  $\chi^2$ -Distanzmaß (vgl. Abschnitt 3.5.2). Die Technik zielt letztlich auf die Varianzzerlegung der  $\chi^2$ -Statistik (Clausen 1998). Die Eigenwerte geben dabei die relative Bedeutung bzw. den Erklärungsanteil der Achsen/Dimensionen an der Gesamtvariabilität im Datensatz an. Sie sind gewöhnlich niedriger als in einer CA durch die Restriktion, nur den Anteil an Variabilität in Matrix Y erklären zu können, welcher durch die Regression an der Matrix X gegeben ist (Jongman et al. 1987).

Die Art-Umwelt-Korrelation (vgl. Abschnitt 3.5.7) ist die multiple Korrelation zwischen den Positionskoordinaten der Probeflächen im Raum der Antwortvariablen (weighted averages of the species scores) und den Positionskoordinaten der Probeflächen im Raum der Erklärungsvariablen (linear combination of the environmental variables). Sie gibt Auskunft über die

Anpassung der Positionskoordinaten der beiden Ordinationen zueinander. Wie bei anderen multiplen Korrelationskoeffizienten ist bei der Interpretation Vorsicht geboten, da die Korrelation sich ständig optimiert, je mehr die Zahl der eingegangenen Umweltparameter  $p$  sich dem Stichprobenumfang  $n$  nähert (ter Braak 1998). In Falle von  $p \geq n-1$  schränken die Umweltparameter die Ordination nicht länger ein und lassen die CCA zu einer indirekten Gradientenanalyse werden (Jongman et al. 1987). Daraus resultiert, dass ein Bogeneffekt (vgl. Abschnitt 3.5.6) auftreten kann. Das Problem lässt sich lösen, indem man überflüssige (annähernd kollineare) Variablen aus der Analyse herausnimmt (Jongman et al. 1987; Kent & Coker 1992; ter Braak 1998). Die Art-Umwelt-Korrelation ist also nur aussagekräftig, wenn die Anzahl der Erklärungsvariablen deutlich geringer als der Stichprobenumfang ist. Dennoch bleibt die Interpretation problembehaftet, da die relative Bedeutung der Achsen unabhängig von der multiplen Korrelation ist. Trotz einer hohen Art-Umwelt-Korrelation kann der Erklärungsanteil der Dimension an der Antwortvariablenmatrix sehr gering sein. Der Eigenwert einer Achse stellt dahingehend ein zuverlässigeres Maß zur Beurteilung der relativen Bedeutung einer Achse gegenüber der Gesamtvariabilität dar (vgl. Abschnitt 3.6.1).

Die Mathematik hinter dem Verfahren geht auf die CA zurück (vgl. Abschnitt 3.5.5) und ist in vielen Schritten annähernd vergleichbar mit der einer Redundanzanalyse.

### Berechnungsschritte einer kanonischen Korrespondenzanalyse

Der RDA-Algorithmus wird wie folgt modifiziert, um eine CCA auszuführen (vgl. Legendre & Legendre 1998, ter Braak 1998):

- (1) Die abhängige Datenmatrix in der Analyse ist nicht Matrix  $Y$ , sondern Matrix  $\bar{Q}$  der Beiträge zur  $\chi^2$ -Statistik wie in der CA. Sie leitet sich von Gleichung 3.5-14 (S. 77) ab.
- (2) Matrix  $X$  wird unter Benutzung der Gewichte  $D(f_{i+})$  standardisiert. Dafür wird das Mittel und die Standardabweichung jeder Spalte der erweiterten Matrix  $X_{\text{infl}}$  berechnet und zur Standardisierung der Erklärungsvariablen herangezogen. In der erweiterten Matrix  $Y_{\text{infl}}$  ist pro Zeile eine Art eingetragen, wodurch sie deutlich mehr Zeilen besitzt als die Originalmatrix  $Y$ . Um die Varianz einer Erklärungsvariablen  $y_j$  zu berechnen, wird der Wert  $f_{++}$  als Divisor benutzt.
- (3) Anstatt einer multiplen Standardregression wird eine gewichtete Form der multiplen Regression angewendet. Die Gewichte in der Diagonalmatrix  $D(p_{i+})^{1/2}$  werden auf die Matrix  $X$  bezogen, wann immer sie in den Regressionsschritten genutzt werden.

**Multiple Regression:** Die Matrix der Regressionskoeffizienten  $B$  und die Matrix der Modellwerte  $\hat{Y}$  ergeben sich aus:

**Gleichung 3.5-17:** 
$$B = [X'D(p_{i+})X]^{-1}X'D(p_{i+})^{1/2}\bar{Q}$$

**Gleichung 3.5-18** 
$$\hat{Y} = D(p_{i+})XB$$

**Eigenwerte und Eigenvektoren:** Die Eigenwertzerlegung (vgl. Gleichung 3.5-16, S. 81) wird an der Matrix  $S_{\hat{Y}\hat{Y}} = \hat{Y}\hat{Y}$  ausgeführt. Die CCA wird dann anhand von Gleichung 3.5-16 (S. 81) berechnet. Auch hier ergeben sich eine Eigenwertmatrix  $\Lambda$  und eine Eigenvektormatrix  $U$ . Die CCA ist somit nur eine gewichtete Form der RDA, angewendet auf Matrix  $\bar{Q}$  (Gleichung 3.5-14, S. 77). Sie verfolgt allerdings einen unimodalen Modellgedanken, anstatt eines linearen. Matrix  $\hat{Y}$  enthält die Vektoren der Ordination, welche den maximalen Bezug zu den gewichteten Linearkombinationen der Erklärungsvariablen aufweisen (Legendre & Legendre 1998; Jongman et al. 1987; ter Braak 1998). Die normalisierte Matrix  $\hat{U}$  ergibt sich aus (s.a. Gleichung 3.5-15, S. 77):

**Gleichung 3.5-19:** 
$$\hat{U} = \bar{Q} U \Lambda^{-1/2}$$

In der CCA enthält diese Matrix nicht die Ladungen der Zeilen in  $\hat{Y}$  zu den Achsen, sondern der Zeilen in  $\bar{Q}$ . Sie wird erneut genutzt, um die Positionskoordinaten der Objekte im Raum der Antwortvariablen der Matrix  $Y$  darzustellen. Die Positionskoordinaten der Objekte im Raum der Linearkombinationen (Matrix  $\hat{Y}$ ) finden sich durch die Nutzung von Matrix  $U$  anstatt der Matrix  $\hat{U}$  (s. Gleichung 3.5-15, S. 77).

Die Projektionen und Ordinationsdiagramme zur Ergebnisrepräsentation einer CCA sind in Abschnitt 3.6.2 erläutert.

#### Kanonische Korrespondenzanalyse in CANOCO®

Die Optionen zur Berechnung der besprochenen CCA-Varianten sind identisch den Optionen für die CA und können dem zugehörigen Abschnitt 3.5.5 entnommen werden.

#### 3.5.9 Räumliche Autokorrelation: Moran's I und partielle Ordination

*"THE BORCARD ET AL. PROCEDURE IS ONE OF THE VERY FEW PROPOSITIONS TO EXPLICITLY INCORPORATE SPATIAL STRUCTURES INTO ECOLOGICAL MODELLING. BECAUSE ECOLOGICAL DATA GENERALLY SHOW STRONG SPATIAL STRUCTURING, THE NEED FOR THIS KIND OF METHOD IS GREAT (MÉOT ET AL. 1998, S. 2)."*

In diesem Abschnitt soll eingangs auf die Bedeutung von Autokorrelation eingegangen werden, da sie die Analyseergebnisse und Modelle stark beeinflussen kann. Danach werden der Autokorrelationskoeffizient Moran's I und die partielle Ordination als Verfahren zu ihrer Messung dargestellt und diskutiert.

In der klassischen Statistik ist eine der fundamentalsten Annahmen beim Testen von Hypothesen die Unabhängigkeit der Beobachtungen (Borcard 1992; Hosmer & Lemeshow 1989; Sokal & Rohlf 1995), da nur so die Anzahl der Freiheitsgrade und damit die Referenzverteilung bestimmbar sind. Die Existenz einer Raumstruktur (Gradienten, Flickenteppiche etc.) im Untersuchungsgebiet widerspricht dieser Forderung, da ein ökologisches Phänomen unterschiedlichster und oft nicht messbarer Natur Einfluss nimmt auf seine Umgebung. Daraus resultiert, dass eng benachbarte Flächen sich oft ähnlicher sind als weit voneinander entfernte. Sie sind nicht unabhängig voneinander. Solche autokorrelierten Beobachtungen finden sich in Variablen mit räumlicher wie mit zeitlicher Ausprägung.

Systematische und Zufallstichprobenpläne (vgl. Abschnitt 3.1.2) sollten Autokorrelationsprobleme vermeiden helfen, können aber nicht als Garanten gegen die Verletzung der Annahme unabhängiger Stichproben angesehen werden. Man weiß inzwischen, dass trotz entsprechender Stichprobenpläne Beobachtungen graduell räumlich voneinander abhängig sein können, wenn die durchschnittliche Distanz zwischen den Probeflächen geringer ist als die Zone des räumlichen Einflusses eines unterliegenden ökologischen Phänomens. Die räumliche Autokorrelation ist somit skalenabhängig, d.h., je nach Erklärungsvariable bilden sich unterschiedliche Distanzklassen der Abhängigkeit aus, und fehlende Autokorrelation auf einer Skalenebene bedeutet nicht, dass diese auf einer beliebigen anderen Ebene nicht anwesend sein kann. Sehr großräumige Gradienten können noch so weit voneinander entfernte Flächen mit Autokorrelationen belegen (Legendre & Fortin 1989).

Eine positive räumliche Autokorrelation zwischen Probeflächen bedeutet konkret, dass die Werte der sie beschreibenden Variablen deutlich ähnlicher sind als zwischen weit entfernt liegenden Probeflächen, welche in der Konsequenz sehr unterschiedlich und damit negativ autokorreliert sind. Folglich korreliert die Distanz mit der abiotischen Ausgestaltung der Probeflächen (Legendre & Fortin 1989) und Werte für Erklärungsvariablen können bis zu einem gewissen Grad aus der Distanz abgeleitet werden (Legendre 1993).



Im modelltechnischen Sinne wird durch positive räumliche Autokorrelation die Vorkommenswahrscheinlichkeit einer vorhandenen Art in ihrer räumlichen Nachbarschaft überschätzt (Legendre & Fortin 1989; Schröder 2000). Klassische Signifikanztests ergeben zu oft Signifikanzen bei Korrelationen und Regressionskoeffizienten. Eine Lösung des letztgenannten Problems findet sich in der Verwendung von Monte Carlo Tests (vgl. Abschnitt 3.6.1), weshalb diese auch in der vorliegenden Arbeit zur Anwendung kommen (vgl. Abschnitt 5.4).

Autokorrelation wird gewöhnlich durch Koeffizienten wie Moran's I oder Geary's c sowie die Mantel Statistik quantifiziert. Für Details zu den Parametern sei auf Legendre & Fortin (1989) oder auch Legendre & Legendre (1998) verwiesen.

Die Gleichung zur Berechnung von **Moran's I** wird angeführt, da bei den Auswertungen in Abschnitt 5.4.1 darauf zurückgegriffen wird:

**Gleichung 3.5-20:**

$$I(d) = \frac{\frac{1}{W} \sum_{h=1}^n \sum_{i=1}^n w_{hi} (y_h - \bar{y})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} \quad \text{für } h \neq i$$

mit:

**I(d):** Moran's I für die Distanzklassen d

**y<sub>h</sub>; y<sub>i</sub>:** Beobachtete Werte auf den Flächen h und i

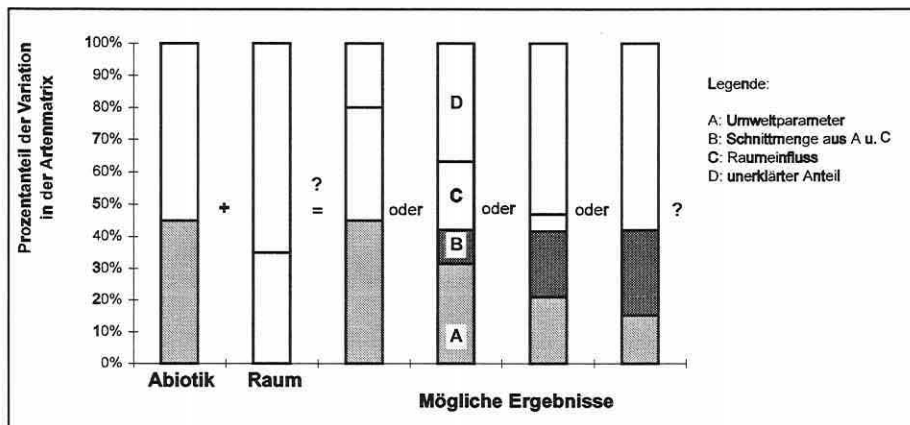
**w<sub>hi</sub>:** Gewichte mit w<sub>hi</sub> = 1 für h und i in der gleichen Distanzklasse und w<sub>hi</sub> = 0 in anderen Fällen (es werden nur Probeflächenpaare in der gleichen Distanzklasse berücksichtigt)

**W:** Summe der w<sub>hi</sub> für eine gegebene Distanzklasse

Die Werte des Autokorrelationskoeffizienten liegen in der Regel im Intervall [-1, +1] und können analog dem Korrelationskoeffizienten nach Pearson interpretiert werden. Positive Autokorrelation korrespondiert zu einem positiven Moran's I-Wert.

Die Visualisierung der Autokorrelationswerte bezogen auf die Distanzklassen erfolgt in Autokorrelogrammen (vgl. Abbildung 5.4-5, S. 170).

Das von Borcard et al. (1992) eingeführte Verfahren der Varianzaufteilung, die **partielle Ordination**, soll es erlauben, die relativen Beiträge verschiedener Erklärungsvariablen – auch Raumeinflüsse – in einer direkten Ordination zu quantifizieren. Es baut auf die existierenden Ordinationsverfahren RDA und CCA auf (Borcard 1992; Legendre & Fortin 1989; Méot et al. 1998; Økland & Eilertsen 1994) und wird als partielle kanonische Korrespondenzanalyse bzw. Redundanzanalyse (pCCA, pRDA) angesprochen (Borcard 1992).



**Abbildung 3.5-11:** Mögliche Erklärungsanteile von Umwelt- und Raumparametern an der Gesamtvariation eines biotischen Datensatzes (Borcard 1992, S. 1047, verändert).

Die Idee ist, die in Abbildung 3.5-11 (S. 86) dargestellten Erklärungsanteile A bis D zu ermitteln, zu quantifizieren und auf ihre Signifikanz zu testen, um den Raumeinfluss und damit die Autokorrelation zu bestimmen. Die Komponenten der Abbildung sind:

- (A) einzigartige Varianz der abiotischen Erklärungsvariablen im Modell
- (B) gemeinsame Varianz beider Variablensätze (Schnittmenge)
- (C) einzigartige Varianz des Raumdatensatzes
- (D) nicht erklärte Varianz im Modell.

Sie lassen sich schreiben als:

- (A) OIS
- (B)  $O \cap S$
- (C) SIO
- (D)  $TI - O \cup S = X$

wobei:

**O und S: Sonstige Variablen (O: other) und Raumvariablen (S: space)**

**I:** Varianz gegeben durch, d.h., OIS bedeutet Varianz gegeben durch {O} ohne {S}

**$\cap$ :** korrespondiert zu der gemeinsamen Varianz in {O} und {S}, der Schnittmenge

**$\cup$ :** korrespondiert zur Gesamtvarianz in {O} und {S}

**TI:** Gesamtvarianz (TI: total inertia) in der Matrix der abhängigen Daten (Arten)

Als synthetische Erklärungsvariablen des Raumeinflusses dienen grundlegend die geographischen Lagekoordinaten des Schwerpunktes einer Fläche. Die Matrix der geographischen Erklärungsvariablen wird aber noch um die Terme einer kubischen Trend-Oberflächen Regression erweitert, um sicherzustellen, dass nicht nur lineare Zusammenhänge, sondern auch komplexere Strukturen extrahiert werden (Lücken etc.), welche die quadratischen und kubischen Ausprägungen der Koordinaten umfassen (Borcard 1992; Legendre & Legendre 1998; Økland & Eilertsen 1994). Man erhält daher neun Raumvariablen als Erklärungsparameter:

**Gleichung 3.5-21:** 
$$\hat{z} = b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3$$

Der Beta-Koeffizient  $b_0$  entfällt bei zentrierten Datensätzen. Eine Vorwärtsselektion (Abschnitt 3.5.3) entscheidet bei der partiellen Ordination über die verbleibenden Variablen im Raumdatensatz.

Zur Berechnung einer partiellen Ordination wird als erstes eine direkte Ordination mit allen Parametern in der Analyse ausgeführt, um die Gesamtvarianz der Arten (TI: total inertia) und die Summe aller kanonischen Eigenwerte ( $O \cup S$ ) zu bestimmen. Letztere geben den Gesamtklärungsanteil der Modellparameter an der Antwortvariablenmatrix. Die nicht erklärbare Komponente in der Antwortvariablenmatrix entspricht  $TI - O \cup S$ .

Im nächsten Schritt werden zwei partielle Ordinationen ausgeführt, in denen einmal die Raumparameter als Erklärungsvariablen eingesetzt werden und die sonstigen Variablen als Kovariablen (Variablen deren Erklärungsgehalt aus der Analyse herausgerechnet wird) in die Analyse eingehen (SIO). Im zweiten Durchlauf umgekehrt (OIS).

Die Summe aller kanonischen Eigenwerte ergibt aus der ersten partiellen Ordination den Wert von SIO und aus der zweiten den Wert von OIS. Der noch fehlende Wert der Schnittmenge ( $O \cap S$ ) lässt sich schließlich berechnen aus:

**Gleichung 3.5-22:** 
$$O \cap S = O \cup S - OIS - SIO$$

Er stellt den Anteil redundanter Information in beiden Datensätzen dar. Anhand der Werte können dann die Prozentanteile der Teilmengen an der Gesamtvarianz im Datensatz (TI) bestimmt werden.

Die partielle Ordination kann prinzipiell auch auf beliebige andere Erklärungsvariablensätze angewendet werden, um unerwünschte Anteile bestimmter Variablenkategorien herauszufiltern (z.B. Beobachter als Kovariablen, s.u.) oder Erklärungsanteile verschiedener Erklärungsvariablensätze (z.B. Bodenparameter) gegeneinander abzuschätzen.

Trotz des einleuchtenden Gedankens der Varianzpartitionierung ist das Verfahren sehr umstritten. Die Ergebnisse sind davon abhängig, wie die Teildatensätze selektiert wurden. Fehlen wichtige Parameter (Umweltparameter / Raumkoordinaten), dann wird die durch einen Teildatensatz bedingte erklärte Variabilität unterschätzt. Die Erklärungsanteile fallen dem unerklärten Varianzanteil oder dem anderen Erklärungsvariablensatz zu. Økland & Eilertsen (1994) belegen außerdem, dass die erklärte Variation überschätzt wird, wenn nicht-signifikante Variablen in der Analyse sind. Daher ist eine Variablenreduktion und -prüfung mit beispielsweise einer Vorwärtsselektion notwendig (Abschnitt 3.5.3). Vergleicht man Teildatensätze unterschiedlicher Variablenanzahl, sind die Fraktionen der Varianz nicht unbedingt verlässlich. Die Frage, ob die Anteile erklärter Varianz zuverlässig sind, wird daraus bedingt, wie komplett ein Variablensatz ist. Nur wenn alle oder fast alle wichtigen Variablen erfasst sind, kann von einer Verlässlichkeit ausgegangen werden (Økland & Eilertsen 1994). Bedenken von Borcard (1992) und Palmer (1998) gegenüber dem Verfahren sind selbst dann nicht unbegründet, da beispielsweise die Eigenwerte einer CCA auf eins limitiert sind (vgl. Abschnitt 3.5.5 und 3.5.8). Stellt man sich einen Teildatensatz mit einem Parameter vor, der die Variabilität eines Antwortdatensatzes, dessen Gesamtvarianz zum Beispiel gleich 4,8 ist, zu 100% erklärt, kann er trotzdem diese verfahrensbedingt nie erreichen. Weitere Probleme bereitet der nicht abschließend verstandene Einfluss des Bogeneffektes (vgl. Palmer 1998 und Abschnitt 3.5.6). Schlussfolgernd aus diesen Kritikpunkten wird das Verfahren nicht für die Prüfung der Autokorrelation in den Daten (Kapitel 5) genutzt, sondern der Weg über den Autokorrelationskoeffizient Moran's I und Autokorrelogramme gewählt.

Die partielle Ordination kann dennoch für nachfolgende Ziele gewinnbringend eingesetzt werden, weshalb sie trotz der Kritik in diesem Abschnitt besprochen wurde:

- (1) Sie stellt einen attraktiven Weg dar, Einzelvariablen – zumindest relativ zueinander – in ihren Erklärungsanteilen an der Gesamtvarianz der biotischen Daten einzuschätzen.  
Nachfolgender Trick ist gleichbedeutend mit  $p$  partiellen Ordinationen (univariaten Modellen) für alle Erklärungsvariablen einzeln und erspart umfangreichen Rechenaufwand in CANOCO®: Die Ergebnisse des ersten Durchlaufs einer Vorwärtsselektion werden aus dem log-File herauskopiert und in einem Tabellenkalkulationsprogramm (z.B. Excel) die Prozentanteile an der Gesamtvarianz berechnet (vgl. z.B. Tabelle 5.4-3, S. 164).
- (2) Verschiedene Personen haben Daten erhoben: In großen Projekten sind häufig mehrere Mitarbeiter für die Datenerhebung verantwortlich. Es können daher Unterschiede in der Betrachtungsweise auftreten: Artenbestimmung, Artmächtigkeiten etc. Weist man jedem Mitarbeiter eine binär kodierte Dummy-Variable zu und setzt diese als Kovariable ein, so kann der Effekt herausgefiltert werden (partialling out).
- (3) Zeitdifferenzen: Jede umfangreiche Datenerhebung fordert Zeit ein. Sind dadurch deutliche Änderungen im Phänotypus der Beobachtungsobjekte auszumachen, können die Erhebungen mit dem Datum belegt werden bzw. bei mehrjährigen Studien die Jahre kodiert werden. In beiden Fällen ist wiederum die Kovariablenfunktion der pCCA oder pRDA anzuwenden.
- (4) Nicht interessierender Gradient: Möchte man einen Gradienten, der bestens bekannt ist, aus den Daten herausnehmen (z.B. Höhengradient einer Gebirgsregion in einer Beweidungsstudie), um die verbleibende Variation zu studieren, kann die Information der Variable oder der Variablen entfernt werden.
- (5) Betrachtung der Residuen: Setzt man alle Erklärungsvariablen als Kovariablen, kann die verbleibende Struktur der Daten betrachtet werden. Sind klare Muster zu erkennen, deutet dies darauf hin, dass ein relevanter Gradient nicht erfasst wurde (Palmer 1998).

### 3.6 Ordinationsdiagnostik und graphische Aufbereitung von Ordinationsergebnissen

Im Abschnitt 3.6 werden Möglichkeiten aufgezeigt, Ergebnisse von Ordinationsverfahren besser zu verstehen, einzuschätzen und darzustellen. Speziell die Ordinationsdiagnostik und Monte Carlo Tests sind eine essentielle Informationsquelle, um Ordinationsergebnisse zu prüfen und im Detail abzusichern. Die Statistiken werden zu dem für die Übertragung der statistischen Modelle in ein GIS-basiertes Prognosesystem benötigt (Abschnitt 4.2 und 5.6). Das exakte Wissen, welche Interpretation ein spezifisches Ordinationsdiagramm zulässt, und die weiterführenden Datenexplorationsmöglichkeiten durch die Software CanoDraw® sind maßgeblich für die abschließende Ergebnisfindung bei einer Datenanalyse (Abschnitte 5.4.2 bis 5.4.4).

#### 3.6.1 Ordinationsdiagnostik

*"USUALLY, AN ORDINATION DIAGRAM IS NOT AN EXACT REPRESENTATION OF THE DATA. (...) CANOCO HAS ORDINATION DIAGNOSTICS TO FIND OUT WHICH SPECIES AND WHICH SAMPLES ARE ILL-REPRESENTED AND WHICH ARE WELL REPRESENTED (TER BRAAK 1998, S. 172)."*

Die Ordinationsdiagnostik, welche CANOCO® bereitstellt, ist umfangreich und neben der Standardzusammenfassung (Ergebnisübersicht s.u.) oft wenig beachtet. Der nachfolgende Überblick zeigt, mit welchen Kennwerten die Ergebnisse in den Abschnitten 5.4.2 bis 5.4.4 begutachtet werden. Alle Statistiken sind den beiden CANOCO®-Ausgabefiles (.log und .sol) bzw. dem sogenannten "Spec\_Env.Tab"-File der durchgeführten Berechnungen entnommen.

**Originaldaten:** Es mag banal klingen, aber eine Quelle an Informationen, die bei der Prüfung von Ergebnissen niemals vernachlässigt werden sollte, sind die Originaldaten selbst. Auch in der Vorbereitung der Analyse dienen sie zum Beispiel dazu, schlecht repräsentierte Arten, die nur selten im Gebiet angetroffen wurden, zu selektieren und passiv zu setzen (vgl. Abschnitte 3.3 und 5.4.1). Bei mehreren Untersuchungsphasen bieten sie Vergleichsmöglichkeiten zu Populationsschwankungen auf einzelnen Flächen oder der Gesamtartenzahl (vgl. Abschnitt 5.3.1).

#### Zum Inhalt des log-Files

**Basisinformationen:** Speziell im Falle von mehreren vergleichend angewandten Analysen, z.B. über verschiedene Untersuchungsphasen hinweg, bekommt man durch die Datensatzbeschreibung einen schnellen Überblick, wie viele Probenflächen, Arten und Erscheinungen vorliegen. Weiterhin wird die Anzahl aktiver und passiver Arten im Datensatz ausgegeben:

#### Übersicht 3.6-1: Basisinformation

No samples omitted			
Number of samples	36	No. of active samples:	36
Number of species	24	No. of passive samples:	0
Number of occurrences	238	No. of active species:	18

**Extremwerte:** Als Erweiterung der multiplen Regression sind direkte Ordinationsverfahren anfällig für Ausreißer. Sie können die Ergebnisse beeinflussen und werden in CANOCO® ermittelt durch die quadrierte Mahalanobis-Distanz einer Probenfläche plus  $1/n$  (Kontrollmaß:  $k$ ). Deren Aussage beschreibt, wie extrem die Position von Ausreißern im Erklärungsvariablenraum ist. Dabei werden drei Testergebnisse ausgegeben:

## Übersicht 3.6-2: Extremwerte

\*\*\*\*\* Check on influence in covariable/environment data \*\*\*\*\*  
 The following sample(s) have extreme values

Sample	Environmental variable	Influence	Covariable influence	Environment space influence
2	24	9.5x		
2				3.7x
5	19	11.4x		
5	24	10.5x		
5				3.8x
10	174	7.3x		
34			3.1x	

Die erste Spalte gibt die Probefläche (Objektnummer), die zweite und dritte prüft einzelne Umweltvariablen (Nummerierung: Variablenkodierung), um univariate Ausreißer aufzuzeigen. Eine Probefläche wird aufgelistet, wenn ihr Wert fünffach höher als der Durchschnitt des Kontrollmaßes  $k$  ist. Solche Probeflächen haben gleichzeitig einen Wert, der drei Standardabweichungen (SD) vom Mittel abweicht. Probefläche 2 in Übersicht 3.6-2 wird durch Variable 24 im Datensatz als Ausreißer erkannt. Sie weicht 4,2 SD vom Mittel der Daten ab:  $(2 \cdot k - 1)^{1/2} = \text{Anzahl der SD vom Mittel} \Rightarrow (2 \cdot 9.5 - 1)^{1/2} = 4,2 \text{ SD}$ .

Die beiden weiteren Spalten zeigen multivariate Ausreißer auf: einmal im Kovariablenraum und im Gesamttraum der Erklärungsvariablen und Kovariablen. Auch wenn keine Kovariablen in der Analyse sind, kann eine Ausgabe erfolgen (Sample 34), falls die Probefläche großen Einfluss auf die Analyse nimmt. Ist ein Ausreißer entdeckt, sollten die Originaldaten auf Plausibilität geprüft werden. Bei korrekten Werten stellt sich die Frage, wie sich dieses Extrem erklärt und ob die Probefläche aus der Stichprobe entfernt werden sollte (ter Braak 1998).

**Korrelationsmatrix:** Die Korrelationsmatrix einer CCA (Abschnitt 3.5.8) wird hier gekürzt dargestellt. Sie ist im Gegensatz zu der Ausgabe in Standardsoftware-Paketen (SPSS®, STATISTICA®) in unimodalen Verfahren gewichtet nach den Artmächtigkeiten/Häufigkeiten der Arten auf den Probeflächen, d.h.  $y_{i+}$  (Palmer 1998; ter Braak 1998).

## Übersicht 3.6-3: Gewichtete Korrelationsmatrix

\*\*\*\* Weighted correlation matrix (weight = sample total) \*\*\*\*

SPEC AX1	1.0000				
SPEC AX2	.0200	1.0000			
ENVI AX1	.9857	.0000	1.0000		
ENVI AX2	.0000	.9035	.0000	1.0000	
FK	.5635	-.4365	.5717	-.4831	
FDAJd99	-.9641	.0756	-.9781	.0837	
SPEC	AX1	AX2	ENVI	AX1	AX2
FK	1.0000				
FDAJd99	-.4800	1.0000			
	FK	FDAJd99			

Die Variablen SPEC AX1 bis  $k$  und ENVI AX1 bis  $k$  beschreiben die Ordinationsachsen. Die Probeflächensignaturen sind einmal abgeleitet von den Antwortvariablen (SPEC) und einmal von den Erklärungsvariablen (ENVI; vgl. Abschnitte 3.5.7, 3.5.8 und 3.6.2). Während in direkten Ordinationen zwischen den SPEC-Achsen Korrelationen auftreten, sind die ENVI-Achsen zueinander unkorreliert, da sie orthogonal zueinander ermittelt werden. In indirekten Analysen sind hingegen die SPEC-Achsen unkorreliert zueinander.

Die Korrelationen zwischen den SPEC-Achsen und den ENVI-Achsen geben die Art-Umwelt-Korrelationen an (vgl. Abschnitt 3.5.7 und 3.5.8).



Die Korrelationen zwischen den einzelnen Erklärungsparametern und den SPEC- bzw. ENVI-Achsen werden als "inter-set"- und "intra-set"-Korrelationen bezeichnet und approximieren den Zusammenhang der Erklärungsvariablen zu den unterschiedlich abgeleiteten Ordinationsachsen. In direkten Ordinationen sind die "intra-set"-Korrelationen im Fokus des Interesses, da sie den Beitrag der Erklärungsparameter zur Formation der Achsen beschreiben. Sie geben dem jeweiligen Gradienten (der Linearkombination) seine Bedeutung (vgl. Abschnitt 5.4.2 und z.B. Tabelle 5.4-12, S. 172). Der untere Teil der Matrix (Übersicht 3.6-3) umfasst die Korrelationen zwischen den Erklärungsvariablen und ermöglicht bei hohen VIF-Werten (s.u.) die Selektion kollinear Variablen (Jongman et al. 1987; ter Braak 1998).

**Verteilungen und Kolinearitätstest:** Die Übersicht 3.6-4 belegt, dass die Mittelwerte aller Positionskoordinaten der Probestellen aufgrund der Zentrierung null sind (in der nicht abgebildeten kompletten Tabelle: Zeile eins bis acht).

**Übersicht 3.6-4: Mittel, Standardabweichung und VIF**

N	name	(weighted) mean	stand. dev.	inflation factor
1	SPEC AX1	.0000	1.6100	
2	SPEC AX2	.0000	1.3520	
5	ENVI AX1	.0000	1.5870	
6	ENVI AX2	.0000	1.2216	
19	FK	97.6140	17.3439	1.7485
195	FDAJd99	100.7319	72.9210	2.3844

Die Standardabweichungen der ENVI-Achsen, d.h. der Positionskoordinaten aller Probestellen aus den linear kombinierten Erklärungsvariablen, sind in der Biplot-Skalierung gleich eins. Bei der in Übersicht 3.6-4 verwendeten Skalierung nach Hill weisen sie ähnliche Werte auf, die meist etwas größer als eins sind (vgl. Abschnitt 3.5.8). Weiterhin werden die gewichteten Mittel und die Standardabweichungen für alle Erklärungsvariablen ausgegeben, welche zur Übertragung der Modelle in ein GIS-basiertes Prognosesystem benötigt werden (vgl. Abschnitt 4.2 und 5.6). Der "Variance Inflation Factor" (VIF) indiziert hochkorrelierte Variablen im Datensatz (vgl. Abschnitt 3.5.3).

**Die Ergebnisübersicht:** Sie wird am Beispiel unimodaler Verfahren erläutert und gibt in allen Ordinationen Auskunft über die Gesamtvarianz (TI: total inertia) in der Matrix Y (Arten · Probestellen). In linearen Verfahren ist sie aufgrund der Skalierung immer gleich 1,0 gesetzt, während sie in Übersicht 3.6-5 1,914 beträgt.

Die Eigenwerte (EV: Eigenvalues) messen den Erklärungsanteil der ersten vier CA-Achsen und liegen zwischen Werten von 0 bis 1. Die Bedeutung eines Eigenwertes ist immer im Zusammenhang mit der Gesamtvarianz zu sehen. Die kumulativen Prozentanteile der erklärten Varianz an der Antwortmatrix drücken diese Relation aus:  $100 \cdot (1. \text{EV}) / \text{TI} = 32,5\%$  bzw.  $100 \cdot (1. + 2. \text{EV}) / 1,914 = 53,9\%$ . Es wird klar, dass bei einer TI von 19,14 anstatt 1,914 der Eigenwert von 0,622 eine völlig andere Aussage trifft: Er erklärt nur noch 3,25% der biotischen Variation. Die Summe aller Eigenwerte ist identisch zur TI. In Übersicht 3.6-5 ist der erste Eigenwert mit 0,622 sehr hoch und impliziert einen eindeutigen Gradienten. Der zweite und dritte Eigenwert ist deutlich niedriger und betont die Dominanz der ersten Achse.

**Übersicht 3.6-5: Ergebnisübersicht einer CA**

**** Summary ****		<u>Korrespondenzanalyse</u>				
Axes		1	2	3	4	Total inertia
Eigenvalues	:	.622	.410	.209	.186	1.914
Cumulative percentage variance of species data	:	32.5	53.9	64.9	74.5	
Sum of all unconstrained eigenvalues						1.914

Als Beispiel stuft Palmer (1998) Eigenwerte um 0,3 bei einer TI von 2,725 (11,0%) noch als strengen Gradienten ein. Auch ter Braak (1998) und Gauch (1982) weisen darauf hin, dass die Eigenwerte für Felddaten durch strukturlöse Variation (vgl. 3.1.7 und 3.5.2) oft niedrig sind; Ordinationen mit geringen Erklärungsanteilen können trotzdem informativ sein (Palmer 1998; Gauch 1982).

Die Ergebnisübersicht einer DCA (Übersicht 3.6-6) weist für den identischen Datensatz Veränderungen auf und ist um die Zeile der Gradientenlänge (Lengths of gradient) ergänzt (vgl. Abschnitt 3.5.6). Die Gradientenlänge wird in Standardabweichungen der mittleren Antwortkurvenspannweite im Datensatz ausgedrückt, da eine DCA immer die Skalierung nach Hill anwendet (ter Braak 1998). In den Abschnitten 2.3, 3.5.5 und 3.5.6 werden die Eigenschaften der Skalierung ausführlich besprochen. Die Gradientenlänge kann zu der einer CCA bei gleicher Skalierung für eine vergleichende Bewertung in Beziehung gesetzt werden.

Allgemein wird die Gradientenlänge als Entscheidungskriterium herangezogen, welches Modell bzw. Analyseverfahren für den vorliegenden Datensatz geeignet erscheint. Als Faustregel gilt (vgl. Abschnitt 3.5.6): Ist die Gradientenlänge der ersten Achse länger als 4 SD, ist ein Artenwechsel entlang des Gradienten zu erwarten, und das unimodale Modell ist vorzuziehen. Umgekehrt sollte ein lineares Modell gewählt werden, wenn der Gradient eine Länge von 3 SD unterschreitet, da wohl lineare Antworten überwiegen. Zwischen 3 und 4 SD gibt es eine Grauzone, in der beide Modelle ähnlich gute Ergebnisse liefern können. Eine interessante Randbemerkung ist, dass die CA/CCA zwei "Gesichter" haben kann und auch bei kurzen Gradienten nicht notwendigerweise zu verwerfen ist (vgl. Abschnitt 5.4.2). Im linearen Kontext (Fokus auf "Inter-Species" und Biplot-Skalierung) kann das Ordinationsdiagramm sowohl nach der Biplot-Regel als auch nach dem Zentroid-Prinzip interpretiert werden (vgl. Abschnitt 3.6.3 und ter Braak 1998, S. 240).

**Übersicht 3.6-6: Ergebnisübersicht einer DCA**

**** Summary ****		<u>Korrespondenzanalyse mit Detrending</u>				
Axes		1	2	3	4	Total inertia
Eigenvalues	:	.622	.133	.076	.051	1.914
Lengths of gradient	:	4.121	2.484	2.089	2.001	
Cumulative percentage variance of species data	:	32.5	39.4	43.4	46.1	
Sum of all unconstrained eigenvalues						1.914

In Fortführung der Eigenwertdiskussion fällt auf, dass nur der erste Eigenwert gleich der einer CA ist. Durch das Detrending (vgl. Abschnitt 3.5.6) wird ab der zweiten Achse der Erklärungsanteil des Bogeneffektes (vgl. ebenfalls Abschnitt 3.5.6) eliminiert und der niedrigere Eigenwert für die reale 2. Achse ausgegeben. Das Detrending-Verfahren ist in der Kritik, da nicht ausgeschlossen werden kann, dass auch reale Information verloren geht (vgl. Abschnitt 3.5.6 und Kapitel 4).

**Übersicht 3.6-7: Ergebnisübersicht einer CCA**

**** Summary ****		<u>Kanonische Korrespondenzanalyse</u>				
Axes		1	2	3	4	Total inertia
Eigenvalues	:	.603	.330	.130	.063	1.914
Species-environment correlations	:	.986	.904	.854	.708	
Cumulative percentage variance of species data	:	31.5	48.7	55.5	58.8	
of species-environment relation:		52.6	81.3	92.6	98.1	
Sum of all unconstrained eigenvalues						1.914
Sum of all canonical eigenvalues						1.147

Die CCA liefert eine Ergebnisübersicht (Übersicht 3.6-7, S. 92), welche – eingeschränkt auf die Erklärungsvariablen in der Analyse – die Erklärungsanteile der Achsen an der Gesamtvarianz der Antwortvariablen wiedergibt. Die Eigenwerte sind daher geringer als in einer CA. Gewöhnlich vergleicht man die Eigenwerte und Ordinationsdiagramme einer direkten mit der zugehörigen indirekten Ordination, um zu testen, ob die relevanten Umweltparameter erfasst wurden (Abschnitt 5.4.2). Sind die Eigenwerte und Muster in den Ordinationsdiagrammen nicht deutlich verschieden, ist der Erklärungsvariablensatz effizient.

Ohne Kovariablen in der Analyse (vgl. Abschnitt 3.5.9) ist die Gesamtvarianz (TI) in der CCA gleich der Summe aller nicht durch die Erklärungsvariablen eingeschränkten Eigenwerte und identisch zu den CA-Ergebnissen. Die Summe aller kanonischen Eigenwerte findet sich in der letzten Zeile. Das Modell erklärt 1,147 von 1,914 (=59,93%) der Gesamtvarianz im biotischen Datensatz. Von der Gesamtvarianz entfallen dabei 1,126 (=58,8%) auf die ersten vier Achsen. Aus diesem Grunde ist in dieser Arbeit in den Ordinationsdiagrammen der Wert EI4% eingeführt worden, der den Erklärungsanteil an der Gesamtvarianz über vier Achsen ausweist (vgl. Abschnitt 5.4.2 und als Beispiel Abbildung 5.4-7, S. 175).

Die Art-Umwelt-Korrelation ist nicht immer aussagekräftig (vgl. Abschnitt 3.5.7 und 3.5.8), daher ziehe man generell die Eigenwerte bei der Interpretation heran (Palmer 1998).

Die Zeile "cumulative percentage variance of species-environment relation" drückt den kumulierten prozentualen Varianzanteil aus, den die Achsen an der durch die Modellparameter erklärbaren Varianz besitzen ( $52,6 = 100 \cdot (1. \text{EV}) / (\sum \text{kanonischer EVs})$  bzw.  $81,3 = 100 \cdot (1. + 2. \text{EV}) / 1,147$ ). Insgesamt entfallen somit 98,1% Erklärungsanteil von der TI aller kanonischen Eigenwerte (1,147) auf die ersten vier Ordinationsachsen.

#### Zum Inhalt des sol-Files (CCA)

**Die Positionskoordinaten:** Die Übersichten 3.6-8 bis 3.6-10 enthalten die Koordinaten für jede Art bzw. Probefläche entlang der Achsen.

Übersicht 3.6-8: Positionskoordinaten der Arten

##### Spec: Species scores

N	NAME	AX1	AX2	AX3	AX4	WEIGHT	N2
	EIG	.6029	.3299	.1298	.0626		
1	Ani_spi	-1.5647	.4016	.5201	.0595	112.48	10.77
2	Ari_arb	-.6132	-.3221	.3637	.7507	4.73	3.73
3	Car_min	.8291	-.8600	.1584	-.0431	49.65	6.94

Übersicht 3.6-9: Positionskoordinaten der Probeflächen im Unähnlichkeitsraum der Arten/Probeflächen

##### Samp: Sample scores

N	NAME	AX1	AX2	AX3	AX4	WEIGHT	N2
	EIG	.6029	.3299	.1298	.0626		
1	1	-.9672	-.0488	.5856	1.8294	40.24	5.45
2	2	-1.9121	-.2298	-2.1835	-.8807	35.86	5.54
3	3	-2.2102	.2255	-1.2244	-.5888	33.03	4.18

Übersicht 3.6-10: Positionskoordinaten der Probeflächen im ökologischen Raum

##### SamE: Sample scores which are linear combinations of environmental variables

N	NAME	AX1	AX2	AX3	AX4	WEIGHT	% FIT
	EIG	.6029	.3299	.1298	.0626		
1	1	-1.3327	.1438	.1362	2.3954	40.24	48.81
2	2	-2.1858	-.2675	-2.2340	-1.3310	35.86	87.68
3	3	-1.9462	-.2395	-1.0114	-.3098	33.03	85.18

Die Koordinaten der Arten dienen sowohl der Einordnung ins Ordinationsdiagramm als auch der Übertragung des Modells in das GIS-Prognosesystem (Abschnitt 4.2 und 5.6). Die Positionskordinaten der Probeflächen korrespondieren zu den zwei Ordinationen einer CCA (vgl. Abschnitt 3.5.8 und Abbildung 3.5-10, S. 82).

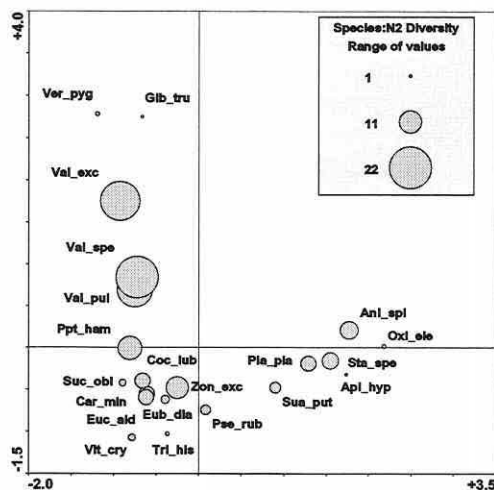
**WEIGHT:** Das Gewicht  $w_k$  der Arten in Übersicht 3.6-8 ist eine nützliche Information zur Interpretation von Ordinationsdiagrammen. In unimodalen Verfahren haben Arten mit geringem Gewicht nur wenig Einfluss auf die Ordination.

In der Übersicht 3.6-9 und 3.6-10 (S. 93) beschreibt es in linearen Verfahren  $w_i = w_i^*$  und in unimodalen  $w_i = w_i^* y_{i+}$ . Der Wert für  $w_i^*$  ist per Voreinstellung eins bzw. null für passive oder temporär gelöschte Flächen (ter Braak 1998). Dies gilt auch für Arten.  $y_{i+}$  ist die gewichtete Gesamtartmächtigkeit/Häufigkeit in einer Probefläche.

**N2:** Das Maß beschreibt die effektive Anzahl der Erscheinungen einer Art im Datensatz. Der Wert ist analog der N2-Diversität nach Hill und steht in Bezug zu einigen anderen Diversitätsmaßen (vgl. Legendre & Legendre 1998). Für Präsenz-Absenz-Daten ist N2 gleich der einfachen Anzahl an Erscheinungen. In unimodalen Verfahren stellt sich der Wert komplexer dar. Man stelle sich vor, eine Art kommt auf drei Probeflächen mit den Häufigkeiten 999, 3 und 3 vor. Das Vorkommen der Art A wird hier in erster Linie durch die Fläche mit dem Wert 999 definiert, da die Spezies in ihrem Schwerpunkt erscheint.

**Gleichung 3.6-1:** 
$$N2 = 1 / \sum_i (w_i^* y_{ik} / y_{+k})^2$$

mit  $w_i^*$ : **Gewicht (s.o., Voreinstellung = 1)**  
 $y_{ik}$ : **Wert einer Art k für eine Probefläche i der Matrix**  
 $y_{+k}$ : **Summe der Werte einer Art k über alle Probeflächen**



**Abbildung 3.6-1: N2-Diversität von Arten (Mollusken).**

N2 in Übersicht 3.6-9 steht für die effektive Anzahl der Arten auf einer Probefläche. In diesem Fall wird ein Multiarten-Vergleich der Probeflächen durchgeführt und in unimodalen Verfahren letztlich die Lage der Signatur für die Fläche im Diagramm von den dominierenden Arten bestimmt. Der Sinngehalt ist ähnlich dem für die Arten aufgezeigten Beispiel (ter Braak 1998).

**Gleichung 3.6-2:** 
$$N2 = 1 / \sum_k (w_i^* y_{ik} / y_{i+})^2$$

mit  $w_i^*$ : **Gewicht (s.o., Voreinstellung = 1)**  
 $y_{ik}$ : **Wert einer Art k für eine Probefläche i der Matrix**  
 $y_{i+}$ : **Summe der Werte aller Arten auf einer Probefläche i**

Der Wert für N2 ist 1,001. Für eine Art B, die mit den Häufigkeiten 100, 100 und 100 auftritt, wird N2 etwa 3,000. In Übersicht 3.6-8 kommt Ari\_arb im realen Datensatz auf vier Flächen mit 3, 1, 1 und 1 Individuen vor und Car\_min auf acht mit 1, 20, 39, 51, 35, 59, 57 und 93 Individuen. Die Interpretation ist daher stark abhängig von der Summe der Werte einer Art über alle Probeflächen, wie Art A und auch Ari\_arb beweisen (s.a. Smilauer 1992, S. 88). Neben dem Diagramm der N2-Diversität (Abbildung 3.6-1) lassen sich auch Shannon Index, N1-Diversität, Richness und Eveness (vgl. Mühlenberg 1993) mit CanoDraw® berechnen und darstellen.

Im Zusammenhang mit den Positionskoordinaten der Arten und Probestellen sollen auch die Positionskoordinaten der Erklärungsvariablen nicht unerwähnt bleiben, obgleich sie nicht direkt zur Rubrik Ordinationsdiagnostik zählen. Sie sind dem sol-File in direkten Ordinationen zu entnehmen. Die Koordinaten der Pfeilspitzen umfasst die Tabelle "BipE: Biplot scores of environmental variables". Die Klassenschwerpunkte für kategoriale Variablen werden besser repräsentiert durch die Koordinaten in der Tabelle "CenE: Centroids of environmental variables (mean.gt.0) in ordination diagram". Sie sollten auch im Ordinationsdiagramm als Punkt-signaturen dargestellt werden. Weiterhin sind in der Tabelle "Regr: Regression/canonical coefficients for standardized variables" die kanonischen Regressionskoeffizienten für die standardisierten Erklärungsvariablen gegeben. Für die Modellextraktion in eine GIS-Umgebung sind diese Koeffizienten erforderlich (vgl. Abschnitt 4.2 und 5.6 sowie Guisan et al. 1999 und ter Braak 1998).

**Die Toleranz der Arten:** Das Gedankengut des unimodalen Modells in der CA, DCA und CCA beinhaltet, dass die Antwortkurve einer Art zur Achse glockenförmig ist. Die Standardabweichung bzw. die Toleranz (Pickett et al. 1994) beschreibt die Spanne zwischen den Wendepunkten der Verteilung (vgl. Abbildung 2.3-3, S. 15). Sie erfasst ca. 68% der Nische einer Art entlang der Ordinationsachse, wenn man zur Interpretation die Normalverteilungskurve heranzieht.

#### Übersicht 3.6-11: Toleranz der Arten

Tol : Species tolerance (root mean squared deviation for species)

N	NAME	AX1	AX2	AX3	AX4	RMSTOL	N2
	FR FITTED	.3150	.1723	.0678	.0327		
1	Ani_leu	1.0683	.7188	1.5401	1.1901	116.68	10.77
2	Ari_arb	1.2482	.7917	.5179	1.2968	101.67	3.73
3	Car_min	.4414	.3776	.5389	.7978	56.22	6.94

Die wahre Toleranz wird unterschätzt, wenn die Skalierung  $\pm 2$  eingestellt wurde (Focus auf Inter-species distances). Im Extremfall ist die Toleranz gleich null, wenn die Art nur einmal im Datensatz vorkommt. Für einen fairen statistischen Vergleich der Nischenbreite muss folgende Korrektur erfolgen, um den Fehler zu beseitigen: Toleranz /  $(1-1/N_2)^{1/2}$ . Ein einfacherer Weg, korrekte Schätzungen zu erhalten, ist, die Ordination erneut mit einer Skalierung auf dem "Inter-sample" Focus zu berechnen. Die so ermittelten Werte dienen der Modellimplementierung ins Prognosesystem (vgl. Abschnitt 4.2 und 5.6).

AX1: Standardabweichung der Nische einer Art gegenüber der ersten Ordinationsachse.

RMSToleranz: Mittlere Toleranz der Art über alle vier Achsen. Die Standardabweichung der Gesamtpopulation in der Tabelle bildet den Divisor zu Berechnung und ist hierbei n und nicht n-1 (ter Braak 1998).

N2: Identisch zu den Werten in der Übersicht 3.6-8 (S. 93).

Beim Diagramm der Artentoleranzen muss berücksichtigt werden, dass die Darstellung nur für Relativvergleiche zwischen den Arten geeignet ist, weil aus technischen Gründen die Originalwerte des CANOCO®-Ergebnisfiles (\*.sol) durch zehn dividiert werden. Es basiert auf den Werten in Übersicht 3.6-11 (S. 95). Ein Beispiel zeigt Abbildung 5.4-30 (rechtes Teilbild, S. 198).

**Die Güte der Repräsentation von Arten im Ordinationsraum:** Die Übersicht 3.6-12 (S. 96) beschreibt die Gesamtvarianz einzelner Arten und wie gut eine Art von einer p-dimensionalen Ordinationslösung repräsentiert wird.

Die Gesamtvarianz (VAR(y)) der Art "Ani\_leu" im Ordinationsraum beträgt 1,48 und wird durch eine zweidimensionale Ordinationslösung zu 72,8% (AX2) erklärt. Davon hat Achse 1 einen Erklärungsanteil von 65,52% und Achse 2 von 72,8% bis 65,52% = 7,28%. Letztlich ist VAR (y) die Summe der Abweichungsquadrate der gewichteten Regression an den Daten für



Art  $k$  über die Ordinationsachsen als Fraktion der Gesamtsumme der Abweichungsquadrate der Art. Die Spalte "%EXPL" listet den Prozentsatz an der Gesamtvarianz der Art auf, welcher durch alle Erklärungsvariablen zusammen, d.h. die maximale Achsenanzahl im Modell, erklärt werden kann an.

#### Übersicht 3.6-12: Güte der Repräsentation von Arten im Ordinationsraum

CFit: Cumulative fit per species as fraction of variance of species

N	NAME	AX1	AX2	AX3	AX4	VAR(y)	% EXPL
	FR FITTED	.3150	.1723	.0678	.0327		
1	Ani_ leu	.6552	.7280	.8867	.8889	1.48	89.22
2	Ari_ arb	.0289	.0424	.0647	.1672	5.16	18.13
3	Car_ min	.2500	.7040	.7240	.7256	1.09	73.42

Bei umfangreichen Datensätzen können mäßig repräsentierte Arten anhand des CFit-Wertes der Dimensionen aus dem Ordinationsdiagramm ausgeblendet werden, um es übersichtlicher zu gestalten. Das Diagramm zur Prüfung der Artenrepräsentationsgüte greift auf die Cfit-Werte zurück und gibt einen schnellen Überblick, wie gut die einzelnen Spezies durch eine Ordinationslösung beschrieben werden (vgl. Abbildung 5.4-29, rechtes Teilbild, S. 198).

#### Der File "Spec\_Env.tab"

Dieser durch CANOCO® automatisch im Arbeitsverzeichnis erstellte File (Environment-by-species table) wird bei jedem Durchlauf überschrieben, so dass er – falls seine Information interessiert – durch Umbenennung gesichert werden muss.

#### Übersicht 3.6-13: Korrelationen zwischen den Arten und einzelnen Modellparametern

Environment-by-species table, analysis 2. 6 Vars 24 Specs 0 Cov

(I5/ 3(8F9.4:/))

24

1

-.3338	.1201	.3794	.0707	.4443	.2946	.4093	-.3271
.0530	-.0538	.5401	.3776	.1328	.0530	-.1281	-.0850
.2543	.2300	.1116	.4907	.5024	.2391	.2222	.4722

...

Ani\_ leu Ari\_ arb Car\_ min Cep\_ hor Coc\_ lub Eub\_ dia Euc\_ ald Oxl\_ ele Pla\_ cor Pla\_ pla  
Ppt\_ ham Pse\_ rub Pun\_ pyg Pyl\_ acu Sta\_ spe Sua\_ put Suc\_ obl Tri\_ his Val\_ exc Val\_ pul  
Val\_ spe Ver\_ pyg Vit\_ cry Zon\_ exc  
FK FDAJd99

In linearen Verfahren umfasst er die Korrelationen zwischen den einzelnen Arten und den Erklärungsvariablen, wenn diese zentriert und standardisiert wurden (Übersicht 3.6-13). Ist die Erklärungsvariablenmatrix nur zentriert worden, sind die Kovarianzen enthalten. Die Informationen entsprechen der Reihenfolge, in der Variablen abgelegt wurden. Sie sind im unteren Teil der Übersicht 3.6-13 aufgelistet. Die Erklärungsvariablen werden ebenfalls nach ihrer Reihenfolge durchnummeriert: Die eins entspricht der Variablen FK. In unimodalen Verfahren finden sich die Koordinaten der gewichteten Schwerpunkte (weighted averages) von Arten gegenüber den standardisierten Erklärungsvariablen in der Tabelle.

Interpretatorisch ergeben sich basierend auf diesen Informationen zwei Möglichkeiten:

Als erstes können die Ergebnisse des Biplots der Arten und Erklärungsvariablen verifiziert werden. Das Ordinationsdiagramm kann zu fehlerhaften Schlussfolgerungen führen, da die graphische Darstellung nur die Hauptmuster der Tabelle wiedergibt. Speziell wenn nur ein geringer Anteil der Art-Umwelt Beziehungen erklärt wird, ist es interessant zu prüfen, ob die Aussagen des Biplots sich in den Originaldaten durch die gewichteten Zentroide bzw. die Korrelationen bestätigen (vgl. Abschnitt 5.4.2 und Tabelle 5.4-22, S. 194).

Als zweites besteht die Möglichkeit, die "Environment-by-species table" mit einer nicht zentrierten PCA zu analysieren (ter Braak 1998). Dies entspricht einer Co-Inertia Analyse nach

Dolédec & Chessel (1994). Weitere Informationen und Beispiele zu dem Analyseverfahren geben: Thioulouse et.al (1997) und Thioulouse & Lobry (1994, 1995).

### Monte Carlo Signifikanztests

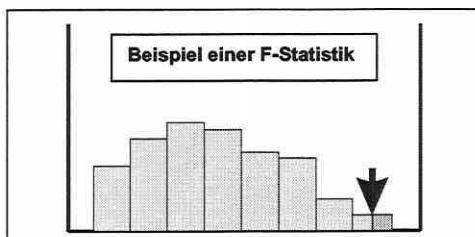
Tests, ob die Antwortvariablen signifikant abhängig sind von den Erklärungsvariablen, legen folgende Hypothesen zu Grunde: Die Nullhypothese lautet, dass die Antwortvariablen unabhängig von den Erklärungsvariablen sind. Entsprechend gegensätzlich ist die Alternativhypothese formuliert: Die Antwortvariablen sind abhängig von den Erklärungsvariablen, d.h., die Arten reagieren auf die Umweltbedingungen.

Für die klassischen statistischen Signifikanztests (Korrelationskoeffizient, t-Statistik oder F-Statistik) wird ein Testwert aus den Daten berechnet und dann eine mathematisch vorformulierte Referenzverteilung für die Nullhypothese gewählt. Sie zeigt die Erwartungswerte für einen nicht bestehenden Zusammenhang zwischen den Arten und der Umwelt. Anhand des Signifikanzlevels wird dann die Wahrscheinlichkeit festgelegt, wie sicher der Testwert oder höhere Werte in der Referenzverteilung auftreten. Das Problem bei dem Einsatz dieser Verfahren ist somit, dass die Referenzverteilung bekannt sein muss und Annahmen trifft: In der F-Statistik ist die Kurve durch die Freiheitsgrade definiert und als Folge dessen wird eine Unabhängigkeit der Erklärungsvariablen, die Normalverteilung der Daten und Varianzhomogenität der Residuen vorausgesetzt. Sind diese Annahmen nicht erfüllt, ist die Statistik ungültig. Die Referenzverteilung der multiplen Regression in einer direkten Ordination ist unbekannt, da sie von der Variablenanzahl, deren Korrelationsstruktur, den Werten der Antwortvariablen und anderen Faktoren abhängt (ter Braak 1998 u.a.).

Die benötigte Referenzverteilung kann aber durch so genannte Monte Carlo Permutationen simuliert werden (Mooney 1997). Man erhält sie, indem man ausgehend von der Nullhypothese den Probenflächen die Werte der Erklärungsvariablen beliebig zuordnet und die Teststatistik berechnet. Wenn die Nullhypothese wahr ist, spielt die Zuordnung keine Rolle. Sonst liegt der Wert der Originalzuordnung höher als die zufallsbedingten Werte und die Alternativhypothese wird angenommen. Das Signifikanzniveau bestimmt sich aus der Gleichung 3.6-3, wobei  $m$  die Anzahl der Permutationen ist, für die der Testwert höher als der Wert aus den Originaldaten liegt. Mit  $n$  wird die Gesamtzahl der durchgeführten Permutationen bezeichnet, die meist auf 199, 499 oder 999 festgelegt werden. Der p-Wert aus Gleichung 3.6-3 kann im Falle von ( $m = 0$ ) für z.B. 999 Permutationen maximal den Wert 0,001 erreichen. Die Signifikanz der Korrelation basiert als Standard in CANOCO® auf einem einseitigen Test.

Gleichung 3.6-3:

$$P = (1 + m) / (1 + n)$$



**Abbildung 3.6-2: Approximierte Teststatistik.** Der Pfeil markiert den Wert der Originaldaten. Werte rechts davon sprechen für die Nullhypothese und Werte links davon für die Alternativhypothese.

Als Bemerkung zur "statistischen Signifikanz" von Monte Carlo Tests muss darauf hingewiesen werden, dass der p-Wert resultiert aus  $n$  Ordinationslösungen, die bezüglich ihrer Qualität miteinander verglichen werden. Ihm fehlt seine enge Bedeutung im Sinne der schließenden Statistik. Ziel des Tests ist festzustellen, ob und wie häufig der Modellwert durch Zufallszuordnungen überschritten wird (vgl. Abbildung 3.6-2). Der Test ist verteilungsfrei und stellt keine Anforderungen an die Daten.

**Gleichung 3.6-4:**

$$F_{\lambda} = \frac{\lambda_1}{RSS_{X+1} / (n - p - q)}$$

mit:  $\lambda_1$ : Eigenwert/Erklärungsanteil der ersten Achse  
 $RSS_{X+1}$ : Summe der Abweichungsquadrate. Differenz der Gesamtvarianz in den Antwortvariablen und der durch die erste Achse erklärten Varianz  
 $n$ : Anzahl der Objekte/Probeflächen  
 $p$ : Anzahl der Erklärungsvariablen  
 $q$ : Anzahl der Kovariablen

**Gleichung 3.6-5:**

$$F_{trace} = \frac{\sum_{i=1}^p \lambda_i / p}{RSS_{X+Z} / (n - p - q)}$$

mit:  $\Sigma \lambda_i$ : Summe aller Eigenwerte in der Analyse  
 $RSS_{X+Z}$ : Summe der Abweichungsquadrate. Differenz der Gesamtvarianz in den Antwortvariablen und der durch das gesamte Modell erklärten Varianz  
 $n$ : Anzahl der Objekte/Probeflächen  
 $p$ : Anzahl der Erklärungsvariablen  
 $q$ : Anzahl der Kovariablen

Der in CANOCO<sup>®</sup> implementierte Algorithmus permutiert in Wirklichkeit nicht die Erklärungsvariablenmatrix sondern die Residuen der Modellwerte in der Antwortvariablenmatrix  $\hat{Y}$ . Weiter kompliziert wird der Algorithmus durch Kovariablen in der Analyse (vgl. hierzu ter Braak 1998). Die Teststatistik in CANOCO<sup>®</sup> kann bezogen werden auf die erste kanonische Achse (Gleichung 3.6-4) oder auf die Summe aller Achsen (Gleichung 3.6-5). Es ist zu bedenken, dass der letztgenannte Test eine geringere Analysekraft hat, da sicherlich einige Achsen nicht signifikant sind (Leps & Smilauer 1999, unveröffentlicht).

Die Optionen in CANOCO<sup>®</sup> für einen Monte Carlo Test können dem Stichprobendesign angepasst werden (vgl. Abschnitt 3.1.2). Es stehen neben der Standardeinstellung (unrestricted permutations) spezielle Tests für räumliche, zeitliche und logische Datenstrukturen zur Verfügung. Diese müssen genutzt werden, wenn die Daten nicht randomisiert und unabhängig voneinander erhoben wurden (ter Braak 1998):

- (1) Blöcke: Blöcke sind Gruppen von Objekten, die zueinander in Beziehung stehen. Sie können durch Kovariablen definiert oder durch eine Klassifikationsregel in einem File eingelesen werden. Die Permutationen erfolgen dann innerhalb der Blöcke, während Objekte aus verschiedenen Blöcken nicht ausgetauscht werden (ter Braak 1998).
- (2) Zeitreihen oder Transekte: Datenstrukturen dieser Art dürfen nicht zufällig permutiert werden, da man eine Autokorrelationsstruktur zwischen den Beobachtungen sowohl in den Antwort- als auch in den Erklärungsvariablen vermuten muss. Das Muster sollte nicht verändert werden, da der Test auf die Beziehungen zwischen den Arten und der Umwelt und nicht auf die Beziehungen innerhalb des jeweiligen Datensatzes zielt. Die Daten werden daher in einer Art zyklischen Struktur fixiert und die Bänder werden gegeneinander rotiert (ter Braak 1998).
- (3) Raster: In diesem Falle wird wie mit einem Transekt verfahren. Man stellt sich das Raster als eine "Matte" vor, die an ihren Enden zusammengeführt wird, so dass eine Rolle entsteht. Die Permutationen erfolgen durch das Rotieren der Rollen.
- (4) Hierarchische Blöcke (Split-Plot-Design): Dieses Modell lässt sich am besten generalisieren und den Datenstrukturen anpassen. Es existieren zwei hierarchische Ebenen (vgl. Abschnitt 3.1.2): die sogenannten "whole plots" umfassen die "split plots". Als Beispiel können Probeflächen innerhalb verschiedener Straten genannt werden. CANOCO<sup>®</sup> bietet Optionen, nur die übergeordneten Blöcke, die untergeordneten Elemente oder beide Hierarchieebenen zu permutieren. Es können in dieser Anwendung keine Blöcke unglei-

cher Größe, d.h. mit nicht identischer Anzahl an untergeordneten Elementen, getestet werden, was bereits in der Versuchsplanung mit berücksichtigt werden sollte.

In dem CANOCO®-Menü "number of split-plots in a whole plot" und "split-plot ordering in data table" wird die Regel erstellt, wie die Hierarchieebenen definiert sind. Bei 36 Probeflächen zu drei homogenen Blöcken wird im ersten Menü die zwölf gewählt und im zweiten die Selektionsreihenfolge im Datensatz festgelegt. Liegen die Blöcke hintereinander, lautet die Regel: Nimm zwölf, überspringe null. Als Resultat finden sich die ersten zwölf Probeflächen im ersten Block wieder. Um die Effekte der Erklärungsvariablen, welche innerhalb der Blöcke variieren, zu testen, werden die Blöcke fixiert (no permutation on whole plot level) und nur die untergeordneten Elemente permutiert (freely exchangeable on split plot level). Die übrigen Kombinationsmöglichkeiten korrespondieren zu den bereits genannten räumlichen und zeitlichen Optionen. Eine Zusatzoption kann aktiviert werden, die den Test anpasst, je nach dem ob die untergeordneten Elemente in einem Block unabhängig oder abhängig voneinander sind (ter Braak 1998).

Eine Möglichkeit, nicht nur die erste Ordinationsachse auf ihre Signifikanz zu prüfen, beschreiben Leps & Smilauer (1999, unveröffentlicht). Man etabliere ein neues CANOCO®-Projekt mit den identischen Einstellungen des Originalprojekts, setze jedoch die Positionskordinaten der Probeflächen (SamE: sample scores) für die erste Achse – Linearkombinationen der Erklärungsvariablen – als Kovariable. Dies ist problemlos, solange keine weiteren Kovariablen in der Analyse sind, da dann einfach der sol-File als Kovariablenfile eingelesen werden kann. Zuvor muss allerdings im File eine Modifikation vorgenommen werden: Man ersetzt den Text "Samp: Sample scores" durch "Xamp: Xample scores", so dass die Tabelle beim Öffnen durch CANOCO® übersprungen wird und die Koordinaten der Tabelle "SamE: Sample scores which are linear combinations of environmental variables" eingelesen werden. Zum Signifikanztest der zweiten Achse wird "AX1" als Kovariable gesetzt, für die dritte Achse "AX1" und "AX2" und so weiter (Leps & Smilauer 1999, unveröffentlicht). Wird der sol-File des Originalprojekts noch benötigt, um in CanoDraw bearbeitet zu werden, ist die Modifikation rückgängig zu machen, da es sonst zum Programmabsturz kommt.

### 3.6.2 Gesichter des Ordinationsdiagramms

**"THE CANODRAW PROGRAM SUPPLEMENTS CANOCO BY DISPLAYING THE RESULTS OF ANALYSIS (...) QUICKLY AND CONVENIENTLY, BY ENABLING INTERACTIVE WORK WITH YOUR DATA, AND BY PRODUCING PRESENTATION QUALITY PLOTS REQUIRED TO DOCUMENT YOUR FINDINGS AND CONCLUSIONS... (SMILAUER 1992, S. 2)."**

Die Erzeugung von Ordinationsdiagrammen basiert auf den Funktionalitäten der Grafiksoftware CanoDraw® (Version 3.10, Smilauer 1992). Sie ist eine direkte Ergänzung der Statistiksoftware CANOCO® und ermöglicht eine Vielzahl von Analysen und Kontrollen der Rohdaten sowie der Ordinationsergebnisse selbst (vgl. auch Abschnitt 3.6.4).

Die Inhalte des eigentlichen Ausgabediagramms einer Ordination können verschieden kombiniert werden. Es stehen drei Stufen zur Verfügung:

1. Scatterplots (vgl. z.B. Abbildung 5.4-6, S.174): Es wird ein Ordinationsdiagramm mit nur einer Informationsbasis erzeugt. Die Auswahl kann auf Arten, Probeflächen oder Erklärungsvariablen gesetzt werden.
2. Bi- oder Jointplots (vgl. z.B. Abbildung 5.4-7, S.175): Es werden zwei Informationsebenen miteinander kombiniert (meist Arten oder Probeflächen mit Erklärungsvariablen).

3. Triplots (vgl. z.B. Abbildung 5.4-23, S. 191): Es entsteht ein Diagramm mit drei Informationsebenen: Arten, Probenflächen und Erklärungsvariablen. Da man im Allgemeinen mit einer großen Menge an Einzelinformationen arbeitet, ist dieses Diagramm häufig informationsüberlastet. Es empfiehlt sich zwei Biplots zu erzeugen und diese dann nebeneinander zu stellen.

Die Ordinationsdiagramme sollen ihre Inhalte im reduzierten Ordinationsraum abbilden. Es sind entsprechende Projektionen gefragt, um den Datensatzeigenschaften gerecht zu werden.

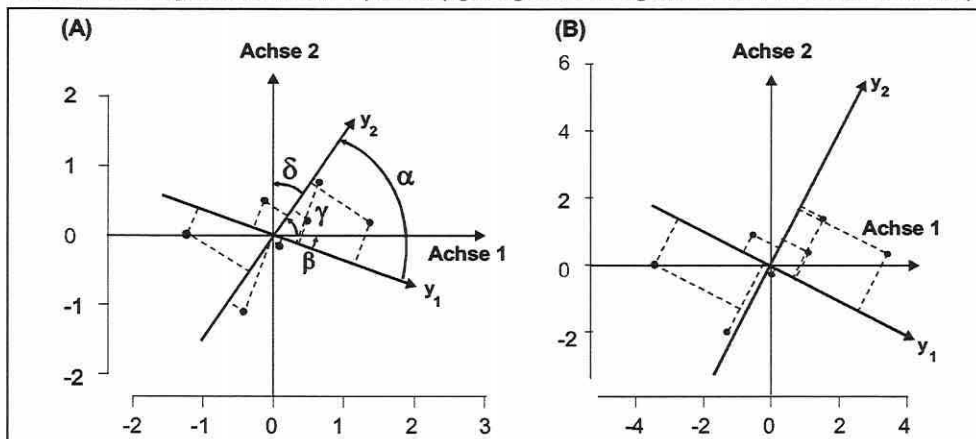
#### Ordinationsdiagramme linearer Verfahren

Es gibt zwei Formen des Biplots zur Repräsentation von PCA-Ergebnissen: den Distanz- und den Korrelationsbiplot (vgl. Abbildung 3.6-3).

Der Distanzbiplot basiert auf der Matrix  $U$  (normierte Eigenvektoren der Länge eins; Koordinaten der Deskriptorachsen) und der Matrix  $F$  (Koordinaten der Objekte; Hauptkomponenten skaliert zur jeweiligen Varianz  $= \lambda_k$ ).

Im Gegensatz dazu basiert der Korrelationsbiplot auf der Matrix  $U\Lambda^{1/2}$  (Eigenvektoren skaliert zur jeweiligen Standardabweichung  $\sqrt{\lambda_k}$ ) für die Deskriptoren und Matrix  $G = F\Lambda^{-1/2}$  für die Objekte. Die Spalten der Matrix  $G$  haben einheitliche Varianzen.

Eine gemeinsame Repräsentation der Eigenvektoren mit den Objektkoordinaten-Matrizen in einem Biplot ist möglich, da  $FU' = Y$  und  $G(U\Lambda^{1/2})' = Y$ . Die Produkte rekonstruieren folglich die zentrierte Originaldatenmatrix  $Y$  perfekt (vgl. Legendre & Legendre 1998, ter Braak 1998 u.a.).



**Abbildung 3.6-3: PCA-Korrelationsbiplot (A):** Die Deskriptoren  $y_1$  und  $y_2$  aus Matrix  $U\Lambda^{1/2}$  sind als Pfeil dargestellt (Kovarianzwinkel  $= \alpha$ ). Die Objekte aus Matrix  $G$  sind als Punkte visualisiert. Die lotrechte Projektion der Objekte (gestrichelt) auf die Deskriptorachsen rekonstruiert die Werte der Objekte entlang der Deskriptoren. **PCA-Distanzbiplot (B):** Die Vektorpfeile von Deskriptor  $y_1$  und  $y_2$  der Matrix  $U$  (Eigenvektoren besitzen die Länge eins) sowie die Objekte der Matrix  $F$  sind im Achsensystem abgebildet. Die Interpretation der Objekt-Deskriptor-Beziehungen basiert auf der rechtwinkligen Projektion der Objekte (gestrichelt) auf die Deskriptorachsen.

**Korrelationsbiplot - Projektion im reduzierten Ordinationsraum der Matrix  $U\Lambda^{1/2}$ :** Die Elemente der Eigenvektoren (Spalten) in der Matrix  $U\Lambda^{1/2}$  stellen die Koordinaten für die Projektion der Deskriptoren (Antwortvariablen, species scores) in Matrix  $Y$  an den verschiedenen PCA- bzw. RDA-Achsen im Korrelationsbiplot dar (Abbildung 3.6-3, Teilbild A). Im Falle einer RDA sind die Positionskoordinaten der Objekte im Raum der Erklärungsvariablen (Matrix  $X$ ) reskaliert zu einheitlichen Varianzen durch die Transformation  $Z\Lambda^{-1/2}$ . Die entsprechenden Koordinaten im Raum der Antwortvariablen (Matrix  $Y$ ) sind reskaliert durch  $F\Lambda^{1/2}$ . Die Matrizen  $Z$



und  $U$  sowie  $F$  und  $U$  können gemeinsam in einem Biplot genutzt werden, weil ihre jeweiligen Produkte die Originalmatrizen rekonstruieren:  $ZU' = \hat{Y}$  und  $FU' = Y$ . Die quantitativen Erklärungsvariablen  $x$  werden in der Projektion unter Nutzung der Korrelation von  $X$  mit den Positionskoordinaten der Objekte im Modell (fitted site scores) dargestellt.

Allgemein ist der Kosinus des Winkels  $\alpha_{ik}$  zwischen den Deskriptoren  $y_i$  und  $y_k$  im multidimensionalen Raum verbunden mit ihrer Korrelation  $r_{ik}$ . Es kann gezeigt werden, dass  $\cos(\alpha_{ik}) = r_{ik}$ . Der Winkel ist gleich dem Kovarianzwinkel, weil die Standardisierung der Zeilen zu einheitlicher Länge nur die Länge der Deskriptorachsen verändert hat, nicht jedoch die Position im multidimensionalen Raum. Gleiches gilt für den Winkel zwischen einem Deskriptor und einer Hauptkomponente im multidimensionalen Raum. Diese Korrelationen werden auch als Ladungen bezeichnet und herangezogen, um zu beurteilen, welche Deskriptoren am meisten durch hohe Korrelationen zu den Eigenvektoren (Hauptkomponenten) beitragen (Legendre & Legendre 1998).

Primäre Eigenschaften eines Korrelationsbiplots:

- (1) Die Hauptkomponente wird durch die Deskriptoren mit dem höchsten Beitrag (Ladung;  $\cos \alpha_{ik}$  = Korrelation) zu der Achse mit interpretierbarer Information belegt.
- (2) Die Projektionskoordinaten einer Deskriptorachse definieren die Position der Spitze dieser Achse im reduzierten Ordinationsraum. Die Deskriptoren sollten als Achsen dargestellt werden, da deren Beziehung zueinander durch die Winkel zwischen ihnen definiert wird.
- (3) Deskriptorachsen, die nur kleine Winkel oder Winkel um  $180^\circ$  zueinander aufweisen ( $\cos 180^\circ = -1$ , d.h. eine perfekte negative Korrelation), können im reduzierten Raum als miteinander korrelierte Gruppen identifiziert werden.
- (4) Die Objekte werden rechtwinklig auf die Deskriptorachsen projiziert, um ihre etwaigen Werte entlang dieser einzuschätzen. Im Korrelationsbiplot sind die Distanzen zwischen den Objekten nicht länger Approximationen ihrer euklidischen Distanz.
- (5) Ergänzend in der RDA: Die Winkel zwischen den Variablen  $x$  und  $y$  lassen deren Korrelation erkennen und durch die rechtwinklige Projektion eines Objektes auf eine Erklärungsvariable kann dessen Wert entlang der Variablen abgeschätzt werden. Binär kodierte Erklärungsvariablen repräsentieren die Zentroide der Objekte, welche die Ausprägung "1" (Eigenschaft vorhanden, zugehörig) der Variable tragen. Die Zentroide können ebenfalls im rechten Winkel auf eine Variable  $y$  projiziert werden, um den Zusammenhang abzulesen. Die Distanzen zwischen den Zentroiden bzw. den Zentroiden und individuellen Objekten stellen keine euklidischen Distanzen dar und können nicht interpretiert werden.

Distanzbiplot - Projektion im reduzierten Ordinationsraum der Matrix  $U$ : Die Elemente der Eigenvektoren (Spalten) in der Matrix  $U$  stellen die Koordinaten für die Projektion der Deskriptoren (Antwortvariablen, species scores) an den verschiedenen PCA- bzw. RDA-Achsen im Distanzbiplot dar (Abbildung 3.6-3, Teilbild B, S. 100). In der RDA erhält man die Positionskoordinaten der Objekte im Raum der Erklärungsvariablen (Matrix  $X$ ) aus Matrix  $Z = \hat{Y}U$  und die Positionskoordinaten der Objekte im Raum der Antwortvariablen (Matrix  $Y$ ) aus Matrix  $F = YU$ . Es können die gleichen Matrixkombinationen wie im Korrelationsbiplot herangezogen werden, um die Originalmatrizen zu rekonstruieren. Eine quantitative Erklärungsvariable  $x$  wird in der RDA-Projektion unter Nutzung der Korrelation von  $X$  mit den Positionskoordinaten der Objekte im Modell dargestellt. Jede Korrelation wird dazu multipliziert mit  $\sqrt{\lambda_k}$ /Gesamtvarianz in  $Y$ , wobei  $\lambda_k$  der Eigenwert der korrespondierenden Dimension ist. Diese Korrektur ist notwendig, da in der Skalierung die Varianzen der Positionskoordinaten zwischen den Dimensionen differieren (Legendre & Legendre 1998; ter Braak 1998).

Generell unterscheidet sich die Projektion der Matrix  $U$  von der Projektion der Matrix  $U\Lambda^{1/2}$  dadurch, dass die Eigenvektoren nicht länger zu ihrer Standardabweichung skaliert sind, die

Deskriptorachsen die gleiche Länge haben und rechtwinklig zueinander im multidimensionalen Raum stehen. Die Winkel zwischen den Deskriptoren (Zeilen in Matrix  $U$ ) und den Hauptkomponenten (Spalten in Matrix  $U$ ) beschreiben die Rotationswinkel der Projektion (berechnet als inverser Kosinus der Eigenvektorkoordinaten in Matrix  $U$ ). Im Gegensatz zum Korrelationsbiplot ist es nicht möglich, die Korrelationen zwischen den Deskriptoren darzustellen bzw. abzuleiten, da sie immer rechtwinklig zueinander stehen.

Haupteigenschaften eines Distanzbiplots:

- (1) Die Distanzen zwischen den Objekten sind Approximationen ihrer euklidischen Distanz im multidimensionalen Raum. Sind die Beziehungen zwischen den Objekten im interpretatorischen Sinne wichtig, ist daher nur der Distanzbiplot die geeignete Darstellung.
- (2) Die lotrechte Projektion der Objekte auf die Deskriptorachsen gibt deren etwaige Position entlang des Deskriptors an.
- (3) Die Winkel zwischen den Deskriptorachsen sind bedeutungslos.
- (4) Ergänzend in der RDA: Die Winkel zwischen den Variablen  $x$  und  $y$  beschreiben wie im Korrelationsbiplot deren Korrelation zueinander. Es erfolgt auch hier die lotrechte Projektion zur Einschätzung der Werte entlang der Variablen und die Visualisierung der binär kodierten Erklärungsvariablen ist gleich dem Korrelationsbiplot. Allerdings approximieren die Distanzen zwischen den Zentroiden sowie den Zentroiden und individuellen Objekten in diesem Diagrammtyp deren euklidische Distanz zueinander.

Schlussfolgernd erlaubt der Distanzbiplot die Interpretation der Distanzen zwischen den Objekten im Raum der Antwortvariablen bzw. der Erklärungsvariablen, während der Korrelationsbiplot fokussiert ist auf die Interpretation der Beziehungen zwischen den Antwortvariablen in Matrix  $Y$ .

#### Ordinationsdiagramme unimodaler Verfahren

Die Inhalte der Biplots einer (kanonischen) Korrespondenzanalyse richten sich nach den gewählten Skalierungsoptionen für die Matrizen  $U$  und  $\hat{U}$ .

Projektion im reduzierten Ordinationsraum: Zur Projektion der Matrizen  $U$  und  $\hat{U}$  im reduzierten Ordinationsraum bzw. in einem gemeinsamen Koordinatensystem sind entsprechende Skalierungen der Zeilen- und Spaltenwerte durchzuführen. Die Matrizen  $U$  und  $\hat{U}$  können gewichtet werden durch die umgekehrt proportionale Wurzel der Zeilen- und Spaltenwerte aus den Diagonalmatrizen  $D(p_{+j})^{-1/2}$  und  $D(p_{i+})^{-1/2}$ . Aus Matrix  $U$  leitet sich Matrix  $V$  mit den Positionskoordinaten der Arten (species scores; Skalierung 1) und aus Matrix  $\hat{U}$  Matrix  $\hat{V}$  mit den Positionskoordinaten der Probenflächen (sample scores; Skalierung 2) ab:

**Gleichung 3.6-6:** 
$$V = D(p_{+j})^{-1/2} U$$

**Gleichung 3.6-7:** 
$$\hat{V} = D(p_{i+})^{-1/2} \hat{U}$$

Der Zusammenhang zwischen den Matrizen  $V$  und  $\hat{V}$ , welche die Ordination der Spalten und Zeilen entlang einer Ordinationsachse beschreiben, wird durch die Wurzel des Eigenwertes einer Achse gemessen (anzusehen als Korrelation zwischen den Ordinationen).

Die Matrix  $F$  mit den Koordinaten der Objekte (Zeilen) aus der Kontingenztafel im reduzierten Ordinationsraum der CA erhält man aus der transformierten Matrix der Eigenvektoren  $V$ . Diese umfasst die Positionskoordinaten der Deskriptoren (Spalten) im Ordinationsraum. Die Gleichung für die Ordinationswerte (component scores) wird auf Matrix  $Q$  angewendet und eine Division durch die Zeilengewichte durchgeführt (Gleichung 3.6-8).

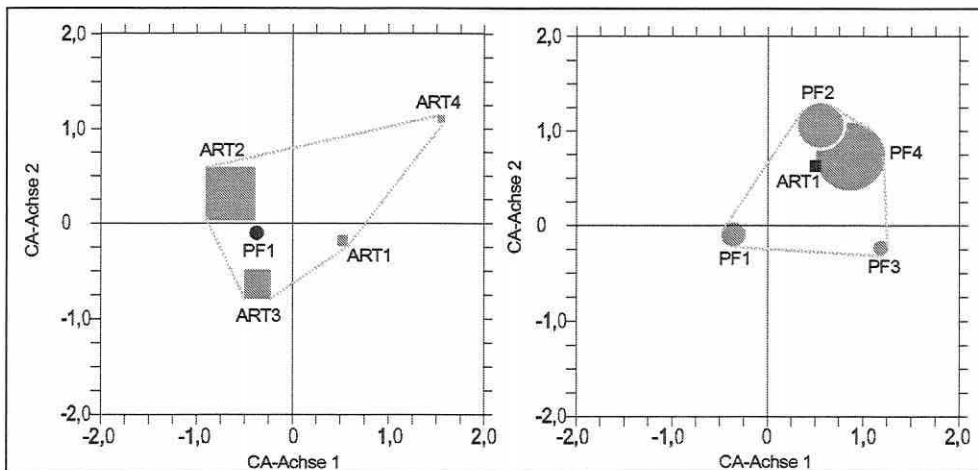
Auf die gleiche Art kann Matrix  $\hat{F}$  abgeleitet werden, welche die Position der Spalten aus der Kontingenztafel definiert. Sie erhält man von der transformierten Matrix der Eigenvektoren  $\hat{V}$ ,

welche die Positionen der Zeilen im Ordinationsraum bestimmt. In diesem Falle wird durch die Spaltengewichte dividiert (Gleichung 3.6-9).

**Gleichung 3.6-8:** (a):  $F = \hat{V} \Lambda^{1/2}$  bzw. (b):  $F = D(p_{i+})^{-1} Q V$

**Gleichung 3.6-9:** (a):  $\hat{F} = V \Lambda^{1/2}$  bzw. (b):  $\hat{F} = D(p_{+j})^{-1} Q' \hat{V}$

In der CCA können die b-Gleichungen nicht genutzt werden, da die Eigenanalyse aus der Kovarianzmatrix berechnet wird, welche aus Matrix  $\hat{Y}$  und nicht aus Matrix  $Q$  hervorgeht. Bei den Skalierungen bilden die Matrizen  $F$  und  $V$  grundsätzlich eine Einheit, so dass die Zeilen gegeben durch  $F$  das Zentroid der Spalten in  $V$  bilden. Umgekehrt formen die Matrizen  $\hat{F}$  und  $\hat{V}$  ein Paar, bei dem die Spalten gegeben durch  $\hat{F}$  die Schwerpunkte der Zeilen in  $\hat{V}$  sind (vgl. Abbildung 3.6-4).



**Abbildung 3.6-4:** Ordinationsdiagramme einer CA. Das Diagramm links repräsentiert den Skalierungstyp 1. Die Probefläche (PF1, Matrix  $F$ ) erscheint im Zentroid der Artsignaturen (Quadratgröße entspricht dem Gewicht des Vorkommens auf PF1, Matrix  $V$ ). Das Bild rechts repräsentiert den Skalierungstyp 2. Die Signatur der Art (ART1, Matrix  $\hat{F}$ ) liegt im Zentroid der Probeflächen (Kreisgröße entspricht dem Gewicht des Vorkommens über die Probeflächen, Matrix  $\hat{V}$ ).

Ergänzend gilt bei der Berechnung einer CCA, dass die Positionskoordinaten der Probeflächen basierend auf den Linearkombinationen der Erklärungsvariablen aus der Matrix  $\hat{Y}$  abgeleitet werden:

**Gleichung 3.6-10:**  $Z_{\text{Skalierung1}} = D(p_{i+})^{-1/2} \hat{Y} U$

**Gleichung 3.6-11:**  $Z_{\text{Skalierung2}} = D(p_{i+})^{-1/2} \hat{Y} U \Lambda^{-1/2}$

Biplots des Skalierungstyps 1 (inter-sample focus) können auf den Matrizen  $F$  und  $V$  oder  $Z_{\text{Skalierung1}}$  und  $V$  (species scores) erstellt werden. Skalierungstyp 2 (inter-species focus) nutzt die Matrizen  $\hat{F}$  und  $\hat{V}$  bzw.  $Z_{\text{Skalierung2}}$  und  $\hat{F}$  (species scores). Aus der CCA gehen letztlich zwei Ordinationen hervor. Die Positionskoordinaten in  $F$  und  $\hat{F}$  entsprechen im übertragenen Sinne einer Näherung zu den Originaldaten eines Streudiagramms, während die Koordinaten in  $Z_{\text{Skalierung1}}$  und  $Z_{\text{Skalierung2}}$  die vorhergesagten Modellwerte darstellen. Gewöhnlich wird die Visualisierung der "Originaldaten" gewünscht. In der CCA verhält es sich umgekehrt: Die  $Z$ -Matrizen beinhalten die durch die Erklärungsvariablen definierten Positionskoordinaten und unterliegen der vollen Einschränkung durch diese (vgl. Palmer 1993). In den Ordinationsdia-

grammen wird deshalb auf die Z-Matrizen zurückgegriffen, obgleich dies in ter Braak (1998) und Smlauer (1992) nicht bzw. nur randlich erwähnt wird.

Haupteigenschaften des Biplots basierend auf den Matrizen  $F / Z_1$  und  $V$  ("Scaling type  $\pm 1$ " / "Focus Scaling on Inter-sample distances" in CANOCO®):

- (1) Es wird ein Diagramm erstellt, in dem die Zeilen der Matrix  $F$  als Zentroide der Spalten in Matrix  $V$  ausgegeben werden, d.h., die von den Arten abgeleiteten Probeflächensignaturen (sample scores) werden im Schwerpunkt der sie definierenden Deskriptoren (Arten; species scores) abgebildet, während die Arten gestreut um den Ordinationspunkt der Probefläche liegen. Diese Skalierung ist optimal, wenn man in erster Linie daran interessiert ist, die Objekte in ihren Beziehungen zueinander zu visualisieren und die Bedeutung der Ordinationsachsen an den Objekten zu messen (Jongman et al. 1987).
- (2) Die Varianz der Probeflächensignaturen entlang jeder Achse reflektiert deren Erklärungsanteil bezogen auf den Eigenwert, während die Varianzen der Artsignaturen identisch bzw. in "Hill's Scaling" etwa identisch sind (ter Braak 1998).
- (3) Als Assoziationsmaß kann die  $\chi^2$ -Distanz (Scaling +1; Biplot scaling) oder entsprechend die generalisierte Mahalanobis Distanz (Scaling -1; Hill's scaling) gewählt werden (vgl. Abschnitt 3.5.2), um Distanzen des Originalraums bestmöglich im reduzierten Ordinationsraum abzubilden (ter Braak 1998). Die Distanzen kennzeichnen den Sachverhalt, dass Punkte, die nahe beieinander liegen, sich bezüglich ihres Musters der relativen Häufigkeiten über die Zeilen sehr ähnlich sind (StatSoft 1997).
- (4) Ergänzend in der CCA: Die Vektoren der Modellparameter aus Matrix  $X$  können mit den Objekten/Variablen in Beziehung gesetzt werden (s.u.).

Haupteigenschaften des Biplots basierend auf den Matrizen  $\hat{F} / Z_2$  und  $\hat{V}$  ("Scaling type  $\pm 2$ " / "Focus Scaling on Inter-species distances" in CANOCO®):

- (1) Die Spalten der Matrix  $\hat{F}$  werden als Zentroide der Zeilen in Matrix  $\hat{V}$  ausgegeben, d.h. die von den Probeflächen abgeleiteten Artsignaturen befinden sich im gewichteten Schwerpunkt der Probeflächen (Probeflächensignaturen), in denen sie auftreten. Die Probeflächen selbst verteilen sich um den Ordinationspunkt der Art. Diese Skalierung ist optimal, wenn der Fokus des Interesses auf den Mustern der Arten zueinander liegt und die Bedeutung der Ordinationsachsen an den Deskriptoren gemessen werden soll (Jongman et al. 1987).
- (2) Die Varianz der Artsignaturen entlang jeder Achse reflektiert dann deren Erklärungsanteil bezogen auf den Eigenwert, während die Varianzen der Probeflächensignaturen identisch bzw. in "Hill's Scaling" etwa identisch sind (ter Braak 1998).
- (3) Auch hier sind die zwei Assoziationsmaße der vorgenannten Skalierung wählbar (Jongman et al. 1987; Legendre & Legendre 1998; ter Braak 1998). Die Distanzen kennzeichnen den Sachverhalt, dass Punkte, die nahe beieinander liegen, sich bezüglich ihres Musters der relativen Häufigkeiten über die Spalten besonders ähnlich sind (StatSoft 1997).
- (4) Ergänzend in der CCA: Die Vektoren der Modellparameter aus Matrix  $X$  können mit den Objekten/Variablen in Beziehung gesetzt werden (s.u.).

Aus den Skalierungsoptionen wird offensichtlich, dass in einem Diagramm nur die Beziehungen zwischen den Deskriptoren (Arten) oder den Objekten (Probeflächen) optimiert dargestellt werden können, nicht aber beides gleichzeitig, da die Ordinationsachsen immer eine lineare Reskalierung der Achsen des umgekehrt skalierten Diagramms sind (ter Braak 1998). Das heißt, die Varianz in einem Positionskoordinatendatensatz definiert den Erklärungsanteil einer Achse, während die Varianzen des anderen Datensatzes entlang der Achse gleich sind. Aus Genanntem resultiert, dass man entweder die Distanzen zwischen den Zeilenpunkten (Probeflächen) oder die zwischen den Spaltenpunkten (Arten), aber nie die Distan-

zen zwischen Zeilen- und Spaltenpunkten interpretieren darf (Clausen 1998). Man kann aber davon ausgehen, dass eine Probefläche, die in Skalierung 2 nahe einem Artpunkt liegt, die Art wahrscheinlicher enthält als eine weit entfernt liegende Probefläche (vgl. Abbildung 3.6-4, S. 103). Anzumerken bleibt hierzu, dass Probeflächen, auf denen jeweils nur eine Art vorkommt, die identische Lage mit der zugehörigen Artsignatur einnehmen, da in unimodalen Verfahren die Probeflächen im Zentroid der auf ihnen auftretenden Arten liegen. Dies zeigt, dass die Korrespondenzanalyse nur die relativen Häufigkeiten/Artmächtigkeiten berücksichtigt. Ungeachtet unterschiedlicher Absolutwerte einer Art wird diese gegenüber dem Gradienten an derselben Position abgebildet. Sind Probeflächen hingegen durch eine Artenkombination von je zwei Arten definiert, bilden sie eine Abfolge zwischen den Signaturen der zugehörigen Arten. Die genaue Lage wird letztlich durch das Verhältnis der relativen Häufigkeiten der Arten zueinander festgelegt.

Generell ergibt die Biplot-Skalierung eine mehr quantitative Interpretationsmöglichkeit durch die Anwendung der Biplot-Regel (Abschnitt 3.6.3) und ist bei kurzen Gradienten mit der Tendenz einer häufigen linearen Antwort zur Achse angebracht. Die Skalierung nach Hill gleicht die durchschnittliche Antwortkurve über alle Achsen an und erlaubt daher die Anwendung der Distanzregel (Abschnitt 3.6.3) bei der Interpretation des Diagramms. Sie ist geeignet für lange Gradienten (>3-4 Standardabweichungen) mit überwiegend unimodalen Antworten. Die Ordinationspunkte der Arten können dann als Optimum der Antwort (ihrer Nische) zum Gradienten angesehen werden (ter Braak 1998).

### 3.6.3 Auswerteregeln für Ordinationsdiagramme

**"IT IS ONE THING TO TRUST THE VALIDITY OF EQUATIONS (...), AND YET ANOTHER TO ENTRUST OUR DATA TO THEM. AN EXAMINATION OF THE CCA ALGORITHM REVEALS THAT IT SHOULD PERFORM WELL. TO THIS ARGUMENT, I RESPOND THAT SEEING IS BELIEVING (PALMER 1993, S. 2225)."**

Abschnitt 3.6.2 erläutert die primären Eigenschaften der verschiedenen Typen von Ordinationsdiagrammen. Im aktuellen Abschnitt werden die drei grundlegenden Auswerteregeln benannt und Übersichten zur vollständigen Interpretation von Ordinationsdiagrammen gegeben.

**Zentroid-Prinzip:** Als Voreinstellung in der CA/CCA (Inter-species distances) liegen die Signaturen der Arten im Zentroid der Probeflächen, auf denen sie erscheinen. Entsprechend umgekehrt verhält es sich, wenn die Option "Inter-sample distances" aktiv ist: Die Probeflächen liegen im Schwerpunkt ihres Artenspektrums (vgl. Abbildung 3.6-4, S. 103). In Fortführung der erstgenannten Einstellung wird die Art im Zentrum ihrer Nische dargestellt. Abbildung 3.6-5 (Teilbild A, S. 106) veranschaulicht dies für die Art *Cirsium arvense*.

Nach der Idee des "Weighted Averaging" (Abschnitt 2.3) ist eine Art am häufigsten in der Nähe ihres Optimums präsent, wenn sie unimodal gegenüber einem Umweltparameter antwortet. Daher ist der durch die Artmächtigkeit/Häufigkeit gewichtete Durchschnitt der Werte einer Umweltvariablen über alle Probeflächen auf der die Art vorkommt, ein begründbarer Indikatorwert für ihre Position entlang des Gradienten (Jongman et al. 1987). In unimodalen Verfahren sind die Positionen der Arten im Diagramm gleichzusetzen mit dem Zentroid der Antwortkurve und speziell bei symmetrischen Kurven, die über ihre volle Spanne beprobt wurden, mit dem Optimum der Kurve (vgl. Abbildung 2.3-6 und Tabelle 2.3-1, S. 18). Arten, die nahe beieinander erscheinen, haben somit ähnliche Lebensraumansprüche und können als ökologische Gruppe angesprochen werden. Diese Tatsache eröffnet die Möglichkeit der Klassifikation von Artengruppen anhand der erklärten Varianz gegenüber einer hypothetischen Variable (CA, DCA) oder einer Linearkombination von Erklärungsvariablen (CCA),



welche die Daten am besten separiert (vgl. z.B. Abbildung 5.4-7 und Tabellen 5.4-23 bzw. 5.4-24, S. 196 und 199). Die Notwendigkeit der Projektion im reduzierten Raum lässt teilweise Verzerrungen auftreten, wodurch die Überprüfung von Gruppierungen durch Klassifikationsverfahren (Clusteranalysen) ratsam ist (vgl. Abschnitt 2.2).

Die Interpretation einer Ordinationsachse hinsichtlich eines unterliegenden Umweltgradienten, also deren Bedeutung, hat in der CA einen eher künstlerischen Charakter, da hier der Nutzer auf sein persönliches Wissen und seine Intuition angewiesen ist. Nur anhand der Artenverteilung im Diagramm können Vermutungen hinsichtlich der das Muster bestimmenden Standortparameter aufgestellt werden. In der CCA ist der Bedeutungsgehalt der Achsen auf die Erklärungsvariablen in Matrix X eingeschränkt (vgl. Abbildung 3.5-5, S. 60). Die Signaturen der Probeflächen zeigen, rechtwinklig auf die Achsen projiziert (Option "Inter-sample distances"), die Lage der Flächen entlang des Gradienten an und lassen so Rückschlüsse auf die achsendifinierenden Standortparameter zu. Dies und der Begriff Ordination (Ordnung) wird in Teilbild B der Abbildung 3.6-5 deutlich.

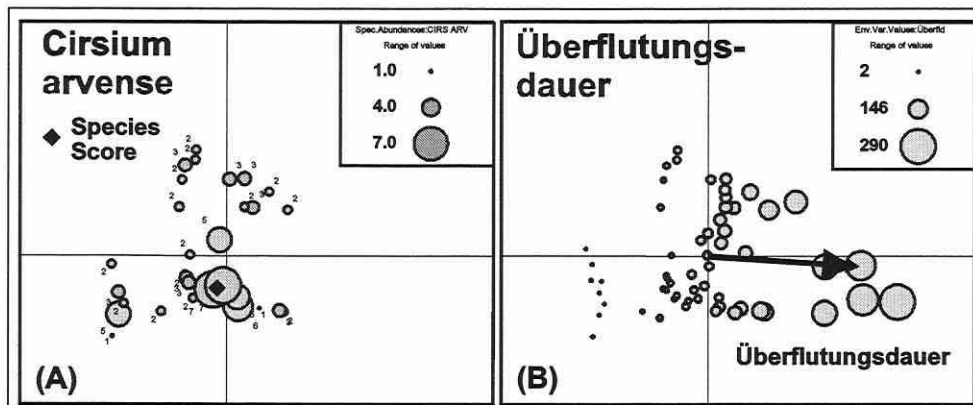


Abbildung 3.6-5: Ordination der Art *Cirsium arvense* (A). Die Signatur der Art (♦) erscheint im Zentroid ihres Vorkommens über die Probeflächen. Ordination der Probeflächen anhand der Standortbedingung Überflutungsdauer (B). In der linken Grafik korrespondieren die Kreisflächen zu den Artmächtigkeiten und in der rechten zu den Werten des Umweltparameters. Ihr Durchmesser steigt mit dem Wert (Rink 1999, S. 19, verändert).

Die Ordination zielt also auf eine Anordnung der Probeflächen oder Arten bezogen auf Umweltparameter bzw. die CCA-Achse. In der Abbildung 3.6-5, Teilbild A, sind die Probeflächen (Kreise) für eine Art korrespondierend zum Achsensystem bzw. in Teilbild B nur nach dem Standortparameter "Dauer der Überflutung" (Pfeil) geordnet. In letzterem lässt sich ein deutlicher Standortgradient visualisieren.

Im Allgemeinen kann "nur" die Zentroid-Regel für Ordinationsdiagramme einer CA/CCA angewendet werden.

**Distanzregel:** Sie ist eine Ergänzung des Zentroid-Prinzips und besagt, dass eine Probeflächensignatur, die nahe einer Artsignatur liegt, wahrscheinlicher die Spezies einschließt als eine weit von dieser entfernte. Eine Rangordnung der Artmächtigkeiten/Häufigkeiten auf den Probeflächen lässt sich somit über die Distanzen zu der Artsignatur für lange Gradienten (ab 3-4 SD) herleiten.

**Biplot-Regel:** Die Positionskoordinaten der Arten in linearen Verfahren (PCA/RDA) beschreiben die Spitze eines Vektors vom Ursprung ausgehend und können als Steigungsparameter einer Geraden angesehen werden. Man verbindet den Ursprung des Diagramms mit den Positionskoordinaten und der resultierende Pfeil zeigt in die Richtung des stärksten Anstiegs der durch das Modell approximierten Artmächtigkeit/Häufigkeit der Art. In der orthogo-

nalen Richtung ist der Wechsel in den Artmächtigkeiten/Häufigkeiten gleich null. Der Vektor kann als Achse gesehen und auf die dem Ursprung gegenüber liegende Seite verlängert werden. Durch die rechtwinklige Projektion der Probeflächensignaturen auf den Vektor einer Art können diese einer Rangordnung zugeführt werden. Das Prinzip der Achsenbildung durch den Ursprung gilt auch für die Probeflächensignaturen, um Arten auf die konstruierte Gerade zu beziehen. Es resultiert die Rangordnung der approximierten Artmächtigkeit/Häufigkeit der Art für die Probefläche. Die Biplot-Regel ist eigentlich nur für Verfahren gültig, die auf einem linearen Modell beruhen (PCA, RDA). Sie kann aber auch in Kombination mit dem Zentroid-Prinzip bei unimodalen Verfahren Anwendung finden, falls die Gradientenlänge kurz ( $< 2\text{-}3\text{ SD}$ ) und die Biplot-Skalierung aktiv ist (ter Braak 1998).

#### **Was lässt sich aus Ordinationsdiagrammen in linearen Verfahren lesen?**

- (1) **Probeflächen/Arten:** Projiziert man die Probeflächensignaturen auf die Achsen der Arten, kann die Ordnung der Artmächtigkeit/Häufigkeit einer Art über die Erhebungseinheiten approximiert werden (PCA). Das gleiche Prinzip in einer RDA angewandt, approximiert die Artmächtigkeit/Häufigkeit bezogen auf die Linearkombination der Erklärungsvariablen, d.h., die Modellwerte anstatt der Originalwerte.
- (2) Werden die Art-Daten zentriert, hat jede Art ihre durchschnittliche Artmächtigkeit/Häufigkeit im Ursprung des Diagramms, und Probeflächen, die zur Spitze des Artenpfeils hin verschoben sind, beherbergen die Art mit einer überdurchschnittlichen Artmächtigkeit/Häufigkeit im Gegensatz zu Probeflächen, die zur entgegengesetzten Seite hin verschoben sind. Werden die Arten nicht standardisiert, zeigt die Länge ihrer Vektoren die unterschiedlichen Artvarianzen an.
- (3) "Focus on sample distances": Die Distanzen zwischen den Probeflächensignaturen approximieren die euklidischen Distanzen der Originaldaten.
- (4) "Focus on species correlations": Der Kosinus des Winkels zwischen zwei Artenachsen ist eine Näherung des linearen Korrelationskoeffizienten zwischen den Arten.
- (5) **Erklärungsvariablen:** Durch die rechtwinklige Projektion der Probeflächensignaturen auf die Achsen der Erklärungsvariablen können deren Werte für die jeweilige Variable eingeschätzt werden. Die Erklärungsvariablen in Matrix X werden in direkten Ordinationen automatisch standardisiert, bevor die Analyse durchgeführt wird, so dass eine Probeflächensignatur nahe dem Ursprung immer zu dem durchschnittlichen Wert der Erklärungsvariablen korrespondiert.
- (6) "Focus on sample distances": Jede Achse einer Erklärungsvariablen lässt sich unabhängig als in die Richtung zeigend interpretieren, in welche die Probeflächensignaturen sich bewegen würden, falls die Werte der zugehörigen Erklärungsvariablen ansteigen. Die Länge eines solchen Vektors ist im Variablenvergleich ein Maß für die Intensität des Effekts.
- (7) "Focus on species correlations": Der Kosinus des Winkels zwischen zwei Erklärungsvariablen ist auch hier als Approximation des Korrelationskoeffizienten zu sehen. In der RDA ist die Schätzung jedoch schlechter als in einer PCA für die abiotischen Daten. Werden Arten und Erklärungsvariablen miteinander verglichen, gilt in beiden Skalierungsoptionen, dass sie positiv korrelieren falls sie in die gleiche Richtung zeigen.

#### **Nominalskalierte Variablen:**

- (8) Die Punktsignaturen lokalisieren die Zentroide der Probeflächen im kanonischen Raum, welche für die nominale Variable bzw. die nominalskalierte Klasse einer kategorialen Variablen den Wert eins annehmen.

- (9) Projiziert man die Signatur einer nominalen Erklärungsvariablen auf eine der Artachsen, kann man die etwaige Artmächtigkeit/Häufigkeit der Art für die Klasse der Variablen abschätzen.
- (10) Bezieht man die Projektion in (12) auf eine quantitative Variable – d.h. deren Achse – so lassen sich die durchschnittlichen Werte der Variablen in der Klasse der kategorialen Variablen einschätzen.
- (11) "Focus on sample distances": Die Distanz zwischen den Zentroiden nominaler Variablen entspricht der euklidischen Distanz und trifft inhaltlich eine Aussage, wie ähnlich die Artenkompositionen der Klassen sind.
- (12) Für beide Skalierungstypen gilt, dass die Distanz zwischen den Zentroiden zugehöriger Klassen und einer Probeflächensignatur deren Wahrscheinlichkeit zu der Klasse zu gehören angibt. Die Fläche ist zu der nächstgelegenen Kategorie zu zählen.

### **Was lässt sich aus Ordinationsdiagrammen in unimodalen Verfahren lesen?**

- (1) **Probeflächen/Arten:** Es wird generell die Optimumposition einer Art entlang eines hypothetischen (CA) oder des durch Erklärungsvariablen definierten Gradienten (CCA) ausgegeben. Das Vorkommen der Art nimmt symmetrisch in alle Richtungen von ihrer Punktsignatur ab. Da die Unähnlichkeiten zwischen den Probeflächen auf Distanzmaßen beruhen, die nur die relativen Artmächtigkeiten/Häufigkeiten einer Art berücksichtigen, sind Probeflächen mit Beispielswerten von 3, 9, 3 und 300, 900, 300 identisch in ihrer Position im Diagramm.
- (2) Es gilt das Zentroid-Prinzip: Die Artsignaturen liegen daher beim Fokus auf den Artdistanzen den Probeflächen am nächsten, auf denen sie mit ihrer höchsten relativen Artmächtigkeit/Häufigkeit auftreten. Beim Fokus auf den Probeflächendistanzen befinden sich die Flächen im Zentroid ihrer Artenkomposition.
- (3) In Erweiterung des Prinzips unter (1) ist die Distanzregel für lange Gradienten attraktiv, da die Probeflächen eine Ordnung erfahren, indem die Artmächtigkeiten/Häufigkeiten einer Art mit der Entfernung von ihrer Signatur abnehmen.
- (4) Für kurze Gradienten kann die Biplot-Regel angewandt werden, falls die Biplot-Skalierung aktiv ist.
- (5) "Focus on sample distances": Wird die Option in der Kombination mit der Skalierung nach Hill angewendet, erscheinen die Distanzen zwischen den Probeflächensignaturen als Standardabweichungen (SD) des Artenwechsels (Turnover) entlang des Gradienten. Flächen, die weiter als 4 SD voneinander entfernt liegen, haben mit großer Wahrscheinlichkeit keine Arten mehr gemeinsam.
- (6) "Focus on species distances": Zusammen mit der Biplot-Skalierung werden die  $\chi^2$ -Distanzen und mit der Skalierung nach Hill die generalisierten Mahalanobis Distanzen zwischen den Artsignaturen repräsentiert.
- (7) Erklärungsvariablen: Projiziert man die Arten rechtwinklig auf die Achsen der Erklärungsvariablen, zeigen sich die Optima bezogen auf die Werte der standardisierten Variablen.
- (8) Die Interpretationsmöglichkeiten der Erklärungsvariablen-Achsen zueinander sind identisch zu Punkt (6) und (7) der linearen Verfahren.

### **Nominalskalierte Variablen:**

- (9) Die Distanz zwischen dem Zentroid einer nominalen Variablen und einer Artsignatur schätzt die relative Gesamtartmächtigkeit/-häufigkeit der Art summiert über alle Flächen in der Probeflächenklasse.
- (10) Der Vergleich von Probeflächensignaturen zu den Zentroiden der nominalskalierten Erklärungsvariablen und zwischen den Zentroiden und quantitativen Erklärungsvariablen entspricht den Prinzipien in linearen Verfahren.

### 3.6.4 Diagramme zur weiterführenden Datenexploration

"(...) THE PROCEDURES ENABLING YOU TO INVESTIGATE THE PROPERTIES OF YOUR MULTIDIMENSIONAL DATA, EITHER IN RELATION TO THE ORDINATION RESULTS (...), OR WITHOUT REFERENCE TO THE ORDINATION PERFORMED (SMILAUER 1992, S. 86)."

Die weiterführende Datenexploration bietet viele Möglichkeiten, die Daten und Ergebnisse der Ordinationen aus verschiedenen Blickwinkeln zu beleuchten. Gerade weil die Ordinationsdiagramme es nicht erlauben, direkt auf die Originaldaten zurückzuschließen, da die Erklärungsvariablen standardisiert und die biotischen Daten oft transformiert werden, bieten alternative Darstellungen ein großes interpretatorisches Potential. Neben der Visualisierung der Originaldaten im Ordinationsraum können die Daten auch auf Klassifikationen bezogen, ihre Verteilung mit der geografischen Raumlage oder jedem anderen Raum in Beziehung gesetzt werden. Gerade die Werkzeuge zur Klassifizierung von Daten in Canodraw® bilden ein effektives Instrument bei der Datenexploration. Es können maximal 16 Klassen etabliert werden (vgl. Smilauer 1992).

#### Vorkommens- und Beteiligungsdiagramm

Das Vorkommensdiagramm (presence plot) stellt die Anzahl der Arten einer vorher festgelegten Gruppe auf den Probestellen im Ordinationsraum als Kreisflächen dar. Die Flächengröße spiegelt deren Zahl optisch wider (Abbildung 3.6-6). Auf diese Weise können die Vertreter ökologischer Gruppen in ihrer Konzentration auf Teilbereiche des Ordinationsraums begutachtet werden (vgl. hierzu auch Abbildung 5.4-32 und Abbildung 5.4-33 auf Seite 200f).

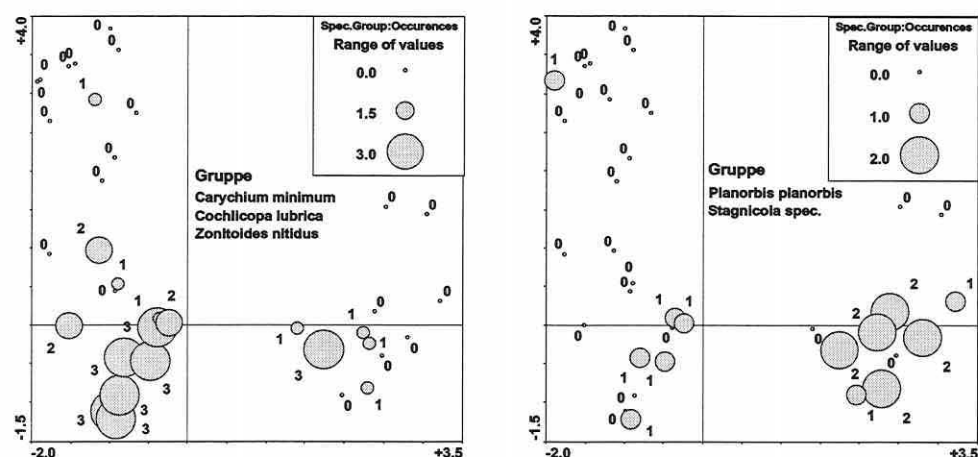


Abbildung 3.6-6: Vorkommensdiagramm für die ökologischen Gruppen "Tieferliegende Feuchtfleichen" (links) und "Tiefere Rinnen, meist länger mit Wasser bestanden" (rechts) der Indikatorgruppe Mollusken (vgl. Abschnitt 5.4.3).

Das Beteiligungsdiagramm (participation plot) ist nur insofern inhaltlich verschieden, als es den Prozentanteil der gewählten Gruppe, gemessen an der Artenkomposition, pro Probestelle darstellt. Die Abbildung 3.6-7 (S. 110) lässt im Vergleich zum Vorkommensdiagramm eine exaktere Einschätzung der Dominanz von Arten/Gruppen auf den Probestellen zu. Es fällt auf, dass die Schirmarten der Gruppe "Tieferliegende Feuchtfleichen" deutlichere Anteile an ihrem Lebensraum besitzen als die zum Vergleich dargestellten Vertreter der Gruppe

"Tiefere Rinnen, meist länger mit Wasser bestanden". Bei der Interpretation ist allerdings die unterschiedliche Artenanzahl pro Gruppe zu berücksichtigen.

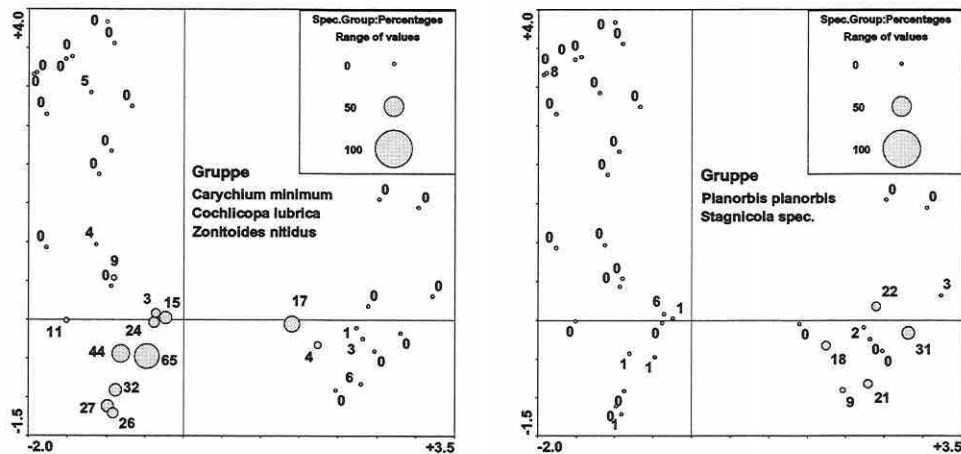


Abbildung 3.6-7: Beteiligungsdiagramm für die ökologischen Gruppen "Tieferliegende Feuchtfleichen" (links) und "Tiefere Rinnen, meist länger mit Wasser bestanden" (rechts) der Indikatorgruppe Mollusken (vgl. Abschnitt 5.4.3).

#### Diagramm zur Beurteilung der klassifizierten Artenkomposition auf Probeflächen und Diagramm zur Partizipation von Arten an den Probeflächenausprägungen

Das Diagramm zur Beurteilung der klassifizierten Artenkomposition (ökologische Gruppen) stellt die Anteile vorgegebener Klassen auf den jeweiligen Probeflächen als "Kuchendiagramme" dar. Man stelle sich vor, dass die Artenkomposition entlang der ersten Ordinationsachse maßgeblich durch eine Erklärungsvariable (z.B. Überflutungsdauer) geprägt wird. Eine Klasse kann jeweils trocken, intermediär und nass lebenden Arten zugeordnet werden. Eine vierte Klasse bilden indifferente Arten mit breiter ökologischer Spanne. Die "Kuchendiagramme" sollten dann am einen Ende der Achse das größte Stück der trockenen Klasse und am anderen Ende der nassen zuweisen (vgl. Abbildung 5.4-28, S. 197).

Das Diagramm zur Partizipation von Arten an den Probeflächenausprägungen ist dem erstgenannten sehr ähnlich. Die "Kuchendiagramme" beziehen sich allerdings auf die Arten und stellen deren Anteile an den definierten Klassen prozentual dar (vgl. Abbildung 5.4-27, S. 196). Zur Überprüfung von Häufigkeiten der Klassenzugehörigkeit für einzelne Arten oder Erklärungsvariablen lassen sich ergänzend Histogramme abrufen.

#### Diagramme zur Visualisierung der Originalwerte von Arten oder Erklärungsvariablen über die Probeflächen

Die Diagramme stellen eine Interpretationshilfe des Ordinationsdiagramms dar. Für jede Probeflächensignatur können die Originalwerte der einzelnen Arten oder Erklärungsvariablen visualisiert werden (Abbildung 3.6-8, S. 111).

Die Größe der Kreisflächen gibt einen optischen Eindruck der Verteilung einer Art über den Ordinationsraum, kann aber leicht beim direkten Vergleich der Diagramme täuschen. Man beachte die Variabilität der Originaldaten in den beiden Grafiken der Abbildung 3.6-8 (S. 111). Bei etwa analogen Kreisflächen werden einmal über 80 und einmal nur drei Individuen numerisch ausgegeben.

Eine generalisierte Darstellung der Erklärungsvariablen in Kombination mit Ordinationsdiagrammen liefern die Konturliniendiagramme in Abbildung 5.4-34 bis Abbildung 5.4-36 (S. 201). Anstatt der Absolutwerte pro Probefläche werden Isolinien für eine Erklärungsvariable erzeugt, um den Informationsgehalt in das Ordinationsdiagramm einzufügen.



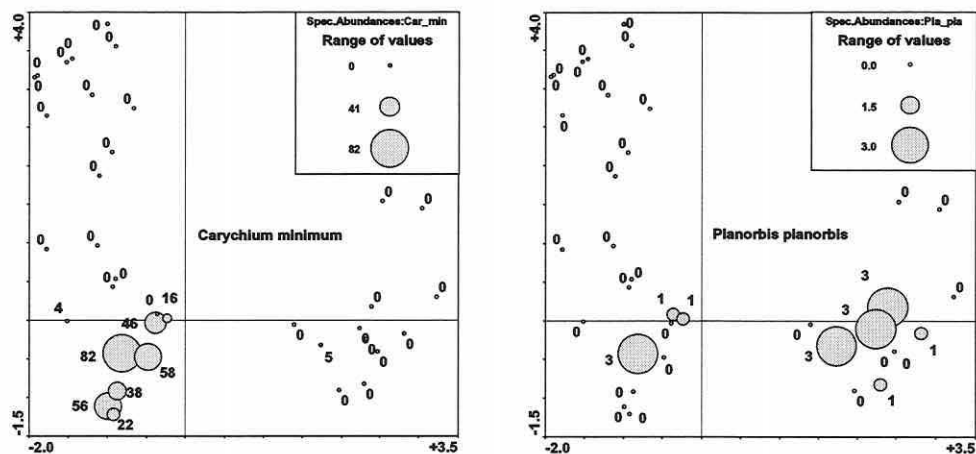


Abbildung 3.6-8: Erhebungsdaten der Molluskenart *Carychium minimum* (links) und *Planorbis planorbis* (rechts).

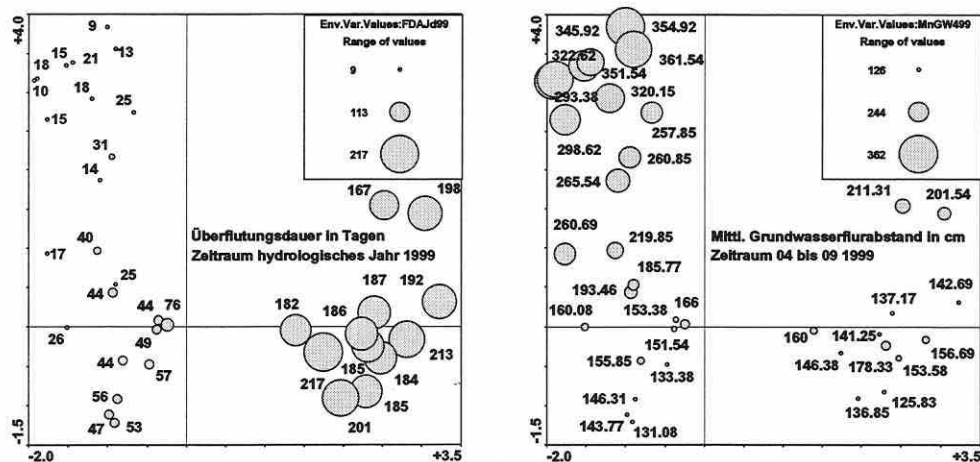
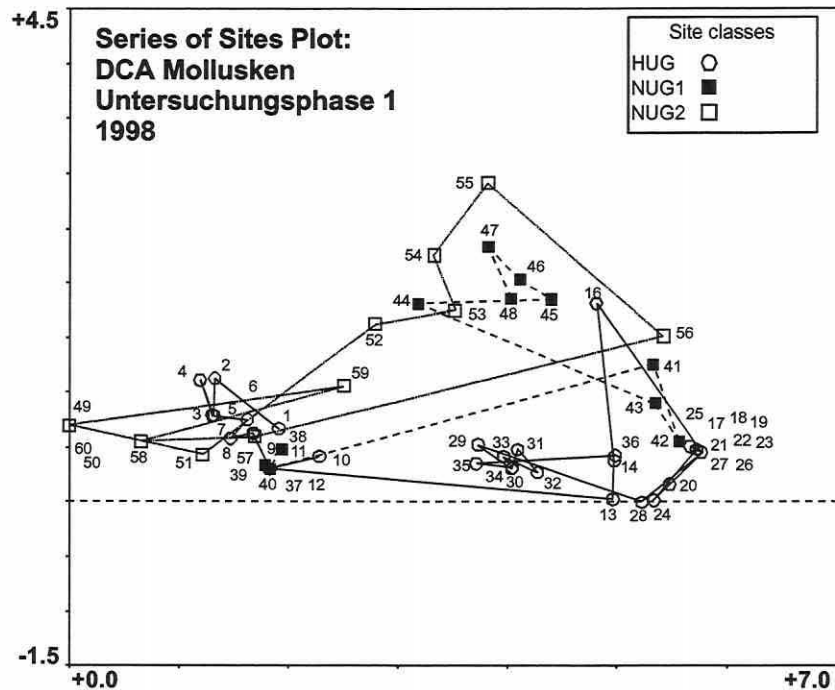


Abbildung 3.6-9: Erhebungsdaten für die Parameter "Überflutungsdauer in Tagen für das hydrologische Jahr 1999" (links) und "Mittlerer Grundwasserflurabstand in Zentimeter über die Monate April bis einschließlich September 1999" (rechts).

### Probeflächendiagramme für Zeit- oder Raumserien

Diese speziellen Diagramme sind in der Lage Zeit- oder Raumserien abzubilden, indem die Probeflächen, welche einer Klasse angehören, durch Linien verbunden sind. Abbildung 3.6-10 (S. 112) zeigt die Raumserien von drei Untersuchungsgebieten, durch die jeweils miteinander verbundenen Probeflächen. An der Lage zur Achse wird deutlich, dass sich NUG2 durch die Verschiebung vieler Flächen nach links in seiner ökologischen Ausstattung der Flächen von den beiden anderen Räumen unterscheidet. Als ergänzendes theoretisches Beispiel stelle man sich vor, dass für die Probeflächen einer Lokalität die Artenkomposition sowie die Erklärungsvariablen über fünf Jahre beobachtet wurden. Die Probeflächen werden nach ihrer Zeitabfolge klassifiziert und im Datensatz geordnet. Das Diagramm hilft dann, die Sukzession innerhalb des Ordinationsraums abzubilden. Technisch limitiert ist die Grafik auf acht Gruppen, da CanoDraw® nicht mehr Linientypen bereitstellt (Smilauer 1992).



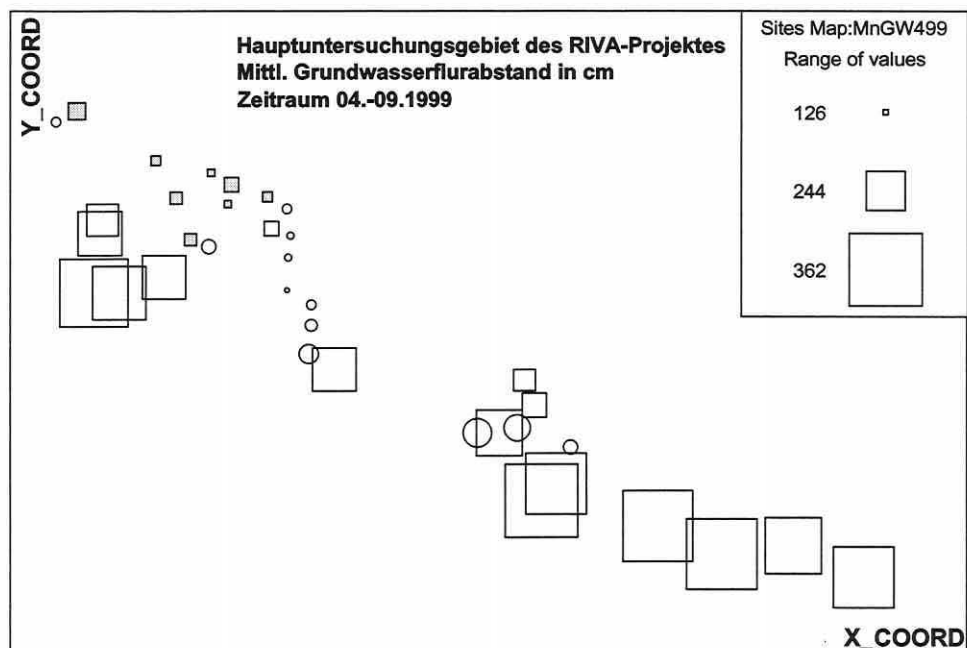
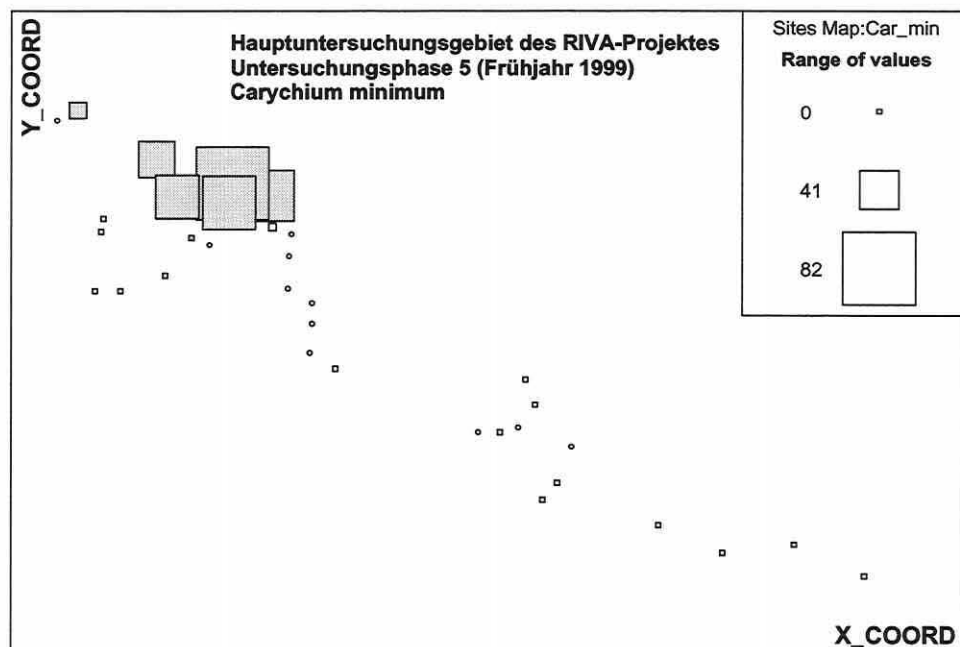
**Abbildung 3.6-10: Diagramm zur Darstellung von Zeit- oder Raumserien. Die Probeflächen der drei Untersuchungsgebiete im RIVA-Projekt sind als Raumserie miteinander verbunden. Die Flächen des NUG2 weichen dabei in ihrer ökologischen Ausstattung deutlich nach links ab.**

#### Raumbezogene Diagramme

Einige der vorgestellten Diagramme können anstatt auf den Ordinationsraum auch auf den realen georeferenzierten Untersuchungsraum bezogen werden (mapping). Die grundlegende Bedingung für einen neuen Raumbezug ist die Existenz eines Datensatzes mit den Lagekoordinaten (CanoDraw®-Map-File; \*.cdm; vgl. Smilauer 1992, S. 108ff).

Die Abbildung 3.6-11 (S. 113) kann mit den Abbildungen 3.6-8 und 3.6-9 (S. 111) verglichen werden, um die Probeflächenverteilung und deren Informationsgehalt mit dem Untersuchungsgebiet in Bezug zu setzen. Eine Plausibilitätsprüfung der Ergebnisse wird so – ohne auf ein Geografisches Informationssystem (vgl. Abschnitt 3.1.5) angewiesen zu sein – deutlich erleichtert.

Für viele der in diesem Abschnitt angesprochenen klassischen Symboldiagramme besteht die Möglichkeit, Isoliniendiagramme durch Kriging zu erzeugen. Diese traditionell in der Geostatistik benutzte Methode zur Regionalisierung von Punktdaten besitzt viele Facetten. Sie ist nicht Gegenstand dieser Arbeit und soll daher nur randlich angesprochen werden. In Jongman et al. (1987) und Legendre & Legendre (1998) finden sich kurzgefasste Einführungen in die Thematik. Smilauer (1992) empfiehlt als Referenz Isaaks & Srivastava (1989). CanoDraw® stellt das "generalized kriging" Verfahren (auch Universal- oder Punkt-Kriging genannt) in Kombination mit einem sphärischen Modell zur Verfügung. Die wählbaren Parametereinstellungen sind in Smilauer (1992, S. 65ff) nachzuschlagen.



**Abbildung 3.6-11:** Darstellung der Erhebungswerte einer Art (oben) und einer Erklärungsvariablen (unten) im georeferenzierten Untersuchungsraum. Die Signaturen symbolisieren die Klassifikation der Probeflächen (Trockenes Grünland: Quadrate; Nasses Grünland: gefüllte Quadrate und Flutrinnen: Kreise).

### 3.7 Generalisierte lineare Modelle

#### 3.7.1 Logistische Regression

*"BEFORE BEGINNING A STUDY OF LOGISTIC REGRESSION IT IS IMPORTANT TO UNDERSTAND THAT THE GOAL OF AN ANALYSIS USING THIS METHOD IS THE SAME AS THAT OF ANY MODELBUILDING TECHNIQUE USED IN STATISTICS: TO FIND THE BEST FITTING AND MOST PARSIMONIOUS, YET BIOLOGICALLY REASONABLE MODEL TO DESCRIBE THE RELATIONSHIP BETWEEN AN OUTCOME AND A SET OF INDEPENDENT VARIABLES (HOSMER & LEMESHOW 1989, S. 1)."*

Bei den statistischen Verfahren der Habitatmodellierung (Lebensraumeignungsmodelle) wird neben den Ordinationsverfahren, speziell für Präsenz-Absenz-Daten, die logistische Regression häufig angewendet (Agresti 1996; Backhaus et al. 1996; Hosmer & Lemeshow 1989; Jongman et al. 1987; Schröder 2000; Smilauer 1992). Durch das Verfahren können Modellanpassungen für einzelne Arten erfolgen, welche die Antwort dieser zu einem Umweltgradienten (univariat oder multivariat) exakter abbilden, als dies mit dem globalen Antwortmodell eines Ordinationsverfahrens (Abschnitt 3.5.4 und 3.5.5) möglich ist. Eine Diskussion der Vor- und Nachteile von Ordination und logistischer Regression findet sich in Abschnitt 4.2.

Die logistische Regression wird in Kapitel 5 angewandt, um, ergänzend zu den Ordinationsmodellen, Modelle für einzelne Schirmarten (vgl. Abschnitte 2.5 und 5.4.3) des RIVA-Projektes zu erhalten. Beide Modelltypen werden schließlich für die Implementierung in das GIS-basierte Prognosemodell des Projektes bereitgehalten (Abschnitt 5.6).

Als Einleitung und zum Verständnis des Verfahrens wird auf die umfassende Literatur über Regressionsverfahren, ihrer Annahmen, Anforderungen und die Vielzahl statistischer Modelle verwiesen: Backhaus et al. (1990, 1996), Berry (1993), Fox (1991), Langbein & Lichtman (1978), Legendre & Legendre (1998), Sokal & Rohlf (1995, 1996), ter Braak & Looman (1987), Zar (1996) und andere.

Im Rahmen der Modellerstellung sollte berücksichtigt werden, dass ausgehend vom Nullmodell ( $\hat{y} = b_0$ ) die Komplexität immer bis zum Vollmodell erhöht werden kann. Letzteres geht durch jeden einzelnen Datenpunkt der Beobachtungswerte und bildet die Realität nicht vereinfacht ab, wie dies von einem Modell erwartet wird. Es ist überangepasst, da man die Datenpunkte nur durch die gleiche Anzahl an Beta-Koeffizienten belegt hat. Somit liefert es gewöhnlich eine stark fehlerbehaftete Vorhersage der nicht beprobten Gesamtpopulation. Es ist nicht übertragbar (vgl. Abschnitt 3.7.1 und 5.4.4). Umgekehrt vereinfacht das Nullmodell die Zusammenhänge so stark, dass dessen Aussage meist wertlos ist. Die gestellte Aufgabe lautet daher: Das einfachste noch aussagekräftige Modell entlang des Komplexitätsgradienten zu finden (vgl. Abschnitt 3.5.3).

Die erste Erweiterungsstufe des linearen Modells führt zu einer Vorstufe des generalisierten linearen Modells, abgekürzt als GLM. Es zeichnet sich dadurch aus, dass qualitative und quantitative Erklärungsvariablen integrierbar sind. Eine qualitative kategoriale Variable kann durch binär kodierte Dummyvariablen ersetzt werden. Für den stochastischen Teil wird ein Mittel von null und eine konstante Varianz angenommen. Eine wichtige Eigenschaft des Modells ist seine Additivität, d.h. die Effekte der einzelnen Erklärungsvariablen sind im Optimalfall unabhängig voneinander. Erhöht man den Wert einer Erklärungsvariablen um eine Einheit, hat dies eine konstante Auswirkung, welche durch den Regressionskoeffizienten zum Ausdruck kommt. Die Veränderung ist unabhängig von den Werten der übrigen Erklärungsvariablen im Modell, wobei Autokorrelation eine entscheidende Rolle spielt (vgl. Ab-

schnitt 3.1.4 und 3.5.9). Generalisierte lineare Modelle (vgl. Agresti 1996; Hosmer & Lemeshow 1989; McCullagh & Nelder 1997) erweitern dieses Regressionsmodell um zwei wesentliche Aspekte:

- (1) Die Annahme, dass die Erwartungswerte der Antwortvariablen immer gleich der Linearkombination der Erklärungsvariablen sind, wird nicht zu Grunde gelegt. Die Skala der Antwortvariable wird bedingt durch die Skala der Erklärungsvariablen, ausgedrückt durch eine parametrische Funktion, die als Link-Funktion angesprochen wird:

**Gleichung 3.7-1:**  $G(EY) = \eta$

Die Notation  $\eta$  steht für den linearen Prädiktor und ist identisch zur systematischen Komponente:

**Gleichung 3.7-2:**  $\eta = \beta_0 + \sum \beta_j X_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$

Der Vorteil solch einer Link-Funktion liegt darin, von der vollständigen realen Werteskala einer Erklärungsvariablen ( $-\infty$  bis  $+\infty$ ) auf ein sinnvolles Intervall dieser Bezug zu nehmen. Als Beispiel können das nur die positiven Häufigkeitswerte oder ein Intervall von 0 bis 1 für Wahrscheinlichkeiten sein (Smilauer 1992).

- (2) GLMs treffen keine so spezifischen Annahmen über den stochastischen Teil des Modells. Die Varianz muss nicht konstant sein und verschiedene Verteilungen der stochastischen Komponente bzw. der Antwortvariablen können berücksichtigt werden (Hosmer & Lemeshow 1989; Smilauer 1992).

Die Annahmen zur Verteilung und die Link-Funktionen können jedoch nicht frei kombiniert werden. So überträgt die logit-Transformation die reale Werteskala auf ein 0/1-Intervall und ist beispielsweise keine gute Link-Funktion für eine Poisson-Verteilung. Einige typische Kombinationen gibt Tabelle 3.7-1 wieder:

**Tabelle 3.7-1: GLM. Variablentypen, Link-Funktionen und Referenzverteilungen (nach Leps & Smilauer 1999, unveröffentlicht).**

Variablentyp	Link-Funktion	Referenzverteilung
Häufigkeiten, Prozentdeckung, Braun Blanquet-Skala	log	poisson
Wahrscheinlichkeiten, Binärdaten, relative Häufigkeiten	logit oder probit	binominal
Dimensionen, Verhältnisse, pos. physikal. Messdaten (Gewicht, Länge etc.)	invers oder log	gamma
Daten auf der Originalskala inkl. neg. Werte	identity	normal (Gauß)

An dieser Stelle soll zur eigentlichen logistischen Regression zurückgekehrt und näher auf die Methode eingegangen werden.

Die multiple logistische Regression zielt darauf, die Antwort einer dichotomen (binär kodierten) Variablen über eine relevante Kombination von Erklärungsvariablen vorherzusagen, und ermöglicht neben der Vorkommensprognose eine Analyse der Relevanz einzelner Umweltparameter für die Erklärung der räumlich verteilten Artvorkommen. Das lineare Modell kann durch die zusätzliche Quadrierung der Erklärungsvariablen und Aufnahme der Terme in die logistische Gleichung an eine unimodale Antwortfunktion angepasst werden (Agresti 1996; Backhaus et al. 1996; Hosmer & Lemeshow 1989).

Letztlich wird die Wahrscheinlichkeit für das Vorkommen einer Art ausgegeben. Diese Werte definieren dann die Zugehörigkeit der abhängigen Variablen zu den zwei Kategorien Vorkommen oder Nichtvorkommen. Ebenso kann eine Klassifikation der Wahrscheinlichkeiten in mehr als zwei Kategorien erfolgen (Guisan et al. 1999).

Der Erwartungswert  $EY$  bleibt in der logit-Transformation auf ein Intervall von null bis eins beschränkt (Gleichung 3.7-3) und beschreibt umgeformt die Vorkommenswahrscheinlichkeit,



anhand von Gleichung 3.7-4 (Agresti 1996; Hosmer & Lemeshow 1989; Menard 1995; Schröder 2000). Sie stellt die eigentliche Modellgleichung dar, welche bei der Implementierung in ein GIS eingesetzt wird.

**Gleichung 3.7-3:** 
$$\text{logit } E(Y/x) = \text{logit } [P(\bar{x})] = \ln \left( \frac{P(\bar{x})}{1-P(\bar{x})} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

**Gleichung 3.7-4:** 
$$P(\bar{x}) = E[Y/\bar{x}] = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}}$$

Es wird iterativ die Kombination der Regressionskoeffizienten gesucht, welche die Log-Likelihood Funktion maximiert (Maximum-Likelihood-Prinzip), um die bestmögliche Trennung zwischen den Ausprägungen der abhängigen Variablen zu bewirken. Sind die Erklärungsvariablen standardisiert, kann anhand der Höhe der einzelnen Beta-Koeffizienten die Trennkraft einer Variablen abgelesen werden. Sie steigt mit der Zunahme des Wertes (Backhaus et al. 1996). Hohe Beta-Koeffizienten zeigen eine intensive Steigung der logistischen Funktion an und beschränken somit den Wahrscheinlichkeitsübergang auf ein enges Wertintervall.

Zur Einschätzung, wie gut die Erklärungsvariablen trennen, d.h., wie gut das Modell ist, wird in generalisierten linearen Modellen – zu denen die logistische Regression zählt – nicht die Summe der Abweichungsquadrate der Residuen als Gütemaß verwendet, sondern das Log-Likelihood-Kriterium LL (Backhaus et al. 1996; Hosmer & Lemeshow 1989; McCullagh & Nelder 1997; Menard 1995; Schröder 2000). Die Größe LL wird meist mit -2 multipliziert, da sie dann asymptotisch  $\chi^2$ -verteilt ist. Die entstandene Größe erhält die Notation -2LL (-2 ln L). Da die Log-Likelihood negativ ist, wird die -2LL-Statistik positiv und höhere Werte stehen für eine schlechtere Modellanpassung (Menard 1995). Ein optimales Modell hätte den Wert null (Backhaus et al. 1996). Ein Problem der Devianz (-2LL) als Gütekriterium liegt in ihrer Beeinflussbarkeit durch die Gruppengröße. Ein Modell basierend auf einer sehr schiefen Verteilung zwischen den Gruppen wird besser eingeschätzt als ein Modell mit gleicher Gruppenverteilung (Backhaus et al. 1996). Das Problem wird durch andere Maße gelöst, die nicht den Wert null, sondern die Devianz des Nullmodells (alle Beta-Koeffizienten sind gleich null) als Referenz heranziehen. Dies entspricht dem Vergleich der Gesamtvarianz (total inertia) mit der Summe der kanonischen Eigenwerte bzw. der Eigenwerte über die signifikanten Achsen in einer direkten Ordination (vgl. Abschnitt 3.6.1).

Der Likelihood Ratio-Test (LR-Test, Gleichung 3.7-5) wird sowohl für diese Aufgabe genutzt als auch zur Bestimmung, ob die Hinzunahme einer Erklärungsvariablen ins Modell dessen prädiktiven Charakter signifikant verbessert. Im letzteren Fall wird die Devianz vor und nach der Implementierung gegenübergestellt, d.h., das Vollmodell wird gegen das um eine Erklärungsvariable reduzierte Modell getestet (Backhaus et al. 1996).

**Gleichung 3.7-5:** 
$$LR = \left[ D(\text{Modell ohne } x_i) - D(\text{Modell mit } x_i) \right] = -2 \ln \left[ \frac{L(\text{Modell ohne } x_i)}{L(\text{Modell mit } x_i)} \right]$$

Bei ausreichend großem Stichprobenumfang ist die LR annähernd  $\chi^2$ -verteilt, so dass der kritische Wert dieser Verteilung zum Vergleich mit dem Wert aus der LR-Statistik zur Signifikanzprüfung herangezogen werden kann. Multiple Pseudo-Bestimmtheitsmaße werden oft als zusätzliche Güteparameter eingesetzt (Agresti 1996; Backhaus et al. 1996; Hosmer & Lemeshow 1989; Knoke & Burke 1980; Menard 1995; Schröder 2000). Von diesen wird Nagelkerkes- $R^2$  (Gleichung 3.7-6, S. 117) in vorliegender Arbeit präferiert, da es einige Kritikpunkte der anderen Maße – der Wert eins wird auch im Optimalfall nicht erreicht – korrigiert und wie das Bestimmtheitsmaß in der linearen Regression interpretiert werden darf. Es be-

schreibt den erklärten Varianzanteil der abhängigen Variablen durch die unabhängigen (Backhaus et al. 1996).

**Gleichung 3.7-6:**

$$R_N^2 = \frac{1 - \left( \frac{L(\text{Nullmodell})}{L(\text{Modell})} \right)^{2/n}}{1 - (L(\text{Nullmodell}))^{2/n}}$$

Durch die Auswertung der univariaten Modelle für sämtliche Erklärungsvariablen lässt sich der durch die jeweilige Variable erklärbare Varianzanteil an der Antwortvariablen bestimmen. Das Prinzip des vorwärts-/rückwärtsschrittweisen Verfahrens ist sinngemäß identisch mit der Vorwärtsselektion in Ordinationsverfahren (vgl. Abschnitt 3.5.3). Ein deutlicher Vorteil ist, dass in SPSS® nicht nur schrittweise die Aufnahme einer neuen Erklärungsvariablen anhand des voreingestellten Signifikanzniveaus erfolgt, sondern auch danach eine Prüfung festlegt, ob durch die Hinzunahme eine andere Variable wieder aus dem Modell entfernt wird (Kolinearität). Das Signifikanzniveau für den Ausschluss ist wählbar. Der Nachteil liegt in dem voll-automatischen Verfahren. Der Bearbeiter hat keinen Einfluss auf die Auswahl während der Selektion. Das häufig genutzte Signifikanzniveau von 0,05 für den Einschluss und 0,1 für den Ausschluss wird in der vorgelegten Arbeit angewendet (Abschnitt 5.5) und die sehr zeitraubende Variablenselektion für die verschiedenen Regressionsmodelle nicht nochmals komplett, sondern nur für den Variablensubdatensatz aus den Ordinationsmodellen durchgeführt (Vorgehen in Abschnitt 4.2 näher diskutiert).

Die Modelldiskriminierung zur Approximation seiner Trennschärfe wird gewöhnlich durch eine Klassifikationsmatrix in Form einer 2x2-Kontingenztafel (Tabelle 3.7-2, vgl. Abschnitt 5.5) vorgenommen (Legendre & Legendre 1998; Morrison et al. 1998; Schröder 2000; Schröder & Richter 2000). Tabelle 3.7-3 stellt Güteparameter zusammen, die aus der Klassifikationsmatrix abgeleitet werden können (Schröder 2000) und auch in Abschnitt 5.5 Verwendung finden.

**Tabelle 3.7-2: Klassifikationsmatrix in einer logistischen Regression**

2x2 Kontingenztafel					
Prognose		Daten			0: Abwesenheit der Spezies 1: Anwesenheit der Spezies
			1	0	
		1	a	b	a+b
		0	c	d	c+d
			a+c	b+d	n
a: Anzahl korrekter Anwesenheitsprognosen			b: Anzahl falscher Anwesenheitsprognosen (Fehler erster Art)		
c: Anzahl falscher Abwesenheitsprognosen (Fehler zweiter Art)			d: Anzahl korrekter Abwesenheitsprognosen		
(a+c)/n: Prävalenz (Anteil der Vorkommen = Anwesenheit)					

**Tabelle 3.7-3: Güteparameter basierend auf der Klassifikationsmatrix einer logistischen Regression (nach Schröder 2000, S. 30).**

Güteparameter	Belegungen aus der Klassifikationsmatrix in Tabelle 3.7-2
Anteil korrekter Prognosen	((a + d) / n) 100
Sensitivität (Prozentanteil korrekter Anwesenheitsprognosen)	(a / (a+c)) 100
Spezifität (Prozentanteil korrekter Abwesenheitsprognosen)	(d / (b+d)) 100

Die Vorkommenswahrscheinlichkeiten in Tabelle 3.7-2 entstammen der Gleichung 3.7-4 (S. 116) und werden zur Prognose der An- oder Abwesenheit einer Art ( $P_1$  oder  $P_0$ ) den realen Beobachtungen gegenübergestellt, wofür eine Klassifikationsregel benötigt wird:

**Gleichung 3.7-7:** **Prognose  $P_0$  = Abwesenheit (0) falls  $P < P_{krit}$**

**Prognose  $P_1$  = Anwesenheit (1) falls  $P \geq P_{krit}$**

Die Felder b und c in Tabelle 3.7-2 beschreiben den fehlerhaft klassifizierten Anteil der Vorkommenswahrscheinlichkeiten durch das Modell. Sie sind von unterschiedlicher Relevanz und werden nach Morrison et al. (1998) in den Fehler 1. Art und den Fehler 2. Art eingeteilt. Die Anzahl falscher Anwesenheitsprognosen (Fehler erster Art) ist rein modelltechnisch weniger dramatisch, da vielerlei Gründe zu dieser Fehleinschätzung führen können: allein die Methodik der Datenerhebung in Studien kann nicht dafür garantieren, dass die tatsächliche Anwesenheit einer Art auch wirklich belegt wird. Zudem können generell geeignete Lebensräume aufgrund verschiedener Ursachen (Ausbreitungsstrategie der Art, Verinselung des Lebensraums etc.) unbesetzt sein. Letztlich darf auch nicht vergessen werden, dass ein Modell selten perfekt ist. Fehler können durch den nicht vom Modell erklärten Anteil der Variabilität in den Daten oder durch Datenrauschen, speziell wenn der Stichprobenumfang niedrig ist, auftreten. Der erste Fall weist häufig darauf hin, dass ein wichtiger Standortfaktor nicht erfasst wurde oder unzureichend beschrieben wird. Gleiches kann umgekehrt für den Fehler zweiter Art (vgl. Tabelle 3.7-2) gelten. Die Anzahl falscher Abwesenheitsprognosen wird durch Migrationsbewegungen mobiler Tierarten gefördert, die einen potenziell ungeeigneten Lebensraum nur queren oder durch Ausstrahlen von geeigneten Flächen zeitweise in den angrenzenden Bereichen auftreten. Ein typisches Beispiel sind Wanderbewegungen von Carabiden in rezenten Auebereichen, die ihren Aufenthaltsort der Überflutungsdynamik anpassen (vgl. Abschnitt 5.4.2). Weniger mobile Indikatorgruppen sind stärker lebensraumgebunden, und der Fehler ist meist dem Modell selbst zuzuschreiben. Diese Fehleinschätzung ist besonders kritisch, wenn durch anthropogene Eingriffe Lebensraumveränderungen vorgenommen werden und als ungeeignet prognostizierte Flächen der Maßnahme zum Opfer fallen (siehe auch Morrison et al. 1998; Schröder 2000).

Insgesamt ist die Belegung/Fehlerrate der Klassifikationsmatrix (Tabelle 3.7-2) von der Festlegung des Klassifikationsschwellenwertes  $P_{krit}$  (Gleichung 3.7-7) stark beeinflusst (Hosmer & Lemeshow

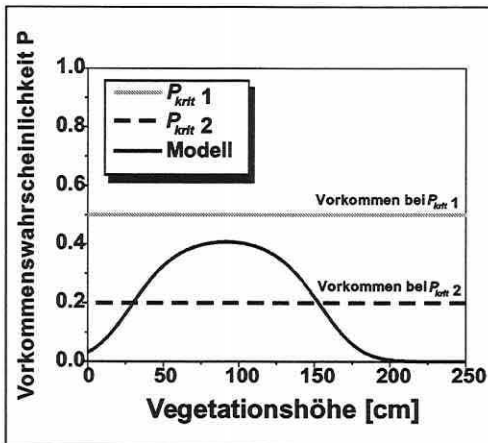


Abbildung 3.7-1: Einfluss des Klassifikationsschwellenwertes auf das Prognoseergebnis.

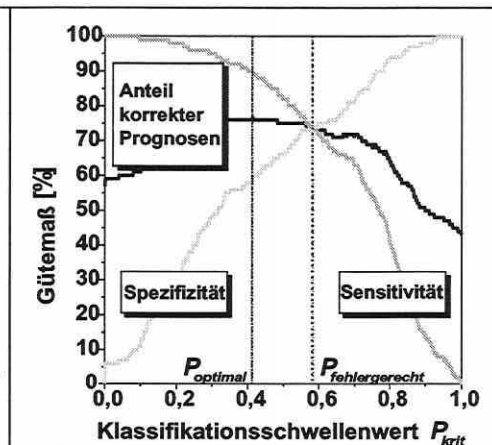


Abbildung 3.7-2: ROC-Kurve und Auswirkung der Schwellenwerte  $P_{opt}$  und  $P_{fehlerrichtig}$  (Schröder 2000, S. 31).

1989; Schröder 2000). Betrachtet man Abbildung 3.7-1 (S. 118) wird die Aussage leicht nachvollziehbar, dass der Klassifikationsschwellenwert neben den Modellparametern selbst eine zentrale Rolle einnimmt. Wird der  $P_{\text{krit}}$  1-Wert gewählt kommt die Art im Vergleich zu  $P_{\text{krit}}$  2 auf keiner Bezugseinheit vor.

Schröder (2000) stellt instruktive Aussagen zur Wahl des Limits verschiedener Autoren vor, von denen die Methode der flexiblen Bestimmung des  $P_{\text{krit}}$  durch eine Technik, die "receiver-operating-characteristic", kurz ROC-Kurve genannt (Abbildung 3.7-2, S. 118), am begründbarsten erscheint (Pearce & Ferrier 2000a; Schröder 2000; Schröder & Richter 2000).

Der Schwellenwert  $P_{\text{opt}}$  maximiert den Anteil korrekter Prognosen, während  $P_{\text{fehlengerecht}}$  so bestimmt wird, dass die Sensitivität und Spezifität (s.o.) den gleichen Wert annehmen. Die ROC-Kurve trägt die beiden Parameter für alle  $P_{\text{krit}}$  des Intervalls von 0 bis 1 gegeneinander auf und betrachtet daher die gesamte Spannweite möglicher Schwellenwerte.

Für die Anwendung der Modelle im späteren GIS-Prognosemodell ist es nicht relevant, welcher  $P_{\text{krit}}$  festgelegt wird, da über die Gleichung 3.7-4 (S. 116) explizit die Vorkommenswahrscheinlichkeiten für jede Bezugseinheit ausgegeben werden. Es erfolgt dann eine klassifizierte Ausgabe der Ergebnisse im GIS (z.B. 0,2er Schritte).

Dennoch soll auch in dieser Arbeit prinzipiell die Bestimmung des fehlergerechten Klassifikationsschwellenwertes zur Modellbewertung verfolgt werden. Er wird aber nicht über die genannte Technik ermittelt, sondern manuell anhand von Kreuztabellen als Näherung an das ROC-Kurven Ergebnis approximiert (vgl. Abschnitt 5.5).

Weiterhin kann die Modellgüte über ROC-Kurven sehr gut eingeschätzt werden, da die Fläche (AUC) zwischen der Kurve und der Diagonale die Klassifikationsgüte unabhängig vom Klassifikationsschwellenwert beschreibt. Der AUC-Wert reicht von 0,5 für ein Zufallsmodell (Abbildung 3.7-3, Teilbild links oben) bis zum Wert 1,0 für ein perfekt angepasstes Modell mit optimalen Trenneigenschaften (Abbildung 3.7-3, Teilbild links unten). Der AUC-Wert von 0,82 in Abbildung 3.7-3 bedeutet, dass in 82% aller Fälle für eine Zufallsflächeneinheit mit nachgewiesener Anwesenheit einer Art  $x$  eine positivere Vorkommenswahrscheinlichkeit berechnet wird als für eine Zufallsfläche mit Abwesenheit der Art (Schröder 2000).

Für eine faire Beurteilung der Modellgüte ist eine Validierung des erstellten Modells an unabhängigen Datensätzen, die zeitlich oder räumlich von den zur Modellerstellung herangezogenen

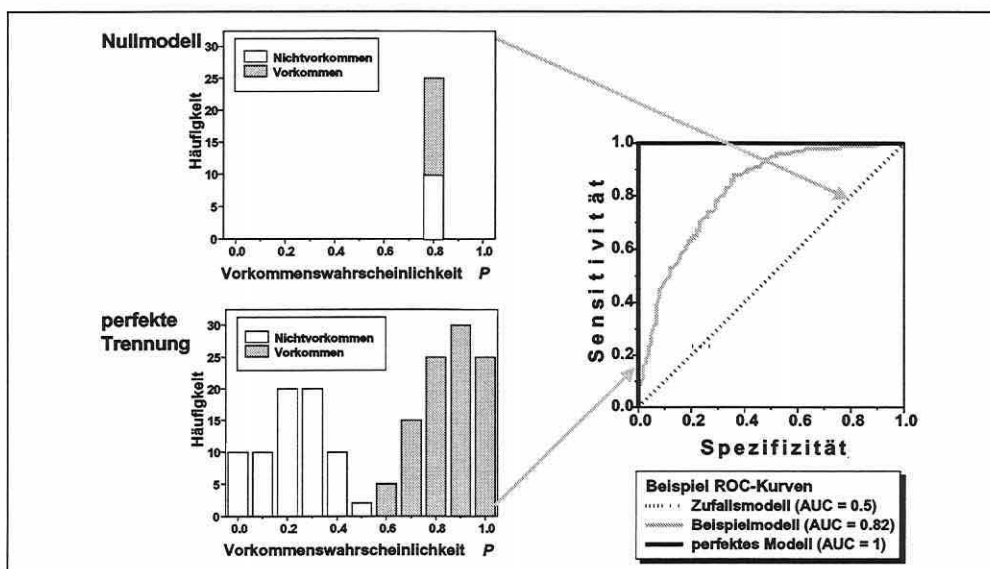


Abbildung 3.7-3: Zusammenhang zwischen ROC-Kurve und AUC-Wert (Schröder 2000, unveröffentlicht).

genen Trainingsdaten abweichen, notwendig und kann auch gleichzeitig dessen Übertragbarkeit veranschaulichen (vgl. Abschnitt 5.5). Die Modellgüte wird generell basierend auf den zur Anpassung verwendeten Modelldaten zu optimistisch eingeschätzt (Fielding & Haworth 1995; Jørgensen 1994; Schröder 2000)

### 3.7.2 Diagramme zur Visualisierung von Antwortkurven gegenüber den Komplexgradienten oder einzelnen Informationsebenen

*"ORDINATION METHODS HAVE CHANGED DRASTICALLY AS A RESULT OF THE RECOGNITION THAT A SPECIES RESPONSE CURVE TO AN ENVIRONMENTAL GRADIENT IS LIKELY TO BE A NONMONOTONIC CURVE AND PROBABLY BELL-SHAPED (AUSTIN 1987, S. 38)."*

Die Positionskoordinaten der Arten in einer unimodalen Ordination repräsentieren den Schwerpunkt ihrer Nischen (vgl. Abschnitte 3.5.5 und 3.5.8), weshalb es erstrebenswert ist, die zugehörige Antwortkurve explizit zu berechnen und auszugeben. In diesen Abschnitt werden daher Diagramme beschrieben, die das Verhalten der Antwortvariablen gegenüber den Ordinationsachsen oder einer Erklärungsvariablen als Antwortkurve visualisieren. CanoDraw<sup>®</sup> ist in der Lage, eindimensionale generalisierte lineare Modelle (GLMs) basierend auf einer Poisson-, Binominal-, Gamma- oder Gauß-Verteilung anzupassen. Es wird automatisch die Link-Funktion für die selektierte Verteilungsfamilie genutzt (vgl. Tabelle 3.7-1, S. 115). Eine CanoDraw<sup>®</sup>-spezifische Einstellung ist, in einem quadratischen Polynommodell mit einem Prädiktor eine Poisson-Verteilung anzunehmen (Smilauer 1992). Daher wird der in einer CCA implementierte unimodale Modellgedanke einer Antwort zum Gradienten repräsentiert und das Optimum sowie die Toleranz geschätzt, falls dies möglich ist.

Grundsätzlich kann der systematische Teil des Modells über ein fixiertes Modell angepasst werden. Das vorbestimmte Modell wird dann ohne Plausibilitätstest aufgebaut. Dies entspricht z.B. der Vorgehensweise im Programmpaket GLIM (Crawley 1993), welches häufiger in der ökologischen Literatur zur Anwendung kommt (Peeters & Gardeniers 1998; Vincent & Haworth 1983 u.a.).

Die Alternative ist die Modellkomplexität schrittweise, basierend auf Devianztests (vgl. Abschnitt 3.7.1), zu ermitteln (Smilauer 1992). CanoDraw<sup>®</sup> führt die schrittweise Regression ausgehend vom Nullmodell durch. Es werden so lange Erklärungsvariablen über den Devianzenvergleich (s. Abschnitt 3.7.1) in das Modell aufgenommen, bis das Stopkriterium, eine signifikante Zunahme des Modellerklärungsanteils, nicht länger erfüllt ist (McCullagh & Nelder 1997; Smilauer 1992). Es besteht hierbei keine Möglichkeit, die Reihenfolge der Aufnahme zu beeinflussen. Im univariaten Fall umfasst der verfügbare Datensatz alle  $x$ - und  $x^2$ -Terme der Erklärungsvariablen; für die Erstellung von Antwortoberflächen (Abschnitt 3.7.3) alle  $x$ -,  $y$ -,  $x \cdot y$ -,  $x^2$ - und  $y^2$ -Terme. Es ist möglich, dass die Interaktionsterme ohne die linearen oder quadratischen gewählt werden (Smilauer 1992).

Es können folgende Diagramme erstellt werden:

- (1) "Ord. axis: env. var. XY plot": Die Werte einer Erklärungsvariablen auf den Probestellen werden gegenüber einer Ordinationsachse dargestellt, d.h., die Beziehung eines Erklärungsparameters zu einer Ordinationsachse kann ausgewertet werden. Zusätzlich wird eine Funktion ausgegeben, welche die angepassten Schätzungen der eigentlichen Probestellenwerte reflektiert.
- (2) "Ord. axis: species XY plot": In diesem Diagramm werden anstatt der Werte einer Erklärungsvariablen die Artmächtigkeiten/Häufigkeiten einer Spezies entlang der Ordinationsachse ausgegeben. Als GLM wird immer ein Modell zweiter Ordnung erstellt. Soweit



möglich wird das Optimum und die Toleranz der Art bezogen auf die gewählte Ordinationsachse berechnet und das Konfidenzintervall ( $p=0,05$ ) für die Optimumschätzung aufgelistet (vgl. Abbildung 5.4-26, S. 195 und Abbildung 5.4-31, S. 198).

- (3) "Env. var.: species XY plot": Es wird ein Streudiagramm mit der kalkulierten Antwortkurve erstellt, welches die Beziehungen zwischen den Artmächtigkeiten/Häufigkeiten einer selektierten Art und den Werten einer Erklärungsvariablen visualisiert. Soweit möglich wird das Optimum und die Toleranz der Art bezogen auf die gewählte Erklärungsvariable berechnet.
- (4) Die Diagramme (2) und (3) können durch eine Gruppierungsfunktion auch für mehrere Arten gleichzeitig erstellt werden (vgl. Abbildung 5.4-11, S. 178).

### 3.7.3 Antwortoberflächen

**"TREND-SURFACE ANALYSIS IS THE OLDEST METHOD FOR PRODUCING SMOOTHED MAPS. IN THIS METHOD, ESTIMATES OF THE VARIABLE AT GIVEN LOCATIONS ARE NOT OBTAINED BY INTERPOLATION, (...), BUT THROUGH A REGRESSION MODEL CALIBRATED OVER THE ENTIRE STUDY AREA (LEGENDRE & LEGENDRE 1998, S. 739)."**

Es ist oft wünschenswert, Antwortoberflächen (trend-surface, response-surface) zu berechnen, um verschiedene Informationen flächig zu betrachten (vgl. Abschnitt 5.4.3). Die Schätzung der Werte erfolgt anhand eines quadratischen Rasters (Grid), dessen Dimension (Zeilen mal Spalten) wählbar ist. Die Konturlinien verbinden Orte gleichen Wertes innerhalb der Oberfläche als Isolinie. Die Software CanoDraw® schlägt Konturlevel für die Isolinien vor, die interaktiv editiert werden können. Der Verlauf der Konturlinien ist streng von der Rasterdimension abhängig. Eine Dimension von 20 erzeugt die Oberfläche anhand von 400 Werten. Die Dimensionalität lässt sich zwar erhöhen, kann jedoch zu einer Destabilisierung führen, wenn die Diskrepanz zwischen den Stützpunkten (z.B. 50 Probeflächen) und der Dimension (z.B. 2500 Werten) zu groß wird. Der sichtbare Effekt sind oszillierende Konturlinien. Es ist ratsam, die Dimension gering zu halten ( $\leq 20$ ) und eine Glättung vorzunehmen (Legendre & Legendre 1998; Smilauer 1992).

In der Ausgabe wird unterschieden zwischen interpolierten Bereichen (durchgehende Linien) und extrapolierten, weniger zuverlässigen Bereichen (gepunktete Linien; vgl. z.B. Abbildung 5.4-34, S. 201). Die attraktivsten Antwortoberflächen bilden die Darstellungen des Vorkommens von Arten oder der Originalwerte von Erklärungsvariablen (z.B. Abbildung 5.4-33 oder Abbildung 5.4-34, S. 201).

In CanoDraw® ist meist nur eine Informationsebene modellierbar, doch durch einen Trick ist die Überlagerung verschiedener Informationsebenen möglich: Die resultierenden Einzeldiagramme werden einfach als Kopien übereinanderprojiziert. Man beachte dabei, dass die Skalen identisch zu wählen sind (Abbildung 3.7-4, S. 122).

Für alle Konturliniendiagramme kann eine Glättungsmethode eingesetzt werden: Der generalisierte Loess-Glättungsalgorithmus (generalized loess smoother) nach Hastie & Tibshirani (1990) baut auf der Standardmethode auf (vgl. Cleveland 1979) und bringt das GLM in den Glättungsalgorithmus ein. Das Verfahren verwendet lokale gewichtete Regressionen und führt zu gewichteten generalisierten linearen Modellen (Smilauer 1992). So werden Antwortoberflächen Binärdaten besser gerecht (Leps & Smilauer 1999, unveröffentlicht). Für nähere Informationen sei auf die Literatur verwiesen (Cleveland 1979; Hastie & Tibshirani 1990; Legendre & Legendre 1998; Smilauer 1992).

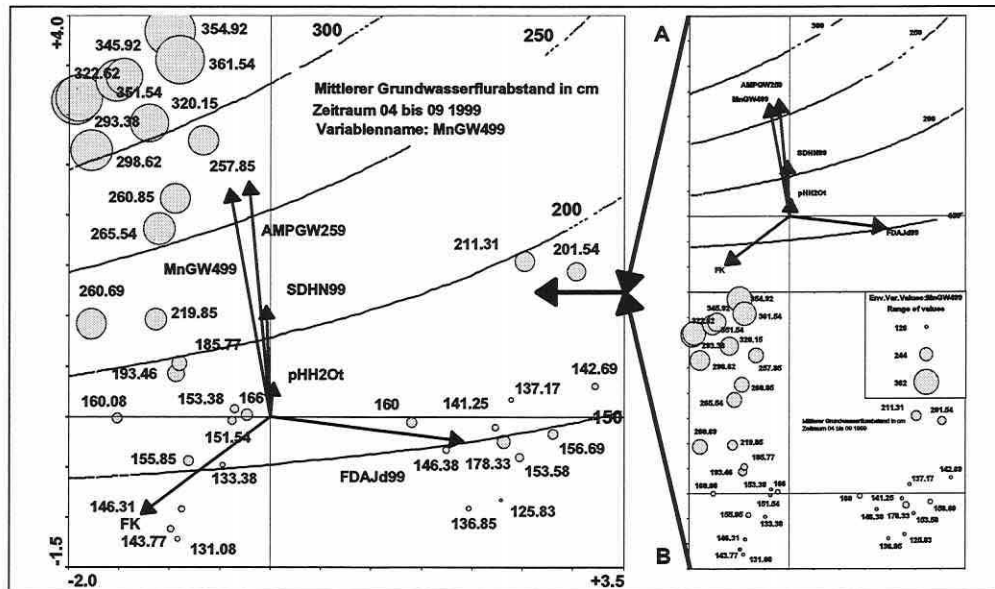


Abbildung 3.7-4: Kombinieren der Informationsebenen A und B: Konturliniendiagramme und Symboldiagramme (vgl. Abbildung 3.6-9, S. 111 und Abbildung 5.4-35, S. 202).

## 4 VERGLEICHENDE BEWERTUNG DER METHODEN

Der erste Teil dieses Kapitels dient der vergleichenden Bewertung der im Methodenteil vorgestellten Ordinationsverfahren und deren Leistungsfähigkeit. Er fasst Anwendungsmöglichkeiten und Kritikpunkte in einem Fazit zusammen, wobei nach indirekten und direkten Verfahren gegliedert wird. Im zweiten Teil erfolgt die Gegenüberstellung von Ordination und logistischer Regression als Techniken zur Erstellung von Lebensraumeignungs-/Prognosemodellen.

### 4.1 Ordination: Gibt es das einzig wahre Verfahren?

*"THE IDEAL ORDINATION METHOD DOES NOT EXIST...  
(PALMER 1998, IDEAL.HTML, S. 1: [HTTP://OKSTATE.EDU/ARTSCI/BOTANY/ORDINATE/](http://okstate.edu/artsci/botany/ordinate/))."*

Die mathematische Logik und Eleganz einer Methode gewährleistet nicht notwendigerweise die Verlässlichkeit ihrer Ergebnisse. Neben der Kenntnis von Arbeitsweise und Berechnungsschritten ist daher auch das Wissen um die Analysekraft bedeutsam. Nur so ist die richtige Einschätzung der Anwendungsmöglichkeiten und Resultate möglich.

Ein fairer Vergleich der Analysekraft von Ordinationsverfahren verlangt nach spezifischen Datensätzen. Felddaten erscheinen hierfür ungeeignet, da die Beurteilung des Verfahrens sich dann eher nach der ökologischen Plausibilität der Ergebnisse richtet. Dies ist subjektiv, solange die wahre unterliegende Struktur unbekannt ist. Es werden simulierte Daten benötigt, welche jegliche Struktur, inklusive Zufallselemente, beinhalten dürfen (Causton 1988; Gauch et al. 1977; Økland 1990; Palmer 1993).

Die Gegenüberstellung der Verfahren anhand simulierter Daten ist selten in der Literatur, dennoch gibt es sie, so dass es in dieser Arbeit nicht notwendig erscheint, die vorliegenden Ergebnisse erneut zu prüfen. Es ist vielmehr von Interesse, die seltenen Literaturstellen zusammenzuführen, offen zu legen und Schlussfolgerungen daraus zu ziehen.

Indirekte Ordinationsverfahren wurden von Gauch et al. (1977) getestet und seine Ergebnisse von Causton (1988) erneut aufgegriffen. Minchin (1987) und Økland (1990) stellten ebenfalls Vergleiche zwischen indirekten Ordinationsverfahren an. Palmer (1993) testete die Robustheit direkter Ordinationsverfahren am Beispiel der kanonischen Korrespondenzanalyse.

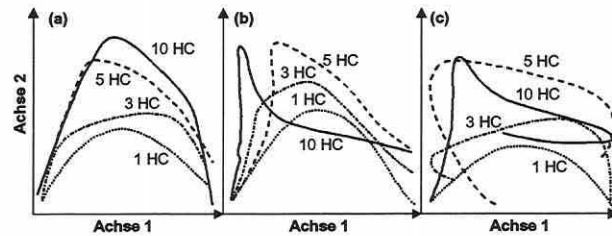
Die Basis aller Tests beruht meist auf Coenoclinesimulation (Gauch 1982a; Gauch & Whitaker 1972; vgl. Abschnitt 2.2). Bei einer perfekten Simulation sind die Antwortkurven von Arten normalverteilt mit homogener Toleranz und die Probeflächen äquidistant entlang des Gradienten verteilt. Die Beta-Diversität (vgl. Abbildung 2.2-2, S. 11) beschreibt die Rate des Artenwechsels und wird hier in "half-changes" (HC) gemessen. Solch eine Einheit definiert sich als Separation entlang des Gradienten, für welche die Ähnlichkeit zweier Probeflächen 50% beträgt (Causton 1988). Das gleiche Prinzip lässt sich auch für eine zweidimensionale Fläche anwenden, in der die Richtungen eine identische oder bewusst verschiedene Beta-Diversität zugeordnet bekommen.

#### Test indirekter Ordinationsverfahren

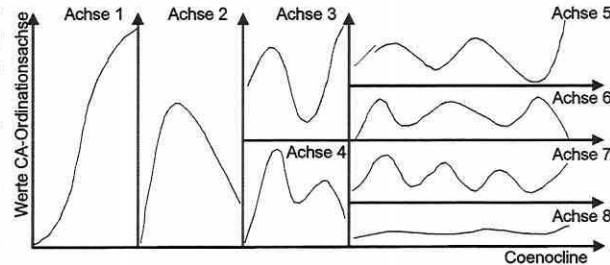
Der Vergleich schließt die Verfahren CA, zentrierte PCA (PCAc) und standardisierte PCA (PCAs) ein. Als Gauch et al. (1977) ihren Test durchführten, war die DCA noch nicht verfügbar. Ihre Analysekraft wird der CA anhand der Untersuchungen von Økland (1990) gegenübergestellt. Die dafür verwendeten Datensätze sind entsprechend in Økland (1990, S. 127) einzusehen.

Die Ergebnisse für Coenoclines verschiedener Beta-Diversität zeigt Abbildung 4.1-1 (S. 124). Das erwartete Ergebnis sollte eine horizontale Linie parallel zur ersten Achse mit äquidistanten Probeflächensignaturen sein. In allen Verfahren zeigt sich eine bogenförmige

Verzerrung bezogen auf die zweite Achse und eine Stauung der Probeflächendistanzen zu den Enden der ersten Achse hin. Der Grad der Verzerrung nimmt mit der Beta-Diversität von der CA über die PCAs zur PCAc hin zu. Letztere zeigt den Hufeiseneffekt (vgl. Abschnitt 3.5.6) mit nach innen gebogenen Enden. Obgleich nur ein Gradient existiert, pflanzt sich die Verzerrung in höhere Dimensionen fort (Abbildung 4.1-2). Das ideale Ergebnis wäre eine Diagonale entlang der ersten Achse und horizontale Linien entlang aller höheren Dimensionen. Es wird deutlich, dass die Dimension und der Grad des Polynoms bis zur achten Achse hin ansteigen, bevor der Einfluss vernachlässigbar wird (Causton 1988; Gauch et al. 1977).



**Abbildung 4.1-1: Ordinationsergebnisse für eine simulierte Coenocline für vier Beta-Diversitätslevel (1, 3, 5 und 10 HC): CA (a), PCAs (b) und PCAc (c) (Causton 1988, S. 215, verändert).**



**Abbildung 4.1-2: Ordinationsergebnis einer CA für eine fünf HC lange Coenocline über acht Dimensionen (Causton 1988, S. 216, verändert).**

Abbildung 4.1-3 (S. 125) visualisiert in den Teilbildern (a) bis (c) die Ordinationsergebnisse einer PCAc für eine 5HC-Coenocline mit 24 äquidistant verteilten Probeflächen. Für die Teilbilder (a) und (b) sollte das Ergebnis eine horizontale Linie sein und für Teilbild (c) müsste ein einfacher Punkt resultieren. Weiterhin bilden die Teilbilder (d) bis (f) Simulationen einer Coenoplane (vgl. Abschnitt 2.2) der Dimension  $8 \cdot 5$  (40 Probeflächen) mit einer Beta-Diversität von  $2,5 \cdot 1,5$  HC ab. Im Idealfall wäre das Ergebnis ein homogenes Raster definiert durch die ersten beiden Ordinationsachsen, welche das Längenverhältnis 2,5:1,5 aufweisen sollten. Die Ergebnisse der Oberflächen aus der ersten und dritten Achse sowie der zweiten und dritten sollten sich als horizontale Linie darstellen. Teilbild (d) ist der Realität schon recht nahe, da die Beta-Diversität gering ist. Das Ergebnis einer CA war ähnlich, aber weniger verzerrt. Für ein Beispiel mit längeren Gradienten ( $4,5 \cdot 4,5$  HC; Abbildung 4.1-4, S. 125) ist das CA-Ergebnis nahezu perfekt (Teilbild a). Die PCAs liefert das Ergebnis in Teilbild (c) und eine PCAc das stark in den Randbereichen nach innen gebogene Ergebnis in Teilbild (b). In einem anderen Vergleich mit den Beta-Diversitäten ( $4,5 \cdot 1,5$  HC) wird die Verzerrung von der langen ersten Achse ausgeübt und verdrängt die Information der zweiten Achse von dieser, d.h., die zweite Achse ist ökologisch nicht aussagekräftig (Causton 1988).

Der Vergleich von CA und DCA wird an den Ergebnissen von Økland (1990) durchgeführt (vgl. auch Minchin 1987). Den Idealfall einer Ordination der simulierten Coenoplane mit  $6 \cdot 1,5$  SD zeigt Abbildung 4.1-5 (S. 125, Teilbild a). Die CA konnte den Gradienten entlang der ersten Achse abbilden, aber es ist auch hier wiederum deutlich ein Bogeneffekt zu erkennen (Abbildung 4.1-5, S. 125, Teilbild b). Die DCA hingegen bildet den ersten Gradienten sehr gut ab und schätzt die Gradientenlänge mit 4,91 SD (real: 6 SD) akzeptabel. Ebenso ist der zweite Gradient zufriedenstellend erfasst und seine Länge mit 1,84 SD (real: 1,5 SD) annehmbar approximiert. Økland (1990) testete weiterhin die Ordinationsergebnisse für eine Coenoplane ähnlicher Beta-Diversitäten ( $2,4 \cdot 2$  SD), welche sowohl in der CA als auch der DCA als gut zu bezeichnen sind. Die Schätzungen der Gradientenlänge durch die DCA liegen bei 2,61 und 2,28 SD.

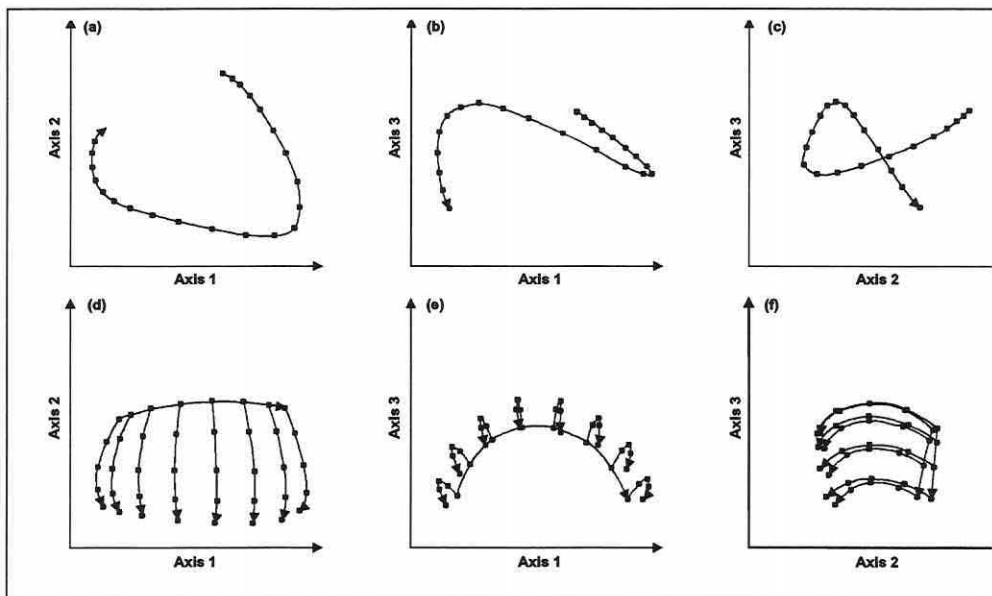


Abbildung 4.1-3: Teilbilder (a) bis (c) visualisieren eine 5HC-Coenocline mit 24 äquidistant verteilten Probenflächen für die Kombinationen der ersten drei PCAc-Achsen. Teilbilder (d) bis (f) die Ordinationsergebnisse für eine Coenoplane der Dimension 2,5 · 1,5 HC mit 40 Probenflächen (Causton 1988, S. 217).

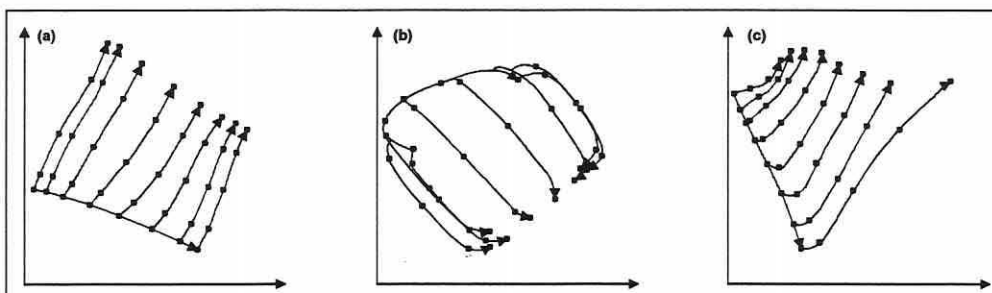


Abbildung 4.1-4: Ordination einer 4,5 · 4,5 HC Coenoplane mit drei Techniken: CA (a), PCAc (b) und PCAs (c) (Causton 1988, S. 218, Auszug).

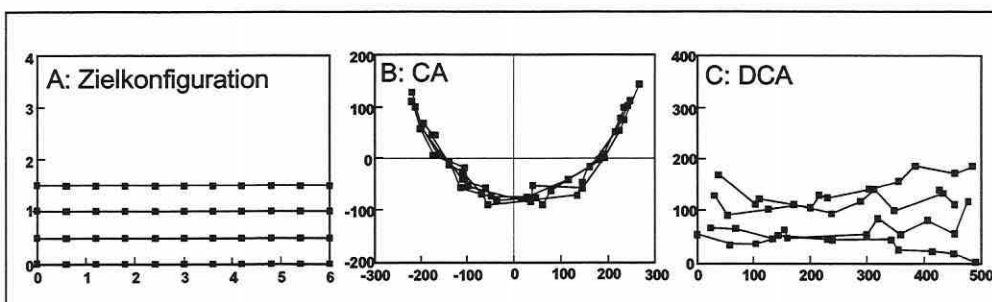


Abbildung 4.1-5: Ordination einer simulierten Coenoplane mit 6 · 1,5 SD. Teilbild (A) zeigt den Idealfall, (B) das Ergebnis einer CA und (C) das einer DCA für die erste und zweite Achse. Skalierung in (A) SD und in (B) SD · 100 (Økland 1990, S. 128f, verändert).



### Diskussion zur PCA

Die PCA ist aufgrund ihrer linearen Modellannahme und des euklidischen Distanzmaßes der Korrespondenzanalyse meist deutlich unterlegen (vgl. Abschnitt 3.5.4 sowie Causton 1988; Gauch 1982a; Økland 1990; Palmer 1998 u.a.). Die reale Datenstruktur wird nicht aufgedeckt, und für Gradientenlängen von mehr als 4 SD ist der Hufeiseneffekt das unvermeidliche Resultat (Austin 1976; Austin & Noy-Meir 1971; Causton 1988; Gauch 1982a; Gauch et al. 1977; Jongman et al. 1987; Kent & Coker 1992; Legendre & Legendre 1998; Økland 1990; Palmer 1998; Swan 1970). Dennoch kann die PCA für sehr kurze Gradienten Anwendung finden ( $< 2$  SD), da in diesem Falle ein größerer Anteil der Antworten monoton ist (vgl. Causton 1988; Gauch et al. 1977; Gauch 1982a; Jongman et al. 1987; Legendre & Legendre 1998; Økland 1990; Palmer 1993; ter Braak 1998). Einigkeit in der Fachwelt findet sich ebenfalls darin, dass die PCA ein effektives Werkzeug ist, um Zusammenhänge zwischen Umweltparametern zu analysieren und sie zu Komplexgradienten zusammenzufassen (Økland 1990; Palmer 1998 u.v.a.). Die Berechnung von Kovarianzen für abiotische Daten, die auf unterschiedlichen Einheiten basieren oder auch für Arten die deutlich dominanter sind als andere (order of magnitude) ist allerdings inkorrekt, solange die Deskriptoren nicht auf eine einheitliche Skala zurückgeführt wurden. Erst für standardisierte Variablen ist es möglich das Maß zu berechnen, da sie eine einheitliche Varianz aufweisen und dimensionslos sind.

Bei der Anwendung des Verfahrens als indirekte Ordinationsmethode für biotische Daten sollten die Arten auf den Probestellen weitgehend identisch sein und sich nur in ihren Artmächtigkeiten/Häufigkeiten unterscheiden, um die Flächen zu differenzieren. Daraus erklärt sich, warum die PCA meist keine effektive Methode für die Studie von Räummustern in der Artenkomposition ist (Legendre & Legendre 1998; Økland 1990). Gradientenlängen zwischen 3 und 4 SD bilden einen Bereich, in der lineare und unimodale Verfahren ähnlich gute Ergebnisse liefern können (ter Braak 1998).

Die nicht zentrierte Variante der PCA mit dem Ursprung in einer hypothetischen Probestelle, in der die Artmächtigkeiten/Häufigkeiten aller Arten gleich null sind, bildet ein inadäquates Modell für natürliche Situationen und zeigt noch intensivere Hufeiseneffekte als die zentrierten Varianten. Es kann generell nicht empfohlen werden. Zu diesem Schluss kommt auch Økland (1990).

Minchin (1987) und Økland (1990) präferieren generell die PCAs, während Jongman et al. (1987) die PCAc im Falle einer sehr geringen Beta-Diversität zur Gradienten-Identifikation empfiehlt, was durch die Studien von Gauch et al. (1977) untermauert wird (vgl. Abbildung 4.1-4, S. 125). In Abbildung 4.1-1 (Teilbild c, S. 124) liefert die PCAc die besten Ergebnisse für eine Beta-Diversität von einer HC. Wann eine PCAc oder PCAs ausgeführt werden sollte, bedingt sich aber nicht nur aus den rein technischen Aspekten, sondern ist auch der Fragestellung der Analyse unterworfen. Um die Beziehungen zwischen den Objekten im reduzierten Ordinationsraum zu studieren, sollte man sich überlegen, ob die Objekte unter Berücksichtigung der Originaldeskriptoren und deren Varianzunterschiede zusammengefasst werden sollen, oder ob alle Deskriptoren gleichgewichtet zu der Gruppierung der Objekte beitragen sollen. Letzteres verlangt nach der standardisierten Matrix. Eine andere Sichtweise kann die Betrachtung der euklidischen Distanz zwischen den Objekten sein. Ist die räumliche Konfiguration der Objekte bezogen auf die Rohdaten (Kovarianzen) oder die standardisierten Daten (Korrelationen) im Mittelpunkt des Interesses?

Trotz einiger befürwortender Argumente für die Methode lehnen manche Autoren die PCA pauschal ab: "Thus the use of PCA for vegetational ordination should be terminated" (Økland 1990, S. 137), oder "There is little justification for the continued application of linear ordination methods to community data for the purpose of indirect gradient analysis" (Minchin 1987, S. 106). Solche Globalurteile sind wenig zuträglich, da gezeigt werden konnte, dass die

PCAc und die PCAs durchaus für kurze Gradientenlängen Verfahren der Wahl sein können (vgl. auch Causton 1988; Palmer 1998; ter Braak 1998 u.a.).

#### **Diskussion zur CA**

Die angeführten Analysen von simulierten Daten haben gezeigt, dass auch die CA unter dem Bogeneffekt und dem Randeffect leidet, aber das Einfalten der Gradientenenden in der Regel ausbleibt (s.a. Austin 1976b; Gauch et al. 1977; Hill 1973, 1979a; Hill & Gauch 1980). Der Effekt ist noch nicht im Detail verstanden. Er wird neben anderen Vermutungen auch auf Inkonsistenzen des Modells zurückgeführt (Gauch 1982a, b; Jongman et al. 1987; ter Braak & Prentice 1988). Eine umfassende Diskussion möglicher Ursachen findet sich in Økland (1990).

Das  $\chi^2$ -Distanzmaß (vgl. Abschnitt 3.5.2) hat Vorteile (Vermeidung des Doppelnullenproblems), aber auch mehrere nicht wünschenswerte Eigenschaften (s. Økland 1990). Es besitzt keine Obergrenze und führt teils zu nur mäßigen Schätzungen der ökologischen Distanz bei Probestellen mit wenigen gemeinsamen Arten (Faith et al. 1987; Minchin 1987). Die Eigenschaften des Distanzmaßes führen zu der Sensitivität des Verfahrens gegenüber Ausreißern (Gauch 1982a; Jongman et al. 1987; Legendre & Legendre 1998; Økland 1990). Causton (1988) zieht als Schlussfolgerung aus seinen Studien, dass die CA robust gegenüber moderaten Ausreißern reagiert, da diese einfach im Zentrum der Ordination aufzufinden sind und die Platzierung der übrigen Probestellen nur geringfügig beeinflussen. Extreme Ausreißer hingegen können sich in der Peripherie des Ordinationsdiagramms finden und führen zu einer massiven Kompression der übrigen Datenwolke. Der Effekt ist leicht auszumachen und die Ursache lässt sich dann aus dem Datensatz entfernen. Eine erneute Ordination der übrigen Daten löst das Problem (Causton 1988; Gauch et al. 1977). CANOCO® prüft die Daten auf Ausreißer anhand einer entsprechenden Statistik (vgl. Abschnitt 3.6.1: VIF und Übersicht 3.6-4, S. 91).

Der Vergleich von Analyseergebnissen aus simulierten Daten legt offen, dass die CA gegenüber der PCA realistischere Ergebnisse in Studien mit natürlichen Daten erwarten lässt, wenn die Datensatzeigenschaften zuvor gewissenhaft geprüft wurden (z.B. die Gradientenlänge in SD durch die Berechnung einer DCA) und die Auswertemethode dann entsprechend optimal gewählt wird (Causton 1988; Gauch 1982a; Gauch et al. 1977; Kenkel & Orłóci 1986; Legendre & Legendre 1998; Økland 1990). Für Gradientenlängen von mehr als 4 SD ist die CA der PCA aufgrund des unterliegenden Modells immer vorzuziehen. Dies wird auch von Palmer (1998), ter Braak (1998) und anderen bestätigt.

#### **Diskussion zur DCA**

Die Auswertungen dieses Abschnitts bescheinigen dem DCA-Verfahren eine höhere Zuverlässigkeit als einer PCA oder CA. Die Analysekraft ist gut, wenn die Gradienten lang sind bzw. sie eine deutlich unterschiedliche Länge haben (vgl. Økland 1990).

Viele Autoren kommen zu dem allgemeinen Schluss, dass die DCA häufig anderen Ordinationstechniken überlegen ist (Gauch 1982a, b; Kent & Coker 1992; Økland 1986a-c, 1990; Palmer 1993; Peet et al. 1988; Pielou 1984), da sie wie die CA auf unimodale Beziehungen der Arten zum Gradienten angepasst ist, aber einige Schwächen von CA und speziell der PCA vermeidet (Hill & Gauch 1980).

Dennoch wird das Verfahren kritisiert: Die DCA liefert instabile Ergebnisse, wenn die ersten beiden Eigenwerte sehr dicht beieinander liegen (Instabilitätsproblem; Oksanen 1988). Es wird durch den Detrending-Prozess eine systematische Modifikation der unterliegenden Datenstruktur impliziert (Kenkel & Orłóci 1986; Wartenberg et al. 1987) und es werden dann Diagramme erzeugt, in denen die Existenz einer realen "kurvigen" Struktur in den Daten verborgen bleibt. Das Problem der Verzerrung wird gelöst durch Tarnung. Diese Täuschung erhöht weder das Verständnis für die Daten noch fördert es die Identifizierung eines Grundes für die vorliegenden Analyseergebnisse (Wartenberg et al. 1987).

Weitere Kritik kommt von Palmer (1993). Er moniert generell die mangelnde Eleganz des Detrending-Algorithmus, die geringe Analysekraft bei schiefen Verteilungen, das Fehlschlagen der Analyse bei komplexen Stichprobendesigns und das gelegentliche Vorkommen einer Kompression des Gradientenendes in eine Zungenform. Vergleichbare Urteile fällen Austin (1985), Dargie (1986), Knox (1989) sowie Minchin (1987) und finden damit Bestätigung in der neuesten Literatur (Legendre & Legendre 1998, S. 472): "Present evidence indicates that detrending should be avoided, except for the specific purpose of estimating the lengths of gradients; such estimates remain subject to the assumptions of the model being true. In particular, DCA should be avoided when analysing data that represent complex ecological gradients."

Trotz aller Kritik sind die Ergebnisse des Verfahrens bei den simulierten Daten aus Abbildung 4.1-6 (S. 128) näher an der Realität als bei der CA und im Allgemeinen bietet das Verfahren die benannte Möglichkeit, Gradientenlängen akzeptabel zu schätzen und damit einen Hinweis zu geben, welches Modell in einer direkten oder indirekten Analyse den Vorzug bekommen sollte. Mit der letztgenannten Aussage gehen auch Jongman et al. (1987), Palmer (1998), ter Braak (1998) und andere konform.

Aus eigener Sicht sollte man Kent (1992, S. 226) zupflichten, der nach seinen Studien zum Stand der Diskussion abschließend bemerkt: "In summary, DCA still remains a widely used and effective indirect ordination technique. Despite the various problems, the method is probably still as good as any other in most situations and better than most in many."

### Test der direkten Ordinationsverfahren

Zunächst geht man davon aus, dass direkte Verfahren die gleichen Vor- und Nachteile besitzen wie ihre indirekten Gegenstücke. Um diese Annahme ansatzweise zu klären, hat Palmer (1993) Tests anhand der CCA durchgeführt. Unmittelbar vergleichende Simulationsergebnisse zur Analysekraft von RDA und CCA – wie für die PCA und CA angeführt – sind mir nicht bekannt.

Palmer simulierte Daten, die in der DCA zu Verzerrungen führen (vgl. Abbildung 4.1-5, S. 125). Seine Zielkonfiguration veranschaulicht Abbildung 4.1-6. Zusätzlich wurden zwei perfekte Erklärungsvariablen  $g_1$  und  $g_2$  sowie vier reine Zufallsvariablen  $r_1$  bis  $r_4$  generiert. Ein perfektes Ordinationsergebnis sollte die Probeflächenanordnung in Abbildung 4.1-6 nachbilden, wobei  $g_1$  entlang der horizontalen und  $g_2$  entlang der vertikalen Achse verläuft.

Die Ergebnisse für vier Stufen von strukturlosem quantitativem Datenrauschen (noise, hier in Prozent-Standardabweichung der Abundanz für jede Art und Probefläche; s.a. Minchin 1987) in den Speziesdaten zeigt Abbildung 4.1-7 (S. 129). Die "WA scores"- und "LC scores"-Teilbilder beziehen sich auf die Matrizen  $F$  und  $Z_{\text{Skalierung } 1}$  (vgl. Abschnitt 3.5.8). Es ist erfreulich, dass die "LC scores" selbst bei extremen Datenrauschen sehr stabile Ergebnisse liefern und das Design nahezu ideal abbilden.

Die Ordination der "WA scores" ist nur minimal schlechter. Auch wenn das Raster zu den Achsen hin leicht gedreht erscheint, ist es parallel zu den aussagekräftigen Erklärungsvariablen  $g_1$  und  $g_2$ . Die Zufallsparameter (Willkürwerte)  $r_1$  bis  $r_4$  spielen keine beeinflussende Rolle.

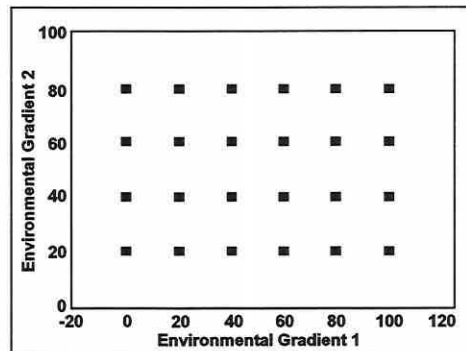


Abbildung 4.1-6: Simuliertes Design und Zielkonfiguration der Ordinationsergebnisse: Rasteranordnung von Probeflächen entlang zweier perfekter Umweltgradienten (nach Palmer 1993, S. 2219).



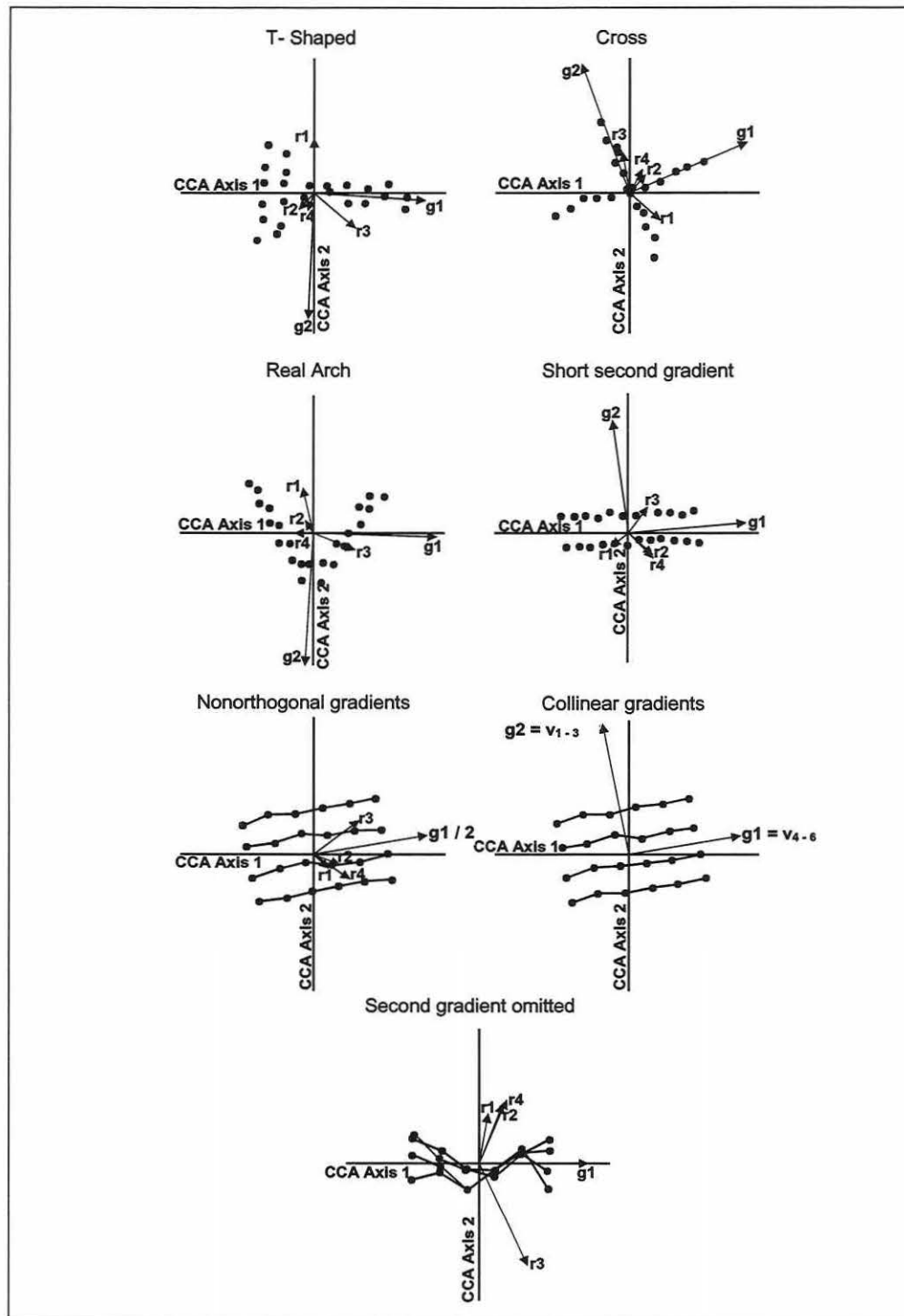


Abbildung 4.1-8: CCA-Ordinationen (LC scores) für komplexe Versuchsanordnungen, hochkorrelierte Gradienten und nahezu kollinearen Variablen pro Gradient; Erklärung im Text (Palmer 1993, S. 2224ff, verändert).



Die Ordinationsergebnisse für verschiedene Fälle ohne Datenrauschen werden in Abbildung 4.1-8 (S. 130) dargestellt. Datenrauschen liefert qualitativ vergleichbare Ergebnisse (Palmer 1993). Komplexe Coenoplanen (T-, Kreuz- und reale Bogenform) sind ebenso durch die CCA sauber erfasst worden wie die Kombination eines langen und eines kurzen zweiten Gradienten (Short 2<sup>nd</sup> gradient). Hier zeigen sich fast keine Verzerrungen, und die Erklärungsvariablen  $g_1$  und  $g_2$  sind korrekt erkannt und ausgerichtet.

Besonders eindrucksvoll ist das Testergebnis für Nichtorthogonalität und Kolinearität. Erstes meint, dass die zwei wichtigen Erklärungsvariablen miteinander hochgradig korreliert sind (Werte der neuen Variablen  $g_2$  entsprechen den von Variable  $g_1$  plus dem sehr gering gewichteten Informationsgehalt der alten  $g_2$ -Variablen: Multiplikationsfaktor 0,01), und letzteres, dass die Variablen  $g_1$  und  $g_2$  nun aus einer Linearkombination von Erklärungsvariablen zusammengesetzt werden. Die drei jeweils zu  $g_1$  ( $v_1 - 3$ ) und  $g_2$  ( $v_4 - 6$ ) führenden Variablen sind in ihren Werten grundlegend identisch. Es wurde nur zur Basisinformation eine verschwindend kleine Zufallskomponente hinzuaddiert. In der Ordination für Nichtorthogonalität kann trotz der extrem kleinen Differenz an Informationsgehalt zwischen  $g_1$  und  $g_2$  das Raster nachgebildet werden. Die Erklärungsvariablen werden in der Ordination übereinander projiziert und weisen beide in die gleiche Richtung. Nimmt man die minimale Zusatzinformation des zweiten Gradienten aus der Analyse, bricht das Raster zusammen (Second gradient omitted; Abbildung 4.1-8). Nahezu kollineare Variablen ( $v_1 - 3$  und  $v_4 - 6$ ) destabilisieren das Raster nicht. Die Variablen werden fast exakt übereinander gelegt und weisen in die korrekte Richtung von  $g_1$  und  $g_2$ . Palmer (1993) kam weiterhin zu ähnlich guten Resultaten für den Test mit drei und vier relevanten Erklärungsvariablen in der Analyse.

### Diskussion der CCA

Aus den Ergebnissen von Palmer (1993) geht hervor, dass die CCA sehr viel realitätstreuer als andere Ordinationsverfahren ist und eine hervorragende Analysekraft in allen von ihm getesteten Situationen zeigt. Das unimodale Antwortmodell lässt sie ihrem linearen Gegenspieler der RDA aus den gleichen Gründen überlegen werden, wie die CA der PCA (vgl. auch ter Braak 1987b). Ein artifizieller Bogeneffekt bleibt aus, solange er nicht real in den Daten vorhanden ist. Dieses direkte Ordinationsverfahren ist sozusagen "immun" gegen die Probleme der CA, DCA und PCA. Die Anwendung des Detrending-Verfahrens, d.h. die Berechnung der bisher nicht erwähnten DCCA, ist nur notwendig, wenn ein detaillierter Vergleich zu einer DCA durchgeführt werden soll (s.a. Palmer 1993; ter Braak 1998), da die Eigenwerte einer DCA ab der zweiten Dimension generell niedriger sind als für eine CA und einen direkten Vergleich zur CCA nicht zulassen (vgl. Abschnitt 3.6.1, Übersicht 3.6-7). Speziell die Neutralität des Verfahrens bei Kolinearität ist vorteilhaft, da somit auch hochkorrelierte Erklärungsvariablen gemeinsam in der Analyse betrachtet werden können, um geringfügige Unterschiede zu erkennen oder einfach nur ihre Korreliertheit zu visualisieren. Palmer (1993, S. 2228) bescheinigt dem Verfahren: "CCA has all the advantages and none of the disadvantages of the DCA."

Die CCA ist beruhend auf Palmers Simulationsergebnissen, den dargelegten theoretischen Eigenschaften und der Prognosemodellfähigkeit (vgl. Abschnitt 4.2 und 5.6) aus eigener Sicht die empfehlenswerteste Ordinationsmethode für ökologische Studien, vorausgesetzt biotische und abiotische Daten sind vorhanden und die Gradientenlänge übersteigt 2 SD. Beachten sollte der Anwender aber, dass die CCA dennoch anfällig gegen Ausreißer (vgl. Abschnitt 3.5.8, Legendre & Legendre 1998 und McCune 1997) und Einflüsse kleinräumiger Heterogenität bleibt (vgl. Abschnitt 2.6 und Palmer & Dixon 1990).

### Fazit

Anhand der vergleichenden Literaturlauswertung konnte gezeigt werden, dass jede Ordinationsmethode ihre spezifischen Stärken und Schwächen besitzt, die eine generalisierte Anwendung von einem Verfahren verneinen. Die Ergebnisse einer multivariaten Datenauswer-

tung werden nicht notwendigerweise validiert, indem fortschrittliche komplexe Verfahren technisch korrekt angewendet werden, sondern die Zuverlässigkeit und Gültigkeit hängt vielmehr von weiter gefassten Umständen, wie der Datenerhebungsstrategie (Abschnitt 3.1.2), den Datenerhebungsmethoden (Abschnitt 3.1.4), der Variablenauswahl (Abschnitte 3.1.7 und 3.5.3), Transformationen (Abschnitt 3.3) und der Kombination eines geeigneten Verfahrens, konsistent zu den Eigenheiten der Datenstruktur, ab (s.a. Abschnitte 3.5.2, 3.5.4 bis 3.5.8 und 3.7.1). Es ist notwendig, differenziert vorzugehen und bereits die grundlegenden Erfordernisse für die Anwendung von Ordinationsverfahren bei der Projektplanung zu berücksichtigen.

Die abschließende Wahl der "besten Ordinationsmethode" wird somit durch die Datensatzeigenschaften bestimmt und ist nicht immer eindeutig (vgl. Causton 1988; Gauch 1982a; Gittins 1979; Gower 1987; Jongman et al. 1987; Kent & Coker 1992; Köhler et al. 1996; Legendre & Legendre 1998; Økland 1990; Orłóci 1978; Palmer 1998; Pielou 1977, 1984; Smilauer 2000; Sokal & Rohlf 1995; Whittaker 1973). Es empfiehlt sich, bei einer neuen Studie immer mit einer DCA zur Einschätzung der Gradientenlänge zu beginnen und anhand des Ergebnisses festzulegen, welches Verfahren für die weiteren Auswertungen geeignet erscheint (unimodales oder lineares Modell). Dann erst erfolgt die Ordination der Daten mit einem indirekten oder – je nach Datenlage – mit einem direkten Verfahren. Es ist zudem angeraten, die Ergebnisse aus einer direkten Ordination mit denen einer indirekten zu vergleichen. So lässt sich prüfen, ob die wichtigsten Erklärungsparameter (definieren die erste CCA-Achse) erfasst wurden. Sind die Erklärungsanteile und Muster der Ordinationsdiagramme speziell entlang der ersten Achse nur geringfügig verschieden, kann von einem guten Modell zur Erklärung der Biotik durch die abiotischen Parameter ausgegangen werden. Eine Übersicht zum Vorgehen bietet der dichotome Entscheidungsschlüssel in Abbildung 4.1-9 (S. 133).

Die Schlussbemerkung des Abschnittes muss lauten: **Das einzig wahre Ordinationsverfahren gibt es bisher nicht**, denn ein solches müsste wenigstens den nachfolgend zusammengefassten Eigenschaften genügen.

- (1) Relevante Standortgradienten werden erkannt und ohne Verzerrungen dargestellt.
- (2) In der Natur existierende ökologische Gruppen werden im Diagramm sichtbar.
- (3) Artifizelle Gruppen treten im Diagramm nicht auf.
- (4) Das Verfahren kann für beliebige Datensätze Anwendung finden: Gradientenlänge, Stichprobenumfang, artenreiche / -arme Systeme etc.
- (5) Ökologische Ähnlichkeit wird durch die Nähe der Signaturen im Ordinationsraum repräsentiert.
- (6) Die Achsenskalierung bezieht sich auf die Beta-Diversität in den Daten.
- (7) Unempfindlichkeit der Ordination gegen Datenrauschen und Ausreißer.
- (8) Strukturierte und unstrukturierte Variation (Datenrauschen) werden getrennt.
- (9) Alle Informationsebenen werden ökologisch objektiv betrachtet (Arten, Erhebungseinheiten und Erklärungsvariablen) und simultan analysiert. Transformationen und Gewichtungen sind überflüssig.
- (10) Die Ergebnisse sind stabil und wiederholbar (identische Lösung bei Mehrfachanalyse).
- (11) Es gibt eine eindeutige Ordinationslösung (direkte Verfahren auf der Basis einer Vorwärtselektion können zu leicht differierenden Lösungen führen).
- (12) Das Verfahren arbeitet mathematisch elegant. Es ist sowohl leicht verständlich als auch jedem Anwender bei Bedarf verfügbar.

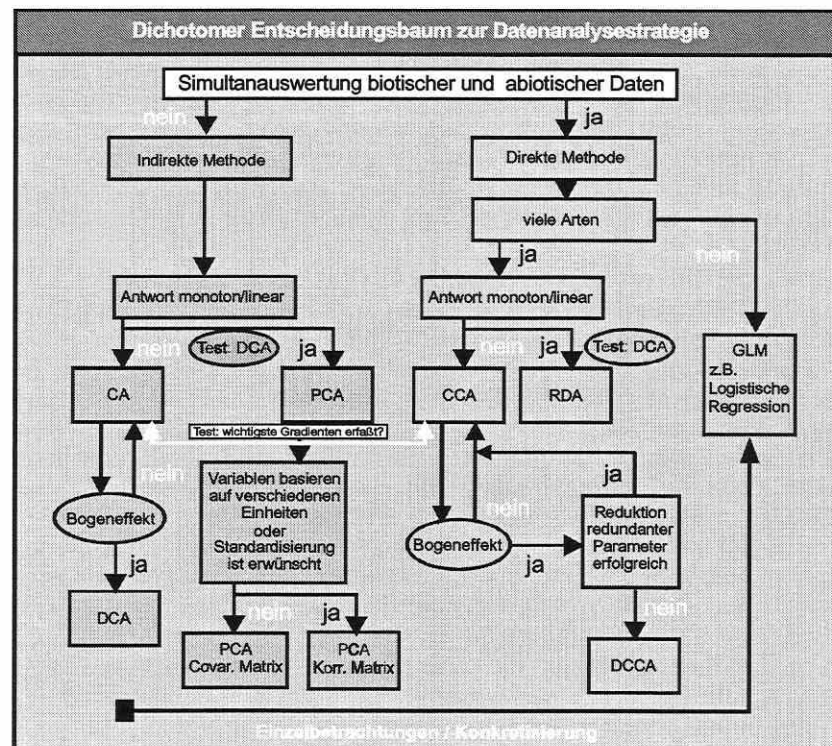


Abbildung 4.1-9: Dichotomer Entscheidungsschlüssel zur Wahl eines geeigneten Verfahrens.

#### 4.2 Direkte Ordination und logistische Regression: Zwei sich ergänzende Verfahren?

*"MOST STUDIES USE ONLY ONE OF THE MANY STATISTICAL TECHNIQUES THAT MAY PROPERLY BE USED, AND LITTLE INFORMATION IS AVAILABLE ON THE RESPECTIVE PREDICTIVE CAPACITY OF EACH APPROACH COMPARED TO THE OTHERS (GUISAN ET AL. 1999, S. 107)."*

Es ist auffällig, dass nur wenige Autoren in ihren Auswertungen verschiedene Verfahren vergleichend einsetzen (z.B. Austin et al. 1984; Guisan et al. 1999 und van de Rijt 1996). Interessante Beiträge zur separaten Anwendung einer Modelltechnik finden sich hingegen häufig (Peeters & Gardeniers 1998; Vincent & Haworth 1983 u.v.a). Pearce & Ferrier (2000a, b) bewerten beispielsweise die Logistische Regression als Werkzeug zur Erstellung von Lebensraumeignungsmodellen.

Wird die Analyse von Daten auf eine Methode beschränkt, kann aber oft ein wichtiger Informationsgehalt unentdeckt bleiben, während durch unterschiedliche Methoden verschiedene Aspekte der Daten ans Licht gebracht werden (Legendre & Legendre 1998; Orlóci 1978). Die Datenanalyse ist daher ein iteratives interaktives Unterfangen. Nimmt man mit verschiedenen Verfahren Einsicht in die Datenstruktur – z.B. Ordination, Klassifikation und logistische Regression – und kommt zu ähnlichen Schlussfolgerungen, so ist auch die Wahrscheinlichkeit, reale Eigenschaften entdeckt zu haben, deutlich gesteigert (vgl. Abschnitte 5.4.2, 5.4.3

und 5.5). Nach der methodenkritischen Betrachtung von Ordinationsverfahren in Abschnitt 4.1 soll daher zusätzlich die logistische Regression als bewährtes Verfahren statischer Modellierung mit in die Diskussion eingeschlossen werden. Die Ziele der Verfahren (direkte Ordination und logistische Regression) sind ähnlich gelagert, so dass es Sinn macht, über die einleitende Frage, ob sich die Verfahren ergänzen, nachzudenken. Die Beantwortung dreier Fragen steht dabei im Mittelpunkt dieses Abschnitts:

- (1) Anpassungsgüte des Modells und Exaktheit der Vorhersage: Sind die Ergebnisse vergleichbar und wird eine mit direkten Ordinationsverfahren gut modellierte Art mit der logistischen Regression ähnlich gut modelliert?
- (2) Gibt es Artengruppen, die durch ein Verfahren besser beschrieben werden, und wie begründet sich dies gegebenenfalls?
- (3) Wie unkompliziert lassen sich die Modelltypen in ein GIS implementieren?

Die Modellerstellung in der CCA und der LR unterscheidet sich. In kurzen Worten: Die CCA integriert alle Antwort- und Erklärungsvariablen in ein Modell, während die LR für jede Art eine neue Modellanpassung mit Selektion einer spezifischen Erklärungsvariablenkombination erfordert (Abschnitte 3.5.7, 3.5.8 und 3.7.1). Die Anpassungsgüte des Modells ist daher meist zwangsläufig in der LR besser, da sie spezifisch für eine Art vorgenommen wird und nicht für alle Arten auf die beiden Globalmodelle "linear" oder "unimodal" zurückgreift. Somit ist auch der Erklärungsanteil an der Varianz der einzelnen Arten in der LR generell höher. Gleiches gilt – in den meisten Fällen – für die Vorhersagegüte. Arten, deren Raummuster gut anhand der CCA modelliert werden, sind auch durch die LR gut beschrieben (vgl. Abschnitt 5.4 und 5.5 sowie Guisan et al. 1999).

Ein Vorteil der CCA ist, dass die Information aus den Artmächtigkeiten/Häufigkeiten der Arten umgesetzt werden kann, da die CCA keine Binärokodierung der Daten voraussetzt (vgl. Abschnitt 3.7.1). Darüber hinaus ist die Modellierung aller Arten ohne großen zeitlichen Aufwand möglich, weil die jeweilige Neuerstellung eines artspezifischen Modells entfällt. Seltene Arten mit wenigen Erscheinungen im Untersuchungsgebiet lassen sich in der LR verfahrensbedingt nicht modellieren. In der CCA werden sie durch die gemeinschaftliche Modellierung in ihren ökologischen Rahmen eingeordnet. Sie besitzen oft ähnliche Lebensraumsansprüche wie häufige Arten im Ordinationsraum, die ihre natürliche Verteilung simulieren, und werden daher, profitierend von diesem Umstand, ihren ökologischen Ansprüchen gerecht modelliert. Der Blick nach globalen Schlüsselgradienten für viele Arten in der CCA lässt das komplexe Zusammenspiel von Umweltparametern oft klarer erscheinen als dies durch die LR für einzelne Schirmarten erreicht werden kann. Die Beschreibung des Raummusters in der Artenkomposition und die ökologische Gruppierung anhand der Erklärungsvariablen wird durch den multivariaten Ansatz erreicht. Die Arten können daher im Kontext ihrer Lebensgemeinschaft betrachtet werden. Das soziologische Gefüge ist anhand der LR nicht interpretierbar. Schlussendlich wird die Beta-Diversität entlang eines Gradienten durch die CCA ausgegeben. In den letztgenannten Punkten ist die CCA der LR klar überlegen, dafür entfällt aber die attraktive Aussage von Vorkommenswahrscheinlichkeiten, welche die LR als Ergebnis liefert (vgl. Abschnitt 3.7.1). Die CCA stellt dafür Distanzen (Toleranzen) zu den gewichteten Schwerpunkten der Nische einer jeweiligen Art bereit (vgl. Abschnitte 3.5.8, 3.6.1 und 3.7.1 sowie Guisan et al. 1999). Die Positionskoordinaten der Probeflächen entlang einer CCA-Achse lassen sich als kanonische Variable einer logistischen Regression zuführen (vgl. Abschnitt 5.5 u. Guisan et al. 1999; Hill 1991).

Für den Vergleich von Ergebnissen wird an dieser Stelle auf erste Auswertungen basierend auf den Daten des RIVA-Projektes (Abschnitt 5.6) zurückgegriffen. Die Anlagen 10 und 12 belegen die weitreichende Analogie der Ergebnisse für die Art *Arrhenatherum elatius*, wobei im Detail durchaus Abweichungen existieren (vgl. auch Ergebnisse von Guisan et al. 1999). Insgesamt können die Ergebnisdifferenzen aber nicht als gravierend eingestuft werden.



Die Beantwortung der eingangs gestellten dritten Frage bezieht sich auf die Möglichkeiten einer GIS-Einbindung der Modelle (vgl. Abschnitt 3.1.5). Speziell im Falle der LR ist dies unproblematisch, da aus der Analyse eine einfache Modellgleichung resultiert (Gleichung 3.7-4, S. 116), anhand der die Antwort einer Art für die jeweilige Rasterzelle im GIS basierend auf den regionalisierten Werten der Modellparameter berechnet werden kann.

Die GIS-Implementierung eines CCA-Modells gestaltet sich weniger einfach, ist aber ebenfalls möglich. Jede Achse der direkten Ordination wird durch eine Linearkombination von Originalerklärungsvariablen beschrieben, deren kanonische Regressionskoeffizienten den Ausgabedateien von CANOCO® zu entnehmen sind (vgl. Abschnitt 3.6.1). Nach einer Destandardisierung der Regressionskoeffizienten (vgl. ter Braak 1998, S. 180) sind sie für die natürlichen Einheiten der Erklärungsvariablen gültig.

Allein die Tatsache, dass die Ordinationsachsen durch eine multiple Regressionsgleichung repräsentiert werden, eröffnet sogleich die Möglichkeit ihrer Implementierung in ein GIS. Jede Achse beschreibt eine Dimension mit einem Schlüsselkomplexgradienten und ihre Kombinationen einen kanonischen Raum. Somit können die Koordinaten jeder Rasterzelle, für die alle Werte der Modellparameter im Untersuchungsraum vorliegen müssen, entlang der Ordinationsachse bestimmt und deren euklidische Distanz zum Zentroid einer Art im kanonischen Raum errechnet werden. Die Werte der abiotischen Modellparameter müssen hierzu in der Regel durch geostatistische Verfahren, Grundwassermodelle etc. approximiert und möglichst durch Referenzmessungen validiert werden. Bei der Simulation des Istzustandes kann speziell durch den Vergleich mit der Vegetationskartierung eine erste Qualitätsprüfung vorgenommen werden.

Die von einer CCA kalkulierten Toleranzen einer Art bestimmen ihre etwaige Realnische in der soziologischen Gemeinschaft gegenüber den Ordinationsachsen (vgl. Abschnitte 3.5.8 und 3.6.1). Die Einteilung der Distanzen in Klassen entspricht sinngemäß der Klassifizierung von Vorkommenswahrscheinlichkeiten (vgl. Abschnitte 3.7.1 und 5.6). Die Distanzklassen können dann in der räumlichen Vorhersage der potenziellen Artenverteilung im GIS angewendet werden.

Um die Übernahme der CCA-Modelle in ein GIS und die notwendigen Sekundärberechnungen zu automatisieren haben Andrew und Stuart Weiss an der Stanford University (USA) ein Modul (CANOGEN) entwickelt. Es importiert die Informationen der CANOCO®-Ausgabedateien und ein ARC/INFO®-Makro (AML) übersetzt diese direkt in eine Karte der räumlichen Verteilung für jede Art im Datensatz. Damit sind in einem Arbeitsschritt alle in der Analyse befindlichen Arten zu Prognosezwecken im Untersuchungsraum nutzbar. Natürlich sind die Prognosen nicht für alle Arten gleich zuverlässig, und die Ergebnisse müssen im Einzelnen hinterfragt werden, da z.B. auch nicht jede Art in der CCA eine optimale Modellanpassung erfährt. Zu berücksichtigen ist weiterhin, dass die zur Umsetzung benötigten Analyseergebnisse aus zwei CCA-Analysen zu kombinieren sind, d.h., die Toleranzen entstammen nicht dem gleichen Projekt wie die übrigen Daten (vgl. Abschnitt 3.6.1 und den Text zur Übersicht 3.6-11, S. 95).

**Das Fazit der Betrachtungen** lautet, dass durch die Möglichkeit einer Simultanauswertung vieler Arten und das generalisierte Modell einer CCA – verglichen zur LR – die Forderungen des Naturschutz-Managements nach vielschichtigen und dennoch einfachen und kostengünstigen Modellen eher zu erfüllen sind. Das aus einer CCA resultierende artenübergreifende Modell minimiert den Aufwand der Modellerstellung bei einer hohen Artenzahl in der Studie, wobei gleichzeitig, wie die Ergebnisvergleiche in diesem Abschnitt erkennen lassen, eine zur LR vergleichbare Aussagekraft und Zuverlässigkeit erzielt wird. Weiterhin ist der Datenerhebungsaufwand, durch wenige allgemeine Schlüsselparameter des Modells gültig für alle Arten, bei der Modellübertragung niedrig gehalten. Die Identifikation ökologischer Gruppen und deren Artenverteilung entlang von Komplexgradienten sowie gegenüber den



Modellparametern ist nur durch die Simultanbetrachtung in Ordinationsverfahren möglich (Zonierungen). Geht es um die Bestimmung von einer Artennische, kann diese Eigenschaft sehr wertvoll sein, da für das Beispiel der Flora belegt ist, dass die Nischeninformation durch verschiedene Lebensgemeinschaften modifiziert werden kann (vgl. Austin et al. 1990; Begon 1998; Økland 1990; Rink 1997c; Walter 1986 sowie Abschnitt 2.3). Die Option einer Artenbetrachtung im Artengefüge entbehrt die LR, da sie auf Einzelartbetrachtung ausgelegt ist. Somit ist es von Bedeutung, welche Beobachtungsobjekte sich im Fokus des Interesses befinden. Lebensgemeinschaften und deren modifizierende Einflüsse auf das Vorkommen stehen bei der Flora und Invertebraten eher im Vordergrund als bei den Vertebraten.

Dennoch besitzt auch die LR Vorteile gegenüber der direkten Ordination. Sie liefert eine klarer definierte Maßeinheit für das Vorkommen von Arten bei räumlichen Prognosen und ist besonders dann anzuraten, wenn die Modellanpassung für eine Art möglichst genau vorgenommen werden soll. Hierdurch ist die Vorhersagegüte eventuell zu steigern. Auch die Ermittlung der Nischenspanne gegenüber den Modellparametern ist direkter möglich, was ein Vorteil aus Indikationssicht darstellt.

Möchte man Rückschlüsse auf die Originalwerte der Schlüsselparameter für eine Art und deren Nischenspanne in der CCA ziehen, ist dies am effektivsten durch eine Abfrage der Gridzellenwerte nach der Modellimplementierung in ein GIS und der Visualisierung des Istzustandes.

Die getätigten Aussagen zeigen schließlich, dass die Vorteile beider Verfahren dahingehend abzuwägen sind, welche Beobachtungsobjekte und Zielstellungen verfolgt werden. Die CCA ist der LR für naturschutzfachliche Prognosen vorzuziehen, solange kein möglichst exaktes Modell für eine spezifische Art gefordert ist, die Lebensgemeinschaft eine bedeutende Rolle spielt und/oder mehr als eine einzelne Art betrachtet werden soll.

Alle bisherigen Schlussfolgerungen zielen auf die separate Anwendung der Verfahren. Es wurde aber schon zu Anfang des Abschnittes herausgestellt, dass Verfahren in Kombination angewendet besondere Effizienz bieten. Die beiden Verfahren können sich letztlich sehr sinnvoll ergänzen, indem die LR für artenspezifische Detailstudien aufbauend auf der CCA eingesetzt wird (Abschnitt 5.5, Austin et al. 1984, Guisan et al. 1999 und van de Rijt 1996).

## **5 GEWINNUNG UND AUSWERTUNG VON ÖKOSYSTEMARER FELDMINFORMATION ZUR ERSTELLUNG EINES PROGNOSE-SYSTEMS IN EINEM AUEÖKOLOGISCHEN VERBUNDPROJEKT**

Das Kapitel umfasst neben einer allgemeinen Einführung in die Untersuchungsgebiete den Weg von der Datenerhebung über die Datenanalyse und Modellierung, mit raum-zeitlicher Robustheitsprüfung der statistischen Modelle, bis hin zur Implementierung der Submodelle in ein GIS-basiertes Prognosesystem. Die Bearbeitungen erfolgen am Beispiel des vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie geförderten Verbundvorhabens "Übertragung und Weiterentwicklung eines Robusten Indikationssystems für ökologische Veränderungen in Auen" – RIVA.

Im Rahmen der dreijährigen Laufzeit des Projektes hat sich RIVA zum Ziel gesetzt, abiotische Schlüsselparameter, hier besonders die bodenphysikalische/-chemische und die hydrologische Situation, für ein Prognosemodell zu ermitteln, welche die strukturelle biotische Ausgestaltung dreier ausgewählter Grünlandstandorte des Deichvorlandes der Mittleren Elbe beeinflussen und letztendlich ausprägen. Die Prämisse des Modellierungsansatzes ist es, die Veränderung der biotischen Ökosystemkomponenten (Arten/ökologische Gruppen) als Reaktion der Änderung abiotischer Umweltfaktoren prognostizieren zu können. Dabei soll eine Vorhersage des Artenverlustes, des Zuwachses oder Rückgangs der räumlichen Vorkommen oder der Verschiebung der artspezifischen Lebensräume bzw. der Organismenkollektive erreicht werden. Das bedeutet, man strebt langfristig an, eine Antwort geben zu können, welche etwaigen Veränderungen im natürliche Gefüge nach anthropogenen Eingriffen zu befürchten sind.

### **5.1 Untersuchungsgebiete**

#### **5.1.1 Allgemeine Vorüberlegungen**

Die Erstellung eines Stichprobenplans basiert auf einer Vielzahl hierarchisch gegliederter Arbeitsschritte und der klaren Definition der Zielstellung (vgl. Abschnitt 3.1 und die Einleitung zum Kapitel 5). Es sind Grundsatzentscheidungen zu treffen, die in den Abschnitten 3.1.1 bis 3.1.7 umfassend erörtert und im Projekt berücksichtigt wurden (Abschnitte 5.1.2 bis 5.2.3 und 5.3.5). Die notwendige Kompromisslösung zwischen arbeitsgruppenübergreifender Fragestellung, statistischer Datenanalysestrategie und Logistik spiegelt sich in der Art des Stichprobenplans wieder (vgl. Abschnitt 5.2.1 sowie Rink et al. 2000).

Als Untersuchungsraum wurde extensiv genutztes Auegrünland festgelegt (vgl. Abschnitt 5.1) und als biotische Indikatoren Carabiden, Mollusken sowie die Flora betrachtet (vgl. Abschnitt 5.3.1). Es wurden drei Untersuchungsgebiete (vgl. Abschnitte 5.1 und 5.1.2) und vier Hauptuntersuchungsphasen (vgl. Abschnitt 5.2.3) über 1998 und 1999 hinweg bearbeitet, um die Modelle (Abschnitte 5.4.1, 5.4.2 und 5.5) hinsichtlich ihrer räumlichen und zeitlichen Robustheit/Übertragbarkeit einer ersten Validierung zu unterziehen (vgl. Abschnitte 3.7.1, 5.2.1, 5.4.1, 5.4.4 und 5.5).

#### **5.1.2 Auswahl der Untersuchungsgebiete**

Anhand der Zielvorstellungen im Projekt, musste ein Kriterienkatalog erstellt werden (s.u.), der eine weitgehend optimierte Auswahl der Untersuchungsgebiete steuert. Dazu war es notwendig, möglichst viele Belange der Teilprojekte und das Gesamtziel des Projektes zu berücksichtigen, um teilweise schwierige Entscheidungen durch das Abwägen ihrer Vor- und Nachteile treffen zu können. Die Kriterien sind nachfolgend diskutiert und zusammengestellt. Die Erfüllung dieser Kriterien durch die letztlich gewählten RIVA-Untersuchungsgebiete listet Tabelle 5.1-1 (S. 139) auf.

### Anzahl der Gebiete

Für die Erstellung des Prognosesystems, das auf einem statistisch validierten Regelsatz ökosystemarer Wirkzusammenhänge im Art-/Standortsystem basieren soll, werden für den Test auf räumliche Modellübertragbarkeit wenigstens drei Untersuchungsgebiete benötigt (vgl. Abschnitte 3.7.1, 5.2.1, 5.4.1, 5.4.4 und 5.5). Die logistischen Einschränkungen, die auch im RIVA-Projekt trotz hohem Beprobungsaufwand ein theoretisch optimales Konzept nicht zuließen, führten dann zu insgesamt sieben Stichprobenplanvarianten, deren Vor- und Nachteile diskutiert wurden. Die Ausgangsbasis für eine weitere Untergliederung in 14 Untervarianten nach Stratenanzahl, Stichprobenumfang pro Stratum, zeitlicher Beprobungsintensität der faunistischen Untersuchungen und Nutzungsintensität bilden:

- (1) **Zwei Untersuchungsgebiete mit gleichverteiltem Stichprobenumfang**
- (2) **Drei Untersuchungsgebiete mit gleichverteiltem Stichprobenumfang**
- (3) **Drei Untersuchungsgebiete mit verschieden gewichtetem Stichprobenumfang**  
(ein Hauptuntersuchungsgebiet und zwei Nebenuntersuchungsgebiete)

Die Vorschläge (1) und (2) mussten verneint werden, da zwei Untersuchungsgebiete die oben genannte Bedingung für den Übertragbarkeitstest nicht akzeptabel erfüllen und drei Untersuchungsgebiete mit gleichverteiltem Stichprobenumfang durch die logistischen Grenzen im Projekt nicht ausreichend beprobt werden konnten. Die Vor- und Nachteile aller Untervarianten mit weiteren Pro- und Contraargumenten sind in Rink (1998, erhältlich bei der Bundesanstalt für Gewässerkunde Koblenz) ausführlich dargestellt.

Als Ergebnis der Gegenüberstellung wurde die Variante (3) gewählt, in der drei Untersuchungsgebiete mit quasihomogenen Teileinheiten (Straten) beprobt werden. Der Gesamtstichprobenumfang wurde auf 60 Probeflächen festgelegt und gewichtet auf das Hauptuntersuchungsgebiet (HUG) und die zwei Nebenuntersuchungsgebiete (NUG1 und 2) verteilt. Im HUG wurden 36 Stichprobenflächen und in den NUGs je zwölf randomisiert innerhalb der Straten eingerichtet. Die Variante wird in Abschnitt 5.2.1 näher erläutert.

### Geografische und topographische Lage

Zur Überprüfung einer räumlichen Übertragbarkeit der Modelle über kurze und längere Entfernungen entlang der Elbe wurde ein NUG in relativer geografischer Nähe zum HUG eingerichtet, während ein zweites in größerer Entfernung verortet wurde (vgl. Abbildung 5.1-1, S. 140). Topographisch liegen die Untersuchungsgebiete in den von einer rezenten Überflutungsdynamik geprägten Flussauen der Elbe (vgl. Anlage 1) und weisen insgesamt ähnliche Randbedingungen auf.

### Repräsentativität/Natürlichkeit

Das Projektziel, ein praxisnahes, übertragbares Prognosesystem für ökologische Veränderungen in Stromauen zu erstellen, schließt eine Flächenauswahl aus, die für weite Abschnitte der Elbe aktuell nicht mehr repräsentativ ist. Sehr seltene, annähernd natürliche Biotopstrukturen mit einem höchst schützenswerten Arteninventar sind wissenschaftlich zwar äußerst interessant und sicherlich durch ihre Zusammensetzung aus meist stenökeren Arten analysetechnisch (Auffindung klarer Zusammenhänge zu Standortfaktoren) erfolgversprechend, nur erfüllt solch ein prognostisches Managementinstrument in unserer anthropogen stark überprägten Kulturlandschaft Mitteleuropas kaum seine Aufgabe.

An der Elbe unterliegen 80-90% aller Grünlandflächen einer mittleren bis intensiven Nutzung. Die Auswahl der Untersuchungsflächen im Projekt orientierte sich daher an Elbauen mit einer repräsentativen Struktur und Ausprägung für die Region "Mittlere Elbe", d.h. Grünländer mit Weidenutzung (vgl. Abschnitt 5.1).

### Nutzung

Die Nutzung ist ein wesentlicher Standortfaktor, der viele Umweltparameter (z.B. natürliche bodenphysikalische und -chemische Bedingungen) deutlich überlagern kann (Scholle 1997).

Diese Eigenschaft birgt Vor- und Nachteile für die durchgeführte Untersuchung. Da die Nutzung einen kodierenden Einfluss auf die übrigen Randbedingungen eines Standorts hat, lassen sich Untersuchungsgebiete durch die bewusste Wahl einer homogenen Nutzungsintensität eher vergleichbar machen. Man muss sich hierbei aber bewusst sein, dass keine absolut identischen Nutzungsbedingungen zwischen Flächen zu erwarten sind. Die Bestrebungen gehen dahin, den Einfluss des Nutzungsgradienten begrenzt zu halten. Die Spanne unterschiedlicher Standortbedingungen wurde somit eingeeengt, um die Übertragbarkeit des Prognosesystems einschließlich einer ersten Validierung bzw. Robustheitsprüfung zu sichern.

### Morphologie

Die morphologische Ausprägung der Untersuchungsgebiete soll ebenfalls aus Vergleichbarkeitsgründen ähnlich sein und muss gewisse untersuchungsrelevante Strukturelemente beherbergen. Zum einen ist eine vergleichbare Morphologie notwendig, da manche auetypischen Spezies an entsprechende Lebensräume – z.B. Flutrinnen – gebunden sind. Die Rinnen wurden als eigenes Stratum aufsummiert (vgl. Abschnitte 3.1.3 und 5.2.1) und müssen eine physikalische Mindestgröße aufweisen, um den notwendigen Stichprobenumfang aufnehmen zu können (Abschnitt 5.2.1). Zum anderen beruhen die primären Stratifizierungskriterien (vgl. Abschnitt 5.2.2) für die Zufallsverteilung teils auf der Morphologie.

Eine ähnliche Lage der Auen zur Elbe und eine vergleichbare Fluss- und Vorlandsmorphologie sollte gegeben sein, um gravierend veränderte hydrologische und sedimentologische Verhältnisse auf den Flächen auszuschließen. Diese Gesichtspunkte wurden bei der Auswahl der RIVA-Untersuchungsflächen berücksichtigt.

### Untersuchungsgebiete des RIVA-Verbundprojektes

Die Begutachtung potenzieller Untersuchungsgebiete umfasste insgesamt elf Geländebegehungen, die letztendlich die Auswahl der Untersuchungsgebiete unter obigen Gesichtspunkten zustande kommen ließ (vgl. Tabelle 5.1-1).

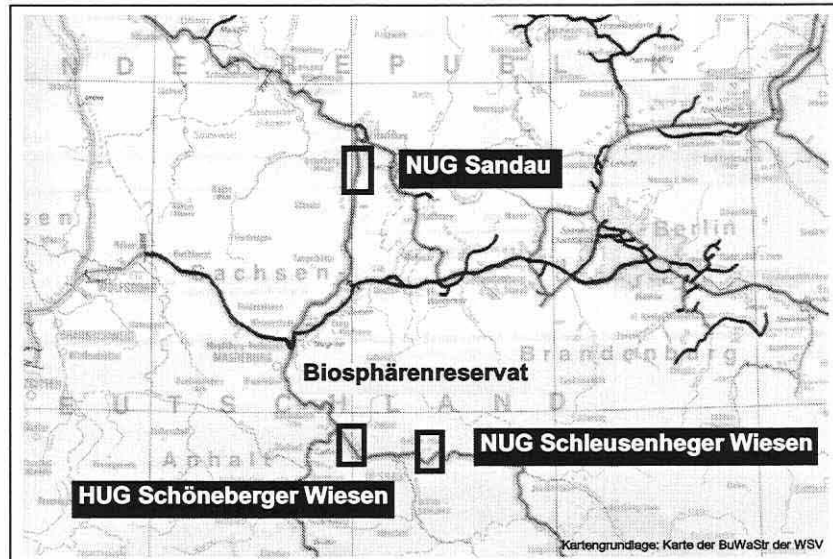
Tabelle 5.1-1: Untersuchungsgebiete des RIVA-Verbundprojektes und Auswahlkriterien

Auswahlkriterien	Untersuchungsgebiete		
	HUG "Schöneberger Wiesen" Elbekilometer 283 bis 285 bei Steckby - ca. 1,3 km <sup>2</sup>	NUG1 "Schleusenheger Wiesen" Elbekilometer 242 bis 243 bei Wörlitz - ca. 0,6 km <sup>2</sup>	NUG2 "Domwerder bei Sandau" Elbekilometer 419 bis 423 bei Sandau - ca. 0,3 km <sup>2</sup>
<b>Topographie</b> (elbangrenzend)	elbangrenzend	elbangrenzend	elbangrenzend
<b>Geografische Lage</b> (HUG nah zu NUG1 und fern zu NUG2)	nahe zu NUG1, oberhalb Saale-Mündung	nahe zu HUG, oberhalb Saale-Mündung	weit entfernt von HUG und NUG1, unterhalb Saale-Mündung
<b>Morphologie</b> (ähnliche Lage, Fluss- und Vorlandmorphologie)	außendeichs rechtselbisch, gewundener, unverzweigter Verlauf, Ufer durchgängig gesichert (Buhnenreihe) hohe Reliefenergie (Flutrinnen, Mulden, Geländestufen, Abgrabung)	außendeichs linkselbisch, gewundener, unverzweigter Verlauf, Ufer durchgängig gesichert (Buhnenreihe) (hohe) Reliefenergie (Flutrinnen, Mulden, Geländestufen)	außendeichs rechtselbisch, gestreckter, unverzweigter Verlauf, Ufer durchgängig gesichert (Buhnenreihe) hohe Reliefenergie (Flutrinnen, Mulden, Altarm Mühlhafen, Geländestufen)
<b>Nutzung / Repräsentativität</b> (Nutzungshomogenität, Elbauen)	wechselfeuchtes Auengrünland, Flutrasen, angrenzend Wald Beweidung eher intensiv vor 1989; extensiv seit 1989 Mahd extensiv seit 1995 typische Nutzung zwischen den Elbdeichen	wechselfeuchtes Auengrünland, Flutrasen, angrenzend Wald Beweidung eher extensiv vor 1991; Portionsweide und Mahd seit 1991 typische Nutzung zwischen den Elbdeichen	wechselfeuchtes Auengrünland, Flutrasen, Pappel-Baumreihen, angrenzend Wald und ähnlich genutzte Flächen Beweidung extensiv, ab 1998 Portionsweide typische Nutzung zwischen den Elbdeichen

### 5.1.3 Zu den Projektuntersuchungsgebieten

#### Geografische und topographische Lage

Das Hauptuntersuchungsgebiet "Schöneberger Wiesen/Steckby" (HUG) und das Nebenuntersuchungsgebiet "Schleusenheger Wiesen/Wörlitz" (NUG1) liegen im ost-westwärts gerichteten Flusslaufabschnitt der Mittleren Elbe.



**Abbildung 5.1-1: Übersichtskarte zur Lage der Untersuchungsgebiete (Kartengrundlage: DBWK2 der Wasser- und Schifffahrtsverwaltung).**

Ein weiteres Nebenuntersuchungsgebiet "Dornwerder bei Sandau" (NUG2) befindet sich daran anschließend im nördlich gerichteten Abschnitt der Mittleren Elbe (vgl. Abbildung 5.1-1). Das HUG befindet sich rechtselbisch bei Elbekilometer 283 bis 285, etwa 40km südlich von Magdeburg. Das Gebiet umfasst ca. 1,3km<sup>2</sup> bei einer mittleren Höhenlage von 51 bis 55m+HN. Es handelt sich um ein Auengrünland im Vordeichland, welches zu dem nordwestlichen Teil des etwa 43.000ha großen Biosphärenreservats "Mittlere Elbe" gehört.

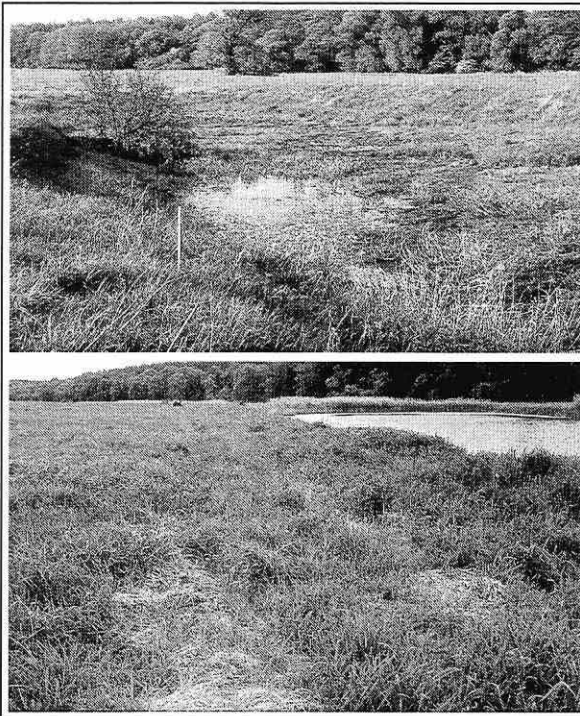
Das flussaufwärts befindliche NUG1 ist linkselbisch zwischen Elbekilometer 242 bis 243 gelegen, nur wenige Kilometer östlich von Dessau auf einer Höhe von 59 bis 62m+HN. Die Gesamtfläche ist mit 0,6km<sup>2</sup> deutlich geringer. Es kann ebenfalls als Auengrünland angesprochen werden, welches sich in der südöstlichen Kulturlandschaftszone des Biosphärenreservats befindet. HUG und NUG1 liegen im zentralen Teil des Mittellaufs der Elbe zwischen den Mündungen der Schwarzen Elster, der Mulde und der Saale.

Oberhalb der Havelmündung zwischen Elbekilometer 417 bis 418, ungefähr 90km nördlich von Magdeburg, ist das rechtselbische NUG2 lokalisiert. Das Auengrünland von nur 0,3km<sup>2</sup> Fläche liegt auf einer Höhe von 25 bis 30m+HN.

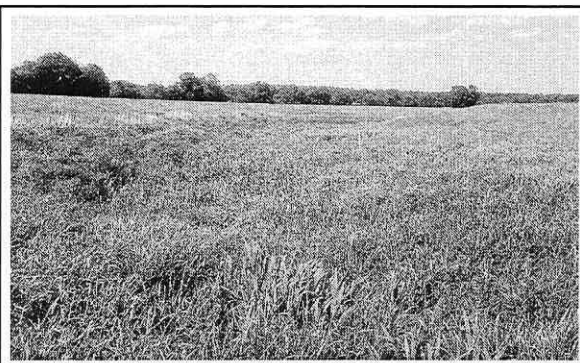
#### Kleinräumige Strukturmerkmale der Untersuchungsgebiete

Das HUG weist einige auffallende morphologische Gegebenheiten auf: Eine Geländestufe im elbnahen Bereich parallel zum Fluss stellt den Uferwall dar. Dieser ist mit einem Anstieg um etwa zwei Meter sehr ausgeprägt. Eine zweite Geländestufe bildet die östliche Begrenzung des Untersuchungsgebiets. Der Bereich zwischen Uferwall und der Begrenzung wird von mehreren in Fließrichtung der Elbe verlaufenden Flutrinnen durchzogen, welche bis auf drei Ausnahmen temporär trockenfallen (vgl. Abbildung 5.1-2, S. 141).





**Abbildung 5.1-2: HUG Schöneberger Wiesen. Probeflächen 11 (oben; Blickrichtung NE) und 33 (unten; Blickrichtung N) im Mai 1998. Im Hintergrund des unteren Fotos ist eine der permanent Wasser führenden Rinnen zu erkennen. Ihr Wasserstand ist im Herbst sehr gering.**



**Abbildung 5.1-3: NUG1 Schleusenheger Wiesen. Probefläche 38 (oben; Blickrichtung N) im Mai 1998.**

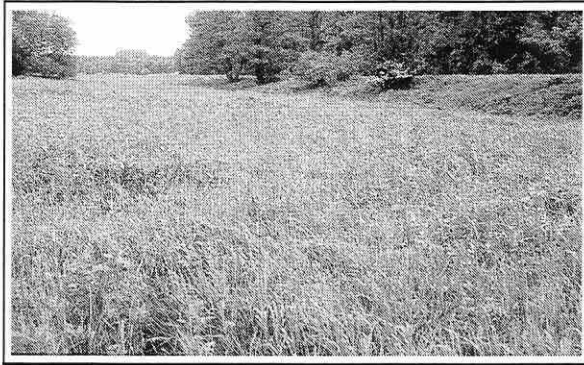
Zwei Auskofferungsflächen von zusammen ca. 500m<sup>2</sup> befinden sich im zentralen Teil des Untersuchungsgebietes und sind gut als rechtwinklige Flächen in Abbildung 5.2-1 (S. 145) auszumachen. Der Fundergraben, ein kleines Fließgewässer, erreicht das Untersuchungsgebiet im Osten und fließt dann parallel zur Geländekante, um an der nördlichen Spitze des Untersuchungsgebietes in die Elbe zu münden.

Das NUG1 erstreckt sich entlang einer Flussbiegung der Elbe. Der Uferwall ist im Norden stärker ausgeprägt als im Süden. Es befinden sich einige Flutrinnen im Gebiet, die teils permanent feucht bzw. versumpft sind (vgl. Abbildung 5.1-3 und Abbildung 5.2-2, S. 146). Im NUG2 erstreckt sich als Nordbegrenzung der sogenannte Mühlhafen, welcher nicht mehr unter wirtschaftlicher Nutzung steht. Dieser ursprüngliche Altarm der Elbe läuft in südöstlich gerichteten Flutrinnen aus. Weiterhin befinden sich im Westteil des Areals einige ständig Wasser führenden Flächen (vgl. Abbildung 5.4-47, S. 213).

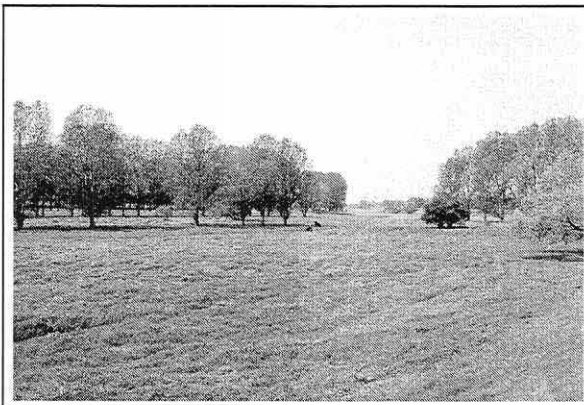
### Böden

Die bodenkundlichen Geländeaufnahmen und Datenerhebungen erfolgten in den Jahren 1998 und 1999 (Meyenburg 1999). Detaillierte Aufnahmen wurden im HUG (vgl. Anlage 1) und dem NUG1 durchgeführt, während für das NUG2 nur probeflächenbezogene Erhebungen vorliegen. Im ufernahen Bereich des HUG gehen Sapropelle aus Organomudden in Wechsellagerung mit Auensand in Ramblen aus Auensanden über. Wei-

ter vom Ufer weg folgt eine Paternia aus Auensanden. Diese tritt auch an einigen Stellen auf den höher gelegenen Bereichen und in näherer Umgebung der Auskofferungsflächen auf. Die höher gelegenen Bereiche im Grünland sind von einer flächenhaften Auelehmdecke überzogen. Es dominieren Vegen aus Auenormallehm bis Auentonschluff über tiefen Auensand. In den tiefer gelegenen Flutrinnen und nahe des Fundergrabens haben sich hauptsächlich Auengleye aus Auensand bis Auentonschluff über Auensande bis Auenschlufftone und Vegagleye aus Auenlehm gebildet.



**Abbildung 5.1-4: NUG1 Schleusenheger Wiesen. Probefläche 46 mit dem Uferwall im Hintergrund (unten; Blickrichtung NE) im Mai 1998.**



**Abbildung 5.1-5: NUG2 Dornwerder bei Sandau. Blick nach NW vom Deich zum Mühlhafen.**

Flugsanddünen legen die Ostbegrenzung des HUG fest. Es sind teils Kolluvialerscheinungen zu beobachten (UFZ 2000, unveröffentlicht). Im NUG1 sind in ufernahen Bereichen Auensande – vorwiegend Rambler – festzustellen. Auf den unteren Geländestufen liegen schwarz-erdeartige Auenböden aus Auenlehm über tiefem Auensand. Auf den sich anschließenden Stufen ist eine flächenhafte Auelehmdecke (Auendeckschicht) auf tiefem Auensand ausgebildet. Diese Deckschicht ist im nördlichen Bereich zwischen eins und zwei Meter mächtig und setzt sich vorwiegend aus Auennormallehm, Sandlehm, Lehmsand sowie Schlufftonen zusammen. Im elbnahen Bereich findet man Sandlehm und Lehmsande, während im elbfernen Bereich Schlufftone auftreten. Im südlichen Bereich des Untersuchungsgebietes ist die Auendeckschicht 0,5 bis 2m mächtig. Im Gegensatz zu den im nördlichen Bereich auftretenden Lehmsanden und Sandlehm treten im Süden Tonlehm und Tonschluffe als typische Zeiger für tiefer gelegene, häufiger vernässte Bereiche auf (Meyenburg 1999).

Die morphologischen Standortbedingungen und die vorherrschenden Bodentypen lassen eine deutliche Beziehung im NUG1 zueinander festzustellen. Die dominierende Bodenform auf höher gelegenen Bereichen ist die Vega aus Auensandlehm über tiefem Auensand. In ausgeprägten Flutrinnen finden sich anmoorige Auengleye aus Auenlehm (UFZ 2000, unveröffentlicht). In intermediärer Lage wurden, abhängig von der hydromorphen Beeinflussung, Gley-Vega, Vega-Gley und Auenwechselgley als Übergangsformen kartiert (Meyenburg 1999). Die Differenzierung der Bodenformen im NUG1 ist wesentlich geringer als im HUG. Die Hauptbodenform im NUG1 ist die Vega aus Auensandlehm über (tiefem) Auensand mit einem Flächenanteil von 67,7%. Im HUG kommt diese nur zu 8,6% vor. Hier ist die häufigste Bodenform die Vega aus Auennormallehm über (tiefem) Auensand mit einem Flächenanteil von 23,9%. Die Untersuchungsichte lässt prozentuale Vergleiche mit dem NUG2 nicht zu. Im NUG2 hat sich aber allgemein im Bereich des Uferwalls eine Paternia aus Auensand und Auenlehm gebildet. Auf den höher gelegenen Geländeabschnitten und Kuppen ist die humose Normal-Vega aus Auenschluff über Auenlehm präsent. In leichten Senken kommt es schon zu Übergängen von Vega-Gleyen zu anmoorigen Auenwechselgleyen aus schlammigem Auenton über Auensand. In den tiefer gelegenen Flutrinnen findet man ausschließlich Auenwechselgleye, an den flachen Ufern des Altarmes "Mühlhafen" tiefe Auengleye. Auffallend ist, dass dieses Untersuchungsgebiet stärker durch das Vorkommen von Auengleyen bzw. Auenwechselgleyen geprägt ist als die beiden flussaufwärts gelegenen Gebiete. Auch

die Substratzusammensetzung tendiert mehr zu lehmigen und tonigen Bodenarten im Vergleich zu den sandigeren Böden des HUG und NUG1 (UFZ 2000, unveröffentlicht).

#### Abflussgeschehen der Elbe und Hydrologie der Untersuchungsgebiete

Das Abflussregime der Elbe ist durch hohen Abfluss im hydrologischen Winterhalbjahr (November-April) und durch geringeren Abfluss im hydrologischen Sommerhalbjahr (Mai-September) gekennzeichnet. So beträgt an den Pegeln der Mittel- und Unterelbe der mittlere Abfluss des Winterhalbjahres ca. 122% und der des Sommerhalbjahres 78% vom mittleren Jahresabfluss. Als Folge der Schneeschmelze sind die Monate März und April die abflussreichsten Monate. Die abflussärmsten sind August bis November. Auf Grund ihrer Durchflussparameter und ihrer Regimekennziffer kann die Elbe als mitteleuropäischer Strom des Regenschnee-Typs eingestuft werden. Die überwiegende Zahl bedeutender Hochwasserereignisse entsteht als Folge intensiver Schneeschmelze in den Mittelgebirgen in Verbindung mit großflächigem ergiebigem Regen (IKSE 1994).

Ansteigendes Hochwasser macht sich zunächst durch die Exfiltration von Druck- oder Qualmwasser in den Senken und Flutrinnen bemerkbar. Dann kommt es zur Überflutung der tiefer gelegenen Bereiche durch Rückstau über Gräben und Altwasserarme und schließlich bei weiterem Flusswasseranstieg zur Überschwemmung des gesamten Deichvorlandes. Einige Flutrinnen im HUG bilden eine Ausnahme hinsichtlich der Überflutungsdynamik, da ihr Wasserstand nicht direkt vom Flusswasserpegel abhängig ist, sondern durch Isolationslage, Bodenverhältnisse und den Einfluss des Fundergrabens die Überflutungsdauer deutlich verlängert ist.

Das Grundwasser in den Untersuchungsgebieten ist überwiegend als nicht flurnah zu bezeichnen. Die Grundwasserflurabstände betragen bis zu drei Metern und in höher gelegenen Bereichen bis fünf Meter. Das Grundwasser steht wegen gut durchlässiger Sande an der Flusssohle in engem hydraulischem Kontakt zum Flusswasser, so dass die Grundwasser-Verhältnisse in der Aue stark vom Flusswasserstand abhängen. Zusätzlich nimmt von den flussbegleitenden Hochlagen einströmendes Grundwasser Einfluss (UFZ 1999, unveröffentlicht).

#### Makroklimatische Kenndaten

Das Gebiet der Mittleren Elbe liegt im Übergangsbereich zwischen dem ozeanisch geprägten Klima Westeuropas und dem kontinentalen Klima Osteuropas. Die mittleren jährlichen Niederschlagssummen sind mit rund 500mm verhältnismäßig gering (vgl. Tabelle 5.1-2). Die jährliche mittlere relative Feuchte beträgt 78%.

**Tabelle 5.1-2: Klimatische Kennwerte für den Großraum der Untersuchungen (Quellen: Hydrologischer Atlas von Deutschland (2000) und Auskünfte des Deutschen Wetterdienstes).**

		Monatliche Niederschläge in Millimeter bei Aken für 1998/1999 (Shj/Whj: Sommer-/Winterhalbjahr)			
		98	99	21 Jahres Durchschnitt	
Potenzielle Evapotranspiration bei Grasoberfläche im Zeitraum 1961 bis 1990 (mittlere Jahressummen):		Jan	335	279	316
		Feb	95	622	298
		Mar	406	420	392
		Apr	413	344	397
		Mai	304	547	479
Gesamtgebiet	575 - 650mm/a	Jun	797	651	619
		Jul	727	247	549
		Aug	509	451	558
		Sep	781	317	422
		Okt	1006	128	314
Mittlere Jahressummen der tatsächlichen Verdunstung (aktuelle Berechnungen):		Nov	418	407	399
		Dez	237	506	490
		Sum (1/10mm)	606	492	523
		Shj (1/10 mm)	353	256	302
		Whj (1/10 mm)	253	236	221
		Mittlerer Jahresniederschlag (1/10 mm)			523
Gesamtgebiet	450 - 500mm/a				
km 240-280 (HUG, NUG1)	450mm/a				
km 415-420 (NUG2)	500mm/a				

Die Tagesmitteltemperaturen liegen im Juli bei + 17,5°C und im Januar bei - 4°C. Die mittlere Anzahl der Eistage liegt bei 24, die der Frosttage bei 81 und die der Sommertage bei 37 nach Müller-Westermeier (1996).

### Landwirtschaftliche Nutzung

Für die Untersuchungsgebiete im RIVA-Projekt stellt sich das Bild der Nutzung wie folgt dar:

#### HUG:

**bis 1989:** Beweidung (200 Großvieheinheiten/ha\*) mit NPK-Düngung und Beregnung mit Elbwasser  
**1989 - 1995:** Beweidung ohne Düngung  
**seit 1995:** unter Vertragsnaturschutz, einschürige Mahd (auch Flutrinnen) nach dem 15. Juni; in manchen Jahren wird im Herbst nachgemäht, keine Düngung.

#### NUG1:

**bis 1991:** Beweidung (50 Großvieheinheiten/ha)  
**1991 - 1997:** Mahd und Beweidung von September bis Mitte Oktober mit drei Großvieheinheiten/ha  
**1998:** Portionsweide (70-80 Großvieheinheiten/ha), mulchen möglich

#### NUG2:

**bis 1998:** extensive Beweidung unter Vertragsnaturschutz  
**seit 1998:** Portionsweide

\*eine Großvieheinheit entspricht etwa einer Kuh (von 500kg) oder 10 Schafen

## 5.2 Datenerhebungsstrategie

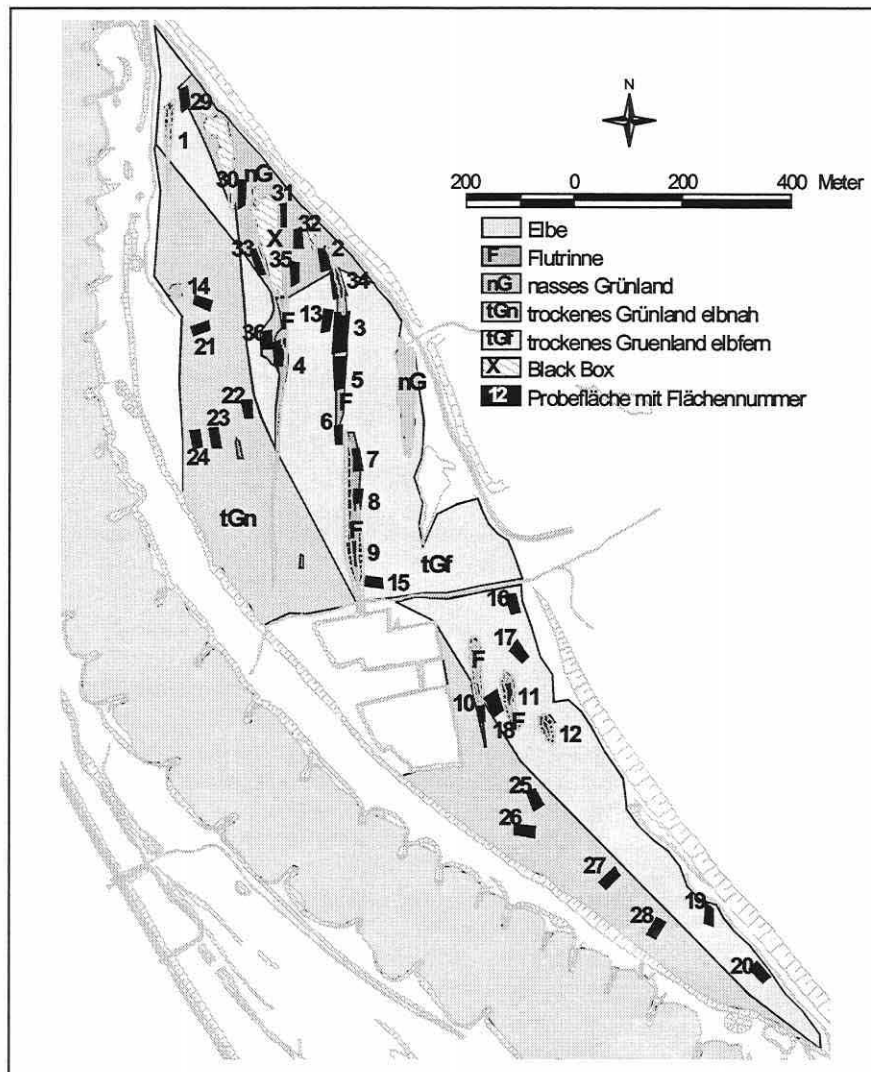
### 5.2.1 Stichprobenumfang, Stichprobenplan und Stratifizierung

Der im RIVA-Verbundprojekt angewendete Stichprobenplan stellt eine den Untersuchungsgebieten angepasste Mischung aus den in Teilbild D und E der Abbildung 3.1-2 (S. 31) dargestellten Prinzipien vor und entstand als Kompromisslösung zwischen logistisch vorgegebenem Stichprobenumfang und der Erfüllung spezifischer statistischer und problembezogener Anforderungen.

Der Stichprobenplan nach dem Prinzip der "stratifizierten Zufallsverteilung" wird als beste Lösung in Anbetracht des maximal leistbaren – aber mit insgesamt nur 60 Probeflächen theoretisch eher eingeschränkten – Stichprobenumfangs und des Projektziels angesehen (vgl. Abschnitte 3.1.2 und 3.1.3), obgleich er für die sehr mobile Indikatorgruppe Carabiden nicht optimal ist (vgl. Abschnitt 5.4.2). Aus der gleichgestalteten Struktur und Nutzungsintensität der Untersuchungsflächen (vgl. Abschnitt 5.1.2) begründet sich die spezifische Stratifizierung des HUG (Abbildung 5.2-1) und des NUG1 (Abbildung 5.2-2, S. 146). Die untere Geländestufe zur Elbe hin ist in beiden Gebieten über fast alle grundlegenden Kriterien hinweg analog und würde sich als weiteres Stratum (vgl. Abbildung 3.1-4, S. 32) neben dem "trockenen höhergelegenen Grünland" und den "Flutrinnen" auch von der räumlichen Abfolge der Straten und dem Gedanken eines Feuchtegradienten sehr sinnvoll in den Versuchsplan integrieren lassen. Im HUG ist sie allerdings auf ihrer gesamten Fläche ungenutzt und daher nicht direkt vergleichbar mit der übrigen extensiv genutzten Formation "Grünland". Da das Wirkungsgeflecht und die funktionalen Verknüpfungen bei der Vielgestaltigkeit der Systemelemente ohnehin nur schwer überschaubar sind, ist eine Trennung/Ausgliederung der relevanten Standortfaktoren für eine zu erwartende Bestandsveränderung der Bioindikatoren ohne den Ausschluss des überprägenden Umweltfaktors Nutzung nicht sicher möglich. Daraus resultierte die Ausgliederung der Teilfläche.

Im NUG1 hingegen ist der nördlichere Abschnitt der elbnahen tief gelegenen Geländestufe im gleichen Umfang genutzt wie die restlichen Flächen. Um die eventuelle Ausdehnung des Existenzgradienten der Bioindikatoren auf diesen Bereich zu erfassen, wurde die genannte Teilfläche der unteren Geländestufe im NUG1 als Stratum "nasses Grünland" in den Stichprobenplan aufgenommen.





**Abbildung 5.2-1: Stratifizierung und Lage der Probestflächen im Hauptuntersuchungsgebiet.**

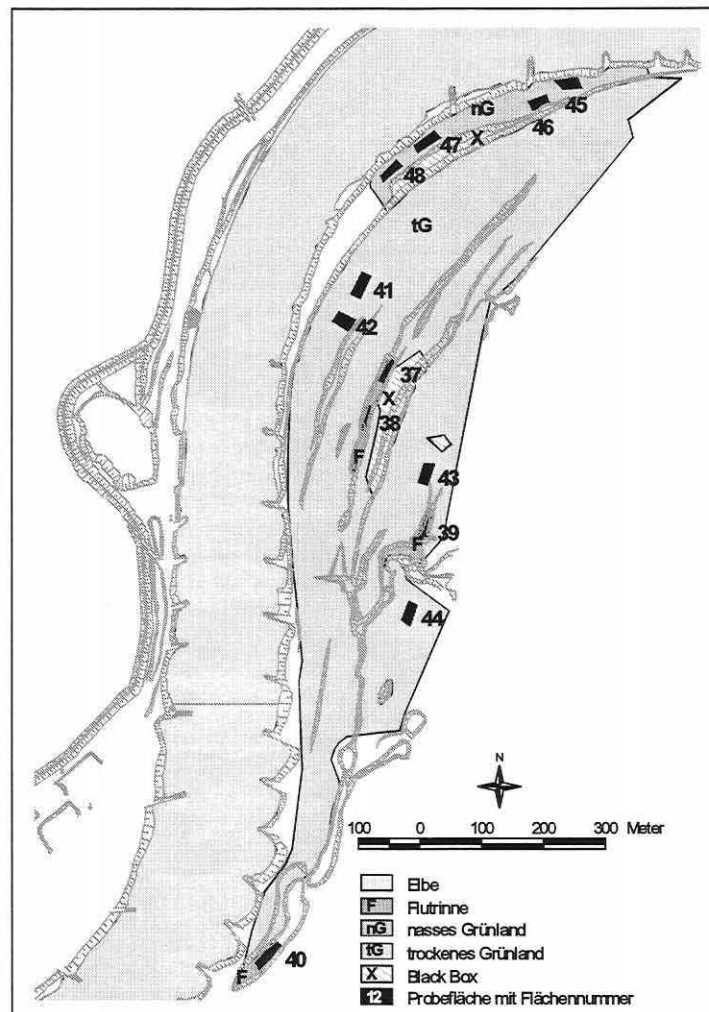
Im HUG wurde dahingegen eine Teilfläche als Stratum "nasses Grünland" definiert, die sich in ihrer spezifischen Eigentümlichkeit des Standorts nicht in den Nebenuntersuchungsgebieten wiederholt (vgl. Tabelle 5.2-1). Sie wurde mit einem reduzierten Stichprobenumfang von acht analysiert. Die Differenz zu den eigentlichen zwölf Stichproben pro Stratum im HUG wurde dem flächig sehr ausgedehnten Stratum "trockenes Grünland" zugeschlagen, das dadurch mit insgesamt 16 Stichprobenflächen belegt ist.

Der Entschluss zu dieser Versuchsanordnung erzeugt eine Abweichung der definierten Straten über die drei Untersuchungsgebiete, d.h., es gibt insgesamt vier Stratentypen, die sich auf die ausgewählten Areale ungleich verteilen (vgl. Tabelle 5.2-1; den Stichprobenplan für das NUG2 zeigt Abbildung 5.4-47, S. 213). Statistisch gesehen ist dieses Design problembehaftet, da die Stratifizierung unter anderem einen Minimumstichprobenumfang pro Stratum gewährleisten soll, um eine ausreichende Wiedererkennbarkeit und damit Ver-



gleichbarkeit der Lebensgemeinschaften in ihrer spezifischen Zusammensetzung unter ähnlichen Lebensbedingungen zu erzielen (vgl. Abschnitte 3.1.2 u. 3.1.3).

Die Lageunterschiede des Stratum "nasses Grünland" in den Untersuchungsgebieten könnten die Vergleichbarkeit aufheben und, falls die Arten sich nicht wiederfinden, eine Modellübertragung scheitern lassen. Umgekehrt bringt der Umstand den Vorteil, dass bei einer Trennfähigkeit der Schlüsselfaktoren im Modell deren räumliche Robustheit bei der Übertragung erste Bestätigungen erfährt, obgleich abgewandelte Standorte von ihrer abiotischen Gesamtausstattung vorliegen (vgl. Abschnitt 5.4.4).



**Abbildung 5.2-2: Stratifizierung und Lage der Probestflächen im ersten Nebenuntersuchungsgebiet.**

Der Grundsatz "ausreichender Stichprobenumfang" wird innerhalb des HUG möglichst konsequent eingehalten, im NUG1 und NUG2 aber nur ein lückenhaftes, allein nicht aussagefähiges Bild erzeugt (jeweils  $n = 12$ ), welches der Validierung oder Falsifizierung der Modellergebnisse aus dem intensiver beprobten HUG dienen soll. Eine Zusammenfassung der Strattentypen pro Untersuchungsgebiet, ihre Abgrenzungskriterien und den Stichprobenumfang gibt Tabelle 5.2-1 (S. 147).

Tabelle 5.2-1: Straten, Stichprobenumfang und Abgrenzungskriterien im Haupt- und den Nebenuntersuchungsgebieten.

Untersuchungs- gebiet	Straten			Probe- flächen	Stratifizierungskriterien	
	Anzahl	Typisie- rung	kurz		primär	sekundär
HUG: Schöneberger Wiesen	3+1 (Anm.)	Flutrinnen	F	12	morphologisches Ge- samtbild der Rinnen- struktur, extensiv genutzt	Eintiefung größer 50cm zum umgebenden Gelän- deniveau
		trockenes Grünland elbfem	tGf	8	morphologisch hochgele- gen zur Elbe (oberhalb Geländestufe) und somit flussferner, etwaige Hal- bierung der Gesamtflä- che "trockenes Grünland" (alleatorische Linie <sup>(Anm.)</sup> zwischen "tGf" und "tGn" orientiert sich teils an der 53m+HN-Isolinie), exten- siv genutzt	trocken (hochgelegen, flussfern), hydrologische Bedingungen am Standort primär durch die Elbe geprägt
		trockenes Grünland elbnah	tGn	8	morphologisch tiefegele- gen zur Elbe (unterhalb Geländestufe) und somit flussnäher; ansonsten wie tGf	wie tGf
		nasses Grünland	nG	8	morphologisch hochgele- gen zur Elbe (oberhalb Geländestufe), extensiv genutzt	nass, hydrologische Bed- ingungen am Standort primär durch den Einfluss des Fundergrabens (Was- serzufuhr) und durch Elbe geprägt
				Σ36		
NUG1: Schleusen- heger Wiesen	3	Flutrinnen	F	4	morphologisches Ge- samtbild der Rinnen- struktur, extensiv genutzt	Eintiefung größer 50cm zum umgebenden Gelän- deniveau
		trockenes Grünland	tG	4	morphologisch hochgele- gen zur Elbe (oberhalb Geländestufe), extensiv genutzt	trocken (hochgelegen, flussfern), hydrologische Bedingungen am Standort primär durch Elbe geprägt
		feuchtes Grünland	fG	4	morphologisch tiefegele- gen zur Elbe (unterhalb Geländestufe), flussnah, extensiv genutzt	feucht bis nass (tiefegele- gen, flussnah), hydrologi- sche Bedingungen am Standort primär durch Elbe geprägt
				Σ12		
NUG2: Dornwerder bei Sandau	3	Flutrinnen	F	4	morphologisches Ge- samtbild der Rinnen- struktur, extensiv genutzt	Eintiefung größer 50cm zum umgebenden Gelän- deniveau
		trockenes Grünland	tG	4	morphologisch hochgele- gen zur Elbe (hier über der 26m+HN-Isolinie), extensiv genutzt	trocken (hochgelegen, flussfern), hydrologische Bedingungen am Standort primär durch Elbe geprägt
		feuchtes Grünland	fG	4	morphologisch tiefegele- gen zur Elbe (hier unter der 26m+HN-Isolinie), extensiv genutzt	feucht bis nass (tiefegele- gen, flussnah), hydrolo- gische Bedingungen am Standort primär durch Elbe geprägt
				Σ12		
Gesamtstich- probenumfang				Σ60		

(Anm.) : Im Hauptuntersuchungsgebiet "Schöneberger Wiesen" wird die sehr große Fläche des Stratum "trockenes Grünland" in die zwei Abschnitte elbnah und elbfem durch eine alleatorische Linie unterteilt. Es erfolgt eine Gewichtung des Stichprobenumfangs von 16:8 zwischen trockenem und nassem Grünland.

### 5.2.2 Arbeitsschritte bei der Erstellung und Etablierung des Stichprobenplans "stratifizierte Zufallsverteilung" im RIVA-Verbundprojekt

Die prinzipielle Umsetzung des im RIVA-Projekt angewendeten Stichprobenplans "stratifizierte Zufallsverteilung" wurde unter Einsatz eines Geografischen Informationssystems (GIS; vgl. Abschnitt 3.1.5) und eines satellitengestützten Navigationssystems zur Ermittlung von hochgenauen Lagekoordinaten (GPS; vgl. Abschnitt 3.1.6) realisiert. Das Vorgehen wird in Detailschritten anhand der Abbildung 5.2-3 (S.149f) erläutert.

Die Recherche nach Kartengrundlagen und Luftbildern zum Aufbau der topographischen und morphologischen Grundinformation zu den Untersuchungsgebieten im GIS bildete den ersten Schritt der Stichprobenplanerstellung (Abbildung 5.2-3, Teilbild 1). Die Stratifizierung des Untersuchungsraumes begann dann mit der Festlegung der Untersuchungsgebietsgrenzen und der Trennung relevanter von nicht untersuchungsrelevanten Flächen (meist Wege oder ähnliches). Nachdem die Grundgesamtheit ausgewählt war, erfolgte mithilfe des GPS die eigentliche Aufteilung der Fläche in Zielgesamtheiten (Straten) (Abbildung 5.2-3, Teilbild 2, ).

Bei allen Messkampagnen in den Untersuchungsgebieten zur Stratifizierung und Probeflächenverortung sind Echtzeit-Empfänger (RTK-GPS: Real Time Kinematic - Global Positioning System) eingesetzt worden, welche die exakten Positionskoordinaten direkt in Gelände anzeigen. Der Einsatz des RTK-GPS bei der Geländestratifizierung erfordert die simultane Nutzung zweier Empfänger (Referenzstation und mobiler Empfänger/Rover; vgl. Abschnitt 3.1.6). Eine detaillierte Beschreibung der Gerätschaften, deren Bedienung sowie Erfahrungen im Gelände sind Rink (1998) zu entnehmen.

Nach der GIS-Übertragung und der Bearbeitung aller Stratifizierungsdaten wurde im GIS auf der Grundlage georeferenzierter Karten die Dreiecksvermaschung mit dem projektspezifischen 10 m-Raster erzeugt, die Teildatensätze der Vermaschungspunkte pro Stratum selektiert und herausgelesen (Abbildung 5.2-3, Teilbild 2A). Neben einer auf die physikalische Größe der kleinsten Stratumsteilfläche angepassten Maschenweite wurde eine dreieckig angeordnete Vermaschung einem quadratischen Raster vorgezogen, um durch den reihenweisen Versatz der Knotenpunkte eine optimale Raumüberdeckung zu erhalten. Ein Zufalls-generator (Programmpaket SPSS<sup>®</sup>) suchte nacheinander einzelne Knotenpunkte aus den Teildatensätzen heraus, so dass diese mit einer numerischen Abfolge belegt werden konnten. Diese Knotenpunkte dienten für die späteren Probeflächen als Basis Eckpunkte (Abbildung 5.2-3, Teilbild 2B). Die Lagekoordinaten des Basis Eckpunktes einer Probefläche wurden anschließend im Gelände aufgesucht, wozu die RTK-GPS-Funktionalität erneut eingesetzt wurde.

Nach der Ortung wurde der Basispunkt der Probefläche dauerhaft gekennzeichnet und die weiteren Eckpunkte verpflockt. Die Flächen wurden zufällig in der Anordnung ihrer Längsseite ausgerichtet. In morphologisch begrenzten Straten (z.B. "Flutrinnen"), musste sich die Fläche in ihrer Lage und Ausdehnung der Geländestruktur anpassen. Die markierten Eckpunkte der Probeflächen wurden dann mit dem GPS zur Visualisierung im GIS eingemessen (Abbildung 5.2-3, Teilbild 3).

Mit dem Datentransfer aller GPS-Lagekoordinaten ins GIS konnten Feldbücher erstellt werden, die Lage und Ausdehnung der Probeflächen zeigen und im Projekt jedem Bearbeiter zur Verfügung gestellt wurden (Abbildung 5.2-3, Teilbild 4). Zur räumlichen Anordnung der Teiluntersuchungsflächen und Messeinrichtungen innerhalb einer Probefläche wurden weiterhin Anordnungsschemata erstellt (vgl. Abschnitt 3.1.4), um den einzelnen Fachwissenschaftlern eine Richtlinie bei der Verortung zu Verfügung zu stellen (Abbildung 5.2-3, Teilbild 5).



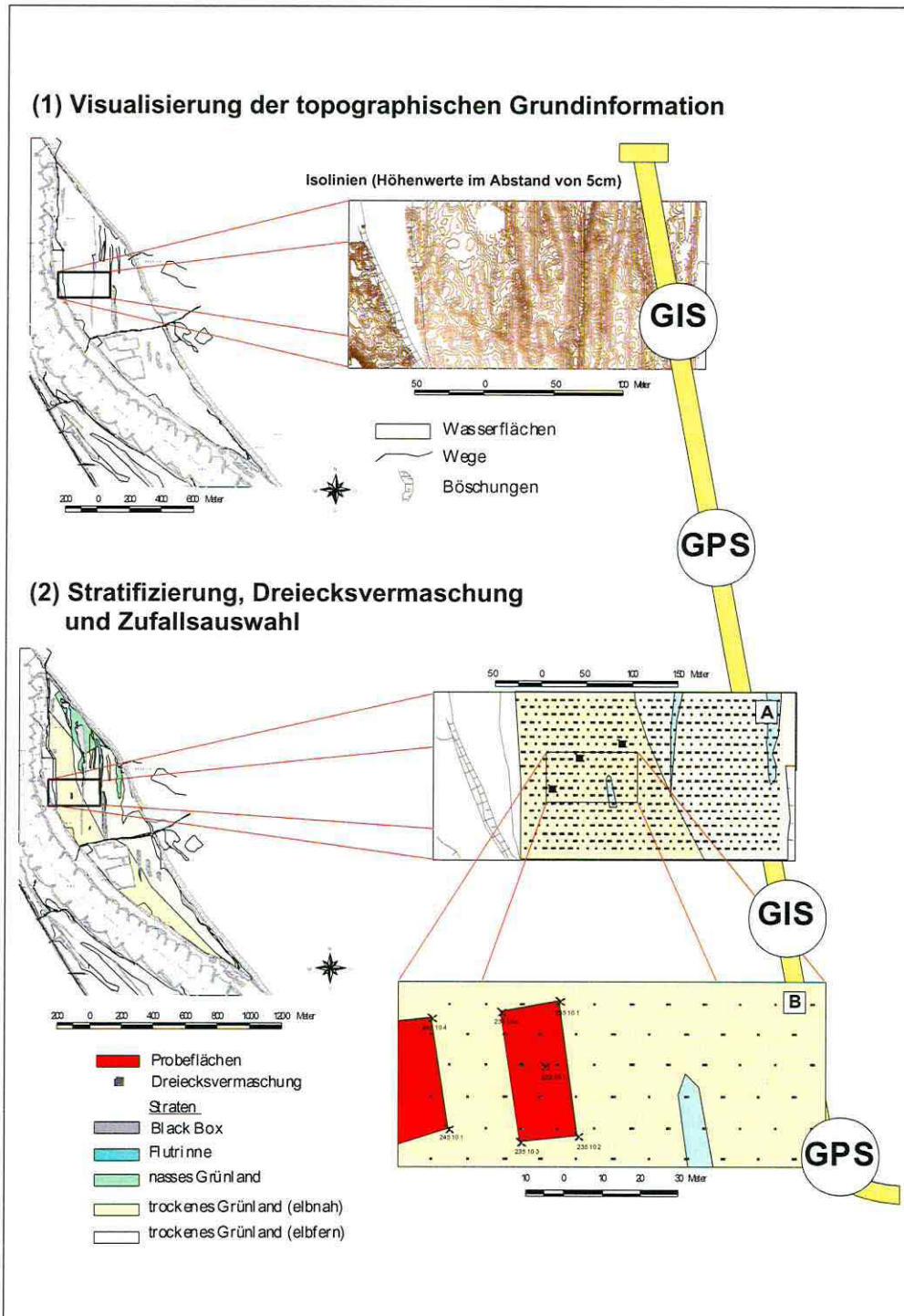


Abbildung 5.2-3: Abbildungshierarchie der Arbeitsschritte bei der Erstellung und Etablierung des Stichprobenplans "stratifizierte Zufallsverteilung" im RIVA-Verbundprojekt. Erläuterungen im Text.

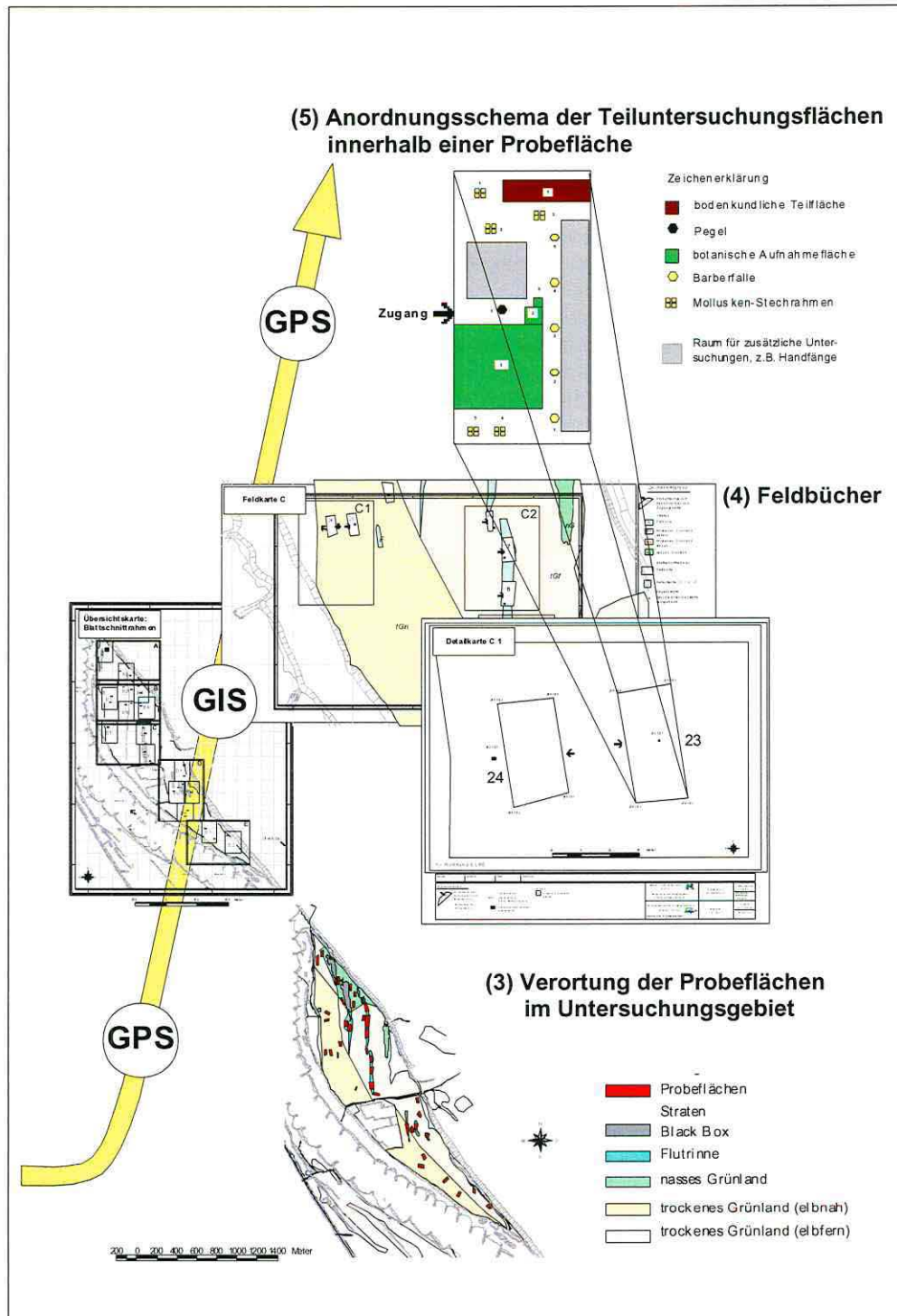


Abbildung 5.2-3 Fortsetzung: Abbildungshierarchie der Arbeitsschritte bei der Erstellung und Etablierung des Stichprobenplans "stratifizierte Zufallsverteilung" im RIVA-Verbundprojekt. Erläuterungen im Text.



Die pflanzensoziologischen Aufnahmeflächen gliedern sich in drei Teilflächen mit 100m<sup>2</sup>, 4m<sup>2</sup> und 1m<sup>2</sup> auf, die durch das Teilprojekt Vegetation (vertreten durch Prof. Klotz, UFZ Halle-Leipzig) definiert wurden. Für das Teilprojekt Laufkäfer (vertreten durch Prof. Gerken, Hr. Schanowski und Hr. Figura – Institut für Landschaftsökologie und Naturschutz Bühl/Gesamthochschule Paderborn-Höxter) wurde eine Reihenanzahl der fünf Barberfallen pro Probefläche vorgegeben und die Flächengröße für die Handfänge bestimmt. Die Weichtierbeobachtung (Teilprojekt vertreten durch Dr. Foeckler, ÖKON Gesellschaft für Landschaftsökologie, Gewässerbiologie und Umweltplanung mbH Regensburg) wurde anhand von je fünf über die Probefläche verteilten Stechrahmenproben durchgeführt. Aus dem summierten Platzbedarf aller Erhebungen und Messeinrichtungen der Teilprojekte ergab sich eine Minimalgröße von 450 bis 500m<sup>2</sup> für eine Probefläche.

Der so genannte Zugangspfeil markiert die Basisseite einer Probefläche, an der die Anordnung der Untersuchungen ausgerichtet wurde. Er sollte weiterhin einen Zutrittspfad festlegen, um die restliche Probefläche zu schonen.

Die hydrologische Messstelle (Grundwasserbeobachtungsrohre, DN 30) ist mittig der Längserstreckung dezentral zum Probeflächenschwerpunkt eingerichtet worden, um eine maximale räumliche Nähe zu den biotischen Untersuchungsflächen zu erreichen (vgl. Abschnitt 2.6). Die bodenkundlichen Standortparameter wurden hingegen randlich der Probefläche entnommen, weil teils erhebliche Störungen bei den Datenerhebungen nicht vermieden werden konnten.

Die Bereitstellung der Anordnungsschemata, der Arbeitskarten und einer Richtlinie zur zeitlichen Abstimmung der Einzelerhebungen (vgl. Abschnitt 5.2.3) sowie die Einrichtung der Teiluntersuchungsflächen im Gelände, eine letzte alternativ zur Handeintragung durchgeführte GPS-Einmessung der Einzelprobenahmestandorte innerhalb der Probeflächen und die georeferenzierte Übernahme ins GIS schlossen die Stichprobenplanerstellung und -umsetzung im Gelände ab. Im GIS waren damit alle Themen (z.B. Probeflächen) und Elemente der Themen (z.B. Teiluntersuchungsflächen und Messeinrichtungen) für weitere Bearbeitungen verfügbar. Die Datenerhebung konnte jetzt nach dem vorgesehenen Zeitrahmen erfolgen (vgl. Rink et al. 2000).

### 5.2.3 Zeitliche Abstimmung der Datenerhebung

Die Laufzeit des Projektes betrug 3,3 Jahre (September 1997 bis November 2000), wobei die Felddatenerhebungen in den Jahren 1998 und 1999 stattfanden. Der eigentliche Beobachtungszeitraum war somit auf zwei Jahre beschränkt. Hinsichtlich des Projektziels ist dies eine sehr knapp bemessene Zeit, da die langfristige natürliche Dynamik von Aueökosystemen einige Probleme für die Untersuchung aufwarf. Wichtige Gesichtspunkte hierzu werden an dieser Stelle besprochen:

Die Untersuchungen im RIVA-Projekt konnten hinsichtlich der hydrologischen Situation in den Elbauen nur einen Ausschnitt erfassen. Da die Abfluss- und Überflutungs-dynamik des Flusses mehrjährigen Zyklen unterworfen ist, war es ein reiner "Glücksfall", dass ein hinsichtlich der Überflutungen nasses und ein trockenes Jahr erfasst wurden. Langfristig angesiedelte Lebensgemeinschaften müssen aber in ihrer Existenzfähigkeit oder Wettbewerbsfähigkeit diesen Extremsituationen generell gewachsen sein.

Für die statistische Datenanalyse im RIVA-Verbundprojekt war es bedeutsam, gemeinsame Erhebungsphasen zu definieren, um neben der räumlichen (vgl. Abschnitt 5.2.1) auch eine zeitlich integre Datenerhebung der Biotik und Abiotik zu gewährleisten. Zeitphasen möglicher Datenerhebungen wurden nach fachlichen Anforderungen in Abbildung 5.2-4 (S. 152) zusammengestellt (Bsp.: 1998, gültig auch 1999).

Die beiden jährlichen Hauptuntersuchungsphasen (UP1 und UP3 in 1998 sowie UP5 und UP7 in 1999) wurden auf Ende April bis Ende Mai und den Monat September festgelegt. Speziell für die Fauna sind zusätzlich Daten im Sommer (Juni-Juli) erhoben worden. Darüber

hinaus galt als zeitliche Grundperiode der Datenerhebung die Kalenderdoppelwoche. Für alle zu beprobenden dynamischen Parameter sollte 14-tägig ein gemessener Datensatz zur Verfügung stehen oder wenigstens gesichert extrapoliert werden können.

Um Störungen für die mobile Gruppe der Laufkäfer auf den Stichprobenflächen möglichst gering zu halten und den Einfluss der verschiedenen Beprobungen auf den natürlichen Bestand und die Fangergebnisse weitgehend auszuschalten, ist die Erhebung des Carabiden-Inventars an den Anfang der Hauptuntersuchungsphasen zu stellen gewesen. Eine trotzdem auftretende Beeinflussung kann aber prinzipiell nicht völlig ausgeschlossen werden, da manche Arten empfindlich auf jegliche Störung reagieren.

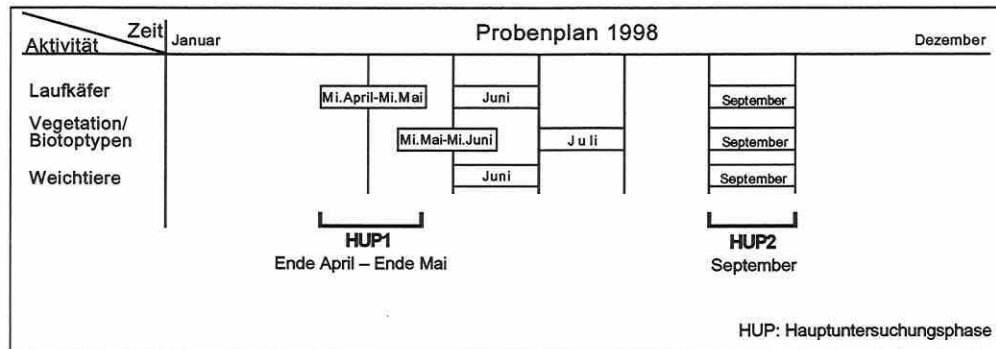


Abbildung 5.2-4: Zeitliche Abstimmung der Einzelerhebungen und der Hauptuntersuchungsphasen (HUP) für 1998 (gültig auch 1999) nach den fachlichen Vorgaben der Teilprojekte.

### 5.3 Datensätze und Datenbankstruktur

#### 5.3.1 Biotik: Die Indikatorgruppen

Für eine Erfassung biozönotisch wirksamer Veränderungen in Auen müssen Organismen (-gruppen) berücksichtigt werden, die auf das Vorhandensein oder Fehlen auenprägender Faktoren besonders empfindlich, gut erkennbar und reproduzierbar reagieren. Als Indikatorgruppen dienten daher Laufkäfer, Weichtiere und die Flora (vgl. auch Spang 1995). Ein allgemeiner Faktor für die Auswahl der drei Indikatorgruppen im RIVA-Projekt war auch deren unterschiedliche Mobilität. Unterschiede zwischen den resultierenden Indikatoreigenschaften der Biodeskriptoren können so erkannt und beurteilt werden (vgl. Abschnitt 5.4 und 5.5).

Die Indikatorgruppe der **Weichtiere** (Mollusken) ist die bestuntersuchte Wirbellosen-Gruppe an der Elbe (Petermeier et al. 1996). Sie wurde und wird gleichermaßen häufig für Indikationssysteme an mehreren europäischen Flüssen eingesetzt, und es ist empirisch erwiesen, dass sich die Tiergruppe – nicht zuletzt wegen ihrer epigäischen Lebensweise – sehr gut eignet, um Standortcharakterisierungen vorzunehmen, Verknüpfungen zu Standortfaktoren abzuleiten und Störungen in Auensystemen nachzuweisen (s. Hügin & Henrichfreise 1992).

Die seit Jahrzehnten intensiv studierte Familie der **Laufkäfer** (Carabiden) zählt zu den wenigen Invertebraten-Taxa semiterrestrischer Standorte, deren ökologische Ansprüche umfangreich dokumentiert sind und die daher auch häufig für Indikationszwecke herangezogen werden. Speziell für den Elberaum liegen teils sehr weit zurückreichende, wohl dokumentierte Artenlisten vor und das qualitative Arteninventar für die Elbe und ihr Einzugsgebiet ist bekannt. Außerdem ist die Biologie und die bodennahe Lebensweise der Carabiden (vgl. Mollusken) besonders geeignet, um den Einfluss von bodenkundlichen Parametern, dem Bodenwasserhaushalt, der Überflutungsdynamik und der Vegetationsstruktur auf Invertebraten beschreiben zu können (Gerken 1989; Thiele 1977).

Ein weiterer Vorteil ergibt sich aus der schnellen Reaktion (Wochen) durch die Mobilität der Tiere auf Veränderungen von hochdynamischen Standortparametern. Eine Beeinflussung oder Störung der Lebensraumdynamik sollte sich daher in kurzen Zeiträumen niederschlagen und damit messbar werden.

Die **Flora** rezenter Auen – es wurden nur höhere Pflanzen erfasst – als Indikatorgruppe dürfte oft über längere Zeitphasen integrieren und zeigt daher eher mittlere Standortbedingungen an. Sie ist darüber hinaus besonders bedeutsam, da sie flächig kartiert (vgl. Anlage 2) und eine Prognose ausgehend von Punktdaten mit dem Istzustand in der Fläche verglichen werden kann. Weiterhin macht sie den Charakter der Aue optisch aus und ist im Naturschutzmanagement stärker öffentlichkeitswirksam als manch andere Indikatorgruppe. Letztere Eigenschaft kann sich für den Einsatz eines Prognoseinstrumentes bei politischen Entscheidungsfindungen als wertvoll erweisen.

Als Folge einer unterschiedlich hohen Toleranz gegenüber Hochwässern, neben anderen Faktoren wie Grundwassereinfluss, Ausprägung des Substrates und mechanischer Beanspruchung, kommt es in natürlichen Flussauen zu einer oft charakteristischen Abfolge von Pflanzengesellschaften vom Fluss in die Aue hinein bzw. mit zunehmender Höhe über Mittelwasser. Das Ausbleiben der Überflutungen verändert die Artenzusammensetzung der Aue grundlegend (IKSE 1994). Solche Zusammenhänge sollen statistisch abgesichert und für die Prognose nutzbar gemacht werden.

#### Datensätze: Weichtiere (Mollusken)

Es wurden Stechrahmenproben nach dem Nasssiebverfahren ausgewertet. Insgesamt wurden 35 Taxa vorgefunden, von denen 26 bis zur Art bestimmt werden konnten. Die 35 Taxa teilen sich in 13 Wassermollusken und 22 Landmollusken auf, bei denen es sich meist um auentypische Arten handelt (Foeckler et al. 2000).

Tabelle 5.3-1: Numerische Eigenschaften der Datensätze Mollusken (5 Stechrahmen pro Probefläche s.u.).

Anzahl der Probeflächen (n)	60 (HUG=36; NUG1=12; NUG2=12)			
Anzahl der Arten	UP1:	UP3:	UP5:	UP7:
<b>HUG</b>				
gesamt (p)	23	22	21	24
Vorkommen auf mehr als zwei Probeflächen	16	15	15	18
<b>NUG1</b>				
gesamt (p)	11	14	17	14
Vorkommen auf mehr als zwei Probeflächen	9	7	7	8
<b>NUG2</b>				
gesamt (p)	13	11	12	17
Vorkommen auf mehr als zwei Probeflächen	5	5	7	6
<i>Anzahl der Elemente in der Datenmatrix (max. mögliche Anzahl an Beobachtungen (ot = n · p)</i>				
HUG (n=36)	828	792	756	864
NUG1 (n=12)	132	168	204	168
NUG2 (n=12)	156	132	144	204
<i>Anzahl der nicht-null Einträge (Anzahl der Beobachtungen: oo)</i>				
HUG	181	203	209	238
NUG1	39	45	50	45
NUG2	31	34	39	47
<i>Anteil der nicht-null Matrixelemente (oolot · 100)</i>				
HUG	22%	26%	28%	28%
NUG1	30%	27%	25%	27%
NUG2	20%	26%	27%	23%
<i>Mittlere Anzahl von Arten auf einer Probefläche (pp = ooln)</i>				
HUG	5	6	6	7
NUG1	3	4	4	4
NUG2	3	3	3	4
<i>Anzahl der Arten mit einem Vorkommen auf über 20 Probeflächen (p20); NUG1 und NUG2 (n&lt;20)</i>				
HUG	3	3	4	3



Zusammen werden sieben Arten nach der Roten Liste Deutschlands (nicht nach der Roten Liste Sachsen-Anhalts) als gefährdet angesprochen. Neben den Lebendfunden wurden auch die Totfunde miterhoben. Diese gehen jedoch nicht in die statistische Datenanalyse ein. Die numerischen Eigenschaften der Datensätze fasst Tabelle 5.3-1 zusammen.

Vergleicht man Tabelle 5.3-1 (S. 153) mit Tabelle 5.3-3 und Tabelle 5.3-4 (S. 155f) fällt auf, dass die vorgefundene Artenzahl der Mollusken deutlich geringer ist als bei den anderen Indikatorgruppen. Innerhalb der Gruppe ist die absolute Artenzahl im HUG durch den differierenden Stichprobenumfang zu den NUGs höher, während der relative Anteil an nicht-null Matrixelementen sehr ähnliche Verhältnisse belegt. Die mittlere Anzahl von Arten erreicht im HUG ihr Maximum. Spezies mit weit gefasster Realnische (Vorkommen auf über 20 Probeflächen) sind selten. Wie sich einige ausgewählte Arten im HUG verteilen, legt Anlage 4 offen.

**Tabelle 5.3-2: Individuenzahl der Art *Anisus spirorbis* für eine und fünf Stechrahmenproben (PF: Probefläche; Z: Anzahl Stechrahmen; UP: Untersuchungsphase 1/5: Frühjahr 98/99, 3/7: Herbst 98/99).**

UP	PF	1	2	3	4	5	6	7	8	9	10	11	12	...	29	30	31	32	33	34	35	36
	Z																					
1	5	185	9	39	7	18	35	6	21	14	205	3	4	...	16	21	1	1	12	6	24	1
1	1	80	6	21	6	12	24	1	10	1	7	1	0	...	2	1	1	0	3	6	4	0
3	5	300	96	124	123	474	65	109	73	141	922	11	7	...	11	13	1	1	5	45	6	1
3	1	40	78	13	11	160	33	56	35	28	306	6	1	...	1	12	0	1	4	44	5	1
5	5	211	30	9	8	12	23	30	18	27	484	1	2	...	11	4	7	4	12	3	3	0
5	1	118	6	4	7	3	20	5	8	2	1	0	2	...	8	1	4	4	11	1	2	0
7	5	208	54	137	39	153	34	39	22	8	431	12	3	...	4	6	1	0	1	7	14	1
7	1	10	32	76	8	114	2	29	3	3	1	4	1	...	0	5	1	0	0	3	14	0

Für die Mollusken wurde weiterhin eine spezielle Auswertung durchgeführt. Es sollte ermittelt werden, inwieweit sich die Anzahl der genommenen Proben je Probefläche auf den Gesamtdatensatz auswirkt. Verglichen wurde der Standarddatensatz und ein Datensatz, welcher pro Probefläche nur eine zufällig gewählte Stechrahmenprobe aus der Menge von fünf berücksichtigt. Die Gesamtzahl der vorgefundenen Arten war in beiden Datensätzen identisch, Differenzen zeigten sich aber für Einzelarten, wie am Beispiel der Art *Anisus spirorbis* in Tabelle 5.3-2 verdeutlicht wird. Einen starken Einfluss nahm die Anzahl der Stechrahmenproben auf die Individuenzahl (vgl. PF1) und das generelle Auffinden der Art auf Probeflächen mit geringen Vorkommen (vgl. z.B. PF 12/UP1 oder PF 36/UP1 und 7). Daraus schlussfolgert sich, dass bei der Definition der ökologischen Nische einer Art auch die Beprobungsintensität pro definierter Flächeneinheit modifizierend wirkt. Modelltechnisch trägt dies zum Fehler 1. Art bei (vgl. Abschnitte 2.4 und 3.7.1). Die Ergebnisse für die Probeflächen 13 bis 28 sind in Tabelle 5.3-2 ausgearbeitet, da die Beispielart auf diesen Flächen nahezu nicht vorkommt.

#### **Datensätze: Laufkäfer (Carabiden)**

Es wurden je fünf Barberfallen innerhalb einer Probefläche aufgestellt und ausgewertet (vgl. Abschnitt 5.3.5 bezüglich der Standardisierung). Insgesamt wurden 133 Spezies bestimmt. Ein Großteil der Spezies sind typische Auenbewohner (Schanowski & Figura 2000). Nach der Roten Liste Deutschlands bzw. der Roten Liste Sachsen-Anhalts sind 22 Arten als gefährdete Arten bewertet. Die numerischen Eigenschaften der Datensätze beschreibt Tabelle 5.3-3 (S. 155). Stellt man erneut Tabelle 5.3-3 der Tabelle 5.3-1 (S. 153) und Tabelle 5.3-4 (S. 155) gegenüber, nimmt die vorgefundene Artenzahl eine intermediäre Stellung zwischen den Mollusken und der Flora ein. Durch die vergleichsweise höhere Mobilität dieser Artengruppe (vgl. auch Häufigkeit des Auftretens) ist die absolute Artenzahl im HUG trotz des differierenden Stichprobenumfangs ähnlich der in den NUGs. Daher übersteigt auch der relative Anteil an nicht-null Matrixelementen in den NUGs den Wert im HUG. Die mittlere Anzahl von Arten lässt Ähnlichkeiten zwischen dem HUG und den NUG1 erkennen. NUG2 weicht im

Frühjahr des zweiten Untersuchungsjahres ins Positive und im Herbst ins Negative ab. Carabidenarten mit weit gefasster Realnische (Vorkommen auf über 20 Probeflächen) sind bei dieser Indikatorgruppe am häufigsten anzutreffen (s.o.). Die Verteilung einiger ausgewählter Arten im HUG zeigt Anlage 5.

**Tabelle 5.3-3: Numerische Eigenschaften der Datensätze Carabiden.**

<i>Anzahl der Probeflächen (n)</i>	<i>60 (HUG=36; NUG1=12; NUG2=12)</i>			
<i>Anzahl der Arten</i>	<i>UP1:</i>	<i>UP3:</i>	<i>UP5:</i>	<i>UP7:</i>
<b>HUG</b>				
<b>gesamt (p)</b>	<b>99</b>	<b>72</b>	<b>90</b>	<b>71</b>
<b>Vorkommen auf mehr als zwei Probeflächen</b>	<b>72</b>	<b>46</b>	<b>60</b>	<b>42</b>
<b>NUG1</b>				
<b>gesamt (p)</b>	<b>80</b>	<b>49</b>	<b>66</b>	<b>48</b>
<b>Vorkommen auf mehr als zwei Probeflächen</b>	<b>52</b>	<b>25</b>	<b>37</b>	<b>27</b>
<b>NUG2</b>				
<b>gesamt (p)</b>	<b>79</b>	<b>37</b>	<b>70</b>	<b>35</b>
<b>Vorkommen auf mehr als zwei Probeflächen</b>	<b>56</b>	<b>20</b>	<b>46</b>	<b>13</b>
<i>Anzahl der Elemente in der Datenmatrix (max. mögliche Anzahl an Beobachtungen (<math>ot = n \cdot p</math>))</i>				
<b>HUG (n=36)</b>	<b>3564</b>	<b>2592</b>	<b>3240</b>	<b>2556</b>
<b>NUG1 (n=12)</b>	<b>960</b>	<b>588</b>	<b>792</b>	<b>576</b>
<b>NUG2 (n=12)</b>	<b>948</b>	<b>444</b>	<b>840</b>	<b>420</b>
<i>Anzahl der nicht-null Einträge (Anzahl der Beobachtungen: oo)</i>				
<b>HUG</b>	<b>1032</b>	<b>643</b>	<b>831</b>	<b>568</b>
<b>NUG1</b>	<b>400</b>	<b>202</b>	<b>267</b>	<b>171</b>
<b>NUG2</b>	<b>377</b>	<b>165</b>	<b>379</b>	<b>125</b>
<i>Anteil der nicht-null Matrixelemente (<math>oo/ot \cdot 100</math>)</i>				
<b>HUG</b>	<b>29%</b>	<b>25%</b>	<b>26%</b>	<b>22%</b>
<b>NUG1</b>	<b>42%</b>	<b>34%</b>	<b>34%</b>	<b>30%</b>
<b>NUG2</b>	<b>40%</b>	<b>37%</b>	<b>45%</b>	<b>30%</b>
<i>Mittlere Anzahl von Arten auf einer Probefläche (<math>pp = oo/n</math>)</i>				
<b>HUG</b>	<b>29</b>	<b>18</b>	<b>23</b>	<b>16</b>
<b>NUG1</b>	<b>31</b>	<b>17</b>	<b>22</b>	<b>14</b>
<b>NUG2</b>	<b>31</b>	<b>14</b>	<b>32</b>	<b>10</b>
<i>Anzahl der Arten mit einem Vorkommen auf über 20 Probeflächen (p20); NUG1 und NUG2 (n&lt;20)</i>				
<b>HUG</b>	<b>19</b>	<b>12</b>	<b>11</b>	<b>9</b>

#### **Datensätze: Flora**

Es liegt eine flächendeckende Vegetations- und Biotopkartierung der Untersuchungsgebiete vor (Amarell 2000). Erstere kann für das HUG in Anlage 2 eingesehen werden. Die pflanzensoziologischen Aufnahmen erfolgten auf drei Teilflächen innerhalb einer Probefläche (vgl. Abbildung 5.2-3, Teilbild 5, S. 149f), wovon nur die Daten der 100m<sup>2</sup>-Fläche für die in Abschnitt 5.4 durchgeführten Analysen verwendet wurden. Als "Stromtalarten" werden in der Literatur 284 von insgesamt 347 Arten der Untersuchungsgebiete angesprochen. Ihr Verbreitungsschwerpunkt liegt in den großen Flussauen. Als Rote Liste Arten werden nach der Roten Liste Deutschlands und Sachsen-Anhalts 41 Arten eingestuft.

Tabelle 5.3-4 (S. 156) listet die numerischen Eigenschaften der Datensätze auf. Die vergleichende Betrachtung dieser Tabelle mit der Tabelle 5.3-1 (S. 153) und Tabelle 5.3-3 (S. 155) weist die Flora als Indikatorgruppe mit den meisten vorgefundenen Arten aus. Die absolute Artenzahl im HUG und den NUGs zeigt den differierenden Stichprobenumfang sehr klar an. Unterschiede in der Artenzahl finden sich auch zwischen den NUGs. Die Artenzahl ist im NUG1 deutlich höher als im NUG2, wenn man den Ausfall von Probeflächen berücksichtigt. Der relative Anteil an nicht-null Matrixelementen ist in den NUGs höher als im HUG. Durch den höheren Stichprobenumfang im HUG werden eher seltene Arten erfasst, die dann die Spanne zwischen der Anzahl der Elemente und der nicht-null Elemente in der Matrix vergrößern. Als Resultat ist der Anteil der nicht-null Matrixelemente niedriger. Die mittlere Anzahl von Arten lässt Ähnlichkeiten zwischen dem HUG und den NUG1 erkennen, während NUG2



geringere Werte besitzt. Die Anzahl von Arten mit weit gefasster Realnische (Vorkommen auf über 20 Probeflächen) ist mit der Indikatorgruppe Mollusken gleichzusetzen, obwohl die Absolutzahl der Pflanzenarten in den Untersuchungsgebieten um das Fünffache höher ist.

**Tabelle 5.3-4: Numerische Eigenschaften der Datensätze Flora.**

<b>Anzahl der Probeflächen (n)</b>	<b>60 (HUG=36; NUG1=12; NUG2=12)</b>			
<b>Anzahl der Arten</b>	<b>UP1:</b>	<b>UP3:</b>	<b>UP5:</b>	<b>UP7:</b>
<b>HUG</b>				
<b>gesamt (p)</b>	117	127	95 (n=34)*	117
<b>Vorkommen auf mehr als zwei Probeflächen</b>	78	77	65	74
<b>NUG1</b>				
<b>gesamt (p)</b>	63 (n=6)*	88	37 (n=6)*	98
<b>Vorkommen auf mehr als zwei Probeflächen</b>	34	40	28	48
<b>NUG2</b>				
<b>gesamt (p)</b>	44	45	59	58
<b>Vorkommen auf mehr als zwei Probeflächen</b>	19	18	27	25
<b>Anzahl der Elemente in der Datenmatrix (max. mögliche Anzahl an Beobachtungen (<math>ot = n \cdot p</math>))</b>				
<b>HUG (n=36)</b>	4212	4572	3230 (n=34)*	4212
<b>NUG1 (n=12)</b>	378 (n=6)*	1056	222 (n=6)*	1176
<b>NUG2 (n=12)</b>	528	540	708	696
<b>Anzahl der nicht-null Einträge (Anzahl der Beobachtungen: oo)</b>				
<b>HUG</b>	671	682	578 (n=34)*	678
<b>NUG1</b>	115 (n=6)*	227	80 (n=6)*	280
<b>NUG2</b>	135	133	193	185
<b>Anteil der nicht-null Matrixelemente (<math>oolot \cdot 100</math>)</b>				
<b>HUG</b>	16%	15%	(18%)*	16%
<b>NUG1</b>	(30%)*	21%	(36%)*	24%
<b>NUG2</b>	25%	25%	27%	27%
<b>Mittlere Anzahl von Arten auf einer Probefläche (<math>pp = ooln</math>)</b>				
<b>HUG</b>	18	19	(17)*	19
<b>NUG1</b>	(19)*	19	(13)*	23
<b>NUG2</b>	11	11	16	15
<b>Anzahl der Arten mit einem Vorkommen auf über 20 Probeflächen (<math>p20</math>); NUG1 und NUG2 (<math>n&lt;20</math>)</b>				
<b>HUG</b>	4	4	4	4

\* Ausfall von Flächen durch Überflutung und Mahd

### 5.3.2 Abiotik: Standortparameter

Basierend auf den Probeflächendaten, Bohrstockproben und Transektkartierungen wurde für das HUG und das NUG1 eine Bodenformenkarte erstellt (siehe Anlage 1). Das NUG2 wurde nicht flächig kartiert. Die Bodenprobennahme und Bestimmung bodenkundlicher Parameter an der Feststoffphase erfolgte nur für die 60 Probeflächen (Rinklebe et al. 2000). Die bodenkundlichen Grundlagendaten umfassen insgesamt 30 physiko-chemische Erklärungsvariablen (vgl. Anlage 6).

Im HUG und dem NUG1 wurden sieben Bohrungen (vier bzw. drei) niedergebracht, die zu Grundwassermessstellen mit kontinuierlicher Datenerfassung ausgebaut wurden. Im NUG2 wurde die behördliche Genehmigung zur Einrichtung solcher Messstellen nicht erteilt. Die 60 Probeflächen wurden weiterhin mit Grundwasserbeobachtungsrohren (14-tägige Ablesung) sowie teils mit Messlatten zur Ablesung von Überflutungswasserständen ausgestattet (Böhnke & Geyer 2000).

Die hydrologische Datenbank der BfG (HYDABA) lieferte Wasserstände von neun untersuchungsgebietsnahen amtlichen Pegelstationen der Elbe (Barby, Coswig, Dessau, Havelberg, Tangermünde, Vockerode, Wittenberg und Wittenberge) für die letzten maximal 30 Jahre. Zusätzlich wurden zur möglichst exakten Kalibrierung der Wasserstände in den Untersuchungsgebieten auf die amtlichen Pegeldaten – in Amtshilfe der Wasser- und Schifffahrtsämter Magdeburg und Dresden – Elbe-Hilfspegel eingerichtet. Pro Untersuchungsgebiet entstanden ein bis zwei Hilfspegelreihen mit bis zu vier einnivellierten Messeinrichtungen.

### Datensätze: Abiotische Rohdaten und Ableitung von Sekundärparametern

Dieser arbeitsaufwändige Schritt der Datenanalysestrategie zielt darauf, eine differenziertere Datengrundlage aus den Basisinformationen abzuleiten. Eine Übersicht über alle primären und sekundären Erklärungsvariablen in RIVA gibt Anlage 6.

#### Bodenparameter

Der Bodendatensatz wurde teils durch Transformationen umgestaltet, gewichtete mittlere kF-Werte für die Auendeckschicht und die Profile bestimmt sowie die Bodenformen der Probe- flächen Fuzzy-kodiert (vgl. Abschnitt 3.3).

Um die Deckschichtmächtigkeit zu bestimmen, musste diese zuerst eine Definition erfahren, da in der Literatur keine eindeutige Beschreibung vorliegt (Meyenburg und Fuchs, persönliche Mitteilung). Weder ist klar, welche Bodenarten zur stratigrafischen Ausscheidung der Deckschicht führen, noch sind Mächtigkeiten der Bodenhorizonte genannt, die zu einer Einbindung oder Ausscheidung einer stratigrafischen Einheit führen (bei stark geschichteten Profilen). Für RIVA wurde dieses Problem durch eine Definition "gelöst", die sowohl projektintern als auch mit Vertretern des Instituts für Bodenkunde in Hamburg (IfB, vertreten durch René Schwartz und Alexander Gröngroft) besprochen und festgelegt wurde. Die Auendeckschicht wird danach stratigrafisch wie folgt ausgeschieden, um Sekundärberechnungen zu ermöglichen:

**(A) Hauptbodenarten zur Ausscheidung der Auensande/Flusssande:**

- Ss** (reiner Sand)
- Su2** (schwach schluffiger Sand)
- Su3** (schluffiger Sand)
- Sl2** (schwach lehmiger Sand)
- St2** (schwach toniger Sand)
- alle Grobbodenarten wie Grus, Kies, Steine.
- Alle anderen Bodenarten gehören zur Auendeckschicht.

**(B) Mächtigkeit eines Zwischenlagenhorizontes in (A), die zur Unterbrechung der stratigrafischen Einheit "Deckschicht" führt: 10 bis 30cm (bei Einzelfallentscheidung unter Berücksichtigung des gesamten Profilaufbaus).**

#### Hydrologische Parameter

Es wurde eine umfangreiche Zahl an tiefenstufen- und zeitphasenbezogenen Sekundärdaten erstellt (vgl. Anlage 6). Darüber hinaus wurden Minima, Maxima, Amplituden, Mittel, Standardabweichungen, absolute und mittlere Frequenzen, Dauer von Zuständen, Wahrscheinlichkeiten bestimmter Zustände etc. für die Grundwasser- und auch die gemessenen Überflutungswasserstände bestimmt. Diese Daten wurden gleichfalls, wenn es sinnvoll erschien, auf einzelne Zeitphasen innerhalb des Jahres bezogen.

#### Distanz- und Strukturparameter

Die Distanzen und Flächenanteile zu und von Raumeinheiten (Biotoptypen, permanente Wasserflächen etc.) wurden mit dem ARC/INFO-Applikation "DIST" berechnet (vgl. Abschnitt 3.2). Sie umfassen Angaben zu (vgl. Abbildung 3.2-3, S. 49):

- Distanz zur gewählten Zieleinheit (z.B. Biotoptyp)
- Fläche der nächstgelegenen Zieleinheit
- Gesamtfläche der Zieleinheit in 100 m Buffer (Einstellung für Mollusken)
- Gesamtfläche der Zieleinheit in 200 m Buffer (Einstellung für Carabiden)
- Distanz zur Zieleinheit größer einer gewählten Minimumfläche (hier nach projektinterner Absprache auf 100 m<sup>2</sup> festgesetzt).
- Fläche der Zieleinheit größer einer gewählten Minimumfläche (100 m<sup>2</sup>, s.o.)

Die Daten wurden für alle Biotoptypen nach Peterson (1998) erzeugt. Weiterhin wurden Distanzen zu hydrologisch definierten Raumeinheiten berechnet:

- Distanz zum Stratum Flutrinne
- Distanz zu permanenten Wasserflächen
- Distanz zur Flussmitte (Strömung; mechanischer Angriff Fluss: Lage in der Aue)

Es wurden außerdem anhand von Feldbegehungen und Recherchen Mäh- und Weidenutzung für die Probeflächen ermittelt und diese entsprechend kodiert in den Datensatz eingebracht. Außerdem wurde die Streuauflage klassifiziert aufbereitet (vgl. Anlage 6).

### 5.3.3 Aufbau des Geografischen Informationssystems

Der Aufbau des Geographischen Informationssystems (GIS; vgl. Abschnitt 3.1.5) zur Visualisierung der Untersuchungsräume, als Instrument zur Erstellung des Stichprobenplans (vgl. Abschnitte 5.2.1 und 5.2.2) und Umgebung für das abschließende Prognosesystem erfordert georeferenzierte Datengrundlagen, für die neben der Verwendung von Karten und Luftbildern, digitale Geländemodelle erstellt wurden.

In den durch rezente Dynamik geprägten Flussauen erscheint es notwendig Karten und Luftbilder aktuelleren Datums zu nutzen. Bei den Kartengrundlagen sollte ein Maßstab von 1:2500 nicht überschritten werden. Zu Projektbeginn lag für keines der Untersuchungsgebiete eine aktuelle amtliche topographische Karte geeigneten Maßstabs vor. Daher wurde ein "Global Positioning System" (GPS) zur Stratifizierung der Untersuchungsgebiete und zur Georeferenzierung der Probeflächen eingesetzt (vgl. Abschnitt 5.2.2). Viele Informationen im Zusammenhang mit dem Stichprobenplan (z.B. Lage der Probeflächen und Messeinrichtungen) lagen somit direkt digital und georeferenziert zur späteren Verschneidung mit den topographischen Karten im GIS vor. Durch die Fertigstellung der Digitalen Bundeswasserstraßenkarte im Maßstab 1:2000 (DBWK2, Bundesanstalt für Wasserbau) während der Projektlaufzeit stand dann nachträglich eine geeignete Kartengrundlage zur Verfügung. Die DBWK2 ist in digitaler Form erhältlich und die Daten können ohne den Schritt des Digitalisierens einer analogen Kartengrundlage in ein GIS übernommen werden.

Überprüfungen zur Höhengenaugkeit des DGM (HUG) mit GPS und Nivellier ( $n = 1176$  DGM-Rasterpunkte) ergaben, dass 47% aller Werte innerhalb eines  $\pm 15\text{cm}$ - und 96% innerhalb eines  $\pm 30\text{cm}$ -Intervalls von der Realität abweichen. Die gemessenen Extrema liegen bei  $+ 52\text{cm}$  und  $- 29\text{cm}$ . Die Lageabweichungen des DGM (HUG) bewegen sich auf wenige Ausreißer in einem  $\pm 5\text{cm}$ -Intervall. Das GPS weist eine Höhengenaugkeit von  $\pm 5\text{cm}$  auf (Peter et al. 1999).

Die für die Befliegung und Erstellung des DGM vom Auftragnehmer vertraglich geforderten Kenngrößen sind in der Tabelle 5.3-5 zusammengestellt. Das DGM des HUG zeigt Anlage 3.

Tabelle 5.3-5: Kenngrößen im Werkvertrag für die Datengewinnung und abschließende DGM-Erstellung

Zielgrößenspezifizierung: Digitales Geländemodell Projekt RIVA		
Gesuchte Größen	Lage- und Höhenkoordinaten der Untersuchungsgebiete	
Bezugssystem	Amtliches Landessystem Sachsen-Anhalt	
Messpunktabstand / - dichte	Punktabstand: 1 m; Punktdichte: 1 pro m² bzw. Ableitung von 5cm-Höhenlinien	
Hydrologische Randbedingungen		
Aufnahmezeitpunkt	Laub- und schneefrei, bei niedrigem Flusswasserstand Bezug: Schöneberger Wiesen: Pegel Aken Schleusenheger Wiesen: Pegel Coswig Dornwerder bei Sandau: Pegel Tangermünde	
Nachqualifizierung		
Lage und Höhe nachzuqualifizierender Flächen	Nachqualifizierung von Flächen (z.B. Schilf, Flutrinnen) analog zu obiger Zielgrößenspezifizierung	
Einbau in DGM	Integration der nachqualifizierten Punkte in das DGM	
Ergebnisse		
Zugelassene absolute Objektpunktunsicherheit (S = 95%, unbekannte Restsystematiken vernachlässigbar)	Lagegenauigkeit uy = ux = 15cm (95%)	Höhengenaugkeit uH = 15cm (95%)

### 5.3.4 Rohdatenaufbereitung und GIS-Visualisierung der abiotischen Variabilität

Im Rahmen der Rohdatenaufbereitung wurde die Integrität der Datenbankinhalte ausführlich überprüft, konstante Variablen entfernt, Verteilungen der Daten gesichtet und alle Formatkonvertierungen zum Import in CANOCO® und andere Software durchgeführt.

Die Visualisierung der Werte einer jeden Erklärungsvariablen über die Probeflächen in den einzelnen Untersuchungsgebieten durch die GIS-Funktionalität half, Fehlwerte und Extrema in den sehr umfangreichen Datenbanken zu entdecken und Unstimmig- oder Auffälligkeiten nachzugehen.

### 5.3.5 Datenbankstruktur und "Shuttles"

Die allgemeine Datenbankstruktur wird in Abbildung 3.4-1 (S. 54) und Abbildung 5.3-1 aufgezeigt. Modul A hält alle Datenbanken der Teilprojekte vor, welche bei Aktualisierungen einfach ausgetauscht werden können. In Modul B werden alle Abfragen erzeugt, welche für die Datenselektion und -reduktion benötigt werden. Diese sogenannten "Shuttles" sind MS-Access® Abfrageroutinen, die sicherstellen, dass spezifische Teildatensätze immer nach dem gleichen Schema abgefragt bzw. erstellt werden.

Als Beispiel seien die Abfrageroutinen zur Erstellung der Carbidendatensätze angeführt. Es musste eine Standardisierung der Daten erfolgen, um die Vergleichbarkeit der probeflächebezogenen Fänge zu gewährleisten. Ausfälle einiger Barberfallen und unterschiedlich lange Expositionszeiten waren zu berücksichtigen. Daher wurden die Fangergebnisse summiert und durch die Gesamtzahl der Expositionstage über alle gültigen Fallen einer Probefläche dividiert. Die resultierenden Individuenzahlen nehmen dadurch sehr niedrige Werte an, so dass sie mit dem Faktor 100 multipliziert wurden, um die abschließenden standardisierten Häufigkeitswerte zu erhalten.

Alle durch die Abfragen in Modul B (Abbildung 5.3-1) extrahierten Daten werden dann in der nachgeschalteten Datenbank des Moduls C abgelegt und stehen zum Export in andere Software bereit.

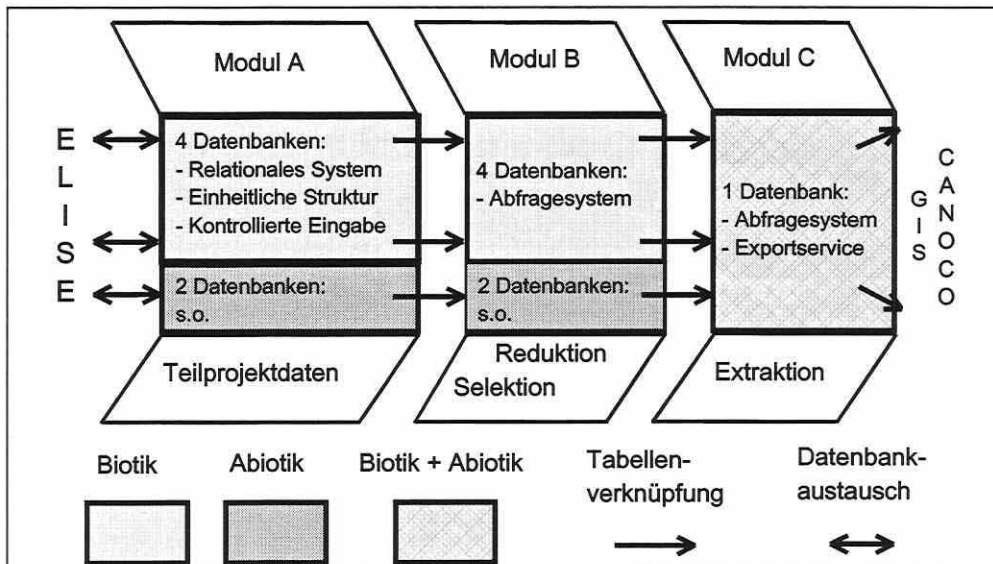


Abbildung 5.3-1: Module der Projektdatenbank in RIVA (Erläuterungen im Text).

## 5.4 Multivariate Datenanalyse: Mustererkennung und Modellbildung

### 5.4.1 Modellbildung: Variablenselektion mit zeitlicher Robustheitsprüfung

Ziel dieses Abschnittes ist es, die Erklärungsvariablen auf ihre Korrelationsstruktur zu testen und eine Selektion relevanter zeitlich robuster Modellparameter vorzunehmen.

Der Weg von den insgesamt 238 zur Verfügung stehenden Erklärungsvariablen (vgl. Anlage 6) zu den resultierenden 13 Modellparametern führt über die Betrachtung der paarweisen Korrelationen, das Ergebnis von Hauptkomponentenanalysen und untersuchungsphasenbezogenen Vorwärtsselektionen.

Der **erste Schritt** zur Datenreduktion ist die **Kontrolle der paarweisen Korrelationen** zwischen allen Erklärungsvariablen. Die Daten erfüllen überwiegend das Datenniveau einer Intervallskala oder Verhältnisskala, so dass der Produkt-Moment-Korrelationskoeffizient nach Pearson genutzt wird, um die Stärke des Zusammenhangs zwischen zwei Variablen zu quantifizieren. Da keine Hypothesentests angestrebt sind, muss die Forderung nach einer Normalverteilung nicht erfüllt sein. Es befinden sich einige ordinalskalierte Variablen im Datensatz, die eine Gegenkontrolle mit dem Rangkorrelationskoeffizienten nach Spearman erfordern. Auszugsweise wird die Korrelationsmatrix einiger Grundwasserkennwerte für 1998 und einige Zeitphasen des Jahres dargestellt. Abbildung 5.4-1 verdeutlicht, dass speziell Parameter verschiedener Zeitphasen kollinear sind (vgl. Variablennamen mit Anlage 6). Die maximalen Grundwasserflurabstände auf den Probeflächen liegen in den Herbstmonaten. Daher sind die Werte für 1998 (nicht dargestellt), die Vegetationsperiode zwischen den Monaten April bis einschließlich September 1998 (MAXGW498) und die Viermonatsphase von Juni bis einschließlich September 1998 (MAXGW698) identisch. Die Daten können für das Beispiel um zwei von drei Variablen reduziert werden.

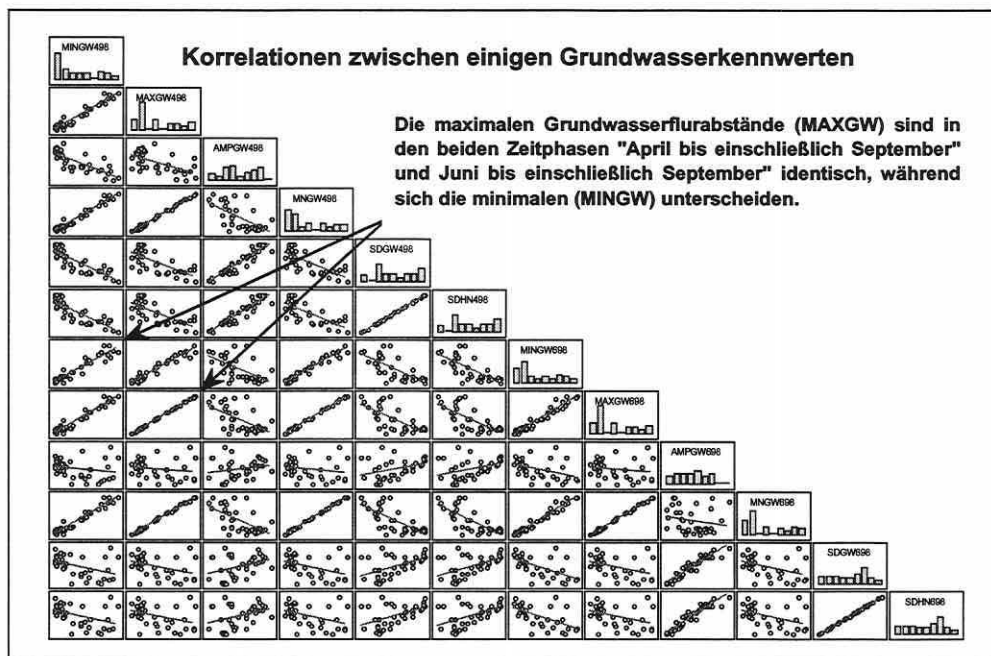


Abbildung 5.4-1: Korrelationen zwischen einigen Grundwasserkennwerten des Jahres 1998 (Variablen: vgl. Anlage 6).



Weitere Beispiele aus dem Bodendatensatz zeigen zusätzliche Möglichkeiten, den Datensatz zu reduzieren. Die pH-Werte der Salzextraktionsmethoden (PHKCLT/PHCACL2T) sind hoch miteinander korreliert und dürfen als kollinear gelten (Abbildung 5.4-2). Die Wassermethode zeigt den aktuellen pH-Wert (PHH2OT), welcher deutlich geringer zu den anderen Werten korreliert ist. Die hohe Korrelation des Gesamtstickstoffs (NT) mit dem Gesamtkohlenstoff (CT) bietet eine weitere Möglichkeit zur Variablenreduktion (Abbildung 5.4-3). Sie kann nur auftreten, wenn nahezu identische Werte für den Gesamtstickstoffgehalt und den organischen Stickstoff auftreten, was auf Kalkabwesenheit im HUG hindeutet. Der Stickstoffgehalt scheint somit hauptsächlich aus der organischen Substanz zu stammen und nicht aus einer mineralisierten Stickstoffquelle.

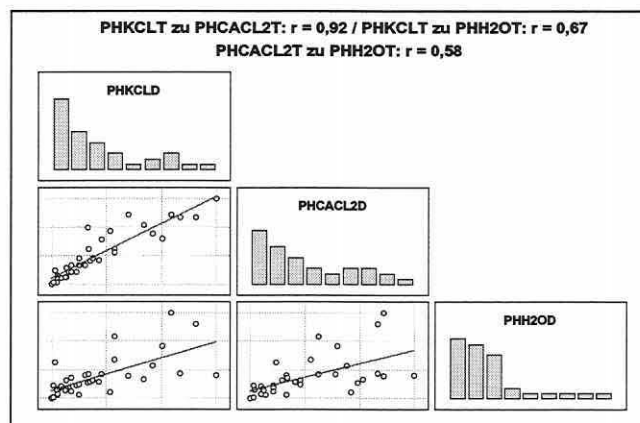


Abbildung 5.4-2: Korrelationen (r) zwischen den verschiedenen pH-Werten (Extraktionsmethoden; Variablen: vgl. Anlage 6).

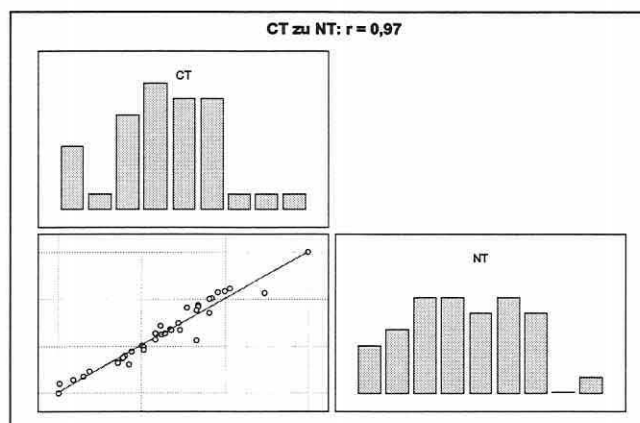


Abbildung 5.4-3: Korrelation (r) zwischen Stickstoff- (CT) und dem Kohlenstoffgehalt (NT).

Tabelle 5.4-1 (S. 162) stellt im Vorgriff die Kolinearitäten (hier paarweise Korrelation  $> 0,9$ ) zwischen den letztlich selektierten Modellparametern (vgl. Abbildung 5.4-4, S. 169) und anderen Parametern im Datensatz zusammen. Die Veränderungen der Korrelationsstrukturen in 1998 und 1999 zeigen, dass Zufälle, Messungenauigkeiten oder andere Einflüsse eine Rolle zu spielen scheinen. Felddaten sind immer durch Datenrauschen belastet, und gerade die Dynamik in den hydrologischen Verhältnissen führt dazu, dass die Korrelationen zwischen den Parametern nicht als absolut stabil angesehen werden dürfen. Sie zeigen eine Tendenz auf, welche kritisch zu prüfen ist. Eine Variablenreduktion ist mit Vorsicht anzuge-

hen, und ein geringer unterschiedlicher Informationsgehalt der Erklärungsvariablen kann dennoch eine unterschiedliche Relevanz für die Antwortvariablen besitzen. Daraus erklärt sich beispielsweise, dass selbst hochkorrelierte Variablen indikatorgruppenspezifisch eine bessere Trennkraft besitzen können als ihr "Pendant".

Tabelle 5.4-1: Paarweise Korrelationen größer 0,9 (nach Pearson) zwischen den Parametern in den Modellen 1998/1999 (kursiv) und anderen Parametern im Datensatz (Variablen: vgl. Anlage 6).

<i>Modellparameter in 1998</i>						
<i>FDAJD98</i>	<i>MAXGW258</i>	<i>MAXFH98</i>	<i>FDFRE98</i>	<i>SDHN98</i>	<i>MNGW498</i>	<i>AMPGW258</i>
FDUP1258	GW140G98	MAXFH258	FDUP1258		GW140M98	SDGW258
PFD98	GW140V98	IFHFD98	PFD98		GW140G98	SDHN258
PFD_UP25	GW140258		PFD_UP25		GW140V98	
MAXFD98	MAXGW98		MAXFD98		GW140258	
	AMPGW98				MAXGW98	
	MEANGW98				AMPGW98	
	MINGW498				MEANGW98	
	MAXGW498				MINGW498	
	MNGW498				MAXGW498	
	MINGW258				MINGW258	
	MNGW258				MAXGW258	
	MINGW698				MNGW258	
	MAXGW698				MINGW698	
	MNGW698				MAXGW698	
					MNGW698	
<i>KAKEFF</i>	<i>PPFL</i>	<i>MSand</i>	<i>FK</i>	<i>PHH2OT</i>	<i>DISTPERM</i>	<i>DISTFLRI</i>
		MTON	MSAND		D3807	
		FK	MTON			
<i>Modellparameter in 1999</i>						
<i>FDAJD99</i>	<i>MAXGW259</i>	<i>MAXFH99</i>	<i>FDFRE99</i>	<i>SDHN99</i>	<i>MNGW499</i>	<i>AMPGW259</i>
FDUP5259	GW140M99	MAXFH259			GW140M99	SDGW259
PFD99	GW140G99				GW140G99	MAXGW499
PFD_UP25	GW140V99				GW140V99	MNGW499
MAXFD99	GW140259				GW140259	MAXGW259
IFHFD99	MAXGW99				MAXGW99	MNGW259
	MEANGW99				MEANGW99	MINGW699
	MINGW499				MINGW499	MAXGW699
	MAXGW499				MAXGW499	MNGW699
	MNGW499				AMPGW259	MEANGW99
	AMPGW259				MAXGW259	GW140M99
	MNGW259				MNGW259	GW140G99
	SDGW259				SDGW259	GW140V99
	MINGW698				MINGW699	GW140259
	MAXGW698				MAXGW699	
	MNGW699				MNGW699	
<i>KAKEFF</i>	<i>PPFL</i>	<i>MSand</i>	<i>FK</i>	<i>PHH2OT</i>	<i>DISTPERM</i>	<i>DISTFLRI</i>
		MTON	MSAND		D3807	
		FK	MTON			

Als **zweiter Schritt** zur Datenreduktion wurde die **Berechnung von Hauptkomponentenanalysen** eingesetzt (vgl. Abschnitte 3.5.3 und 3.5.4). Zur Übersichtlichkeit werden nur die Ergebnisse für den Bodenparameterdatensatz dargelegt.

Die Ergebnisse in Tabelle 5.4-2 (S. 163) zeigen die Zusammenfassung der bodenphysikalischen, der den Säure-Basen-Haushalt beschreibenden und der den Nährstoffgehalt beschreibenden Parameter zu jeweils einer Komponente (PCA-Achse), d.h., es entstehen Variablengruppen mit ähnlichen Erklärungsgehalten. Greift man die Beispiele pH-Werte und Stickstoff-/Kohlenstoffgehalte der paarweisen Korrelationen (s.o.) erneut auf, so finden sich die Parameter auf gleichen Komponenten wieder, und die Korrelationsstrukturen zwischen den Parametern werden von den Ladungen auf die Komponenten nachgezeichnet. Die vierte

Komponente der PCA erscheint wenig eindeutig. Das pflanzenverfügbare Kalium (KPFL) bildet mit dem pflanzenverfügbaren Phosphor (PPFL) die fünfte Komponente. Hervorzuheben ist, dass das pflanzenverfügbare Phosphor auf die erste und fünfte Komponente mit recht hohen Anteilen lädt. Dies zeigt eine uneindeutige Korrelationsstruktur. Der Parameter trägt entweder zu zwei Wirkkomponenten bei oder ist nur artifiziell mit den Parametern einer bzw. der Komponenten korreliert.

Die Korrelationsstruktur des Phosphors zeichnete sich bereits in den paarweisen Korrelationen ab: PPFL ist dabei höher zu den bodenphysikalischen Parametern Luftkapazität (LK: 0,67), mittlerer Sandgehalt (MSAND: 0,67), mittlerer Tongehalt (MTON: -0,66) und der Feldkapazität (FK: -0,65) als zum Kalium (0,41) korreliert.

**Tabelle 5.4-2: Bodenparameter mit hohen Ladungen auf die Hauptkomponenten einer PCA (Variablen: vgl. Anlage 6).**

Ladungen				
Komponente 1	Komponente 2	Komponente 3	Komponente 4	Komponente 5
LK: -0,95	PHKCLT: 0,97	CT: 0,97	CT/NT_VE: 0,77	KPFL: 0,92
KF: -0,91	PHCACL2T: 0,91	NT: 0,92	LF: -0,70	PPFL: 0,47
MSAND: -0,89	PHH2O: 0,75	KAKPOT: 0,77		
FK: 0,88		KAKEFF: 0,76		
MTON: 0,87				
PPFL: -0,76				

Die hier nicht tabellarisch dargestellten PCA-Analyseergebnisse für den Gesamtdatensatz an Erklärungsvariablen in RIVA führen zu einer Verlagerung der Bodenparameter auf die dritte und höhere Komponenten. Die Überflutungs- und die Grundwasserparameter werden auf den ersten zwei Komponenten zusammengefasst und weisen meist Ladungen größer 0,7 auf. Die dominante Bedeutung der hydrologischen Parameter für die Erklärung des biotischen Musters in den Auegebieten wird durch dieses Ergebnis unterstrichen und durch die Ergebnisse aus Schritt vier nochmals deutlich rückbestätigt.

Die Parametersélection für die Ordinationsmodelle in RIVA (vgl. Abschnitt 5.4.2) resultiert weitgehend aus den Schritten eins, drei und vier. Die Funktionalität der Hauptkomponentenanalyse (PCA) als Datenreduktionsmethode sollte jedoch durch den Vergleich mit den anderen Verfahren bestätigt werden. In keinem Modell (vgl. Abbildung 5.4-4, S. 169) treten mehrere Variablen auf, die der gleichen Komponente in Tabelle 5.4-2 angehören.

Der **dritte Schritt** reduziert den Satz der Erklärungsvariablen um **konstante Parameter**, da diese praktisch keine Trennkraft besitzen. Zu den Parametern zählt die Lagerungsdichte (LD) mit Werten von generell 1,5 und 1,6. Ebenso weisen einige Variablen des Distanzdatensatzes konstante bzw. nahezu konstante Werte auf.

Den **vierten Schritt** bildet die **Vorwärtsselektion** (vgl. Abschnitt 3.5.3). Mit dem Basiswissen aus den Schritten eins bis drei werden partielle Ordinationen für die drei Indikatorgruppen anhand der vier Untersuchungsphasen berechnet. Der erste Durchlauf ist univariat (vgl. auch Abschnitt 3.5.9) und zeigt in Tabelle 5.4-3 (S. 164) und Tabellen 5.4-7 bis 5.4-8 (S. 167) die Bedeutung der Variablen für die Erklärung des biotischen Raummusters ohne den Einfluss bereits im Modell befindlicher Parameter. Nach dem Einschluss des effektivsten Parameters (höchster Eigenwert – EV) in die Modellgleichung verändert sich die additive Erklärungsrelevanz der übrigen Parameter (Parameterabfolge) durch verschiedene Kolinearitätsanteile zu den bestehenden Modellparametern. Beim Vergleich der Tabelle 5.4-3 (S. 164) mit Tabelle 5.4-5 (2. Durchlauf, S. 165) wird dieser Sachverhalt deutlich.

Um weiterhin eine zeitlich robuste Parameterauswahl zu treffen, werden die einzelnen Untersuchungsphasen (UP1 u. UP5: Frühjahr 1998/99; UP3 u. UP7: Herbst 1998/99) parallel zueinander bewertet. Es zeigt sich bereits in den univariaten Tabellen des ersten Durchlaufs, dass die Auswertung einer Einzelphase die "echte" Relevanz der Erklärungsvariablen nicht

verlässlich wiedergeben kann. Hervorgehoben wird dies durch die jeweils in die Übersichten des ersten Durchlaufs übertragen Markierungen in den Tabellen 5.4-3 und 5.4-7 bis 5.4-8, S. 167ff), welche die schrittweisen bzw. abschließenden Ergebnisse der Vorwärtsselektion für die Indikatorgruppen darlegen.

Das Prinzip dieses Vorgehens wird am Beispiel der Mollusken ausführlich erläutert (vgl. z.B. Tabelle 5.4-5, S. 165).

### Mollusken:

Im ersten Durchlauf der Vorwärtsselektion wird die jährliche Überflutungsdauer in Tagen (FDAJd98/99) in allen Phasen als Parameter mit dem höchsten Erklärungsanteil am biotischen Muster der Mollusken ausgegeben. FDAJd98/99 hat einen äußerst dominanten Einfluss auf die Artenverteilung der Mollusken im Raum. Im Mittel erklärt sie 32,8% bei einer Gesamtvariabilität von durchschnittlich 1,76 und ist zeitlich hochstabil (Tabelle 5.4-3). Tests bestätigen, dass sie auch die Artenverteilung in den Herbstphasen definiert, obwohl Überflutungen nur im Frühjahr auftreten. Der "Memory-Effekt" weitgehend ortsgebundener Indikatorgruppen (Mollusken und Vegetation), d.h. die meist langsame Reaktionszeit, führen zu dieser Besonderheit.

**Tabelle 5.4-3: Erster Durchlauf der Vorwärtsselektion (univariat) für das Molluskenmodell (UP: Untersuchungsphase; EV: Eigenwert und erklärter %-Anteil an der Gesamtvarianz (TI) in den biotischen Daten). Die nach sechs Durchläufen als Modellparameter ermittelten Variablen wurden in die Liste des ersten Durchlaufs übertragen und grau markiert (Variablen: vgl. Anlage 6).**

UP1	Partielle Ordination EV	% zu 1.821TI	UP3	Partielle Ordination EV	% zu 1.618TI	UP5	Partielle Ordination EV	% zu 1.661TI	UP7	Partielle Ordination EV	% zu 1.914TI
gewmkDe	0.03	1.6 Kpfl	0.06	3.7 pHH2O2	0.06	3.6 SDGW499	0.05	2.6			
LK	0.08	4.4 GWDS698	0.06	3.7 gewmkDe	0.08	4.8 Kpfl	0.08	4.2			
GWDSv98	0.08	4.4 pHACI2t	0.06	3.7 DistFluM	0.08	4.8 MDeckS	0.08	4.2			
gewmkFr	0.09	4.9 pHH2O2	0.06	3.7 gewmkFr	0.09	5.4 SDHN499	0.1	5.2			
GWDSm98	0.09	4.9 MDeckS	0.06	3.7 MDeckS	0.1	6.0 pHH2O2	0.1	5.2			
SDGW98	0.1	5.5 gewmkFr	0.07	4.3 DeckSUn	0.11	6.6 pHACI2t	0.1	5.2			
DeckSUn	0.1	5.5 GW140698	0.08	4.9 pHACI2t	0.11	6.6 ChtNVerh	0.11	5.7			
ChtNVerh	0.11	6.0 ChtNVerh	0.09	5.6 GWDSv99	0.11	6.6 LK	0.12	6.3			
GWDSg98	0.12	6.6 DistFluM	0.09	5.6 GWDSm99	0.12	7.2 gewmkFr	0.13	6.8			
pHCACI2t	0.12	6.6 DeckSUn	0.09	5.6 LK	0.12	7.2 AMPGW499	0.14	7.3			
MDeckS	0.12	6.6 LK	0.12	7.4 SDHN259	0.13	7.8 AMPHN99	0.15	7.8			
Kpfl	0.13	7.1 GWDSv98	0.12	7.4 AMPHN259	0.13	7.8 DeckSUn	0.15	7.8			
nFK	0.15	8.2 AMPHN98	0.12	7.4 SDHN99	0.13	7.8 SDHN99	0.15	7.8			
pHKClt	0.15	8.2 SDGW98	0.12	7.4 GWDSg99	0.14	8.4 DistFluM	0.16	8.4			
DistFluM	0.16	8.8 GWDSm98	0.13	8.0 Kpfl	0.15	9.0 GWDSv99	0.17	8.9			
GW140m98	0.16	8.8 GWDSg98	0.14	8.7 ChtNVerh	0.15	9.0 nFK	0.17	8.9			
AMPHN98	0.16	8.8 nFK	0.16	9.9 AMPGW99	0.15	9.0 GWDSm99	0.17	8.9			
pHH2O2	0.17	9.3 AMPGW499	0.17	10.5 AMPGW499	0.15	9.0 AMPHN499	0.19	9.9			
MinGW98	0.19	10.4 MinGW98	0.17	10.5 DistFIRI	0.17	10.2 AMPGW99	0.22	11.5			
AMPGW98	0.2	11.0 SDHN98	0.19	11.7 nFK	0.17	10.2 DistFIRI	0.22	11.5			
SDGW258	0.2	11.0 SDHN98	0.19	11.7 MinGW259	0.18	10.8 Ppfl	0.23	12.0			
SDHN258	0.2	11.0 SDGW498	0.2	12.4 KAKeff	0.18	10.8 SDHN699	0.23	12.0			
AMPGW258	0.21	11.5 SDHN498	0.2	12.4 MaxGW99	0.19	11.4 SDGW699	0.23	12.0			
GWDS258	0.21	11.5 SDGW698	0.2	12.4 Ctotal	0.19	11.4 GWDSg99	0.23	12.0			
mTon	0.22	12.1 KAKeff	0.21	13.0 AMPHN499	0.21	12.6 KAKeff	0.24	12.5			
KAKeff	0.22	12.1 GW140m98	0.22	13.6 MinGW99	0.21	12.6 AMPGW699	0.24	12.5			
SDGW498	0.23	12.6 AMPGW698	0.22	13.6 GWDS259	0.24	14.4 MinGW99	0.24	12.5			
SDHN498	0.23	12.6 DistFIRI	0.23	14.2 SDGW99	0.24	14.4 mTon	0.25	13.1			
GW140v98	0.24	13.2 mTon	0.25	15.5 mTon	0.24	14.4 GW140699	0.25	13.1			
Ppfl	0.25	13.7 Ctotal	0.25	15.5 DistPerm	0.26	15.7 MaxGW99	0.25	13.1			
SDHN98	0.25	13.7 AMPGW98	0.26	16.1 Ntotal	0.26	15.7 SDGW99	0.27	14.1			
GW140g98	0.25	13.7 Ppfl	0.26	16.1 MinGW499	0.26	15.7 DistPerm	0.28	14.6			
MinGW498	0.26	14.3 MinGW498	0.27	16.7 GW140259	0.26	15.7 Ctotal	0.29	15.2			
Ctotal	0.26	14.3 GW140v98	0.27	16.7 MnGW259	0.27	16.3 FK	0.3	15.7			
MnGW498	0.27	14.8 MinGW698	0.28	17.3 MeanGW99	0.27	16.3 mSand	0.31	16.2			
MaxGW498	0.27	14.8 MaxGW698	0.29	17.9 mSand	0.27	16.3 MeanGW99	0.31	16.2			
MaxGW98	0.27	14.8 MaxGW498	0.29	17.9 GW140g99	0.28	16.9 Ntotal	0.34	17.8			
MaxGW258	0.27	14.8 MaxGW98	0.29	17.9 FK	0.28	16.9 GW140g99	0.34	17.8			
FK	0.27	14.8 MnGW698	0.29	17.9 GW140v99	0.29	17.5 FDFre99	0.34	17.8			
MnGW258	0.27	14.8 MnGW498	0.29	17.9 GW140m99	0.29	17.5 GW140v99	0.34	17.8			
MeanGW98	0.28	15.4 GW140g98	0.29	17.9 SDGW259	0.3	18.1 GW140m99	0.34	17.8			
mSand	0.29	15.9 FK	0.29	17.9 MaxGW499	0.3	18.1 MinGW499	0.35	18.3			
GW140258	0.29	15.9 MeanGW98	0.3	18.5 MaxGW259	0.3	18.1 MaxGW699	0.36	18.8			
MinGW258	0.3	16.5 mSand	0.31	19.2 MnGW499	0.3	18.1 MaxGW499	0.36	18.8			
Ntotal	0.32	17.6 Ntotal	0.31	19.2 AMPGW259	0.3	18.1 MnGW699	0.36	18.8			
DistPerm	0.33	18.1 DistPerm	0.32	19.8 Ppfl	0.32	19.3 MinGW699	0.36	18.8			
KAKpot	0.39	21.4 KAKpot	0.39	24.1 KAKpot	0.32	19.3 MnGW499	0.36	18.8			
FDFre98	0.52	28.6 FDFre98	0.43	26.6 FDFre99	0.32	19.3 KAKpot	0.39	20.4			
maxFH98	0.53	29.1 maxFH98	0.45	27.8 maxFH99	0.4	24.1 maxFH99	0.47	24.6			
FDAJd98	0.64	35.1 FDAJd98	0.58	35.8 FDAJd99	0.5	30.1 FDAJd99	0.58	30.3			

Der zweite Durchlauf ergibt erste Schwankungen in der Listenposition des Parameters mittlerer Grundwasserflurabstand für die Vegetationsperiode (MnGW498 bzw. MnGW499; Tabelle 5.4-3, S. 164 und Tabelle 5.4-6, S. 166). Die Unterschiede können sowohl jahreszeitlich als auch erhebungsjahrspezifisch sein. Die univariate Auswertung für diesen Parameter zeichnet die Tendenz der multivariaten nach. Zur Übersicht werden die beiden Ergebnisse in Tabelle 5.4-4 gegenübergestellt.

**Tabelle 5.4-4: Univariate und multivariate Eigenwerte sowie prozentuale Erklärungsanteile des Parameters mittlerer Grundwasserflurabstand für die Vegetationsperiode (MNGW498/499; U: univariat; M: multivariat).**

UP1				UP3				UP5				UP7			
Eigenwert		%Anteil		Eigenwert		%Anteil		Eigenwert		%Anteil		Eigenwert		%Anteil	
U	M	U	M	U	M	U	M	U	M	U	M	U	M	U	M
0,27	0,24	14,8	13,2	0,29	0,27	17,9	16,7	0,3	0,28	18,1	16,9	0,36	0,31	18,8	16,2

Es ist leicht zu verstehen, dass die multivariaten Ergebnisse insgesamt geringere Erklärungsanteile ab dem zweiten Durchlauf aufzeigen, da die Parameter in ökologischen Studien selten gänzlich unabhängig voneinander sind (vgl. Einführung). Die Hinzu- oder Herausnahme einer Erklärungsvariablen aus dem Modell ändert somit auch sämtliche Beta-Koeffizienten der Modellgleichung (vgl. Abschnitt 3.5.3). Dieses Problem lässt die Vorwärtsselektion zu einem streng hierarchischen Verfahren werden. Schwankungen der Erklärungsanteile von statischen Parametern wie der Feldkapazität (FK) sind auf die Variabilität des biotischen Datensatzes zurückzuführen. Die Artenzahl bleibt nicht konstant und bewirkt geringfügige Differenzen.

Bei der Gegenüberstellung der Untersuchungsphasen verhält sich der Parameter (MnGW498 bzw. MnGW499) am stabilsten hinsichtlich seiner Erklärungsanteile (Tabelle 5.4-5) und wird deshalb nach einer Signifikanzprüfung als zweiter Modellparameter selektiert. Durch dieses Vorgehen wird eine Relevanz der Modellparameter über die Zeit angestrebt. Es bleibt zu beachten, dass an dieser Stelle keine zeitliche Robustheit im eigentlichen Sinne gemeint ist: Die Werte eines Modellparameters aus einem Erhebungszeitraum können nicht direkt zur Prognose bei der zeitlichen Modellübertragung eingesetzt werden. Dieser Selektionsprozess soll nur den Weg für den späteren Test der zeitlichen Übertragung von Ordinationsmodellen bereiten (Abschnitt 5.4.4).

**Tabelle 5.4-5: Zeitliche Robustheitsprüfung am Beispiel des Molluskenmodells: Erster und zweiter Durchlauf der Vorwärtsselektion (Variablen: vgl. Anlage 6).**

UP1	Eigenwert	UP3	Eigenwert	UP5	Eigenwert	UP7	Eigenwert
<b>1. DURCHLAUF</b>							
...	...	...	...	...	...	...	...
FDFre98	0,52	FDFre98	0,43	FDFre99	0,32	KAKpot	0,39
maxFH98	0,53	maxFH98	0,45	maxFH99	0,4	maxFH99	0,47
FDAJd98	0,64	FDAJd98	0,58	FDAJd99	0,5	FDAJd99	0,58
<b>2. DURCHLAUF</b>							
...	...	...	...	...	...	...	...
MaxGW98	0,23	MaxGW498	0,26	MnGW259	0,26	MeanGW99	0,27
MnGW498	0,24	MaxGW98	0,26	GW140g99	0,26	GW140g99	0,28
GW140v98	0,24	MinGW498	0,26	MaxGW499	0,27	MaxGW699	0,29
MeanGW98	0,25	MnGW698	0,26	GW140v99	0,28	MaxGW499	0,29
MinGW498	0,25	GW140g98	0,27	GW140m99	0,28	GW140v99	0,3
MinGW258	0,26	MnGW498	0,27	MnGW499	0,28	GW140m99	0,3
MnGW258	0,26	MinGW698	0,27	MaxGW259	0,29	MnGW699	0,3
MaxGW258	0,26	MeanGW98	0,27	SDGW259	0,3	MnGW499	0,31
GW140258	0,27	GW140v98	0,27	AMPGW259	0,3	MinGW699	0,32



Vergleicht man die Ergebnisse der Vorwärtsselektion für UP1 und UP5 ohne und mit zeitlichem Abgleich über die Untersuchungsphasen (Tabelle 5.4-6), wird deutlich, dass die Modellparameter zum Teil absolut sowie in der Reihenfolge der Aufnahme ins Modell differieren, wodurch die Entscheidung für eine Parallelisierung zur zeitlich relevanten Parameterauswahl nochmals gestützt wird. Untersuchungsphasenbedingte Bedeutungsschwankungen, die über die zeitlich ohnehin immer leicht differierenden Werte von Parametern hinausgehen, führen sonst zu mangelnder Prognosekraft der Modelle und wirken sich für eine echte zeitliche Übertragbarkeit nachteilig aus.

**Tabelle 5.4-6: Differenz der Ergebnisse einer Vorwärtsselektion in UP1 und UP5 ohne (links) und mit zeitlichem Abgleich (rechts) des Molluskenmodells. Die Signifikanzprüfung (P-Wert) beruht auf Monte-Carlo-Tests mit 999 Permutationen (Variablen: vgl. Anlage 6).**

ohne Abgleich			UNTERSUCHUNGSPHASE 1 (1998)			mit Abgleich		
Parameter	Eigenwert	P-Wert	Parameter	Eigenwert	P-Wert	Parameter	Eigenwert	P-Wert
FDAJD98	0,64	0,001	FDAJD98	0,64	0,001	FDAJD98	0,64	0,001
GW140258	0,27	0,001	MNGW498	0,24	0,001	MNGW498	0,24	0,001
MINGW258	0,16	0,001	SDHN98	0,10	0,004	SDHN98	0,10	0,004
PHH2OT	0,07	0,005	AMPGW258	0,14	0,001	AMPGW258	0,14	0,001
SDHN98	0,05	0,017	FK	0,02	0,679	FK	0,02	0,679
DECKSUN	0,05	0,022	PHH2OT	0,07	0,002	PHH2OT	0,07	0,002
GWDSM98	0,05	0,021						
Parameter	Eigenwert		Parameter	Eigenwert		Parameter	Eigenwert	
FDAJD98	0,50		FDAJD98	0,50		FDAJD98	0,50	
AMPGW259	0,29		MNGW498	0,28		MNGW498	0,28	
SDHN99	0,09		SDHN98	0,08		SDHN98	0,08	
SDGW259	0,06		AMPGW258	0,07		AMPGW258	0,07	
PHH2OT	0,04		FK	0,04		FK	0,04	
AMPHN499	0,05		PHH2OT	0,05		PHH2OT	0,05	

Die Abstimmung der Modelle ist mit einem Verlust an Erklärungsanteilen in den einzelnen Untersuchungsphasen verbunden, der aber weniger gravierend ist als eine geringe Stabilität der Modellparameter. Durch die Abstimmung können gelegentlich Parameter die Signifikanzschranke ( $\alpha=0,05$ ) für die Aufnahme ins Modell unterschreiten. Trotz geringfügiger Trennkraft solcher Variablen (vgl. P-Wert der Feldkapazität - FK in UP1, Tabelle 5.4-6) werden diese aus Konsistenzgründen in den multivariaten Modellen mitgeführt, da sie insgesamt über die Untersuchungsphasen eine Relevanz anzeigen.

#### Carabiden:

Deutliche jahreszeitliche Schwankungen der Erklärungsanteile von Modellparametern weist das Carabidenmodell auf. Es prägt sich schon im ersten Durchlauf kein Parameter durch, welcher über alle UP eine maximale Trennkraft aufweist. Der mittlere Grundwasserflurabstand (MnGW498 bzw. MnGW499) ist zeitlich am stabilsten in seiner Trennkraft, welche aber verglichen zu den anderen Modellen mit durchschnittlich 21,9% Erklärungsanteil bei einer mittleren Gesamtvariabilität von 1,14 geringer ausfällt. Aus der insgesamt schlechteren Trennkraft der Einzelparameter resultiert auch die geringere Anzahl der Selektionsdurchläufe. Nach der Aufnahme von fünf Erklärungsvariablen ins Modell wird die gewählte Signifikanzschranke von  $\alpha = 0,05$  bereits erreicht (vgl. auch Abbildung 5.4-4, S. 169).

Der Modellparameter Distanz zu permanenten Wasserflächen (DistPerm) ist in den trockenen Herbstphasen von viel höherer Relevanz als im feuchten Frühjahr, da viele der Flutrinnen zu der Zeit wassererfüllt sind (Distanz zu Flutrinnen: DistFIRi) und erst zum Herbst hin austrocknen. Gleiches zeigt die Überflutungsdauer, die im Frühjahr bedeutender ist. Sie spielt generell bei einer mobilen Artengruppe eine weniger entscheidende Rolle, da die Individuen ihr Raumverhalten dem Überflutungswasserstand oder allgemein gesagt den verän-

dernten Bedingungen anpassen können. Der "Memory-Effekt" weniger mobiler Artengruppen entfällt zum Teil (vgl. Mollusken und Flora).

**Tabelle 5.4-7: Erster Durchlauf der Vorwärtsselektion (univariat) für das Carabidenmodell (UP: Untersuchungsphase; EV: Eigenwert und erklärter %-Anteil an der Gesamtvarianz (TI) in den biotischen Daten). Die nach fünf Durchläufen als Modellparameter ermittelten Variablen wurden in die Liste des ersten Durchlaufs übertragen und grau markiert (Variablen: vgl. Anlage 6).**

UP1	Partielle Ordination		UP3	Partielle Ordination		UP5	Partielle Ordination		UP7	Partielle Ordination		UP7	Partielle Ordination		UP7
EV	% zu 1.412TI		EV	% zu 1.370TI		EV	% zu 1.405TI		EV	% zu 1.519TI		EV	% zu 1.519TI		EV
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
LK	0.03	2.1 pHH2O	0.02	1.5 gewmktDe	0.03	2.1 pHH2O	0.03	2.1 pHH2O	0.03	2.0		0.03	2.0		2.0
gewmktDe	0.04	2.8 GWDS698	0.02	1.5 pHH2O	0.04	2.8 GWDS698	0.03	2.8 GWDS698	0.03	2.0		0.03	2.0		2.0
pHH2O	0.04	2.8 Kpfl	0.03	2.2 Kpfl	0.04	2.8 MDeckS	0.03	2.8 MDeckS	0.03	2.0		0.03	2.0		2.0
Kpfl	0.04	2.8 MDeckS	0.04	2.9 LK	0.04	2.8 Kpfl	0.04	2.8 Kpfl	0.04	2.6		0.04	2.6		2.6
nFK	0.05	3.5 gewmktDe	0.05	3.6 SDHN499	0.04	2.8 gewmktDe	0.05	2.8 gewmktDe	0.05	3.3		0.05	3.3		3.3
SDHN98	0.05	3.5 LK	0.05	3.6 SDGW499	0.05	3.6 SDGW499	0.05	3.6 SDGW499	0.05	3.3		0.05	3.3		3.3
gewmktPr	0.05	3.5 pHKCl	0.06	4.4 nFK	0.05	3.6 LK	0.06	3.6 LK	0.06	3.9		0.06	3.9		3.9
Ppfl	0.06	4.2 SDHN98	0.06	4.4 Ppfl	0.05	3.6 SDHN499	0.06	3.6 SDHN499	0.06	3.9		0.06	3.9		3.9
pHCACl2	0.06	4.2 gewmktPr	0.07	5.1 mTon	0.05	3.6 pHKCl	0.06	3.6 pHKCl	0.06	3.9		0.06	3.9		3.9
mSand	0.06	4.2 Ppfl	0.07	5.1 Ctota	0.05	3.6 AMPHN99	0.07	3.6 AMPHN99	0.07	4.6		0.07	4.6		4.6
mTon	0.06	4.2 AMPHN98	0.08	5.8 AMPHN99	0.06	4.3 Ppfl	0.08	4.3 Ppfl	0.08	5.3		0.08	5.3		5.3
FK	0.06	4.2 nFK	0.08	5.8 FK	0.06	4.3 mTon	0.08	4.3 mTon	0.08	5.3		0.08	5.3		5.3
pHKCl	0.07	5.0 pHKCl	0.08	5.8 mSand	0.06	4.3 mSand	0.09	4.3 mSand	0.09	5.9		0.09	5.9		5.9
Ctota	0.07	5.0 mTon	0.08	5.8 KAKpot	0.06	4.3 nFK	0.1	4.3 nFK	0.1	6.6		0.1	6.6		6.6
DeckSUn	0.07	5.0 mSand	0.09	6.6 FDFre99	0.06	4.3 FK	0.1	4.3 FK	0.1	6.6		0.1	6.6		6.6
MDeckS	0.08	5.7 DistFluM	0.09	6.6 Ntotal	0.06	4.3 DistFluM	0.11	4.3 DistFluM	0.11	7.2		0.11	7.2		7.2
Ntotal	0.08	5.7 FK	0.1	7.3 pHKCl	0.06	4.3 SDHN99	0.11	4.3 SDHN99	0.11	7.2		0.11	7.2		7.2
KAKpot	0.09	6.4 GW140698	0.1	7.3 KAKeff	0.07	5.0 DistFIRI	0.12	5.0 DistFIRI	0.12	7.9		0.12	7.9		7.9
KAKeff	0.09	6.4 PFD98	0.11	8.0 gewmktPr	0.08	5.7 Ctota	0.12	5.7 Ctota	0.12	7.9		0.12	7.9		7.9
AMPHN98	0.1	7.1 FDAJd98	0.11	8.0 MDeckS	0.08	5.7 PFD99	0.13	5.7 PFD99	0.13	8.6		0.13	8.6		8.6
AMPGW258	0.1	7.1 Ctota	0.11	8.0 DeckSUn	0.08	5.7 FDAJd99	0.13	5.7 FDAJd99	0.13	8.6		0.13	8.6		8.6
CtNtVerh	0.1	7.1 DeckSUn	0.11	8.0 CtNtVerh	0.09	6.4 FDFre99	0.14	6.4 FDFre99	0.14	9.2		0.14	9.2		9.2
SDHN258	0.1	7.1 SDGW98	0.11	8.0 DistPerm	0.09	6.4 KAKpot	0.14	6.4 KAKpot	0.14	9.2		0.14	9.2		9.2
SDGW258	0.1	7.1 AMPGW698	0.11	8.0 AMPHN499	0.1	7.1 gewmktPr	0.14	7.1 gewmktPr	0.14	9.2		0.14	9.2		9.2
GWDSm98	0.11	7.8 maxFH98	0.11	8.0 GWDSv99	0.12	8.5 GWDSv99	0.15	8.5 GWDSv99	0.15	9.9		0.15	9.9		9.9
GWDSv98	0.11	7.8 GWDSv98	0.11	8.0 GWDSm99	0.12	8.5 GWDSm99	0.16	8.5 GWDSm99	0.16	10.5		0.16	10.5		10.5
maxFH98	0.12	8.5 FDFre98	0.12	8.0 SDHN99	0.13	9.3 Ntotal	0.16	9.3 Ntotal	0.16	10.5		0.16	10.5		10.5
FDFre98	0.12	8.5 GWDSm98	0.12	8.8 AMPGW499	0.14	10.0 SDGW99	0.16	10.0 SDGW99	0.16	10.5		0.16	10.5		10.5
SDGW98	0.12	8.5 DistFIRI	0.13	9.5 DistFluM	0.14	10.0 AMPGW499	0.17	10.0 AMPGW499	0.17	11.2		0.17	11.2		11.2
DistPerm	0.12	8.5 CtNtVerh	0.13	9.5 GWDSg99	0.16	11.4 AMPHN499	0.17	11.4 AMPHN499	0.17	11.2		0.17	11.2		11.2
DistFluM	0.13	9.2 KAKpot	0.13	9.5 SDHN259	0.17	12.1 DeckSUn	0.17	12.1 DeckSUn	0.17	11.2		0.17	11.2		11.2
GWDSg98	0.14	9.9 AMPGW498	0.13	9.5 MinGW259	0.22	15.7 AMPGW99	0.17	15.7 AMPGW99	0.17	11.2		0.17	11.2		11.2
AMPGW498	0.15	10.6 Ntotal	0.13	9.5 SDGW99	0.23	16.4 KAKeff	0.18	16.4 KAKeff	0.18	11.8		0.18	11.8		11.8
PFD98	0.19	13.5 GWDSg98	0.14	10.2 AMPGW99	0.23	16.4 CtNtVerh	0.18	16.4 CtNtVerh	0.18	11.8		0.18	11.8		11.8
FDAJd98	0.19	13.5 MinGW98	0.14	10.2 DistFIRI	0.24	17.1 DistPerm	0.19	17.1 DistPerm	0.19	12.5		0.19	12.5		12.5
MinGW98	0.19	13.5 SDHN698	0.15	10.9 GWDS259	0.24	17.1 GWDSg99	0.19	17.1 GWDSg99	0.19	12.5		0.19	12.5		12.5
SDGW498	0.2	14.2 SDGW698	0.16	11.7 PFD99	0.24	17.1 GW140698	0.22	17.1 GW140698	0.22	13.2		0.22	13.2		13.2
SDHN498	0.2	14.2 DistPerm	0.16	11.7 FDAJd99	0.24	17.1 MaxGW99	0.22	17.1 MaxGW99	0.22	14.5		0.22	14.5		14.5
DistFIRI	0.21	14.9 KAKeff	0.16	11.7 MinGW99	0.25	17.8 maxFH99	0.22	17.8 maxFH99	0.22	14.5		0.22	14.5		14.5
GWDS258	0.24	17.0 SDGW498	0.2	14.6 MaxGW99	0.27	19.2 MinGW99	0.23	19.2 MinGW99	0.23	15.1		0.23	15.1		15.1
GW140m98	0.26	18.4 SDHN498	0.21	15.3 maxFH99	0.28	19.9 SDHN699	0.24	19.9 SDHN699	0.24	15.8		0.24	15.8		15.8
AMPGW98	0.27	19.1 AMPGW98	0.23	16.8 SDGW259	0.28	19.9 SDGW699	0.25	19.9 SDGW699	0.25	16.5		0.25	16.5		16.5
MinGW258	0.28	19.8 GW140m98	0.23	16.8 MinGW499	0.29	20.6 MinGW499	0.26	20.6 MinGW499	0.26	17.8		0.26	17.8		17.8
GW140v98	0.29	20.5 MaxGW698	0.26	19.0 MeanGW99	0.3	21.4 AMPGW699	0.27	21.4 AMPGW699	0.27	17.8		0.27	17.8		17.8
GW140258	0.3	21.2 MaxGW498	0.26	19.0 MinGW259	0.3	21.4 GW140v99	0.3	21.4 GW140v99	0.3	19.7		0.3	19.7		19.7
MaxGW498	0.31	22.0 MaxGW98	0.26	19.0 GW140259	0.3	21.4 GW140m99	0.3	21.4 GW140m99	0.3	19.7		0.3	19.7		19.7
MinGW498	0.31	22.0 MinGW498	0.26	19.0 AMPGW259	0.32	22.8 GW140g99	0.31	22.8 GW140g99	0.31	20.4		0.31	20.4		20.4
MaxGW98	0.31	22.0 MaxGW98	0.26	19.0 GW140v99	0.32	22.8 MaxGW699	0.32	22.8 MaxGW699	0.32	21.1		0.32	21.1		21.1
GW140g98	0.31	22.0 MeanGW98	0.26	19.0 GW140m99	0.32	22.8 MaxGW499	0.32	22.8 MaxGW499	0.32	21.1		0.32	21.1		21.1
MinGW258	0.32	22.7 GW140g98	0.26	19.0 GW140g99	0.32	22.8 MnGW699	0.33	22.8 MnGW699	0.33	21.7		0.33	21.7		21.7
MaxGW258	0.32	22.7 MnGW498	0.27	19.7 MaxGW499	0.33	23.5 MeanGW99	0.33	23.5 MeanGW99	0.33	21.7		0.33	21.7		21.7
MinGW498	0.32	22.7 GW140v98	0.27	19.7 MaxGW259	0.33	23.5 MinGW499	0.33	23.5 MinGW499	0.33	21.7		0.33	21.7		21.7
MeanGW98	0.33	23.4 MinGW98	0.28	20.4 MnGW499	0.33	23.5 MinGW699	0.33	23.5 MinGW699	0.33	21.7		0.33	21.7		21.7

## Flora:

Das von den Mollusken bekannte Bild, dass die Überflutungsdauer dominierender Faktor der Artenverteilung ist (mittlerer Erklärungsanteil von 21,68% bei einer insgesamt sehr hohen durchschnittlichen Gesamtvariabilität von 3,84), findet sich auch für die Flora als nicht mobile Indikatorgruppe (Tabelle 5.4-8, S. 168). Betrachtet man das Erklärungspotential des zweiten und dritten Modellparameters (MaxGW258/698/259/699 und KAKeff; vgl. Abbildung 5.4-4, S. 169) erkennt man die annähernde Konstanz über die Untersuchungsphasen.

Die Häufigkeit der Überflutung (FDFre98/99) nimmt 1998 größeren Einfluss auf die Artenverteilung als 1999, da die geringeren Überflutungen dieses Jahres in erster Linie die Flutrinnenflächen von den übrigen Probestellen (Null-Werte) separieren.

Eine Besonderheit des Vegetationsmodells stellt der Erklärungsparameter pflanzenverfügbarer Phosphorgehalt (Ppfl) dar. Er ist anhand der Vorwärtsselektion dem Parameter mittlerer Sandgehalt (mSand) vorzuziehen, muss aber eventuell für die Modellanwendung nachträg-

lich noch verworfen werden, falls man ihn nicht geostatistisch regionalisieren kann, wodurch er für ein Prognosemodell ungeeignet wird.

**Tabelle 5.4-8: Erster Durchlauf der Vorwärtsselektion (univariat) für das Floramodell (UP: Untersuchungsphase; EV: Eigenwert und erklärter %-Anteil an der Gesamtvarianz (TI) in den biotischen Daten). Die nach sieben Durchläufen als Modellparameter ermittelten Variablen wurden in die Liste des ersten Durchlaufs übertragen und grau markiert (vgl. Anlage 6 bezüglich der Variablennamen).**

UP1	Partielle Ordination			UP3	Partielle Ordination			UP5	Partielle Ordination			UP7	Partielle Ordination			EV	% zu 3.700TI
	EV	% zu 3.886TI			EV	% zu 3.847TI			EV	% zu 3.907TI			EV	% zu 3.907TI			
gewmktDe	0.12		3.1	GWDS698	0.1		2.6	MDeckS	0.11		2.8	GWDS699	0.11		3.0		3.0
MDeckS	0.12		3.1	AMPGW698	0.12		3.1	gewmktDe	0.12		3.1	gewmktDe	0.12		3.2		3.2
pHH2Ot	0.13		3.3	gewmktDe	0.12		3.1	pHH2Ot	0.14		3.6	pHH2Ot	0.13		3.5		3.5
gewmktPr	0.16		4.1	pHH2Ot	0.13		3.4	AMPHN99	0.15		3.8	SDGW499	0.13		3.5		3.5
Kpfl	0.18		4.6	MDeckS	0.14		3.6	SDHN499	0.16		4.1	MDeckS	0.15		4.1		4.1
AMPGW258	0.19		4.9	GW140698	0.15		3.9	gewmktPr	0.17		4.4	SDHN499	0.16		4.3		4.3
SDHN98	0.19		4.9	SDHN98	0.16		4.1	CINVerh	0.19		4.9	CINVerh	0.17		4.6		4.6
DistPerm	0.2		5.1	SDGW698	0.16		4.1	SDGW499	0.2		5.1	gewmktPr	0.17		4.6		4.6
SDHN258	0.21		5.4	Kpfl	0.18		4.6	DistPerm	0.19		4.9	Kpfl	0.18		4.9		4.9
SDGW258	0.21		5.4	gewmktPr	0.18		4.6	AMPHN499	0.21		5.4	LK	0.19		5.1		5.1
CINVerh	0.21		5.4	SDHN98	0.19		4.9	Kpfl	0.22		5.6	AMPHN99	0.19		5.1		5.1
DeckSUn	0.22		5.7	CINVerh	0.2		5.2	DistFIRI	0.22		5.6	AMPHN499	0.21		5.7		5.7
LK	0.23		5.9	AMPGW498	0.2		5.2	SDHN98	0.22		5.6	DistPerm	0.22		5.9		5.9
AMPGW498	0.23		5.9	DistPerm	0.2		5.2	DeckSUn	0.24		6.1	DeckSUn	0.22		5.9		5.9
GWDSm98	0.24		6.2	LK	0.21		5.4	GWDSv99	0.25		6.4	SDHN98	0.22		5.9		5.9
pHCACi2t	0.24		6.2	DeckSUn	0.23		5.9	GWDSm99	0.25		6.4	AMPGW499	0.23		6.2		6.2
pHKClt	0.24		6.2	GWDSm98	0.24		6.2	pHCACi2t	0.26		6.7	GWDSv99	0.23		6.2		6.2
GWDSv98	0.25		6.4	GWDSv98	0.24		6.2	LK	0.27		6.9	pHCACi2t	0.24		6.5		6.5
nFK	0.28		7.2	pHKClt	0.25		6.5	AMPGW499	0.28		7.2	GWDSm99	0.24		6.5		6.5
mTon	0.28		7.2	pHCACi2t	0.26		6.7	nFK	0.28		7.2	mTon	0.27		7.3		7.3
AMPHN98	0.28		7.2	nFK	0.26		6.7	DistFluM	0.29		7.4	DistFIRI	0.29		7.8		7.8
DistFIRI	0.29		7.5	SDGW498	0.27		7.0	GWDSg99	0.29		7.4	GWDSg99	0.29		7.8		7.8
SDGW498	0.3		7.7	SDHN498	0.28		7.2	Cttotal	0.29		7.4	nFK	0.3		8.1		8.1
Cttotal	0.3		7.7	mTon	0.28		7.2	MinGW259	0.3		7.7	SDHN699	0.31		8.4		8.4
SDHN498	0.3		7.7	AMPHN98	0.29		7.5	SDHN259	0.31		7.9	SDGW699	0.31		8.4		8.4
GWDSg98	0.3		7.7	Cttotal	0.29		7.5	Ntotal	0.33		8.4	DistFluM	0.32		8.6		8.6
KAKeff	0.32		8.2	DistFIRI	0.3		7.7	KAKeff	0.33		8.4	Cttotal	0.32		8.6		8.6
SDGW98	0.32		8.2	GWDSg98	0.3		7.7	mTon	0.33		8.4	mSand	0.32		8.6		8.6
DistFluM	0.33		8.5	DistFluM	0.32		8.3	mSand	0.34		8.7	FK	0.33		8.9		8.9
mSand	0.33		8.5	Ppfl	0.32		8.3	FK	0.35		9.0	Ppfl	0.33		8.9		8.9
FK	0.34		8.7	Ntotal	0.32		8.3	MinGW99	0.36		9.2	AMPGW699	0.34		9.2		9.2
Ntotal	0.34		8.7	KAKeff	0.32		8.3	GWDS259	0.37		9.5	KAKeff	0.34		9.2		9.2
Ppfl	0.34		8.7	SDGW98	0.33		8.5	KAKpot	0.37		9.5	Ntotal	0.35		9.5		9.5
KAKpot	0.38		9.8	mSand	0.33		8.5	FDFre98	0.38		9.7	FDFre99	0.38		10.3		10.3
GWDS258	0.41		10.6	FK	0.33		8.5	MeanGW99	0.4		10.2	KAKpot	0.39		10.5		10.5
AMPGW98	0.42		10.8	KAKpot	0.37		9.6	MnGW259	0.41		10.5	GW140699	0.4		10.8		10.8
GW140m98	0.42		10.8	AMPGW98	0.43		11.1	GW140259	0.42		10.7	MinGW99	0.41		11.1		11.1
MinGW98	0.43		11.1	GW140m98	0.43		11.1	Ppfl	0.43		11.0	MeanGW99	0.44		11.9		11.9
GW140v98	0.45		11.6	MinGW98	0.45		11.6	AMPGW99	0.43		11.0	AMPGW99	0.46		12.4		12.4
MinGW498	0.46		11.8	GW140v98	0.45		11.6	SDGW259	0.43		11.0	GW140g99	0.48		13.0		13.0
GW140g98	0.48		12.4	MinGW498	0.46		11.9	GW140g99	0.44		11.3	MinGW499	0.49		13.2		13.2
GW140258	0.49		12.6	GW140g98	0.48		12.4	MinGW499	0.45		11.5	MinGW699	0.5		13.5		13.5
MaxGW258	0.5		12.9	MinGW698	0.51		13.2	MaxGW99	0.45		11.5	GW140v99	0.5		13.5		13.5
MnGW258	0.51		13.1	MeanGW98	0.52		13.4	MaxGW259	0.47		12.0	GW140m99	0.5		13.5		13.5
MeanGW98	0.51		13.1	MnGW498	0.53		13.7	MnGW499	0.47		12.0	MaxGW99	0.5		13.5		13.5
MnGW498	0.52		13.4	MnGW698	0.54		13.9	AMPGW259	0.47		12.0	MnGW499	0.51		13.8		13.8
MinGW258	0.52		13.4	MaxGW698	0.54		13.9	GW140v99	0.47		12.0	SDGW99	0.5		13.5		13.5
MaxGW498	0.53		13.6	MaxGW498	0.54		13.9	GW140m99	0.47		12.0	MnGW699	0.51		13.8		13.8
MaxGW98	0.53		13.6	MaxGW98	0.54		13.9	MaxGW499	0.49		12.5	MaxGW699	0.52		14.1		14.1
FDFre98	0.69		17.8	FDFre98	0.68		17.6	SDGW99	0.56		14.3	MaxGW499	0.52		14.1		14.1
maxFH98	0.72		18.5	maxFH98	0.7		18.1	maxFH99	0.66		16.9	maxFH99	0.7		18.9		18.9
FDAJd98	0.84		21.6	FDAJd98	0.83		21.4	FDAJd99	0.82		21.0	FDAJd99	0.84		22.7		22.7

### Die abiotischen Modellparameter:

Nachfolgend werden die Erklärungsvariablen der Ordinationsmodelle aufgelistet und in Abbildung 5.4-4 (S. 169) zusammengefasst. Die Abfolge spiegelt deren Bedeutungsgehalt, d.h. die Reihenfolge der Aufnahme ins Modell, wider.

#### Molluskenmodell:

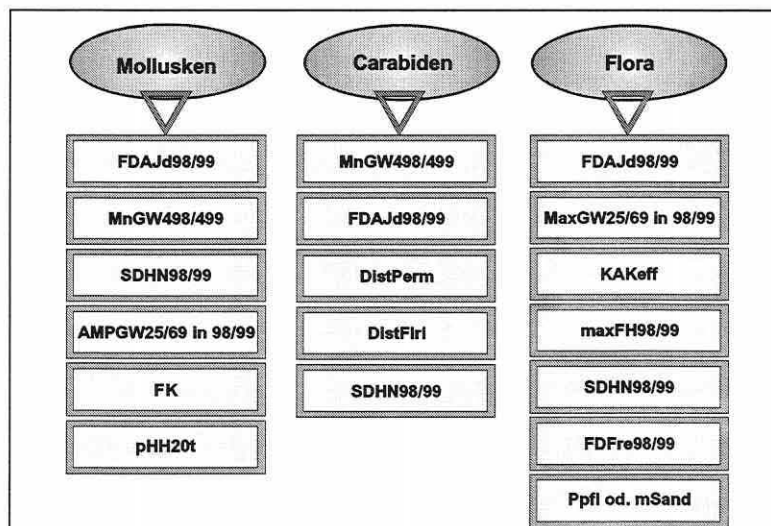
1. **FADJD98/FDAJD99:** Jährliche Überflutungsdauer in Tagen (Abflussjahr).
2. **MnGW498/499:** Mittlerer Grundwasserflurabstand in der Vegetationsperiode (01.04.-30.09.) eines Kalenderjahres.
3. **SDHN98/99:** Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand.
4. **AMPGW258/698/259/699:** Amplitude der Grundwasserflurabstände zwischen dem 01.02.-31.05. bzw. 01.06.-30.09. eines Kalenderjahres.
5. **FK:** Feldkapazität [mm]; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994).
6. **pHH2Ot:** pH-Wert H<sub>2</sub>O delogarithmisiert (Wasserstoffionenkonzentration); Transformation: Multiplikation mit 1.000.000.

**Carabidenmodell:**

1. **MnGW498/499:** Mittlerer Grundwasserflurabstand in der Vegetationsperiode (01.04.-30.09.) eines Kalenderjahres.
2. **FADJD98/FDAJD99:** Jährliche Überflutungsdauer in Tagen (Abflussjahr).
3. **DistPerm:** Distanz zu permanenten Wasserflächen.
4. **DistFlri:** Distanz zum Stratum Flutrinne.
5. **SDHN98/99:** Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand.

**Floramodell:**

1. **FADJD98/FDAJD99:** Jährliche Überflutungsdauer in Tagen (Abflussjahr).
2. **MaxGW258/698/259/699:** Maximum Grundwasserflurabstand zwischen dem 01.02.-31.05. bzw. 01.06.-30.09. eines Kalenderjahres.
3. **KAKeff:** Effektive Kationenaustauschkapazität [cmolc/kg]; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994).
4. **maxFH98/99:** Maximale Überflutungshöhe in m (höchstes Überflutungsereignis im Abflussjahr).
5. **SDHN98/99:** Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand.
6. **FDFre98/99:** Häufigkeit/Anzahl Überflutungsereignisse im Kalenderjahr.
7. **Ppfl:** pflanzenverfügbare Phosphor [mg/100 g] bzw. **mSand:** Mittlerer Sandgehalt [%]; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994).



**Abbildung 5.4-4:** Die Modellparameter der Ordinationsmodelle RIVA (Kürzel im Text und Anlage 6 erläutert).

Inwieweit die Modellparameter multikollinear sind, wird durch den VIF-Wert überprüft (vgl. Abschnitt 3.6.1). Das Ergebnis der Multikollinearitätsprüfung (Tabelle 5.4-9) offenbart, dass in einzelnen Untersuchungsphasen Überflutungsparameter und Grundwasserparameter hohe Werte aufweisen (kursiv in Tabelle 5.4-9, S.170). Im Molluskenmodell der UP5 betrifft dies den mittleren Grundwasserflurabstand in der Vegetationsperiode eines Kalenderjahres (20,47) und die Amplitude der Grundwasserflurabstände zwischen dem 01.02. und 31.05. eines Kalenderjahres (19,14) sowie im Floramodell der UP3 die Parameter maximale Überflutungshöhe in Meter (19,92) und Häufigkeit/Anzahl der Überflutungsereignisse im Kalenderjahr (18,52). Die kritische Grenze von 20 (ter Braak 1998) wird jedoch nur einmal geringfügig überschritten (kursiv, unterstrichen) und als Ausreißer akzeptiert (vgl. Abschnitt 3.5.3).

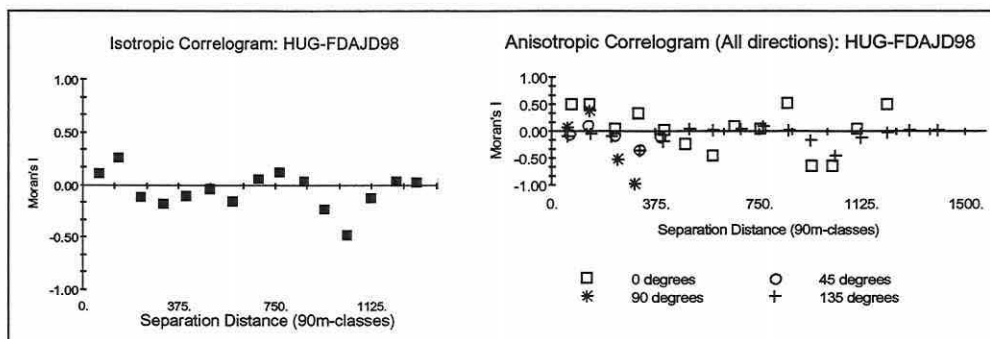
**Tabelle 5.4-9: Multikolinearitäten der Modellparameter (Erläuterungen im Text; Parameterabfolge wie in Abbildung 5.4-4, S. 169 – \* Werte für Ppfl).**

Mollusken				Carabiden				Flora			
UP1	UP3	UP5	UP7	UP1	UP3	UP5	UP7	UP1	UP3	UP5	UP7
2,48	2,27	2,35	2,38	2,58	2,47	3,00	2,78	13,67	5,32	9,82	7,56
1,88	1,48	<u>20,47</u>	3,64	1,70	1,61	1,92	1,72	2,97	3,02	6,11	10,33
1,72	1,43	1,89	1,61	1,64	1,70	1,52	1,57	1,94	2,50	1,93	1,94
2,74	2,24	19,14	2,89	2,35	2,14	2,34	2,14	4,04	19,92	10,52	13,33
2,14	2,53	2,30	1,75	1,28	1,21	1,46	1,43	1,60	1,47	1,77	2,24
1,43	1,59	1,36	1,52					12,34	18,52	2,33	2,33
								1,73*	1,76*	2,29*	1,77*

### Ergebnisse: Analyse der räumlichen Autokorrelation

Die Analyse, ob die Probestellen eine räumliche Autokorrelationsstruktur aufweisen, wurde mit der Software GS<sup>+</sup> (Robertson 1998) durchgeführt. Zur Prüfung wurde der Autokorrelationskoeffizient Moran's I verwendet (vgl. Abschnitt 3.5.9). Die isotropen und anisotropen Korrelogramme ergeben sowohl für die 45m- und 90m-Distanzklassen nur eine mäßige bis niedrige Autokorrelation (Werte in der Regel zwischen  $\pm 0,5$ ). Wie erwartet sind in den niedrigen Distanzklassen positive und in den hohen Distanzklassen negative Moran's I-Werte anzutreffen. Eine starke Beeinflussung der Modelle ist somit nicht anzunehmen.

Beispielkorrelogramme für die Überflutungsdauer in Tagen des Jahres 1998 (HUG) visualisiert Abbildung 5.4-5. Eine klare Tendenz ist nicht zu erkennen. Die Prüfung für den 90°-Winkel scheint der Aussage zu widersprechen, es ist jedoch zu bemerken, dass die 300m-Klasse nur mit vier Paaren besetzt ist und der hohe Korrelationswert nicht als zuverlässig gelten darf. Die Werte für Moran's I schwanken für das Beispiel im isotropen Diagramm zwischen maximal +0,27 und minimal -0,47.



**Abbildung 5.4-5: Korrelogramme für die Überflutungsdauer in Tagen 1998 (HUG).**

Die Grundwasserparameter (Flurabstände) zeigen eine deutlichere Tendenz innerhalb der  $\pm 0,5$ -Spanne, was nicht einer gewissen Logik entbehrt, da die Wasserstände sich träge verhalten und nur eine kleinräumige sehr bewegte Morphologie diesem entgegenhalten könnte. Die Bodenparameter unterliegen ebenfalls einer gerichteten, aber erheblich schwächeren Tendenz.

### 5.4.2 Mustererkennung: Ordination

Der Prozess der Mustererkennung umfasst neben der direkten Ordination anhand der Modellparameter (CCA; vgl. Abschnitt 3.5.8) Kontrollberechnungen zur Artenverteilung im Ordinationsraum mit indirekten Verfahren (CA/DCA; vgl. Abschnitte 3.5.5 und 3.5.6). Es werden die Resultate beider Ordinationstypen angeführt, um zu prüfen, ob die relevanten, das Antwortmuster der Biotik erklärenden Standortparameter erfasst wurden. Ist dies der Fall, dürfen



keine starken Abweichungen der Eigenwerte und Verteilungen in den direkten und indirekten Ordinationen auftreten.

### Ergebnisse Mollusken:

DCA: Die Länge der Gradienten (Tabelle 5.4-10) rät zu der Anwendung von unimodalen Modellen (vgl. Abschnitt 3.5.6).

Tabelle 5.4-10: DCA-Ergebnisse für die Indikatorgruppe Mollusken im Hauptuntersuchungsgebiet.

**** Summary **** HUG UP1					Probefläche 15 ausgefallen (leer)
Axes	1	2	3	4	Total inertia
Eigenvalues	.670	.106	.060	.032	1.821
Lengths of gradient	4.228	2.891	2.301	2.530	
Cumulative percentage variance of species data	36.8	42.6	45.9	47.7	
**** Summary **** HUG UP3					
Axes	1	2	3	4	Total inertia
Eigenvalues	.609	.245	.075	.030	1.618
Lengths of gradient	2.608	2.489	1.790	2.266	
Cumulative percentage variance of species data	37.6	52.8	57.4	59.2	
**** Summary **** HUG UP5					
Axes	1	2	3	4	Total inertia
Eigenvalues	.560	.154	.073	.037	1.661
Lengths of gradient	3.562	1.420	1.824	1.486	
Cumulative percentage variance of species data	33.7	43.0	47.4	49.6	
**** Summary **** HUG UP7					
Axes	1	2	3	4	Total inertia
Eigenvalues	.622	.133	.076	.051	1.914
Lengths of gradient	4.121	2.484	2.089	2.001	
Cumulative percentage variance of species data	32.5	39.4	43.4	46.1	

CA: Die CA-Ergebnisse (Tabelle 5.4-11) weisen generell hohe Eigenwerte für die erste Achse auf und werden dann sprunghaft niedriger. Dies ist bereits ein Indiz für die klare Dominanz bestimmter Umweltfaktoren, ausgedrückt in der biotischen Variabilität der Arten.

Tabelle 5.4-11: CA-Ergebnisse für die Indikatorgruppe Mollusken im Hauptuntersuchungsgebiet.

**** Summary **** HUG UP1					Probefläche 15 ausgefallen (leer)
Axes	1	2	3	4	Total inertia
Eigenvalues	.670	.377	.301	.121	1.821
Cumulative percentage variance of species data	36.8	57.5	74.0	80.6	
**** Summary **** HUG UP3					
Axes	1	2	3	4	Total inertia
Eigenvalues	.609	.386	.165	.155	1.618
Cumulative percentage variance of species data	37.6	61.5	71.7	81.3	
**** Summary **** HUG UP5					
Axes	1	2	3	4	Total inertia
Eigenvalues	.560	.387	.165	.113	1.661
Cumulative percentage variance of species data	33.7	57.0	67.0	73.7	
**** Summary **** HUG UP7					
Axes	1	2	3	4	Total inertia
Eigenvalues	.622	.410	.209	.186	1.914
Cumulative percentage variance of species data	32.5	53.9	64.9	74.5	

Die prozentualen Erklärungsanteile an der Artenverteilung sind für ökologische Studien hoch (vgl. Gauch 1982). Die Schwankungen der Gesamtvariabilität (TI) im Molluskendatensatz über die jeweiligen Untersuchungsphasen könnten – falls nicht zufallsbedingt – Reaktionen der Indikatorgruppe auf veränderte Umweltbedingungen anzeigen.

UP1 und UP5 weisen unterschiedliche Überflutungsbedingungen auf, wobei das Frühjahr 1999 deutlich stärker durch Überflutung geprägt war. Umgekehrt fielen im Herbst 1999 ungewöhnlich hohe Niederschläge, die sich in den Unterschieden der UP3 und UP7 wieder spiegeln könnten (vgl. Tabelle 5.1-2, S. 143).

CCA: Ausgehend von den gewichteten Korrelationsmatrizen des Verfahrens wird der Beitrag der Erklärungsvariablen zu den Ordinationsachsen interpretiert (Tabelle 5.4-12). Die Korrelationen der Modellparameter zu den Ordinationsachsen im Artenraum (SPEC AX) sind generell etwas niedriger als die Korrelationen zu den Achsen im Umweltraum (ENVI AX; vgl. Abschnitt 3.5.1). Für gewöhnlich wird in direkten Ordinationsverfahren der Fokus auf letztere gelegt, da die Achsen basierend auf den Umweltvariablen interessieren. Diese Aussage ist auch für die auf den nächsten Seiten besprochenen anderen Indikatorgruppen gültig.

**Tabelle 5.4-12: Gewichtete Korrelationsmatrizen der Molluskenmodelle für UP1 bis UP7 (maximale/relevante Korrelationen zu den Achsen sind grau unterlegt; Variablen: vgl. Anlage 6).**

HUG UP1		**** Weighted correlation matrix (weight = sample total) ****						Mollusken	
FK	-0,5841	-0,1752	-0,3124	0,0233	-0,589	-0,1955	-0,366	0,0397	
pHH2Ot	0,3324	0,354	-0,4016	0,1425	0,3352	0,3949	-0,4705	0,2428	
SDHN98	-0,4983	-0,4088	0,2105	0,3988	-0,5025	-0,4561	0,2467	0,6794	
MnGW498	-0,282	0,6469	0,4424	0,192	-0,2844	0,7216	0,5183	0,3272	
AMPGW258	0,1805	0,185	0,756	0,1229	0,182	0,2063	0,8857	0,2094	
FDAJd98	0,9738	0,1461	-0,0258	0,0471	0,982	0,163	-0,0302	0,0802	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	

HUG UP3		**** Weighted correlation matrix (weight = sample total) ****						Mollusken	
FK	0,6263	-0,2859	-0,3675	0,0063	0,6377	-0,3278	-0,4043	0,0132	
pHH2Ot	-0,2689	0,0787	-0,275	0,0815	-0,2738	0,0903	-0,3026	0,1718	
SDHN98	0,4473	0,0219	0,6148	0,0966	0,4555	0,0252	0,6765	0,2036	
MnGW498	0,2686	0,8068	0,1114	0,0903	0,2735	0,925	0,1226	0,1904	
AMPGW698	0,3955	-0,5221	-0,2723	0,2942	0,4027	-0,5986	-0,2997	0,6201	
FDAJd98	-0,9724	0,0499	-0,1092	0,0108	-0,9901	0,0572	-0,1202	0,0227	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	

HUG UP5		**** Weighted correlation matrix (weight = sample total) ****						Mollusken	
FK	-0,6258	-0,3665	0,1774	-0,2243	-0,6431	-0,4061	0,2248	-0,3112	
pHH2Ot	0,0187	0,1258	0,556	0,0638	0,0192	0,1394	0,7043	0,0885	
SDHN99	-0,0188	0,4127	-0,6217	-0,1086	-0,0193	0,4573	-0,7876	-0,1506	
MnGW499	-0,1982	0,8463	-0,0779	0,1701	-0,2037	0,938	-0,0987	0,2361	
AMPGW259	-0,1067	0,877	-0,1172	-0,0089	-0,1096	0,972	-0,1485	-0,0124	
FDAJd99	0,9383	-0,0891	0,0825	0,0393	0,9642	-0,0988	0,1045	0,0545	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	

HUG UP7		**** Weighted correlation matrix (weight = sample total) ****						Mollusken	
FK	0,5635	-0,4365	-0,3283	0,1731	0,5717	-0,4831	-0,3846	0,2445	
pHH2Ot	-0,2806	0,1639	-0,4503	-0,1734	-0,2847	0,1814	-0,5275	-0,2448	
SDHN99	0,3193	0,2457	0,572	-0,0788	0,324	0,2719	0,67	-0,1113	
MnGW499	0,4212	0,7884	0,0209	-0,1201	0,4273	0,8726	0,0245	-0,1696	
AMPGW699	-0,0611	-0,7369	-0,1332	-0,357	-0,062	-0,8156	-0,156	-0,5041	
FDAJd99	-0,9641	0,0756	-0,0764	0,053	-0,9781	0,0837	-0,0895	0,0749	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	

Die sehr hohe Korrelation der Überflutungsdauer (FDAJd98/99) bestimmt in den Molluskenmodellen konstant die erste Ordinationsachse. Sie beschreibt daher in erster Linie die Arten-trennung anhand des Überflutungsregimes der Aue. Weiterhin korreliert die Feldkapazität (FK) stabil und berücksichtigungswert hoch auf selbige Achse. Die zweite Achse wird in UP3

bis UP7 klar durch die Grundwasserverhältnisse charakterisiert. Der mittlere Grundwasserflurabstand in den Monaten April bis einschließlich September in 1998 und 1999 (MnGW498/499) trägt die eigentliche Bedeutung der Achse. Die Amplitude der Grundwasserflurabstände in den Frühjahrs- und Herbstmonaten (AMPGW25/69 für 1998 und 1999) verhält sich weniger konstant. Im Frühjahr 1998 und 1999 trennt sie zwar in beiden Fällen die Artenzusammensetzung, besitzt aber dennoch einen unterschiedlichen Bedeutungsgehalt. In 1998 lädt sie auf die dritte Achse und bildet keine Einheit mit dem mittleren Grundwasserflurabstand aus. Während die mittleren Grundwasserflurabstände in 1998 auf allen Probestellen durchweg höher ausfallen, ist die Amplitude in den Monaten Februar bis einschließlich Mai in 1998 für die trockenen Probestellen meist niedriger. Dies ist eine Folge der nicht vorhandenen Überflutung dieser Probestellen in 1998. Frühjahrs- und Herbstphasen zeigen jeweils eine gleich- bzw. gegenläufige Korrelation der beiden Grundwasserparameter zueinander. Dies erklärt sich durch die unterschiedlichen Bezugszeiträume. Der mittlere Grundwasserflurabstand integriert über sechs Monate und bleibt in seinem Wert konstant. Im Gegensatz dazu ist die Amplitude für lange überflutete Probestellen im Frühjahr geringer als für die trockeneren, da das Grundwasser eine Zeit an der Geländeoberfläche ansteht. Im Herbst drehen sich die Verhältnisse um. Die Flutrinne und die nassen Flächen erfahren eine relativ stärkere Grundwasserabsenkung als die übrigen Flächen in der gegebenen Zeitphase. Die dritte Achse ist nicht mehr so eindeutig anzusprechen, obgleich die Standardabweichung der Gesamtamplitude Überflutungswasserstände zu Grundwasserflurabständen als Dynamikzeiger diese Achse noch am ehesten zu beschreiben vermag. Im ersten Modell verdrängt die Amplitude der Grundwasserflurabstände in den Frühjahrsmonaten den Parameter auf die vierte Achse. Der pH-Wert der Wasserextraktionsmethode (pHH2Ot) ist insgesamt von geringerer Trennkraft, wie Tabelle 5.4-3 (S. 164) zeigte. Er wurde dennoch in die Modelle aufgenommen, da die Monte Carlo Tests ihm einen signifikanten Beitrag zur Artenentrennung zuweisen.

Es mag auffallen, dass die Vorzeichen der Korrelationen über die Modelle wechseln. Dies ist bedeutungslos und aus technischen Gründen zufällig. Das resultierende Ordinationsdiagramm erscheint bei einem Vorzeichenwechsel nur gespiegelt.

Neben den Korrelationsmatrizen geben die CA- und CCA-Ordinationsdiagramme der Molluskenmodelle (Abbildungen 5.4-6 bis 5.4-10, S. 174f) ergänzende Informationen speziell zur biotischen Datenstruktur. Sowohl das Ähnlichkeitsverhältnis der Arten hinsichtlich ihrer allgemeinen Lebensraumsprüche (CA) als auch bezogen auf die Modellparameter (CCA) lässt sich studieren (vgl. Abschnitte 3.5.5 und 3.5.8). Eine Überprüfung, ob die Erklärungsvariablen die biotische Struktur trennen können, liefert der Vergleich von indirekten mit den direkten Ordinationsdiagrammen (Palmer 1998), welcher exemplarisch für UP1 besprochen wird (Abbildung 5.4-6 verglichen mit Abbildung 5.4-7, S. 174f). Es ist offensichtlich, dass die Trennkraft der direkten Ordination etwas schwächer ist (65,4% erklärte Variabilität zu 80,6% über vier Achsen bezogen auf die rein biotische Variabilität von 1,82 TI in den Daten der UP1), die Strukturen der biotischen Lebensgemeinschaften aber sauber durch die wenigen abiotischen Modellparameter nachgezeichnet werden (vgl. Artenmuster im indirekten und direkten Ordinationsdiagramm). Dies ist die Bestätigung, dass keine wirklich relevante Erklärungsvariable unerfasst blieb.

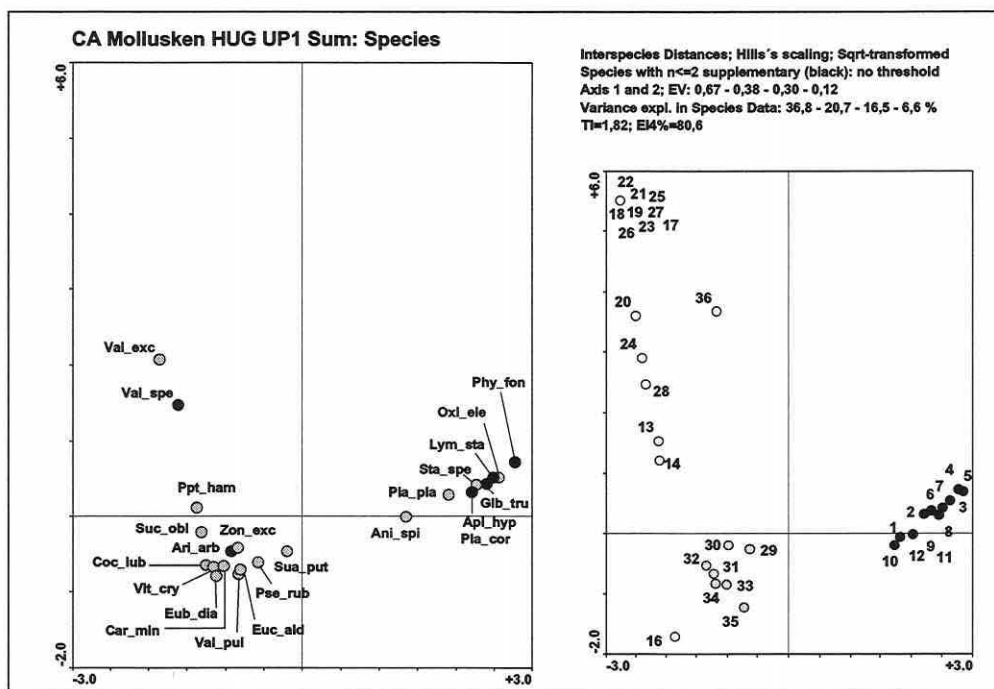
**Tabelle 5.4-13: Eigenwertvergleich der CA- mit den CCA-Modellen für die Indikatorgruppe Mollusken (UP: Untersuchungsphase; EV: Eigenwert der Achse; %: Erklärter Prozentanteil der Achse an der biotischen Variabilität).**

Achse	UP1		UP3		UP5		UP7	
	EV	%	EV	%	EV	%	EV	%
	CA/CCA	CA/CCA	CA/CCA	CA/CCA	CA/CCA	CA/CCA	CA/CCA	CA/CCA
1	0,67/0,65	36,8/36,1	0,61/0,59	37,6/36,3	0,56/0,53	33,7/31,8	0,62/0,60	32,5/31,5
2	0,38/0,29	20,7/15,9	0,39/0,29	23,9/17,8	0,39/0,31	23,3/18,8	0,41/0,33	21,4/17,2

Der gleiche Sachverhalt wird auch ausgedrückt durch die Eigenwerte und die prozentualen Erklärungsanteile der CA- und CCA-Ergebnisse für die ersten beiden Achsen zueinander. Es sind generell nur geringe Differenzen auszumachen (Tabelle 5.4-13, S. 173).

Es bleibt anzumerken, dass die Aussagen nur auf die aktuell erfasste biotische Variabilität von 1,82 im Molluskendatensatz der UP1 anzuwenden sind. In diesem Wert ist das allgemeine biotische Datenrauschen (z.B. Erhebungsungenauigkeiten) integriert. Das indirekte Modell erklärt über eine zu bestimmende Anzahl von Dimensionen letztlich 100% der Gesamtvariabilität im biotischen Datensatz, ist aber in höheren Dimensionen meist nicht mehr eindeutig interpretierbar, weshalb man nur vier oder weniger Dimensionen analysiert und vergleichend zum direkten Modell bewertet. Die direkten Modelle sind auf wenige Modellparameter begrenzt und ignorieren Langzeitumweltbedingungen, d.h. die Historie der Umweltfaktoren. Viele Molluskenarten besitzen einen mehrjährigen Lebenszyklus und integrieren so frühere Standortbedingungen in ihre aktuelle Präsenz im Untersuchungsgebiet. Diese und andere Faktoren lassen den Erklärungsanteil wohl generell geringer ausfallen.

Bei der zeitlichen Modellübertragung wird genanntes ebenfalls deutlich: Die Modellgleichung einer UP zeigt eine geringere Trennkraft bei ihrer unveränderten Anwendung zur Prognose in einer anderen UP (vgl. Abschnitt 5.4.4).



**Abbildung 5.4-6: CA-Molluskenmodell der UP1 basierend auf allen Stechrahmenproben einer Probefläche (Sum); erste und zweite Achse (Axis 1 und 2). Darstellung der Arten links und der Probeflächen rechts. Probeflächendiagramm mit Straten: weiß - trockenes Grünland, grau - nasses Grünland und schwarz - Flutrinnen. Arten mit Vorkommenshäufigkeit geringer oder gleich zwei sind passiv ins Modell eingefügt und schwarz im Artendiagramm gekennzeichnet (Species with  $n \leq 2$  supplementary). Analyseoptionen: Fokus entspricht Distanzen zwischen den Arten (Interspecies Distances), Skalierung erfolgte nach Hill und die biotischen Daten wurden Quadratwurzel transformiert (Sqrt-transformed). Keine Sichtbarkeitsgrenze nach Modellanpassungsgüte für die Arten gewählt (no threshold). Modellresultate: Eigenwerte mit der ersten Achse beginnend (EV), prozentuale erklärte Variabilität in den biotischen Daten mit der ersten Achse beginnend (Variance expl. in Species Data), Gesamtvariabilität der biotischen Daten (TI) und durch das Modell erklärte prozentuale Variabilität der TI über vier Achsen (EI4%).**

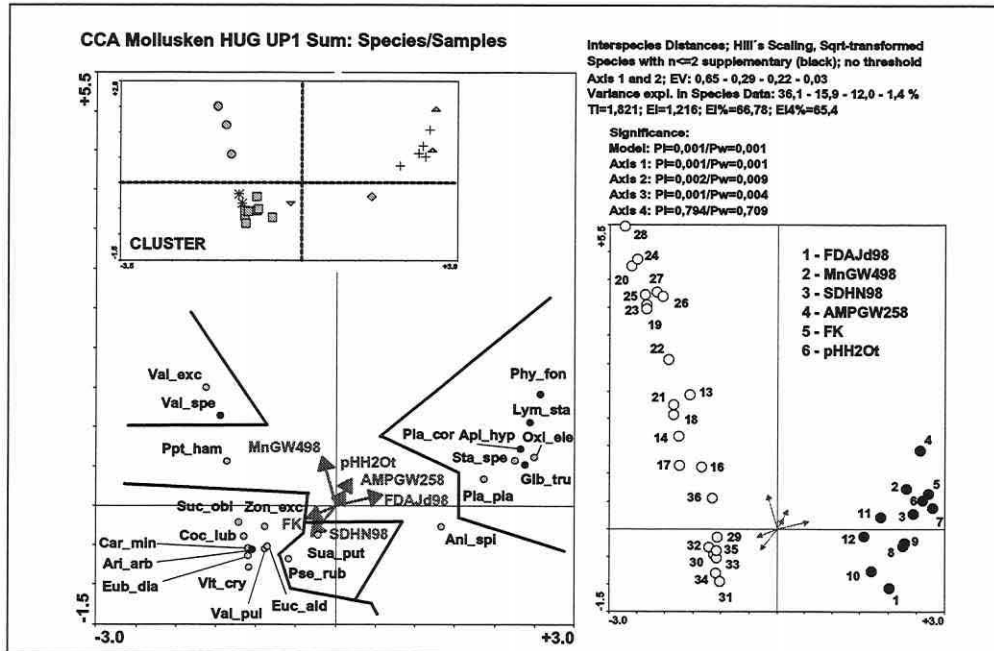


Abbildung 5.4-7: CCA-Molluskenmodell der UP1 (Erläuterungen in Abbildung 5.4-6, S. 174). Ergänzende Angaben: Maximal durch das Modell erklärbare Variabilität (EI) und prozentuale Variabilität an der TI (EI%). Signifikanz der Trennkraft für das Modell und die Achsen: zwischen den Probenflächen innerhalb (Pi) und über die Straten (Pw). Ergebnis der Clusteranalyse (Cluster) und fachliche Abgrenzung der ökologischen Gruppen (Linien im Artendiagramm). Aufnahmeabfolge der Parameter ins Modell (1 – 6).

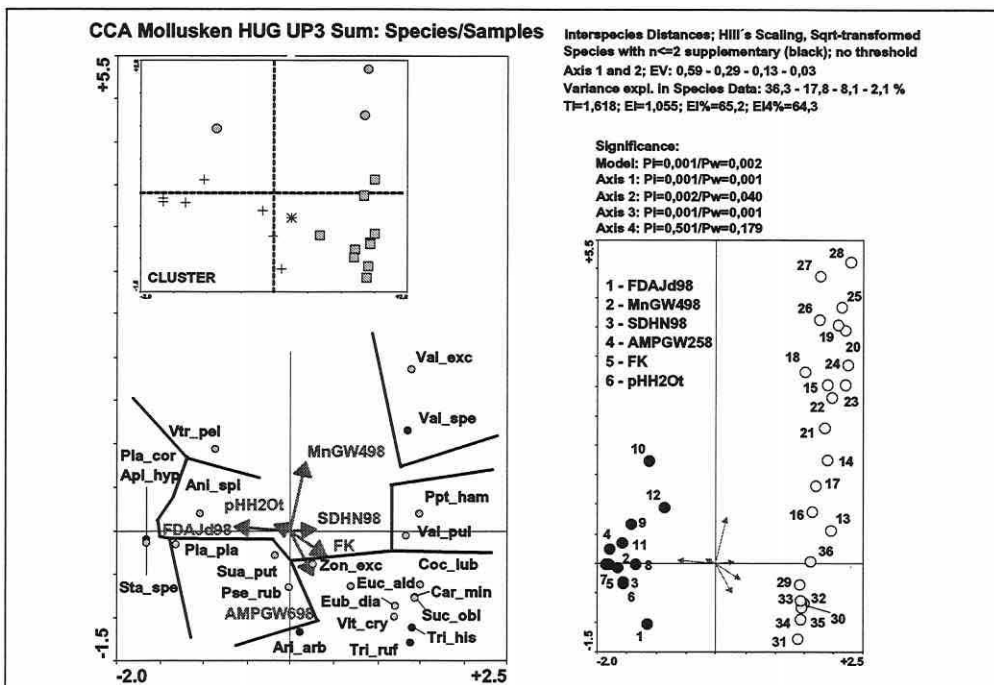


Abbildung 5.4-8: CCA-Molluskenmodell der UP3 (Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).



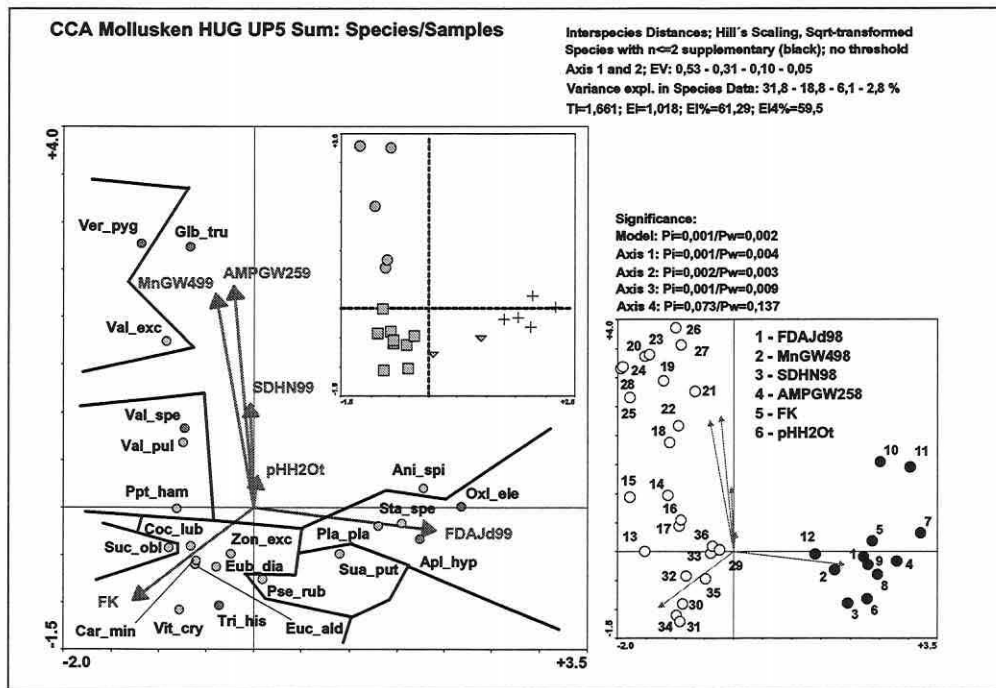


Abbildung 5.4-9: CCA-Molluskenmodell der UP5 (Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).

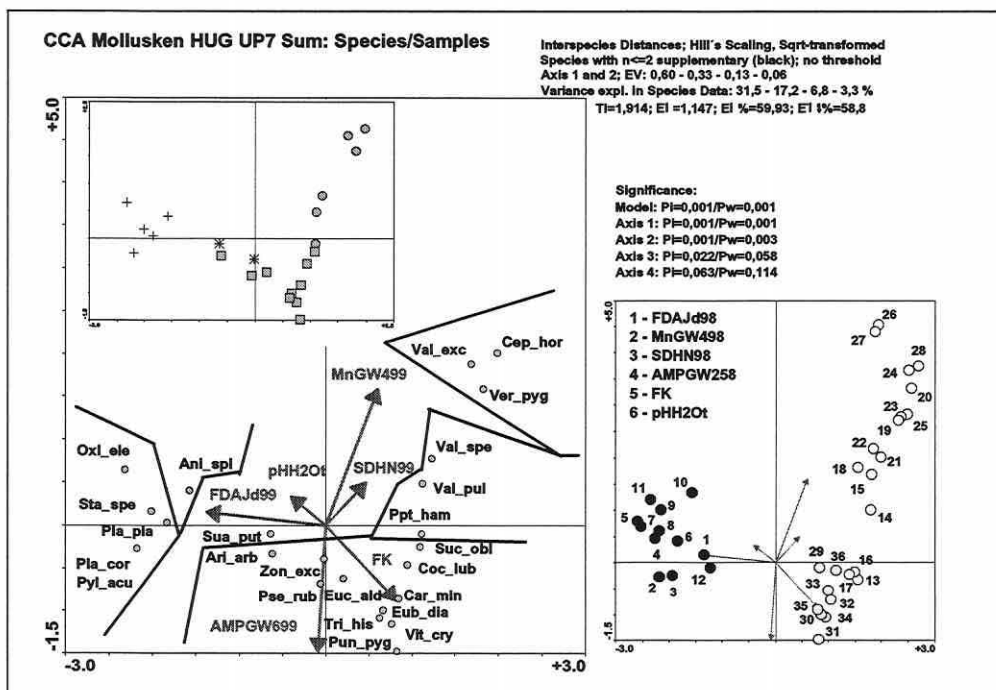


Abbildung 5.4-10: CCA-Molluskenmodell der UP7 (Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).

In allen direkten Ordinationen wird der Eigenwert (EV) und der korrespondierende prozentuale Erklärungsanteil an der Gesamtvariation in den biotischen Daten (Variance explained in Species Data) von der ersten zur zweiten Achse sprunghaft geringer. Er fällt um das 1,7 bis 2,2-fache. Die Gesamtvariabilität im biotischen Datensatz (TI), der durch das Modell erklärte Anteil an der Gesamtvariabilität im biotischen Datensatz (EI) sowie die jeweiligen Prozentwerte des erklärten Anteils an der Gesamtvariabilität des Modells (EI%) und über die ersten vier Achsen (EI4%) ermöglichen die Einordnung der Modellgüte. Alle direkten Modelle sind in der Lage, um die 60% der biotischen Variabilität zu erklären, und dürfen als ausgesprochen gut bezeichnet werden (vgl. EI4%-Wert in den Abbildungen 5.4-7 bis 5.4-10, S. 175f).

Die Signifikanz des Gesamtmodells sowie der ersten vier Achsen wurde durch Monte-Carlo-Tests für Split-Plot-Designs (999 Permutationen) bestimmt (vgl. Abschnitt 3.5.3). Sie bestätigt die Trennkraft der Erklärungsvariablen im Modell sowohl zwischen den Probeflächen innerhalb der Straten (Pw) als auch zwischen den Straten selbst (Pi) für die ersten drei Achsen. Die Signifikanz der dritten Achse bleibt in der UP7 leicht unter dem  $\alpha$ -Niveau von 0,05 für den Pw-Wert. Die vierte Achse trennt nicht signifikant.

Überwiegend in der Peripherie der Diagramme treten Arten mit extremeren Lebensraumansprüchen auf, die oft den selteneren Arten im Datensatz entsprechen (schwarze Signatur in den Artendiagrammen: passive Arten mit einer Vorkommenshäufigkeit von  $n \leq 2$ ).

Die Arten lassen ökologische Gruppierungen erkennen, die durch die Nähe ihrer Signaturen zum Ausdruck kommen (vgl. Abschnitt 3.6.3). Auch eine recht hohe Probeflächentreue ist über die UPs unter der Annahme erkennbar, dass die Arten auf ihnen nahe gelegenen Flächen im Ordinationsdiagramm wahrscheinlicher auftreten (höhere Artmächtigkeiten/Individuenzahlen als auf weiter entfernten). Arten, die über vier Standardabweichungen der Achsen voneinander entfernt im Ordinationsraum liegen, kommen in der Regel nicht gemeinsam auf Probeflächen vor. Deutlich zu bemerken ist eine Trennung der Flutrinnearten und -flächen von der übrigen Datenwolke entlang der ersten Achse. Es entstehen disjunkte Gruppen (vgl. auch Abschnitt 5.4.3). Der Sprung entlang des Gradienten wird durch die sehr unterschiedliche Überflutungsdauer verursacht, da gerade in den isoliert liegenden Flutrinne häufig das Oberflächenwasser sehr lange stehen bleibt. Mangels Abfluss fallen die Flächen erst durch Versickerung und Evapotranspiration über das Jahr hin wieder trocken. Die Trennung der Arten und Probeflächen entlang der zweiten Achse beruht auf dem durch das Grundwasser definierten Feuchteregime. Weiterführende Auswertungen und die Ausscheidung von ökologischen Gruppen sind Themen des Abschnitts 5.4.3.

### Ergebnisse Carabiden:

Tabelle 5.4-14: DCA-Ergebnisse für die Indikatorgruppe Carabiden im Hauptuntersuchungsgebiet.

**** Summary **** HUG UP1					
Axes	1	2	3	4	Total inertia
Eigenvalues	.381	.090	.050	.034	1.412
Lengths of gradient	2.971	1.432	1.246	0.929	
Cumulative percentage variance of species data	27.0	33.4	36.9	39.3	
**** Summary **** HUG UP3					
Axes	1	2	3	4	Total inertia
Eigenvalues	.344	.102	.054	.040	1.370
Lengths of gradient	2.728	1.477	1.258	1.260	
Cumulative percentage variance of species data	25.1	32.6	36.5	39.5	

## Fortsetzung der Tabelle 5.4-14

**** Summary **** HUG UP5					
Axes	1	2	3	4	Total inertia
Eigenvalues	.396	.101	.057	.037	1.405
Lengths of gradient	2.616	1.506	1.307	1.071	
Cumulative percentage variance of species data	28.2	35.4	39.4	42.1	
**** Summary **** HUG UP7					
Axes	1	2	3	4	Total inertia
Eigenvalues	.423	.106	.073	.049	1.519
Lengths of gradient	2.694	1.498	1.336	1.278	
Cumulative percentage variance of species data	27.8	34.8	39.6	42.9	

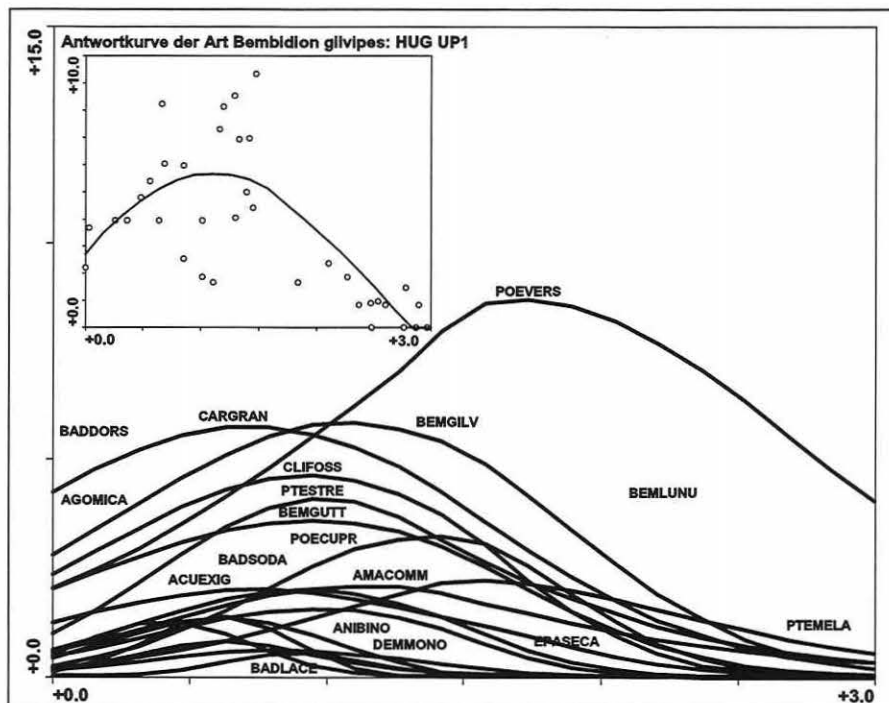


Abbildung 5.4-11: Die unimodale Antworten von Carabiden gegenüber der ersten DCA-Achse (Skalierung in SD) in UP1 des HUG. Es wurde ein fixiertes Generalisiertes Lineares Modell (GLM) zu Grunde gelegt und mit einer Kurvenglättung über 35 Nachbarn sowie einer log-link-Funktion und einer Poisson-Verteilung gearbeitet. Das Teilbild zeigt für die Art *Bembidion glivipes* die Quadratwurzel transformierten Häufigkeiten über die Probeflächen.

DCA: Die durch die DCA approximierten Gradientenlängen (Tabelle 5.4-14) unterschreiten die Grauzone für eine gleichberechtigte Anwendbarkeit von linearen und unimodalen Modellen (3 bis 4 SD) knapp. Dies spricht eigentlich für die Anwendung eines linearen Modells (Abschnitt 3.5.6 und 4.1). Es wurden, um eine endgültige Entscheidung zu treffen, Hauptkomponenten- und Redundanzanalysen basierend auf der nicht zentrierten und der zentrierten Kovarianzmatrix sowie zusätzlich der Korrelationsmatrix berechnet (vgl. Abschnitte 3.5.4 und 3.5.7). Deren Ergebnisse sowie die der kanonischen Korrespondenzanalyse in der Biplot-Skalierung (vgl. Abschnitte 3.5.5 und 3.5.8) wurden mit den Fachwissenschaftlern auf

Plausibilität hin diskutiert. Weiterhin wurden für alle Arten der UP1 die Antworten gegenüber der ersten DCA-Achse geprüft (Abbildung 5.4-11, S. 178).

Aus Abbildung 5.4-11 wird ersichtlich, dass die Beta-Diversität (vgl. Abschnitt 2.2) im Untersuchungsgebiet für die Indikatorgruppe gering ist. Eine Vielzahl der Arten, deren Existenzoptimum mittlere Bedingungen erfordert, weist eine unimodale Antwort zum Gradienten auf (nur diese sind in der Abbildung dargestellt). Die teils hohe Streuung der Daten wird am Beispiel der Art *Bembidion gilvipes* (kleines Teilbild in Abbildung 5.4-11) visualisiert. Die Zuverlässigkeit der Antwortkurven ist daher im Einzelnen zu hinterfragen.

Die vergleichende Auswertung der Ergebnisse verschiedener PCA-Varianten (vgl. Abschnitt 4.1) zeigt keine drastischen Verzerrungen der Artenverteilung bezogen auf das Gesamtmuster im Diagramm. Bei der Einzelbetrachtung von Arten kommt es zu Verschiebungen in den ökologischen Gruppierungen. Die Gegenüberstellung der Ergebnisse aus der CA und der PCA auf der Korrelationsmatrix lassen in der PCA deutlichere Parallelen des Verteilungsmusters für Artenpaare ablesen, während diese in der CA eine feinere Trennung in ihrer Lage zum Gradienten erfahren. Die unterschiedliche Aussagefähigkeit linearer und unimodaler Methoden führte nach ausführlichen Tests und Modellvergleichen zu der Entscheidung eine CCA in der Biplot-Skalierung für die weiteren Auswertungen anzuwenden. Das Verfahren in Kombination mit dieser Skalierung weist die Besonderheit auf, zusätzlich wie ein RDA-Ordinationsdiagramm interpretiert werden zu dürfen (ter Braak 1998).

Tabelle 5.4-15: CA-Ergebnisse für die Indikatorgruppe Carabiden im Hauptuntersuchungsgebiet.

**** Summary **** HUG UP1					
Axes	1	2	3	4	Total inertia
Eigenvalues	.381	.160	.132	.087	1.412
Cumulative percentage variance of species data	27.0	38.3	47.6	53.8	
**** Summary **** HUG UP3					
Axes	1	2	3	4	Total inertia
Eigenvalues	.344	.147	.120	.092	1.370
Cumulative percentage variance of species data	25.1	35.9	44.7	54.4	
**** Summary **** HUG UP5					
Axes	1	2	3	4	Total inertia
Eigenvalues	.396	.153	.117	.078	1.405
Cumulative percentage variance of species data	28.2	39.1	47.4	53.0	
**** Summary **** HUG UP7					
Axes	1	2	3	4	Total inertia
Eigenvalues	.423	.176	.131	.101	1.519
Cumulative percentage variance of species data	27.8	39.4	48.0	54.7	

CA: Der vorliegende Stichprobenplan scheint für eine Artengruppe mit hoher Mobilität wenig geeignet zu sein. Als Folge werden erheblich kürzere Gradienten abgebildet und der Erklärungsanteil/der Eigenwert für die Artentrennung sinkt verglichen mit den Molluskenergebnissen drastisch. Während die Coenocline der ersten Achse bei den Mollusken zwischen 32,5% und 37,6% der Variabilität der Arten zu erklären vermag, erreichen die Carabidenmodelle unabhängig von den Erklärungsvariablen der direkten Ordinationen nur zwischen 25,1% und 28,2%. Ein Vergleich höherdimensionaler Achsen ist wegen des auftretenden Bogeneffekts (vgl. Abbildung 5.4-12, S. 182) nicht ratsam. DCA- und CCA-Resultate lassen sich durch den Detrending-Prozess der DCA nicht direkt vergleichen.

CCA: Die Korrelationsstrukturen in Tabelle 5.4-16 lassen schnell erkennen, dass die Überflutungsdauer (FDAJd98/99) für mobile Artengruppen, und das dürfte nicht nur für Laufkäfer gelten, eine weniger dominante Rolle zu spielen scheint. Diese Aussage kann jedoch zum Trugbild werden, da durch die Möglichkeit, mit der Flutwelle zu wandern und einen Ortswechsel zu Flächen mit angepassten Lebensbedingungen durchzuführen, die eigentliche Bedeutung des Parameters und die Existenzspanne der Arten nicht mit dem vorliegenden Stichprobenplan in RIVA klar erfasst werden konnte. Da multivariate Analysemethoden nicht unbedingt auf Kausalzusammenhänge zurückschließen lassen, darf die indirekte Wirkung nicht unterschätzt werden. Wiederkehrende langandauernde Überflutungen können die Vegetationsstruktur dauerhaft verändern und damit Verschiebungen in der Carabiden-Besiedlungsstruktur fördern. Ebenso ist ab einer bestimmten kritischen Schwelle der Überflutungsintensität mit mangelnden Refugienflächen und zu großen Rückzugsdistanzen zu rechnen. Die Bedeutungsschwankung des Parameters zwischen den Frühjahrs- und den Herbstphasen könnte diese These stützen. Der höhere Erklärungsanteil im Frühjahr mag Bezug nehmen auf die Flächenverfügbarkeit, während im Herbst eher die Erreichbarkeit feuchter Flächen Relevanz gewinnt. Dies fällt auch auf, in dem anstatt Boden- erstmals Distanzparameter (zweite Achse) ins Modell Eingang finden. Sowohl die Distanz zu den Flutrinnen (DistFIRi) als auch zu permanenten Wasserflächen (DistPerm) korreliert hoch mit den ersten beiden Ordinationsachsen. Die saisonalen Modellaspekte finden sich dabei in der Tatsache, dass die Distanz zu den Flutrinnen, welche im Herbst bis auf wenige Ausnahmen trockenfallen, in den feuchten Frühjahrphasen einen höheren Erklärungsanteil an der Artenverteilung aufweist als im Herbst.

**Tabelle 5.4-16: Gewichtete Korrelationsmatrizen der Carabidenmodelle für UP1 bis UP7 (maximale/relevante Korrelationen zu den Achsen sind grau unterlegt; Variablen: vgl. Anlage 6).**

HUG UP1 **** Weighted correlation matrix (weight = sample total) ****								
							Carabiden	
DistFIRi	0,6554	0,5711	-0,0678	0,292	0,6861	0,6171	-0,0763	0,3732
DistPerm	0,2146	-0,7964	-0,2784	0,2549	0,2246	-0,8605	-0,3136	0,3258
SDHN98	0,2127	0,0068	0,3244	-0,127	0,2226	0,0074	0,3653	-0,1624
MnGW498	0,9164	0,0241	-0,1687	-0,1319	0,9592	0,026	-0,19	-0,1686
FDAJd98	-0,5664	-0,2588	-0,6676	0,0075	-0,5928	-0,2796	-0,7519	0,0095
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4

HUG UP3 **** Weighted correlation matrix (weight = sample total) ****								
							Carabiden	
DistFIRi	0,4669	-0,5051	-0,2419	-0,0013	0,4974	-0,5666	-0,3089	-0,0015
DistPerm	0,5098	0,6478	0,1105	-0,3199	0,5431	0,7267	0,1411	-0,3734
SDHN98	0,0288	-0,3186	-0,5862	-0,2124	0,0307	-0,3575	-0,7484	-0,2479
MnGW498	0,8583	-0,3019	-0,0864	0,1192	0,9144	-0,3387	-0,1104	0,1391
FDAJd98	-0,1813	0,778	-0,1492	0,3333	-0,1931	0,8728	-0,1905	0,3891
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4

HUG UP5 **** Weighted correlation matrix (weight = sample total) ****								
							Carabiden	
DistFIRi	0,6915	0,549	0,0072	-0,1158	0,7202	0,6198	0,0085	-0,1394
DistPerm	0,0636	-0,4979	-0,6632	-0,2122	0,0662	-0,562	-0,7776	-0,2555
SDHN99	0,4891	-0,1441	0,1252	-0,5904	0,5094	-0,1626	0,1467	-0,7108
MnGW499	0,9079	0,0559	-0,1765	0,0128	0,9455	0,0631	-0,207	0,0154
FDAJd99	-0,7401	0,0722	-0,468	-0,0426	-0,7708	0,0815	-0,5487	-0,0513
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4

HUG UP7 **** Weighted correlation matrix (weight = sample total) ****								
							Carabiden	
DistFIRi	0,393	-0,4151	0,0478	0,4898	0,4203	-0,4869	0,0592	0,6272
DistPerm	0,5291	0,664	-0,0215	-0,0062	0,5658	0,7788	-0,0266	-0,0079
SDHN99	0,4233	-0,1954	0,3109	0,2739	0,4527	-0,2292	0,3852	0,3508
MnGW499	0,8614	-0,3238	0,0641	0,0148	0,9213	-0,3798	0,0795	0,0189
FDAJd99	-0,3532	0,5233	0,4952	-0,2275	-0,3777	0,6137	0,6136	-0,2913
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4



In der trockenen Jahreszeit gewinnen daher die permanenten Wasserflächen an Gewicht. Die Standardabweichung des Wasserstands im Kalenderjahr (bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand: SDHN) für 1998 und 1999 wird zwar basierend auf den Signifikanztests ins Modell aufgenommen, ist aber wenig effizient. Die mittleren Grundwasserflurabstände (MnGW) hingegen integrieren über eine längere Zeitphase im Jahr (April bis einschließlich September) und bringen so die weniger dynamische Feuchtkomponente der Aue zum Ausdruck. Sie werden zum wichtigsten Erklärungsparameter im Modell (erste Achse).

Verglichen mit der Modellerstellung in der RDA zeigt sich kein deutlich abweichendes Bild. Es werden die gleichen Parameter selektiert. In zwei Untersuchungsphasen zeigt sich der Parameter "Grundwasser niedriger als 140cm unter Flur, längste Periode in Tagen innerhalb der Monate 4 bis 9 in 1998 bzw. 1999" als trennstärker gegenüber dem "mittleren Grundwasserflurabstand". Die Informationsgehalte der beiden Parameter sind jedoch recht ähnlich. Je nach Vorbehandlung der biotischen Matrix (Standardisierung ja/nein) zeigen sich weiterhin Verschiebungen im Bedeutungsgehalt einzelner Parameter. Wird die biotische Matrix standardisiert, verliert die Distanz zu den Flutrinnen merklich an Erklärungsanteilen. Daraus lässt sich ableiten, dass einige dominante Arten im Datensatz eine hohe Affinität zu diesem Parameter aufweisen. Nach der Gleichgewichtung der Arten kommt es zum Bedeutungsverfall. Zudem ändert sich die Reihenfolge der Aufnahme von Erklärungsvariablen ins Modell. Diese Erkenntnisse bestätigen erneut, dass die Wahl des Analyseverfahrens beachtenswerten Einfluss auf die Ergebnisse nimmt.

Die vergleichende Bewertung der CA- und CCA-Ergebnisse demonstriert Tabelle 5.4-17, Abbildung 5.4-12 und Abbildung 5.4-13 (S. 182). Auch in den Carabidenmodellen zeigen die direkten Modelle gegenüber den indirekten eine geringere Trennkraft.

Tabelle 5.4-17 führt nur die erste Achse der Ordinationen an, da die CA einen starken Bogeneffekt aufweist und die zweite Achse als quadratische Funktion der ersten gelten muss.

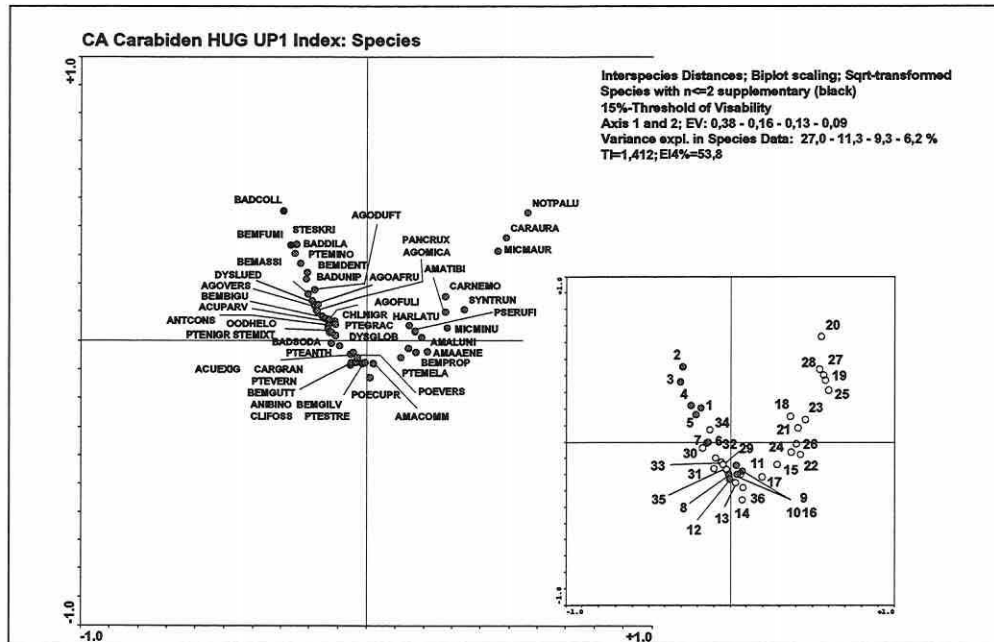
Vergleicht man das CCA-Ergebnis mit den hier nur schriftlich eingefügten RDA-Ergebnis, so erhält man für die erste Achse einen Erklärungsanteil an der biotischen Variabilität von 23,8% anstatt 24,4% (Tabelle 5.4-17, UP1, % CCA).

**Tabelle 5.4-17: Vergleich der CA- mit den CCA-Modellen für die Indikatorgruppe Carabiden (EV: Eigenwert der Achse; %: erklärter Prozentanteil der Achse an der biotischen Variabilität).**

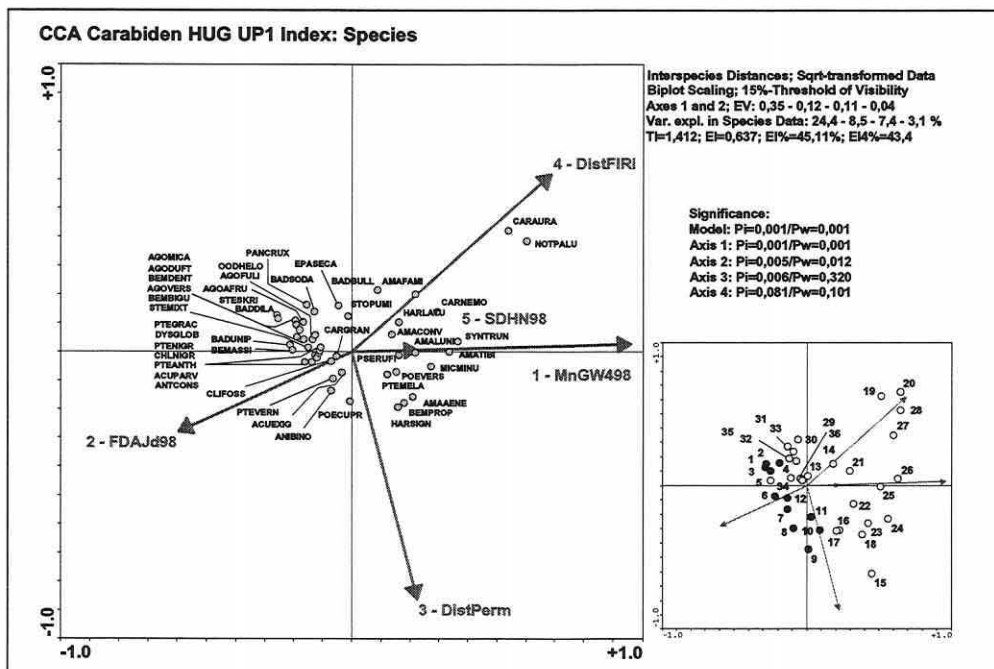
Achse	UP1		UP3		UP5		UP7	
	EV CA/CCA	% CA/CCA	EV CA/CCA	% CA/CCA	EV CA/CCA	% CA/CCA	EV CA/CCA	% CA/CCA
1	0,38/0,35	27,0/24,4	0,34/0,30	25,1/21,9	0,40/0,36	28,2/25,8	0,42/0,37	27,8/24,1

Die Trennung der Arten durch die Erklärungsvariablen in der CCA ist im Rahmen des biotischen Datenmaterials (Problematik der Mobilität in Bezug auf die Datenerhebungsstrategie) zwar gegeben, bleibt allerdings hinter den Erwartungen zurück. Eine klare Separierung der Spezies in stabile ökologische Gruppen ist äußerst schwierig, da die Artengemeinschaften über die Jahreszeiten und die Untersuchungsjahre Verschiebungen erfahren (vgl. Abbildung 5.4-13 bis 5.4-16, S. 182f). Der Eigenwertsprung in der CCA von der ersten zur zweiten Achse ist für diese Indikatorgruppe noch gravierender (vgl. Mollusken und Flora). Die Werte unterscheiden sich um das 2,7 bis 3,1-fache. Die Gesamterklärungsanteile der Modelle bewegen sich um 43% (vgl. EI% in den folgenden Abbildungen) und sind damit erheblich schlechter als für die Mollusken, aber insgesamt gut für ökologische Modelle.

Trotz der insgesamt hochsignifikanten Modelle gibt es einige Besonderheiten bei den Signifikanzen der Achsen anzusprechen.



**Abbildung 5.4-12: CA-Carabidenmodell der UP1 basierend auf den standardisierten Fangergebnissen einer Probefläche (Index vgl. Abschnitt 5.3.5; Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).**



**Abbildung 5.4-13: CCA-Molluskenmodell der UP1 basierend auf den standardisierten Fangergebnissen einer Probefläche (Index vgl. Abschnitt 5.3.5; Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).**

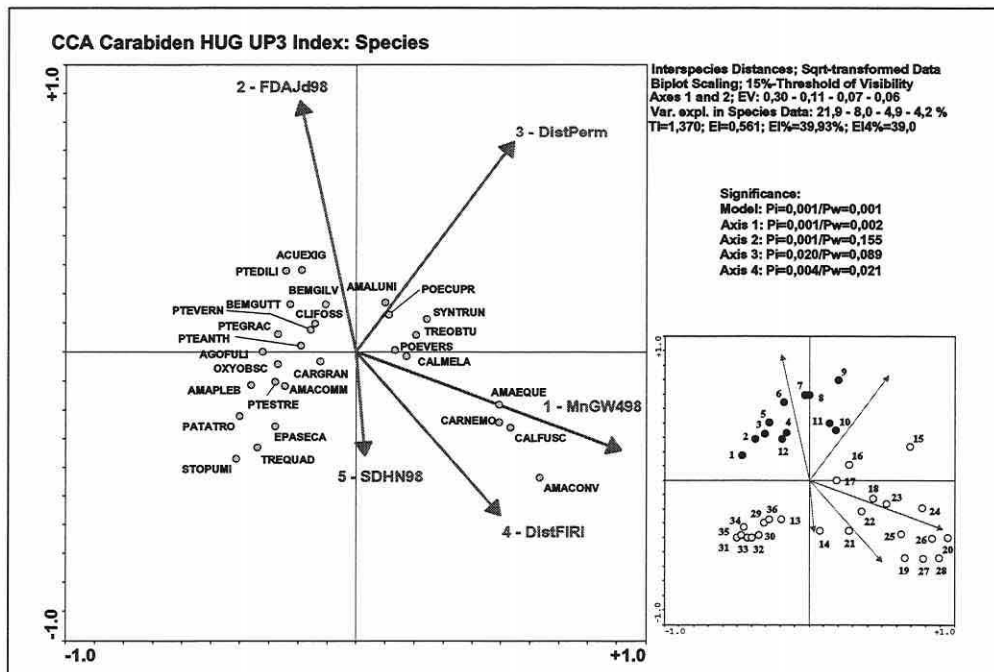


Abbildung 5.4-14: CCA-Molluskenmodell der UP3 basierend auf den standardisierten Fangergebnissen einer Probefläche (Index vgl. Abschnitt 5.3.5; Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).

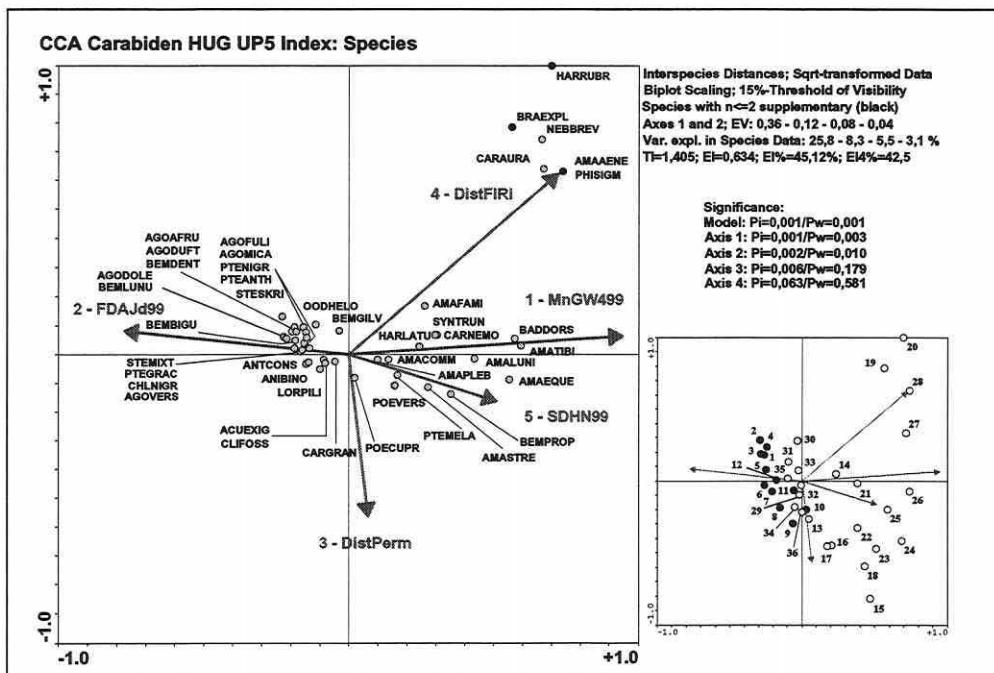


Abbildung 5.4-15: CCA-Molluskenmodell der UP5 basierend auf den standardisierten Fangergebnissen einer Probefläche (Index vgl. Abschnitt 5.3.5; Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).

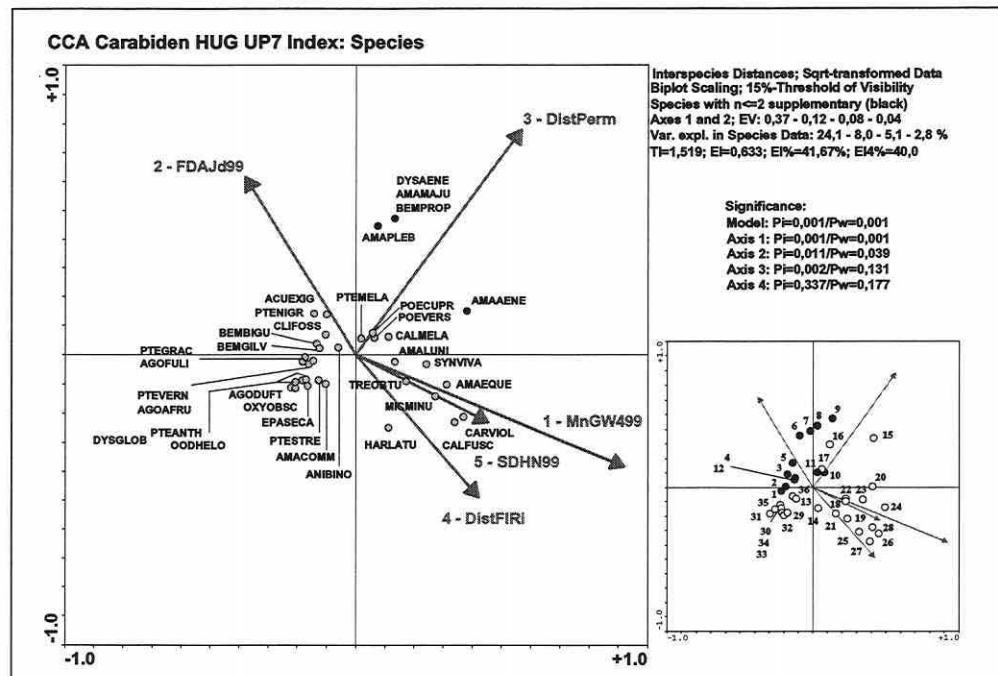


Abbildung 5.4-16: CCA-Molluskenmodell der UP7 basierend auf den standardisierten Fangergebnissen einer Probestfläche (Index vgl. Abschnitt 5.3.5; Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).

Die erste Achse ist in allen Fällen hochsignifikant, während die zweite Achse sich schon an die Signifikanzschranke ( $P_w$ ) annähert bzw. diese in den Herbstmodellen nicht erreicht. Auch die nachgeschalteten Achsen weisen keine konstant verlässliche Trennkraft auf. Durch die verstärkte Ladung der Überflutungsdauer auf die zweite Achse in UP3 und UP7 verändert sich das Gesamtbild der Ordinationen. Anhand der Probestflächenanordnung lässt sich der Einfluss der Änderung erkennen. Weitgehend gegenläufig korreliert (Kosinus des Winkels nähert sich in Abbildung 5.4-13, S. 182, und Abbildung 5.4-15, S. 183, dem Wert -1 an) zum mittleren Grundwasserflurabstand bildet sich eine Abfolge der Flutrinnenflächen über die nassen bis zu den trockenen Flächen entlang der ersten Achse aus. In den Herbstphasen ist dies zwar immer noch nachvollziehbar, die Abfolge verläuft aber nicht parallel mit der ersten Achse. Die Distanz zu permanenten Wasserflächen als dritte ins Modell aufgenommene Variable zeigt durch den 90°-Winkel zum mittleren Grundwasserflurabstand ihre Unkorreliertheit zu diesem an. Durch die Migrationsbewegungen der Tiere ergeben sich in den Modellen keine disjunkten Gruppen (vgl. Mollusken und Flora).

#### Ergebnisse Flora:

DCA: Gradientenlängen zwischen 5,3 und 6,3 SD (Tabelle 5.4-18, S. 185) lassen nur unimodale Ordinationsverfahren zur Analyse der Flora in Frage kommen.

CA: Es zeichnet sich erneut ein sehr dominanter erster Gradient für die Artentrennung ab (Eigenwertsprung, Tabelle 5.4-19, S. 185). Die Gesamtvariabilität im Datensatz ist über die vier Erhebungsphasen sehr ähnlich, obgleich in UP5 zwei Probestflächen ausgefallen sind, da sie bereits gemäht waren. Somit scheinen die über die beiden Jahre veränderten Umweltbedingungen keine merkliche Auswirkung auf die Flora ausgeübt zu haben. Dies spricht für eine längere Reaktionszeit der Indikatorgruppe. Es besteht wohl eine hohe zeitliche Autokorrelation.

Tabelle 5.4-18: DCA-Ergebnisse für die Indikatorgruppe Flora im Hauptuntersuchungsgebiet.

**** Summary **** HUG UP1					
Axes	1	2	3	4	Total inertia
Eigenvalues	.882	.274	.182	.135	3.886
Lengths of gradient	5.906	2.421	2.378	1.573	
Cumulative percentage variance of species data	22.7	29.8	34.4	37.9	
**** Summary **** HUG UP3					
Axes	1	2	3	4	Total inertia
Eigenvalues	.878	.272	.141	.103	3.803
Lengths of gradient	5.303	3.033	1.647	2.206	
Cumulative percentage variance of species data	23.1	30.3	34.0	36.7	
**** Summary **** HUG UP5					
Probeflächen 2 u. 4 ausgefallen (nicht beprobbar)					
Axes	1	2	3	4	Total inertia
Eigenvalues	.865	.336	.131	.085	3.907
Lengths of gradient	6.327	3.609	1.527	1.586	
Cumulative percentage variance of species data	22.1	30.7	34.1	36.3	
**** Summary **** HUG UP7					
Axes	1	2	3	4	Total inertia
Eigenvalues	.870	.292	.131	.073	3.700
Lengths of gradient	5.479	2.964	1.602	1.928	
Cumulative percentage variance of species data	23.5	31.4	35.0	36.9	

Tabelle 5.4-19: CA-Ergebnisse für die Indikatorgruppe Flora im Hauptuntersuchungsgebiet.

**** Summary **** HUG UP1					
Axes	1	2	3	4	Total inertia
Eigenvalues	.882	.450	.288	.254	3.886
Cumulative percentage variance of species data	22.7	34.3	41.7	48.2	
**** Summary **** HUG UP3					
Axes	1	2	3	4	Total inertia
Eigenvalues	.878	.433	.322	.237	3.803
Cumulative percentage variance of species data	23.1	34.0	42.2	48.6	
**** Summary **** HUG UP5					
Probeflächen 2 u. 4 ausgefallen (nicht beprobbar)					
Axes	1	2	3	4	Total inertia
Eigenvalues	.865	.478	.389	.345	3.907
Cumulative percentage variance of species data	22.1	34.0	43.9	52.6	
**** Summary **** HUG UP7					
Axes	1	2	3	4	Total inertia
Eigenvalues	.870	.423	.349	.253	3.700
Cumulative percentage variance of species data	23.5	34.9	44.4	51.2	



CCA: Durch die Korrelationsstrukturen in Tabelle 5.4-20 wird kenntlich, welche Parameter die Ordinationsachsen mit Information belegen. Die erste Achse ist bestimmt durch Parameter des Überflutungsregimes. Sowohl die jährliche Überflutungsdauer (FDAJd98/99) als auch die maximale Überflutungshöhe im Untersuchungsjahr (maxFH98/99) zeigen sich konstant hoch korrelierend. Die Überflutungshäufigkeit pro Jahr (FDFre98/99) wird bedingt durch die Überflutungsdauer und erfährt in 1999 mit sehr lang andauernden intensiven Überflutungen eine Informationsveränderung. Speziell die in 1998 selten bis nicht von Überflutung betroffenen Flächen werden nun mehrmals kurzzeitig erreicht, so dass deren Werte von den niedrigsten in 1998 zu den höchsten in 1999 werden. Dies geht einher mit der Verlagerung ihrer Trennkraft auf die zweite Ordinationsachse. Das ebenfalls auf die erste Achse ladende pflanzenverfügbare Phosphor (Ppfl) zeigt sich schon in den Voranalysen als weitgehend unkorreliert zu anderen Parametern im Erklärungsdatensatz. Die moderate Korrelation zur ersten Achse und zu den Überflutungsparametern erklärt sich ansatzweise dadurch, dass die Löslichkeit unter reduzierenden Bedingungen zunimmt. Die Phosphorzufuhr erfolgt überwiegend über die Zersetzung von Biomasse. Der nachgeordnete zweite Gradient repräsentiert die grundwasserbedingten Feuchtverhältnisse (MaxGW258/698 bzw. 259/699: maximale Grundwasserflurabstände in den Drei-Monatsphasen 1998 und 1999) und die effektive Kationenaustauschkapazität des Bodens (KAKeff). Die KAKeff integriert die Bodenart, den Humusgehalt sowie den pH-Wert und steht gleichzeitig als Summenparameter für die Kationen Mangan, Magnesium, Kalium, Calcium und Eisen. Phosphor ist unabhängig von ihr.

**Tabelle 5.4-20: Gewichtete Korrelationsmatrizen der Floramodelle für UP1 bis UP7 (maximale/relevante Korrelationen zu den Achsen sind grau unterlegt; Variablen: vgl. Anlage 6).**

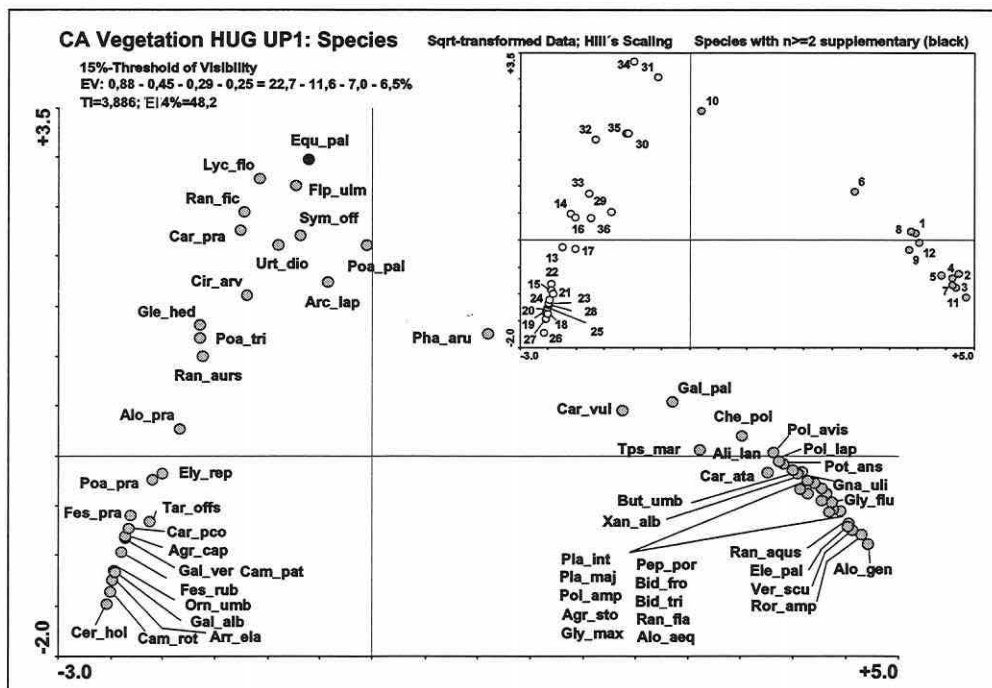
HUG UP1								Flora	
**** Weighted correlation matrix (weight = sample total) ****									
Ppfl	0,5175	-0,2839	-0,232	0,2567	0,524	-0,3059	-0,2418	0,2802	
KAKeff	-0,1778	0,7409	-0,1028	0,4483	-0,18	0,7983	-0,1072	0,4895	
SDHN98	-0,2114	0,2002	0,372	0,4191	-0,2141	0,2158	0,3876	0,4576	
MaxGW258	-0,5505	-0,7075	0,0188	0,2584	-0,5573	-0,7624	0,0195	0,2821	
FDAJd98	0,9695	-0,1481	-0,0876	-0,02	0,9816	-0,1596	-0,0913	-0,0218	
FDFre98	0,9425	0,0698	0,0119	0,0009	0,9542	0,0752	0,0124	0,001	
maxFH98	0,8727	-0,1567	0,3754	0,0971	0,8836	-0,1689	0,3912	0,106	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	
HUG UP3								Flora	
**** Weighted correlation matrix (weight = sample total) ****									
Ppfl	0,5037	-0,2459	-0,1597	0,2618	0,512	-0,2653	-0,1744	0,2836	
KAKeff	-0,1584	0,8013	-0,0942	0,3533	-0,161	0,8646	-0,1029	0,3827	
SDHN98	-0,2255	0,2268	0,367	0,4007	-0,2292	0,2447	0,4008	0,4341	
MaxGW698	-0,6409	-0,6197	-0,0466	0,2778	-0,6514	-0,6886	-0,0509	0,301	
FDAJd98	0,9667	-0,1429	-0,08	0,0039	0,9825	-0,1542	-0,0874	0,0043	
FDFre98	0,8463	-0,0371	0,3992	0,1314	0,8602	-0,04	0,436	0,1423	
maxFH98	0,8599	-0,1472	0,3553	0,1226	0,874	-0,1588	0,388	0,1328	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	
HUG UP5								Flora	
**** Weighted correlation matrix (weight = sample total) ****									
Ppfl	0,6133	-0,3024	0,3485	0,067	0,6256	-0,3268	0,3943	0,0812	
KAKeff	-0,1092	0,8191	0,1394	0,3391	-0,1114	0,8851	0,1577	0,4107	
SDHN99	-0,2321	-0,376	-0,2536	0,4137	-0,2367	-0,4062	-0,2869	0,5011	
MaxGW259	-0,5418	-0,6718	0,1636	0,1793	-0,5526	-0,7259	0,1851	0,2171	
FDAJd99	0,9774	-0,0059	-0,0571	0,0189	0,9971	-0,0063	-0,0647	0,0229	
FDFre99	-0,4148	0,6783	-0,1695	-0,226	-0,4231	0,7329	-0,1917	-0,2738	
maxFH99	0,8251	0,3451	-0,2733	0,0758	0,8416	0,3728	-0,3092	0,0919	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	
HUG UP7								Flora	
**** Weighted correlation matrix (weight = sample total) ****									
Ppfl	0,5305	-0,218	0,193	0,2638	0,5375	-0,2343	0,2242	0,2913	
KAKeff	-0,1773	0,8444	0,2042	0,2302	-0,1797	0,9075	0,2373	0,2541	
SDHN99	-0,2739	-0,3074	-0,1416	0,6341	-0,2776	-0,3304	-0,1645	0,7001	
MaxGW699	-0,6221	-0,6298	0,2244	0,1674	-0,6304	-0,6768	0,2608	0,1849	
FDAJd99	0,986	-0,0104	0,0068	0,0003	0,9992	-0,0112	0,0079	0,0003	
FDFre99	-0,4843	0,5526	-0,1941	-0,2312	-0,4908	0,5939	-0,2256	-0,2553	
maxFH99	0,8635	0,2739	-0,28	0,121	0,875	0,2944	-0,3254	0,1336	
	SPEC AX1	SPEC AX2	SPEC AX3	SPEC AX4	ENVI AX1	ENVI AX2	ENVI AX3	ENVI AX4	

Die KAKeff beschreibt somit den Boden als Komplex-Gradienten. Meist sind feuchte Standortverhältnisse sowie Bodenarten mit höheren Schluff- und Tongehalten mit hohen Austauschkapazitäten assoziiert, daher zeigt sich die KAKeff negativ korreliert zu den höchsten Grundwasserflurabständen. Die dritte Achse ist nicht eindeutig definiert und die Standardabweichung des Wasserstandes im Kalenderjahr 1998 und 1999 gewinnt erst für die vierte Achse an Bedeutung.

Der Eigenwertvergleich zwischen den CA- und CCA-Ergebnissen bestätigt, dass die wichtigen Parameter erfasst wurden (Abbildung 5.4-17 und 5.4-18, S. 188). Das direkte Modell vermag das biotische Muster der Flora im Untersuchungsgebiet zu erklären. Die Erklärungsanteile zwischen den indirekten und direkten Analysen sind nur geringfügig niedriger (Tabelle 5.4-21). Es werden trotz des vermeintlichen Bogeneffekts in der CA zwei Achsen in der Tabelle verglichen, da das Muster der jeweils korrespondierenden CCA gegen eine echte Verzerrung spricht.

**Tabelle 5.4-21: Vergleich der CA- mit den CCA-Modellen für die Indikatorgruppe Flora (EV: Eigenwert der Achse; %: erklärter Prozentanteil der Achse an der biotischen Variabilität)**

Achse	UP1		UP3		UP5		UP7	
	EV	%	EV	%	EV	%	EV	%
1	0,88/0,86	22,7/22,0	0,88/0,85	23,1/22,0	0,87/0,82	21,9/21,0	0,87/0,84	23,5/22,8
2	0,45/0,37	11,6/9,4	0,43/0,35	10,9/9,2	0,48/0,36	12,1/9,1	0,42/0,35	11,4/9,4



**Abbildung 5.4-17: CA-Floramodell der UP1 basierend auf der größten Erhebungseinheit (100m²) einer Probesträhle (FL01, vgl. Seite 151 und Abbildung 5.2-3, S. 150; Erläuterungen in Abbildung 5.4-6, S. 174 und 5.4-7, S. 175).**

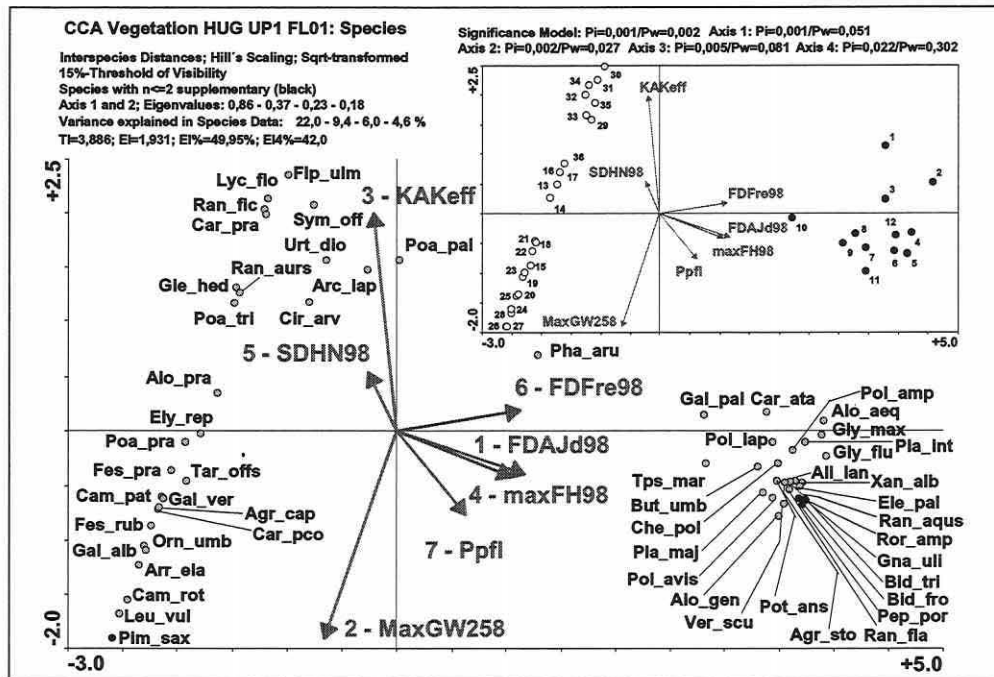


Abbildung 5.4-18: CCA-Floramodell der UP1 basierend auf der größten Erhebungseinheit (100m<sup>2</sup>) einer Probe-fläche (FL01, vgl. Seite 151 u. Abbildung 5.2-3, S. 150; Erläuterungen in Abbildung 5.4-6, S. 174 u. 5.4-7, S. 175).

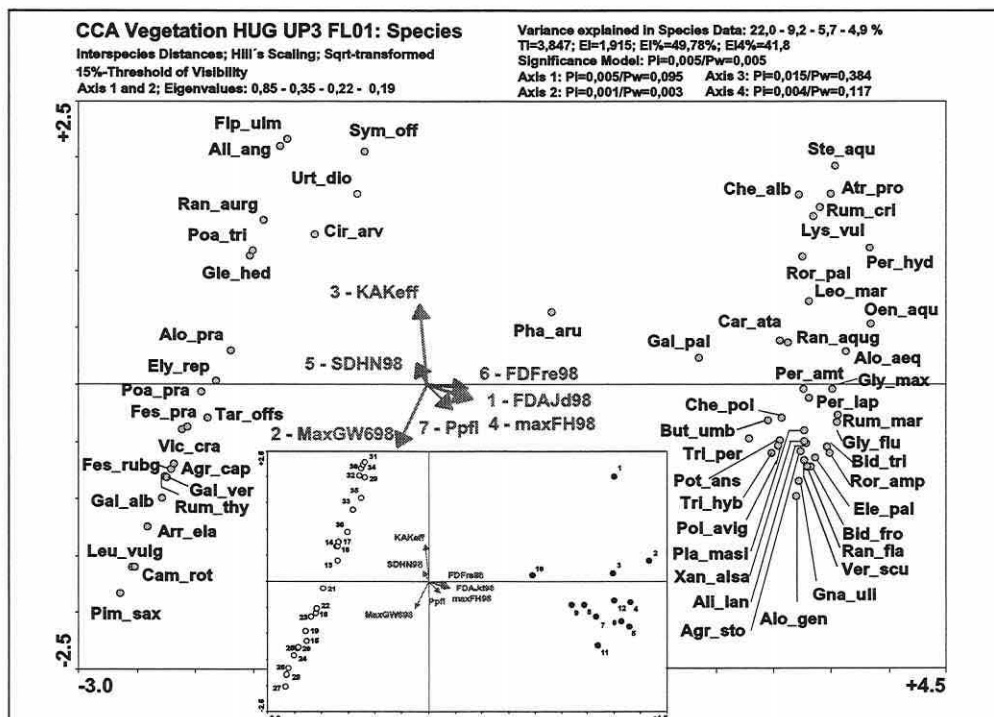


Abbildung 5.4-19: CCA-Floramodell der UP3 basierend auf der größten Erhebungseinheit (100m<sup>2</sup>) einer Probe-fläche (FL01, vgl. Seite 151 u. Abbildung 5.2-3, S. 150; Erläuterungen in Abbildung 5.4-6, S. 174 u. 5.4-7, S. 175).

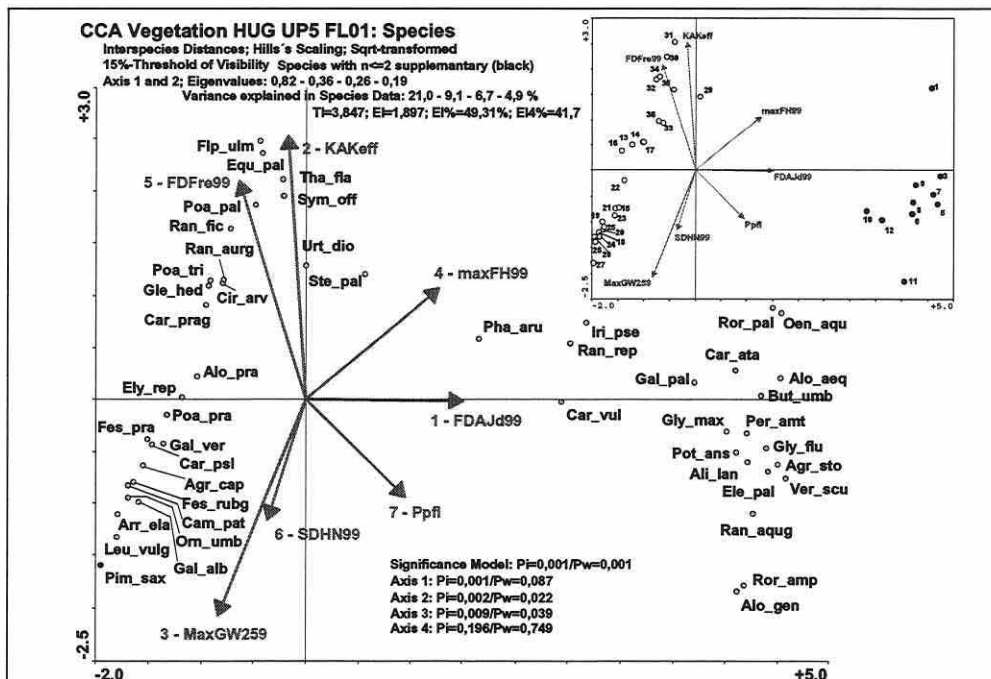


Abbildung 5.4-20: CCA-Floramodell der UP5 basierend auf der größten Erhebungseinheit (100m<sup>2</sup>) einer Probe-fläche (FL01, vgl. Seite 151 u. Abbildung 5.2-3, S. 150; Erläuterungen in Abbildung 5.4-6, S. 174 u. 5.4-7, S. 175).

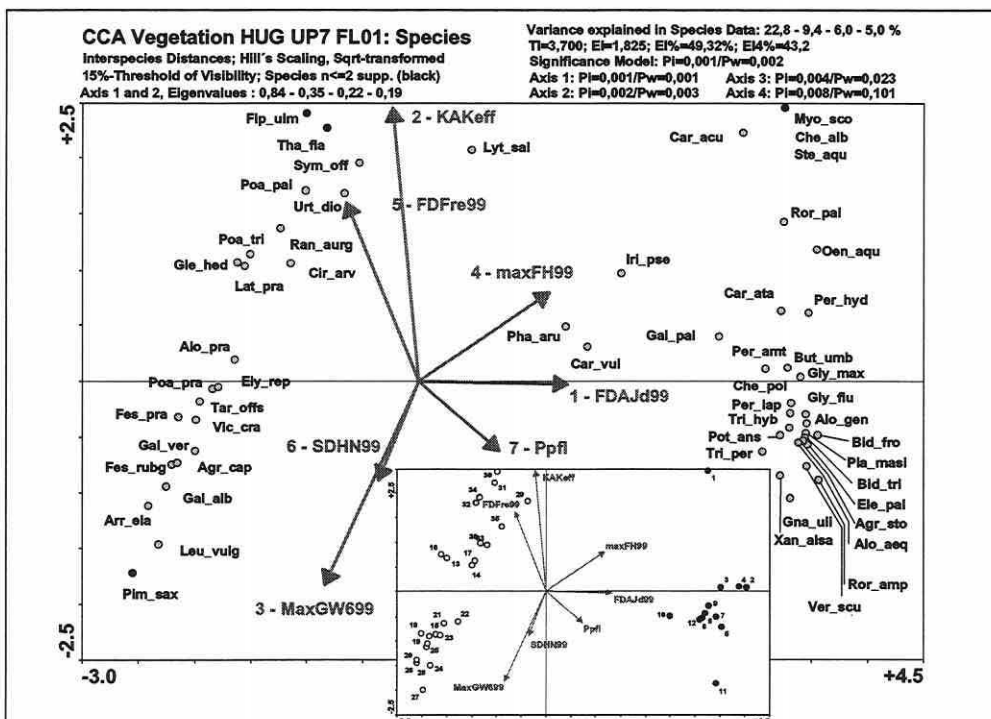


Abbildung 5.4-21: CCA-Floramodell der UP7 basierend auf der größten Erhebungseinheit (100m<sup>2</sup>) einer Probe-fläche (FL01, vgl. Seite 151 u. Abbildung 5.2-3, S. 150; Erläuterungen in Abbildung 5.4-6, S. 174 u. 5.4-7, S. 175).

In den Floramodellen finden sich vergleichbare Ergebnisse zu den anderen Modellen (Fauna). Hohe Eigenwerte der ersten Achse und um 2,3 bis 2,4-fach niedrigere für die zweite Achse. Die sehr ausgeprägte Dominanz der Überflutungsdauer, welche in 1999 nahezu exakt der ersten Ordinationsachse folgt, teilt die Biotik erneut in disjunkte Gruppen (vgl. Molluskenmodell). Die Trennung der feuchten bis trockenen Flächen erfolgt hauptsächlich entlang der zweiten Ordinationsachse. Dies ist an der Verteilung der Probeflächen klar nachvollziehbar. Die Eigenwerte der ersten beiden Achsen sind beachtlich und die Gesamterklärungsanteile von über 49% (vgl. EI% in den vorangegangenen Abbildungen) für die Datensätze mit der höchsten Variabilität (TI) in der Untersuchung gut. Die Signifikanztests verraten, dass das Gesamtmodell hochsignifikant ist, aber die Achsen differenziert zu betrachten sind. Die Parameter der ersten Achse sind zwar in der Lage, zwischen den Straten hochsignifikant zu trennen, scheitern allerdings an der Trennung innerhalb der Straten.

Dies bedeutet, dass die Flora speziell durch die Überflutungsdauer in ökologische Großgruppen getrennt werden kann, aber eine Feindifferenzierung erst durch weitere Erklärungsvariablen im Modell möglich ist (zweite Achse). Die dritte und vierte Achse erreicht ebenfalls fast ohne Ausnahme keine Signifikanz des Pw-Wertes.

Insgesamt belegen alle Modelle über die Indikatorgruppen hinweg den dominanten Einfluss des Feuchtreghimes für die Erklärung des biotischen Musters in einer rezenten, unter Überflutungsdynamik befindlichen Aue. Alle weiteren Erklärungsvariablen treten in den Hintergrund.

#### 5.4.3 Weiterführende Auswertungen zu den Ordinationsergebnissen, Ableitung ökologischer Gruppen und Schirmartenselektion

Die erste explorative Informationsebene von Ordinationsverfahren bilden die im vorangegangenen Abschnitt dargestellten Standard-Ordinationsdiagramme. Sie dienen der Mustererkennung im Gesamtdatensatz und stellen die Basismodelle dar. Die zweite Informationsebene liefert darauf aufbauend Detailinformationen zu den Modellparametern und den Beobachtungsobjekten. Sie dient weiterhin der Plausibilitätskontrolle von Einzelergebnissen und prüft die Repräsentationsgüte des Modells auf der Ebene einzelner Beobachtungsobjekte. Erst diese Überprüfung trennt zuverlässige Detailergebnisse von weniger zuverlässigen und ermöglicht so eine auch statistisch besser abgesicherte Grundlage zur Bildung ökologischer Gruppen und der Selektion von Schirmarten, die stellvertretend für eine Gruppe näher betrachtet werden. Einige bisher noch offene Fragen können so beantwortet werden.

##### **Disjunkte Gruppen:**

In den Flora- und Molluskenmodellen entstehen entlang des ersten Gradienten durch den sprunghaften Anstieg der Überflutungsdauer von den Flutrinnenflächen zu den übrigen Probeflächen disjunkte Gruppen, welche separat betrachtet werden. Es ist von Interesse, ob diese Teildatensätze eine bessere Trennung in der Einzelanalyse erfahren.

Dazu wurden zuerst nur die Arten selektiert, welche in der Ordination der Flora (UP1 des HUG; vgl. Abbildung 5.4-18, S. 188) Koordinaten größer +2,5 entlang des ersten Gradienten aufweisen. Außerdem wurden nur die Probeflächen des Stratum 'Flutrinne' (Flächen 1-12) aktiv in der Analyse betrachtet. Das Ergebnis des neu gebildeten Modells zeigt Abbildung 5.4-22 (Teilbild A, S. 191). Im Gegensatz zum originalen Gesamtmodell erlangt die Amplitude des Grundwasserstandes über die Monate April bis September 1998 (AMPGW498) für den Teildatensatz ein höheres Gewicht als der maximale Grundwasserflurabstand der dreimonatigen Frühjahrsphase (MaxGW258) und geht daher ersatzweise ins Modell ein. Vergleicht man den univariaten Erklärungsanteil der Parameter AMPGW498 und MaxGW258 für das Gesamtmodell (Tabelle 5.4-8, S. 168) wird offensichtlich, dass der Parameter für dieses nicht die gleiche Relevanz erlangt. Im Gesamtmodell besitzt der MaxGW258 univariat 12,9% und die AMPGW498 nur 5,9% Erklärungsanteil an der biotischen Variabilität.



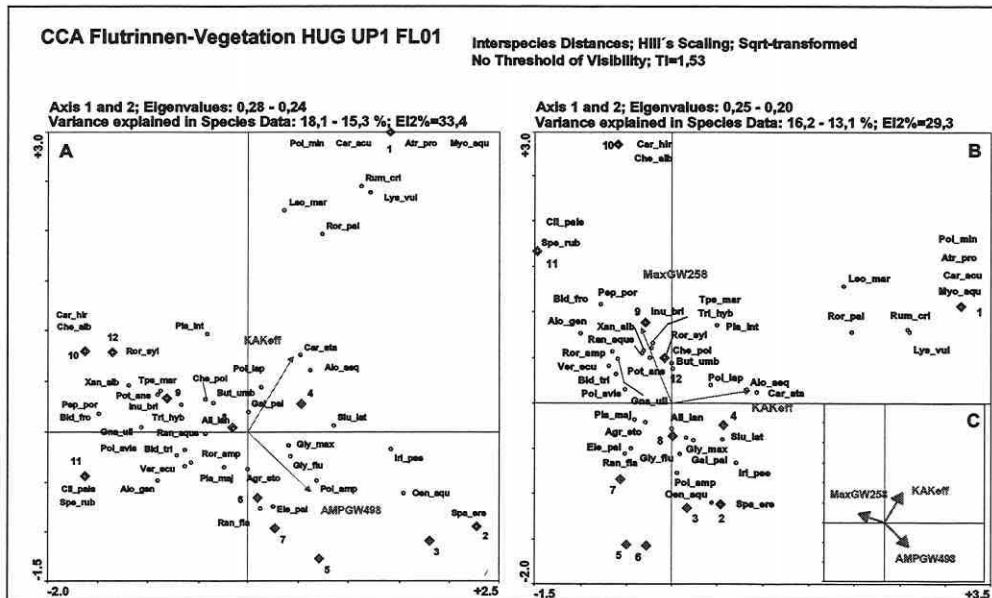


Abbildung 5.4-22: Ordination des Teildatensatzes 'Flora der Flutrinnen' für die UP1 des HUG. Neu erstelltes Modell (Teilbild A), Austausch des neu selektierten Modellparameters Amplitude des Grundwasserstandes über die Monate April bis September 1998 (AMPGW498) mit dem maximalen Grundwasserflurabstand der dreimonatigen Frühjahrsphase (MaxGW258) (Teilbild B) und Korrelationsstruktur der Modellparameter zueinander (Teilbild C, KAKeff: effektive Kationenaustauschkapazität).

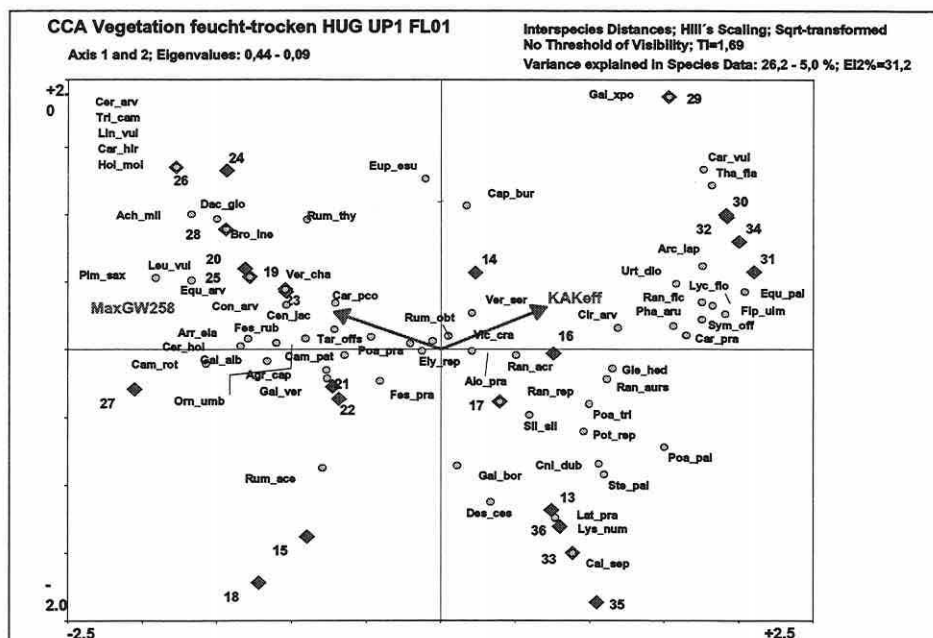


Abbildung 5.4-23: Ordination des Teildatensatzes 'Flora der trockenen bis nassen Probestellen' für die UP1 des HUG (MaxGW258: maximaler Grundwasserflurabstand der dreimonatigen Frühjahrsphase und KAKeff: effektive Kationenaustauschkapazität).

Teilbild C der Abbildung 5.4-22 (S. 191) verrät die negative Korrelationsstruktur der Parameter MaxGW258 und AMPGW498 (Kosinus des Winkels; vgl. Abschnitt 3.6.3) zueinander, falls man sie beide ins Modell aufnimmt. Trotz des geringeren Erklärungsanteils von - 4,1% über zwei Achsen (vgl. EI2%: 33,4 zu 29,3 in Abbildung 5.4-22, Teilbild A und B) ist die getroffene Aussage durch den MaxGW258 annähernd gleich zur AMPGW498 (Teilbild A zu B). Das Muster der Artenverteilung wird durch den Austausch der beiden Variablen im Modell, abgesehen von einer Drehung, kaum beeinflusst. Die separate Ordination der disjunkten Gruppe 'Flutrinne' mit einem neuen Parameter in der Auswahl erbringt für keine Untersuchungsphase einen zusätzlichen Informationsgewinn.

Für den verbleibenden Teil des Datensatzes (trockene bis feuchte Flächen) wurde die gleiche Kontrollberechnung durchgeführt (Abbildung 5.4-23, S. 191). Die Modellbildung führt zu einem reduzierten, aber identischen Erklärungsvariablensatz bezogen auf das Gesamtmodell (vgl. Abbildung 5.4-18, S. 188). Weitere Parameter mit geringen Erklärungsanteilen und die kombinierte Betrachtung der ersten und dritten Achse in einem zweidimensionalen Diagramm lassen keine relevante Zusatzinformation für die Artentrennung erkennen. Die dritte Achse ist bei nur zwei Erklärungsvariablen in der Analyse nicht durch Umweltparameter eingeschränkt, sondern ordiniert die Arten rein anhand ihrer biotischen Information wie bei einer CA. Dies dient dem Test, ob eine interpretierbare Struktur sichtbar wird, die nicht von den Erklärungsvariablen in der Analyse beschrieben werden kann. Die Analyse führt auch für den Molluskendatensatz zu keinen weiterführenden Erkenntnissen.

#### Univariate Ergebnisse für wichtige Modellparameter:

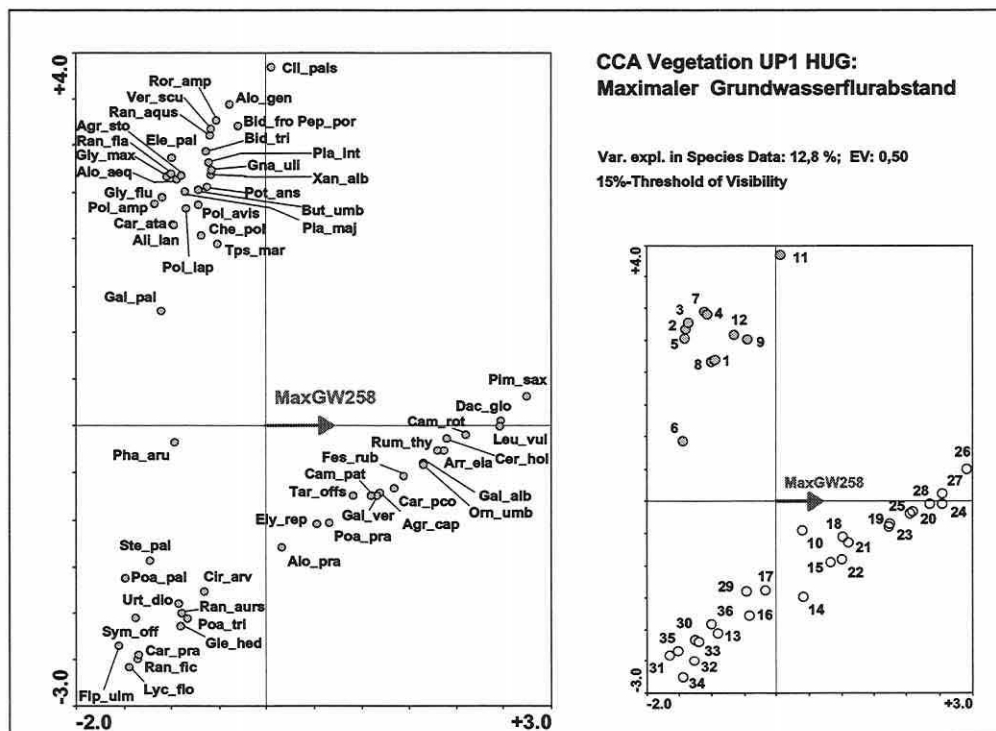


Abbildung 5.4-24: Univariate Ordination der Flora in UP1 HUG: Maximaler Grundwasserflurabstand der dreimonatigen Frühjahrsphase.

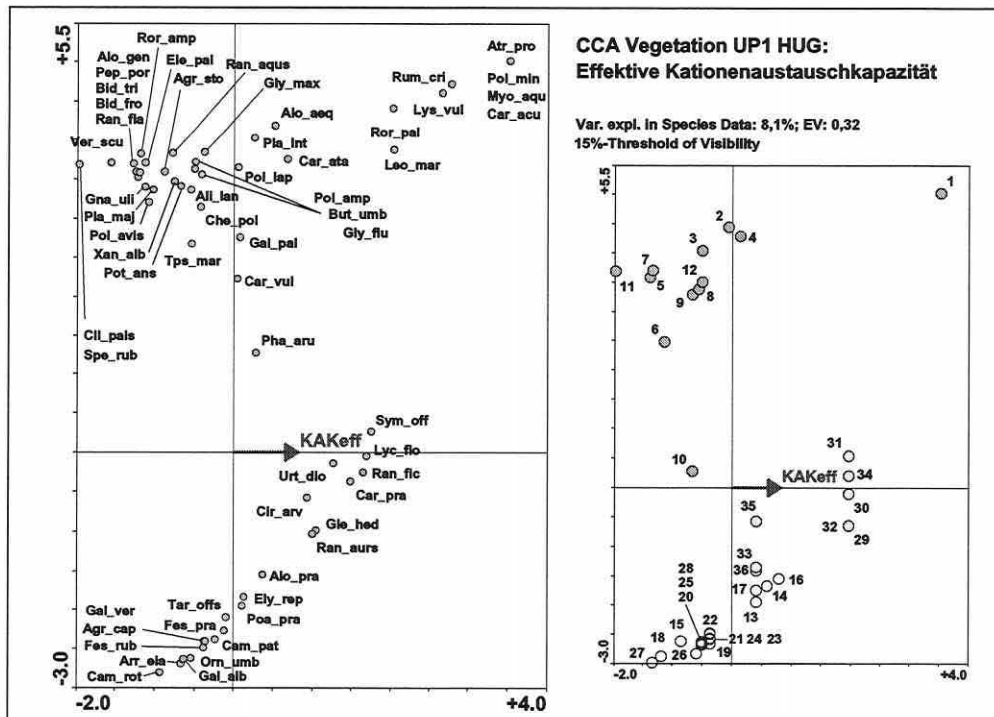


Abbildung 5.4-25: Univariate Ordination der Flora in UP1 HUG: effektive Kationenaustauschkapazität.

Es ist häufig sehr hilfreich, für die Modellparameter univariate Ordinationen zu berechnen, um den alleinigen Einfluss auf das biotische Muster und die Abfolge der Arten und Probenflächen entlang des Gradienten mit dem Gesamtmodell in Beziehung zu setzen. Beim Beispiel der Flora bleibend werden die Parameter MaxGW258 und KAKeff für die Darstellung gewählt (Abbildung 5.4-24, S. 192, und 5.4-25). Mit etwa 12,8% und 8,1% univariatem Erklärungsanteil am biotischen Muster (EV: 0,5 bzw. 0,32) zeigen beide Parameter eine gute Trenneigenschaft (Signifikanz:  $\alpha < 0,05$ ). Das Gefälle der maximalen Grundwasserflurabstände von den trockeneren Probenflächen weg (Abbildung 5.4-24, S. 192) und die maximalen Werte der effektiven Kationenaustauschkapazität auf den feuchten Probenflächen (Abbildung 5.4-25) ist klar ausgeprägt (Grobansprache der Flächen nach Stratifizierung des Stichprobenplans). Im Vergleich zu Abbildung 5.4-18 (CCA-Gesamtmodell, S. 188) sind einige strukturierende Eigenschaften der beiden Parameter für die Biotik bei der univariaten Betrachtung klarer zuzuordnen.

Der trennkraftigste Parameter des Gesamtmodells, die Überflutungsdauer in Tagen, bedingt eher eine Grobdifferenzierung der Arten, da der kontinuierliche Übergang der Überflutungswerte größtenteils auf die räumlich sehr schmalen Flankenbereiche der Flutrinnen zusammengedrängt ist. Sie wird hier nicht als Beispiel verwendet, da die Daten der 36 Probenflächen im HUG den Artenwechsel der Flankenbereiche nicht in Feinabstufung widerspiegeln können. Es entsteht eine Lücke im Überflutungsgradienten, welche sich auch für die Ausbildung der disjunkten Gruppen im Mollusken- und Floramodell verantwortlich zeigt (s.o. und vgl. z.B. Abbildung 5.4-18, S. 188).

#### Korrelationen zwischen den Arten und den Modellparametern:

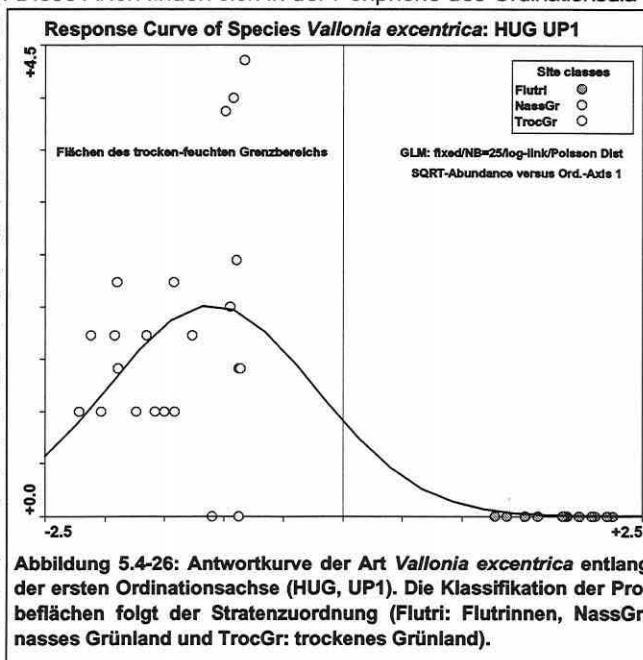
Die Ordinationsdiagramme vermögen nur die Hauptmuster der Korrelationen offen zu legen. Möchte man exaktere Werte, kann der File "Spec\_Env.tab" analysiert werden (vgl. Abschnitt 3.6.1). Dem File sind bei linearen Verfahren (zentriert und standardisiert) die Korrelationen



Die Clusterergebnisse können für die Gruppe der Mollusken den Abbildungen 5.4-7 bis 5.4-10 (S. 175f) entnommen werden. Diese Indikatorgruppe wird erneut als Beispiel ausgewählt, da durch die geringe Artenzahl eine bessere Übersichtlichkeit gewährt ist. Für die Flora und die Carabiden werden im Folgenden die Gruppen nur benannt (vgl. Tabelle 5.4-23, S. 196).

Wertet man die beiden Informationsbasen 'Ordination und Clusteranalyse' gegeneinander aus, kann man leichte Abweichungen feststellen, die bei genauerer Überprüfung und dem Vergleich über die Untersuchungsphasen hinweg schnell als mathematisches Artefakt des zu starren Klassifikationsalgorithmus erkannt werden. Die Art *Perpolita hammonis* (Ppt\_ham) wird in der UP1 durch die Clusteranalyse der sehr trockene Standorte besiedelnden Gruppe zugeordnet (Abbildung 5.4-7, Teilbild oben, S. 175), während sie in der Frühjahrsphase 1999 feuchter erscheint, was hinsichtlich des stärkeren Überflutungsgeschehens in diesem Jahr zu erwarten ist (Abbildung 5.4-9, Teilbild oben, S. 176). Ein analoges Bild zeigt sich auch im Vergleich der UP3 und UP7 zueinander (Abbildung 5.4-8 und 5.4-10, S. 175f). Die mit den Fachwissenschaftlern vorgenommene Gruppierung basierend auf den Ordinationsdiagrammen berücksichtigt die differierenden Randbedingungen und zeigt sich somit flexibler (Abbildung 5.4-7 bis Abbildung 5.4-10 auf S. 175ff, Liniendarstellung in Hauptbild). Sie ordnet die Art bezogen auf ihr räumliches Vorkommen und den Ordinationsergebnissen allgemein der Gruppe 'Übergang trockene/feuchte Flächen' zu (vgl. Tabelle 5.4-23, S. 196). Trotz der Abweichungen belegen die Analysen, dass die Klassifizierung im reduzierten Ordinationsraum die Strukturen weitgehend verzerrungsfrei auf die dimensional nicht limitierte Clusteranalyse bezogen wiedergibt. Speziell die Berücksichtigung von vier einzelnen Erhebungsphasen kann die Stabilität der Gruppen demonstrieren und Ausreißer lokalisieren helfen. Es ist anfangs erstaunlich, dass eine Gruppe in UP3 sich über den gesamten oberen Teil der Grafik erstreckt. Diese Erscheinung tritt auf, da *Vitrina pellucida* (Vtr\_pel), ein absoluter Einzelfund, nur auf Probestfläche 10 vorkommt, welche durch *Anisus spirorbis* (Ani\_spi) mit 922 Individuen zurecht in den feuchten Sektor verschoben ist. Gleichzeitig finden sich auf dieser Fläche nur noch die Arten *Vallonia excentrica* (Val\_exc, 14 Individuen) und *Vallonia spec.* (Val\_spe, ein Individuum). Diese Arten finden sich in der Peripherie des Ordinationsdiagramms, da sie als einzige auch auf den trockeneren Probestflächen existieren können. Ihr Optimum finden sie allerdings eher im trocken-feuchten Grenzbereich (Abbildung 5.4-26), solange die Überflutungswasserstände nicht längere Zeit diese Flächen wie in 1999 erreichen. Sie ist zwar in letzterem Falle noch präsent, aber mit erheblich geringeren Individuenzahlen. Diese Interpretationen sind nur unter Rückgriff auf die Originaldaten möglich, da in der Ordination die relative Individuenzahl bzw. Artnächtigkeit zugrunde gelegt wird.

Interessiert die Partizipation der einzelnen Arten an den Probestflächenausprägungen, kann diese der Abbildung 5.4-27 (S. 196)



**Abbildung 5.4-26: Antwortkurve der Art *Vallonia excentrica* entlang der ersten Ordinationsachse (HUG, UP1). Die Klassifikation der Probestflächen folgt der Stratenzuordnung (Flutri: Flutrinnen, NassGr: nasses Grünland und TrocGr: trockenes Grünland).**



anschaulich entnommen werden. Die typische Kontinuität mit Häufungen entlang der Gradienten (vgl. Abschnitt 2.2) bilden die Grafiken über die Untersuchungsjahre (UP1 und UP5) sehr gut ab. Die Artenkomposition wechselt entlang der Flutrinnen über das nasse zum trockeneren Grünland hin. Eine annuelle Variabilität ist vorhanden, ändert aber die Grundaussage nicht. Als konkretes Beispiel sei wiederum die Art *Vallonia excentrica* (Val\_exc) in Abbildung 5.4-27 betrachtet. Sie partizipiert in 1998 durch die geringere Überflutung mit deutlich höheren Anteilen an den Probeflächen des nassen Grünlandes (vgl. auch Abbildung 5.4-26, S. 195) als in 1999. Hier verlagert Sie Ihren Vorkommensschwerpunkt auf die höher gelegenen Flächen des Stratum trockenes Grünland (Abbildung 5.4-27, rechtes Teilbild für UP5).

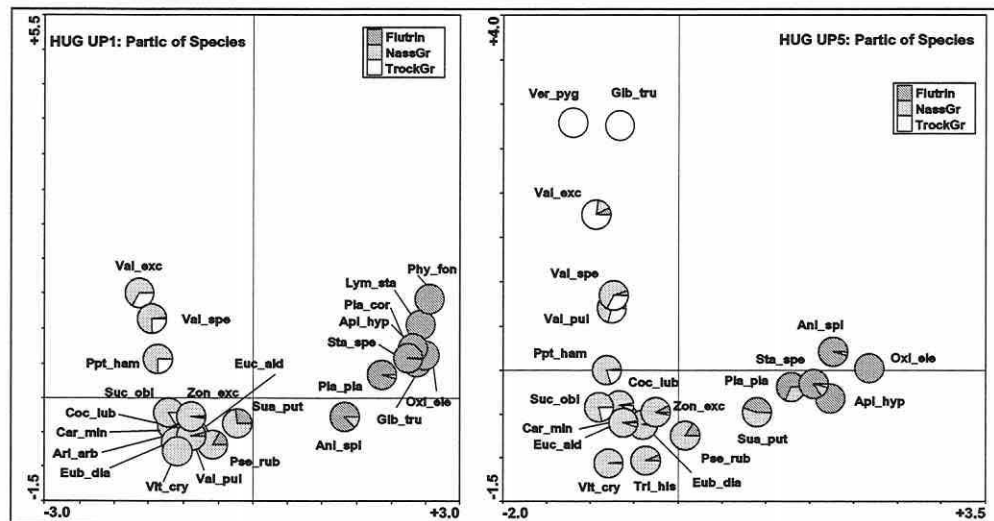


Abbildung 5.4-27: Partizipation der Arten an den nach der Stratifizierung geordneten Probeflächenausprägungen (HUG, UP1 und UP5).

Tabelle 5.4-23: Ökologische Gruppen und verbale Typisierung anhand der statistischen Ergebnisse.

Gruppe	Mollusken	Carabiden	Flora
1	Tiefere Rinnen, meist länger mit Wasser bestanden	Tiefe, lange überschwemmte Flutrinnen	<i>Persicaria amphibia</i> -Gruppe: Nasse, häufiger überschwemmte Standorte
2	Seichte, länger trocken liegende Rinnen	Kürzere Zeit überschwemmte, flache Rinnen, Phalarisröhrichte und feuchte bis frische Mähwiesen	<i>Galium palustre</i> -Gruppe: Nasse, etwas weniger häufig überschwemmte Standorte
3	Übergang von den Feuchflächen zu den seichteren Rinnen	Phalarisröhrichte	<i>Phalaris arundinacea</i> -Gruppe: Feuchte bis nasse Standorte
4	Tieferliegende Feuchflächen		<i>Urtica dioica</i> -Gruppe: Feuchte, nährstoffreiche, wenig genutzte Standorte
5	Übergang trockene/feuchte Flächen	Wechselfeuchte bis mäßig trockene Mähwiesen	<i>Alopecurus pratensis</i> -Gruppe: Frische Standorte
6	Höhergelegene, die meiste Zeit trocken liegende, genutzte Wiesenflächen	Trockene, allenfalls sehr kurz überschwemmte Mähwiesen	<i>Arrhenatherum elatius</i> -Gruppe: Trockenere Standorte

Die am Beispiel der Mollusken aufgezeigte Vorgehensweise bei der Bildung ökologischer Gruppen wurde ebenfalls für die Carabiden und die Flora angewandt und führt zu den abschließend in Tabelle 5.4-23 (S. 196) angeführten Ergebnissen. Die Gruppen korrespondieren zu den Klassen in Abbildung 5.4-28, welche die Verteilung über die Probestellen als Diagramm zur Beurteilung der klassifizierten Artenkomposition auf den Probestellen wiedergibt. Der Wechsel in den Anteilen der Gruppen an den Probestellen ist anhand der Kuchendiagramme nachzuvollziehen. Die Gruppierung der Arten bildet die Grundlage des nächsten Auswerteschrittes: der Selektion von Schirmarten für die Gruppen.

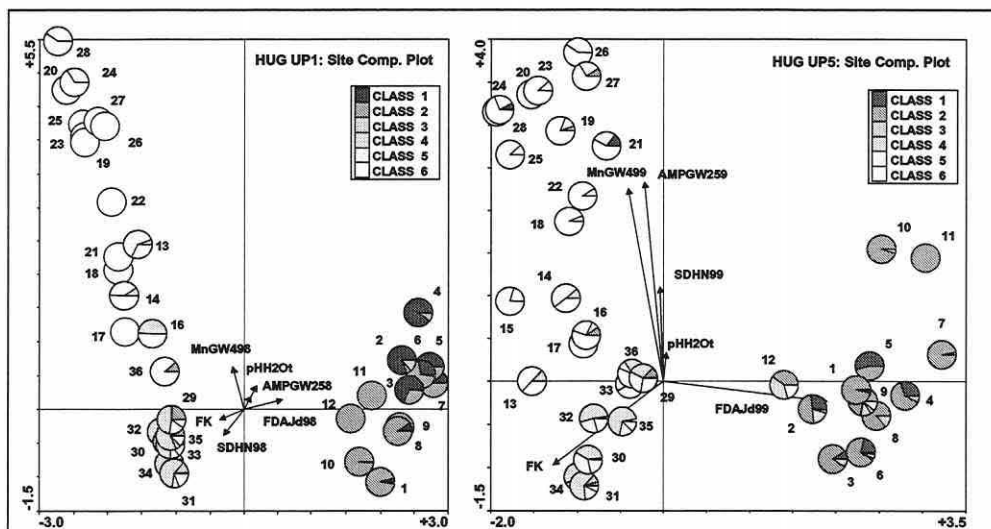


Abbildung 5.4-28: Verteilung der ökologischen Gruppen (vgl. Tabelle 5.4-23) über die Probestellen (HUG, UP1 und UP5).

#### Selektion der Schirmarten:

Die Selektion der Schirmarten soll wenigstens zwei typische Vertreter pro ökologischer Gruppe ausweisen, welche durch Mitnahmeeffekte stellvertretend für die Ansprüche der Gruppenmitglieder an ihre jeweilige Nische stehen. Neben den in Abschnitt 2.5 benannten Kriterien kamen weiterhin statistische Kennwerte bzw. Gütewerte zum Tragen. Neben der Häufigkeit von Erscheinungen, dem Datensatzprofil der Originaldaten, dem Vorkommen in möglichst allen Untersuchungsphasen im HUG und weiterhin einem Auftreten auch in den NUGs (Übertragbarkeit/Modellvalidierung) wurden Gesamtvarianz der jeweiligen Art, deren durch die ersten beiden Achsen des Modells erklärter Anteil daran (Cfit) und deren Toleranz (Standardabweichung der Antwortkurve) bewertet.

Gerade einige der typischen Arten mit geringer Existenzspanne werden schnell durch einen anthropogenen Eingriff in ihrer Erscheinung und Verbreitung betroffen. Aus modelltechnischen Gründen bzw. im Rahmen des Prognosegedankens sollten daher auch die Folgebesiedler bestimmbar sein. Konkurrenzstarke ubiquistische Arten müssen deshalb ebenfalls im Modell integriert sein. Die statistischen Kennwerte werden auszugsweise für die Mollusken in der nachfolgenden Abbildungen 5.4-29 und 5.4-30 (S. 198) offengelegt. Das mit CanoDraw® erstellte Diagramm (Abbildung 5.4-29, rechts) kann zwar durch das Modell gut von weniger gut repräsentierten Arten anhand des Kreisdurchmessers unterscheiden, die Aufschlüsselung nach Achsen liefert ein Balkendiagramm aber differenzierter (Abbildung 5.4-29, links). *Planorbis planorbis* (Pla\_pla) ist verglichen mit *Vallonia excentrica* (Val\_exc) etwa gleich gut durch das Modell erklärt, die Repräsentationsgüte ist aber auf verschiedene Achsen aufgeteilt, d.h. maßgeblich durch verschiedene Erklärungsvariablen definiert (vgl. Abbildung 5.4-7, S. 175).

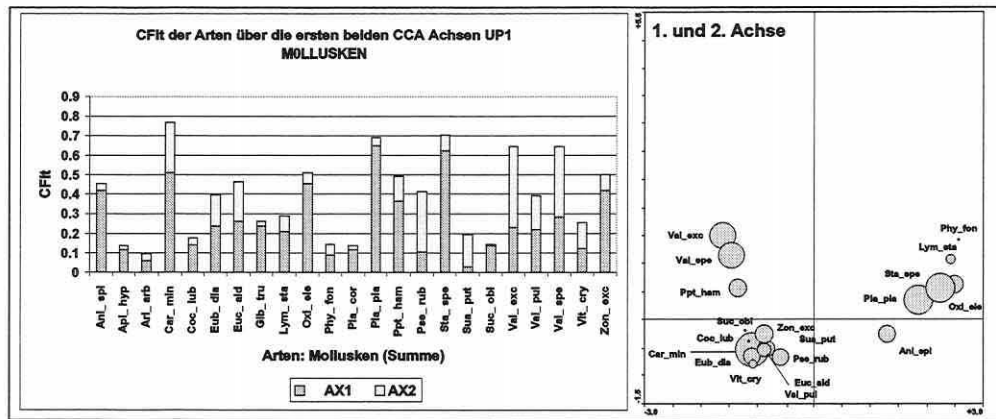


Abbildung 5.4-29: Güte der Modellanpassung (CFit entspricht Prozentzahl: 0-90) für die Molluskenarten (HUG, UP1); nach Anteilen der Ordinationsachsen AX1 und AX2 (links) und als Summe (rechts).

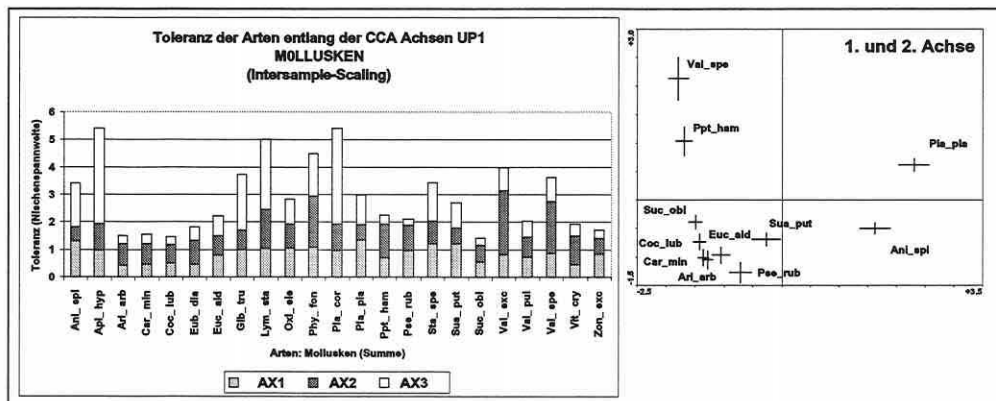


Abbildung 5.4-30: Toleranzen der Molluskenarten (HUG, UP1). Im Vergleich über drei Achsen (AX1 bis AX3, links) und als relative Werte über zwei Achsen (rechts, vgl. Erläuterungen im Text). Das Teilbild rechts umfasst aus darstellungstechnischen Gründen nur einen Auszug aus dem Gesamtdatensatz.

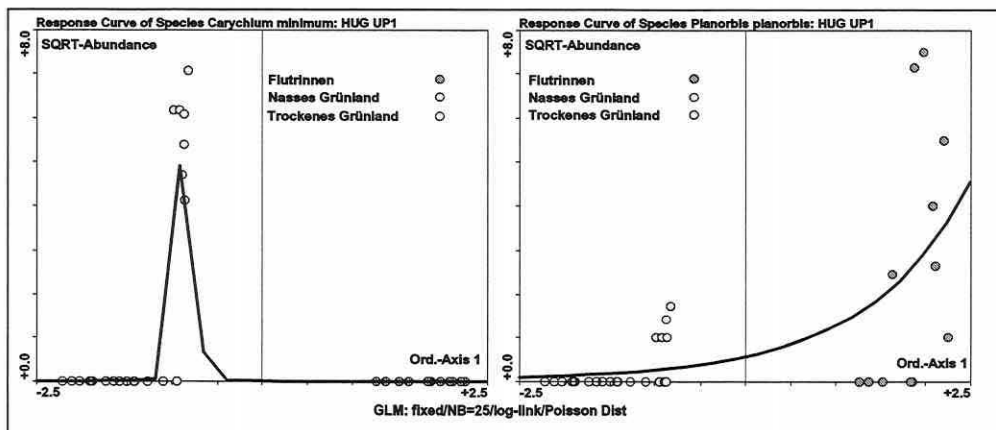


Abbildung 5.4-31: Antwortkurve der Arten *Carychium minimum* (links) und *Planorbis planorbis* (rechts) entlang der ersten Ordinationsachse (HUG, UP1).

Die Toleranz der Arten gegenüber dem Gradienten (Achsen) kann auf ähnliche Weise visualisiert werden (Abbildung 5.4-30, S. 198). Bei der Interpretation des Toleranzausdrucks aus CanoDraw® ist zu berücksichtigen, dass die Kreuze nur einen Relativvergleich zulassen (Abbildung 5.4-30, rechts). Sie stellen nicht die wirklich durch die Ordination approximierten Toleranzwerte dar, wie das Stapelbalkendiagramm. Zudem können in diesem mehr als zwei Achsenwerte simultan visualisiert werden. Die sehr geringe Toleranz von *Carychium minimum* (Car\_min) entlang der ersten Ordinationsachse kann die Antwortkurve in Abbildung 5.4-31 (S. 198) nochmals anschaulich bestätigen. Die Antwort von *Planorbis planorbis* (Pla\_pla), in der gleichen Abbildung rechts dargestellt, muss eine vorsichtige Interpretation des Toleranzwerts vor Augen führen. Als eine Art Standardabweichung der Nische ist diese nur gültig, wenn die Antwortkurve nahezu komplett beprobt wurde. Dieses Wissen ist essentiell zur korrekten Einschätzung von Prognoseergebnissen bei der Anwendung des ARC/INFO®-Moduls "CANOGEN" (vgl. Abschnitt 5.6).

Tabelle 5.4-24: Auswahl der Schirmarten innerhalb der ökologischen Gruppen aus Tabelle 5.4-23 (S. 196).

Gruppe	Mollusken		Carabiden		Flora	
1	<i>Planorbis planorbis</i> <i>Stagnicola spec.</i>	Pla_pla	<i>Stenolophus skrimshirani</i>	Steskri	<i>Oenanthe aquatica</i>	Oen_aqu
		Sta_spe	<i>Bembidion biguttatum</i>	Bembigu	<i>Agrostis stolonifera</i>	Agr_sto
			<i>Agonum versutum</i>	Agovers	<i>Eleocharis palustris</i>	Ele_pal
			<i>Agonum Duftschmidi</i>	Agoduft	<i>Persicaria amphibia</i>	Per_amt
			<i>Pterostichus gracilis</i>	Ptegrac	<i>Rorippa amphibia</i>	Ror_amp
			<i>Agonum fuliginosum</i>	Agufuli	<i>Glyceria maxima</i>	Gly_max
					<i>Glyceria fluitans</i>	Gly_flu
2	<i>Anisus spirorbis</i>	Ani_spi	<i>Carabus granulatus</i>	Cargran	<i>Carex acuta</i>	Car_ata
			<i>Bembidion gilvipes</i>	Bemgilv	<i>Galium palustre</i> <i>Iris pseudacorus</i>	Gal_pal Iri_spe
3	<i>Succinea putris</i>	Sua_put	<i>Amara communis</i>	Amacomm	<i>Phalaris arundinacea</i>	Pha_aru
	<i>Pseudotrichia rubiginosa</i>	Pse_rub	<i>Epaphius secalis</i>	Epaseca		
4	<i>Cochlicopa lubrica</i>	Coc_lub			<i>Glechoma hederacea</i>	Gle_hed
	<i>Carychium minimum</i>	Car_min			<i>Symphytum officinale</i>	Sym_off
	<i>Zonitoides nitidus</i>	Zon_exc			<i>Urtica dioica</i>	Urt_dio
5	<i>Vallonia pulchella</i>	Val_pul	<i>Syntomus truncatellus</i>	Syntrun	<i>Alopecurus pratensis</i>	Alo_pra
	<i>Perpolita hammonis</i>	Ppt_ham	<i>Amara lunicollis</i>	Amaluni	<i>Elymus repens</i>	Ely_rep
			<i>Calathus melanocephalus</i> ( <i>Poecilus versicolor</i> )	Calmela (Poever)		
			<i>Amara strenua</i>	Amastre		
6	<i>Vallonia excentrica</i>	Val_exc	<i>Carabus auratus</i>	Caraura	<i>Ornithogalum umbellatum</i>	Orn_umb
			<i>Amara equestris</i>	Amaeque	<i>Arrhenatherum elatius</i>	Arr_ela
			<i>Calathus fuscipes</i>	Calfusc	<i>Galium verum</i> <i>Galium album</i>	Gal_ver Gal_alb

Tabelle 5.4-24 (S. 199) gibt einen Überblick über die selektierten Arten, welche für die ökologisch gegliederten Gruppen von den Fachwissenschaftlern und durch den statistischen Abgleich bestimmt wurden (Mollusken: F. Föckler & O. Dechner – ÖKON Gesellschaft für Landschaftsökologie, Gewässerbiologie und Umweltplanung mbH Regensburg; Carabiden: A. Schanowski & W. Figura – Institut für Landschaftsökologie und Naturschutz Bühl/Gesamthochschule Paderborn-Höxter und Flora: U. Amarell – Umweltforschungszentrum Leipzig/Halle).

#### Ökologisches Verhalten der Schirmarten:

Die Mollusken werden weiterhin verwendet, um Detailauswertungen exemplarisch vorzustellen. *Anisus spirorbis* (Schirmart der Gruppe 2) zeigt in UP5 ein den Arten *Planorbis planorbis* und *Stagnicola spec.* (Schirmarten der Gruppe 1) sehr ähnliches Verhalten (Abbildung 5.4-32). Die Vorkommenswahrscheinlichkeit (Isolinienwerte in Abbildung 5.4-32) nimmt jeweils in Richtung der feuchten Flächen und Flutrinnen zu. Der Schwerpunkt der Arten weicht von der 1.0-Linie negativ ab, da bei der Ordination eine Gewichtung nach der Individuenzahl eingeht. Die Vorkommenswahrscheinlichkeit hingegen richtet sich nach dem Präsenz-Absenz-Verhalten der Arten. *Anisus spirorbis* ist in allen jenseits der 1.0-Linie befindlichen Flutrinnen anzutreffen. Ihr Vorkommen streut jedoch bis auf Fläche 27 und 16 mit jeweils einem Individuum. Die maximalen Häufigkeiten finden sich auf den Flächen 10 (484 Tiere) und 1 (211 Tiere). Die Auswertung der Grafiken in Abbildung 5.4-32 belegt einerseits die sehr ähnliche Lage der Schwerpunkte aller Arten und andererseits die Tendenz der Gruppe 2 zu noch nasserem Bedingungen, ersichtlich an der geringeren Distanz zwischen 1er- und 0.5er-Isolinie in Abbildung 5.4-32 (rechts). Während *Planorbis planorbis* auf keiner der Flächen 13 bis 18 auftritt, gibt es einen Einzelfund von *Stagnicola spec.* auf Fläche 24 (ein Individuum).

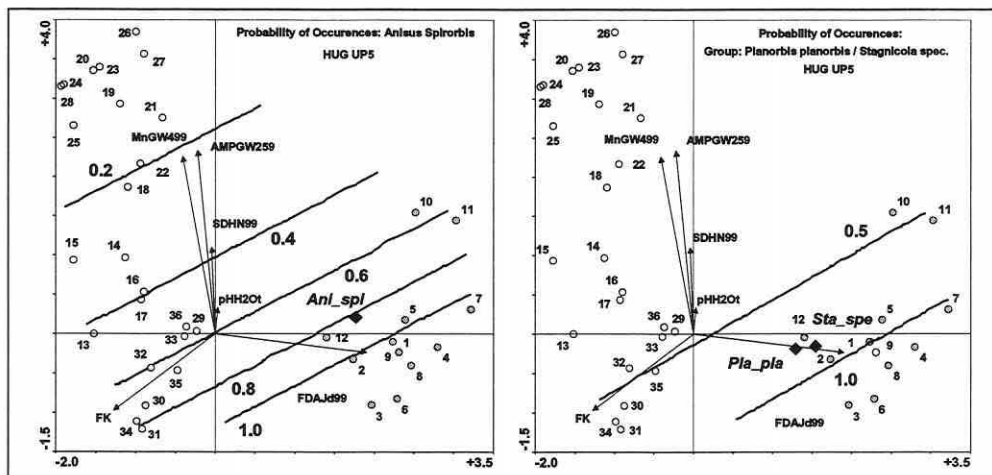


Abbildung 5.4-32: Isoliniendarstellung der Vorkommenswahrscheinlichkeiten für die Schirmarten der Gruppen 1 (links) und 2 (rechts). Probeflächen: weiß - trockenes Grünland, hellgrau - nasses Grünland und dunkelgrau – Flutrinnen.

Alle ausgewählten Arten der Molluskengruppe 4 (vgl. Tabelle 5.4-23, S. 196) haben einen sehr ähnlichen Schwerpunkt im Ordinationsraum der ersten und zweiten Achse (Abbildung 5.4-33, S. 201). Der Schwerpunkt der Art *Zonitoides nitidus* (Artcode: Zon\_exc) weicht nur geringfügig positiv entlang der ersten Achse ab. Anhand der Einzelbetrachtung (Abbildung 5.4-33, Teilbilder rechts) wird erkenntlich, wie weit und in welche Teilräume der Ordinations-ebene die Arten ausstrahlen.



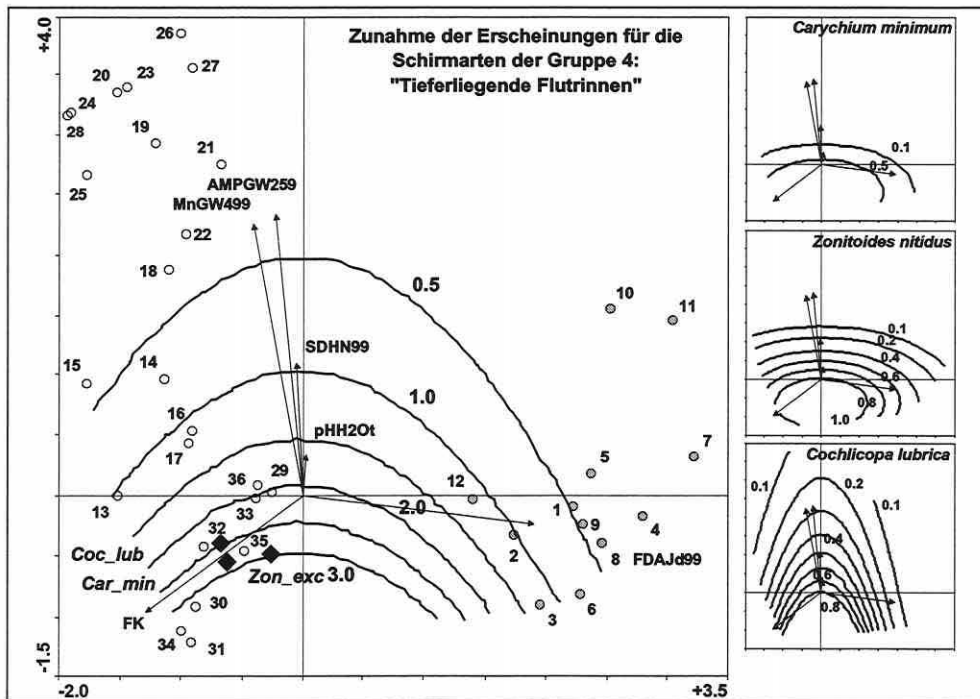


Abbildung 5.4-33: Schirmarten der Gruppe 4. Zunahme der Erscheinungen für die Schirmarten der Gruppe 4: "Tieferliegende Flutrinnen" (Hauptbild) und Vorkommenswahrscheinlichkeiten der einzelnen Arten (Teilbilder). Probeflächen: weiß - trockenes Grünland, hellgrau - nasses Grünland und dunkelgrau – Flutrinnen.

Da die Ordinationsmodelle für die wurzeltransformierten biotischen Daten und standardisierten Erklärungsvariablen erstellt wurden, interessiert es weiterhin, die Artinformationen auf die Originalwerte der Umweltparameter zurückzuführen, um die Nischen (-optima) konkreter ablesen zu können. Zu diesem Zweck werden die Ordinationsdiagramme mit der Originalinformation der einzelnen Modellparameter überlagert (Abbildungen 5.4-34 bis 5.4-36).

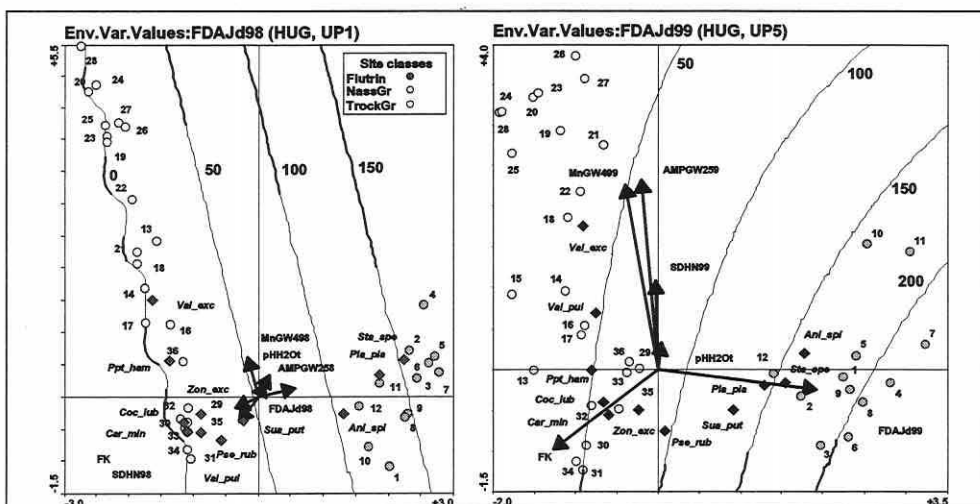


Abbildung 5.4-34: Isolinien der Überflutungsdauer in Tagen für die Molluskenmodelle: HUG, UP1 und UP5. Probeflächen: weiß - trockenes Grünland, hellgrau - nasses Grünland und dunkelgrau – Flutrinnen.

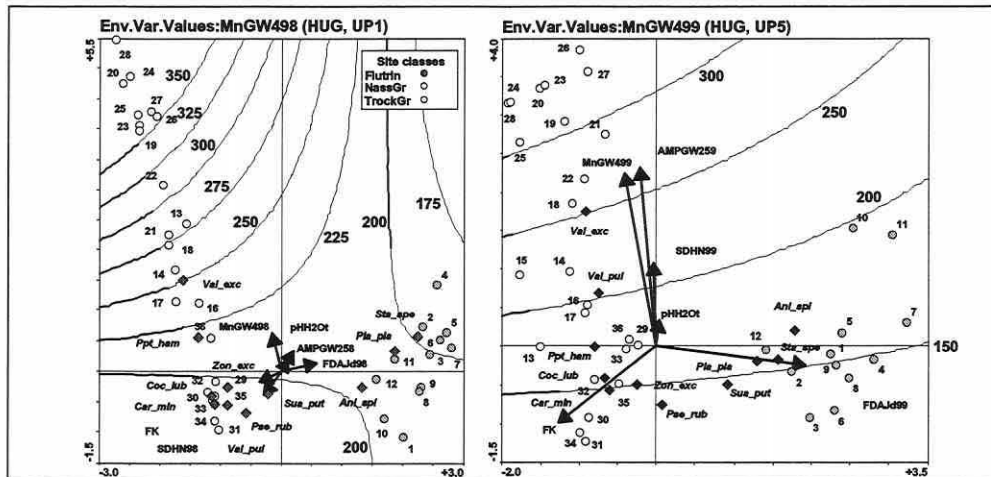


Abbildung 5.4-35: Isolinien des mittleren Grundwasserflurabstandes für die Monate April bis einschließlich September (in cm) für die Molluskenmodelle: HUG, UP1 und UP5. Probeflächen: weiß - trockenes Grünland, hellgrau - nasses Grünland und dunkelgrau – Flutrinnen.

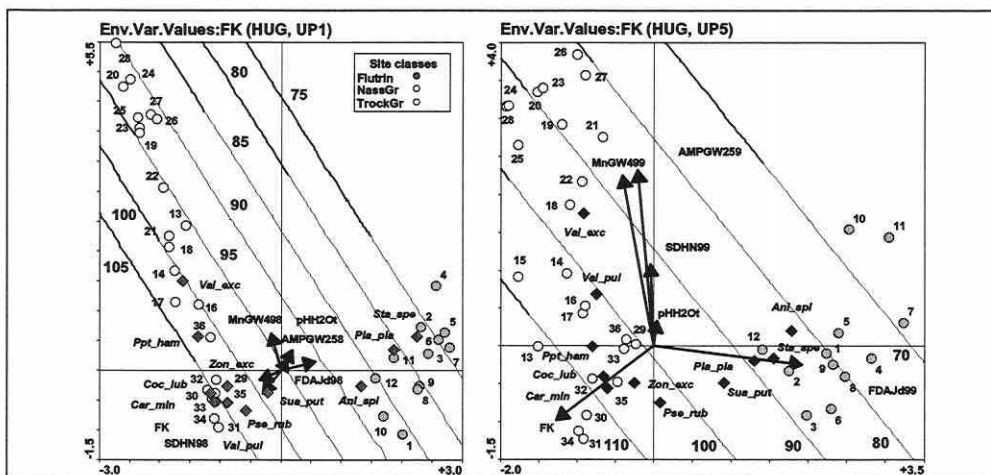


Abbildung 5.4-36: Isolinien der Feldkapazität (in mm) für die Molluskenmodelle: HUG, UP1 und UP5. Probeflächen: weiß - trockenes Grünland, hellgrau - nasses Grünland und dunkelgrau – Flutrinnen.

Beispielsweise findet sich das Optimum der Nische von *Planorbis planorbis* (Pla\_ple) in Abbildung 5.4-34 (S. 201) für UP1 und UP5 im Bereich der 150-Tage-Überflutungsdauer-isolinie und, um ein weiteres Beispiel für Abbildung 5.4-35 zu nennen, *Vallonia excentrica* (Val\_exc) auf der 250cm-Isolinie des mittleren Grundwasserflurabstandes für die Monate April bis einschließlich September. Vergleichbare Auswertemöglichkeiten bietet Abbildung 5.4-36 für die Feldkapazität.

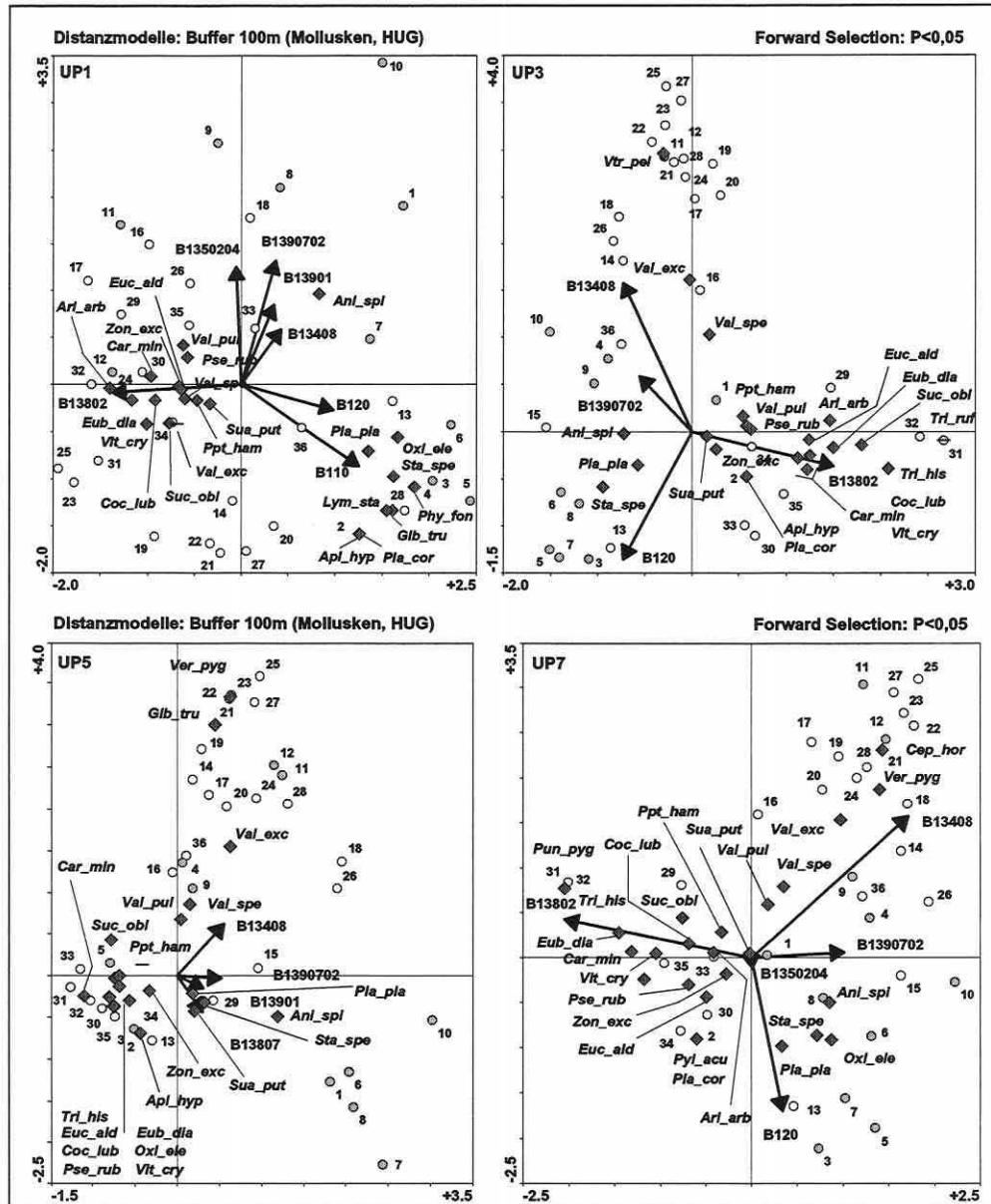
Eine mögliche Kombination der Abbildungen definiert dann die multidimensionale Nische der Arten. Die Darstellungsweise ermöglicht eine deskriptive Beschreibung der Optima und über die Toleranzinformation zu den einzelnen Arten (vgl. Abbildung 5.4-30 links, S. 198) ihrer approximierten Realnische. Das ARC/INFO®-Modul CANOGEN arbeitet nach diesem Schema und wird in Abschnitt 5.6 näher erläutert. Daher erübrigt sich die manuelle Entwicklung einer Tabelle mit den Optima und Nischen für alle Schirmarten der ökologischen Gruppen in dieser Untersuchung.

### Distanzparametermodelle:

Die Distanzparametermodelle eröffnen einen Ausblick, wie basierend auf der räumlichen Verteilung von Vegetationseinheiten, die implizit einen Lebensraum mit seinen abiotischen Bedingungen definieren, auch die Fauna des Raumes in ihrem Muster teilweise erklärt werden kann. Basierend auf den durch die ARC/INFO®-Applikation "DIST" berechneten Daten (vgl. Abschnitt 3.2) wird die Molluskenfauna in ihrer Struktur zu erklären versucht (Abbildung 5.4-37, S. 204). Die CCA-Ordinationsdiagramme wurden basierend auf den gleichen Grundeinstellungen wie die Modelle in den Abbildungen 5.4-7 bis 5.4-10 (S. 175f) erstellt. Die Modellparameter wurden jedoch nicht über die Untersuchungsphasen hinweg abgeglichen, sondern phasenspezifisch durch den Prozess der Vorwärtsselektion bestimmt. Alle Parameter sind auf dem Niveau  $\alpha=0,05$  (999 Permutationen) signifikant trennend. Die Erklärungsvariablen sind in ihren 1998er- und 1999er-Werten konstant, so dass Veränderungen auf das Verhalten der Arten zurückgeführt werden müssen. Um die Modelle interpretieren zu können, werden die Distanzparameter in ihrer Bedeutung aufgelistet:

- 10: vegetationsarme Sandfläche bzw. vegetationsarme Fläche mit bindigem Substrat
- 20: wechselfeuchte Rasenschmielen-Auenwiesen des *Deschampsion cespitosae* an großen Flüssen
- 30: Gebüsche und Feldgehölze nasser bis feuchter Standorte mit überwiegend autochthonen Arten
- 3408: artenarmes Intensivgrünland frischer Standorte
- 3802: Schilfröhricht
- 3805: Wasserschwadenröhricht
- 3806: Rohrglanzgrasröhricht
- 3807: sonstige Röhrichte
- 3901: krautige Ufersäume und -fluren an Gewässern
- 3906: Neophyten-Staudenflur
- 340403: ausdauernde Sandtrockenrasen mit geschlossener Narbe
- 350204: Flutrasen
- 370202: nährstoffreiches, rasiges Seggenried
- 390702: trocken-warme Ruderalstandorte auf bindigem Boden
- 430404: Weichholzauewälder
- 430405: Hartholzauewälder

Die so genannten 100m-Buffer (Kennzeichnung durch B1 am Anfang der Variablenamen in Abbildung 5.4-37, S. 204) beschreiben die Anteile der Biotoptypen innerhalb eines 100m-Radius um den Schwerpunkt einer Probeffläche. Unter Absprache mit den Fachwissenschaftlern wurde der Bufferradius für die Mollusken auf 100m und für die hier nicht dargestellten Carbiden auf 200m festgelegt (vgl. Abschnitt 5.3.2). Die Auswertung der Tabelle 5.4-25 (S. 205) zeigt eine generelle Verschlechterung der Modellgüte verglichen mit den Standardmodellen (abiotische Parameter). Die Erklärungsanteile an der biotischen Variabilität werden generell geringer. Befindet sich artenarmes Intensivgrünland frischer Standorte anteilig in den Buffern (B13408), so ist dies immer ein klarer Indikator für die ökologische Gruppe "höher gelegene, die meiste Zeit trocken liegende, genutzte Wiesenflächen", vertreten durch die Schirmart *Vallonia excentrica* (Val\_exc). Der Erklärungsparameter ist über alle Untersuchungsphasen stabil. In UP1 und UP3 ist Variable B13408 als zweite und in UP5 und UP7 als erste Variable ins Modell eingegangen. Diese Variable ist hochgradig kollinear mit den Variablen "Wasserschwadenröhricht" (B13805) und "Rohrglanzgrasröhricht" (B13806), welche negativ korreliert sind und somit das entgegengesetzte Ende der durch den Parameter B13408 gebildeten Achse darstellen. Der in den Untersuchungsphasen gegenläufig zu der letztgenannten Variable als erster oder zweiter in die Modelle aufgenommene Parameter "Schilfröhricht" (B13802; Ausnahme UP5: "sonstige Röhrichte", B13807) trennt immer die Gruppe "tieferliegende Feuchtfächen", vertreten durch die Schirmarten *Cochlicopa lubrica* (Coc\_lub), *Carychium minimum* (Car\_min) und *Zonitoides nitidus* (Zon\_exc), ab. Trotz der Auswahl des Parameters B13807 in UP5 durch die Vorwärtsselektion belegt der Austausch durch B13802 die Kongruenz der Aussage im Modellvergleich (Abbildung 5.4-38, S. 205).



**Abbildung 5.4-37: CCA-Distanzparametermodelle (100m-Buffer) für den Molluskendatensatz: UP1 bis UP7. Probestellen: weiß - trockenes Grünland, hellgrau - nasses Grünland und dunkelgrau – Flutrinnen.**

Die Gruppe "tiefere Rinnen, meist länger mit Wasser bestanden" zeigt eine Affinität zu der Einheit "wechselfeuchte Rasenschmielen-Auenwiesen des *Deschampsion cespitosae* an großen Flüssen" (B1120).

Es ist also möglich, anhand des Modells der Flora Rückschlüsse auf die Besiedlungsmuster der Mollusken zu ziehen. Dies bedeutet, dass die abiotischen Parameter, die das Floramodell aufbauen, in Verkettung mit den Distanzparametermodellen der Mollusken – ohne zusätzlichen Erhebungsaufwand – Prognosen unter einem gewissen Güteverlust durchführbar sind.





#### 5.4.4 Zeitliche und räumliche Übertragbarkeit der Ordinationsmodelle

Dieser Abschnitt geht der Frage nach, ob eine Modellübertragung sowohl zeitlich als auch räumlich möglich ist. Zuerst wird die zeitliche Übertragung (Untersuchungsphasen im Hauptuntersuchungsgebiet), danach die räumliche (Nebenuntersuchungsgebiete) angesprochen. Die zeitliche Übertragung erfolgt mehrstufig:

- (1) Bleibt das modelltechnisch approximierte Optimum der Nische für die Arten gleich über die vier Untersuchungsphasen?
- (2) Wo werden passiv in die Modelle der UP1 und UP5 eingeordnete Erhebungen der Biotik anderer Untersuchungsphasen abgebildet?
- (3) Bestehen Unterschiede in den Erklärungsanteilen zwischen den Optimalmodellen und den über die Untersuchungsphasen abgestimmten Modellen?
- (4) Bestehen Unterschiede in den Erklärungsanteilen zwischen dem Optimalmodell der UP1 bei dessen Übertragung auf die übrigen Untersuchungsphasen und den separat gebildeten Optimalmodellen der jeweiligen Untersuchungsphase?

##### Nischenoptima der abgestimmten Modelle:

Die Auswertung, ob Artenoptima über die Untersuchungsphasen einer Dynamik unterliegen, ist von Bedeutung, um zu beurteilen, wie stark die einzelnen Modelle von zeitlichen Veränderungen in den Umweltbedingungen betroffen sind. Ist die Nische über die Zeit auf eine Wertespanne der Erklärungsvariablen festgelegt, kann von einer saisonalen und annuellen Übertragbarkeit der Modelle ausgegangen werden. Die Darstellung der Untersuchungsergebnisse erfolgt anhand der Mollusken. Am Beispiel zweier wesentlicher Modellparameter wird deutlich, dass die Optima für die meisten Arten eine saisonale und über die Untersuchungsjahre auftretende Verschiebung aufweisen (Abbildung 5.4-39 und 5.4-41, S. 207). Sie sind für 1999 gegenüber 1998 häufig ins Positive (feuchtere Bedingungen bezogen auf die Originalwerte) verschoben.

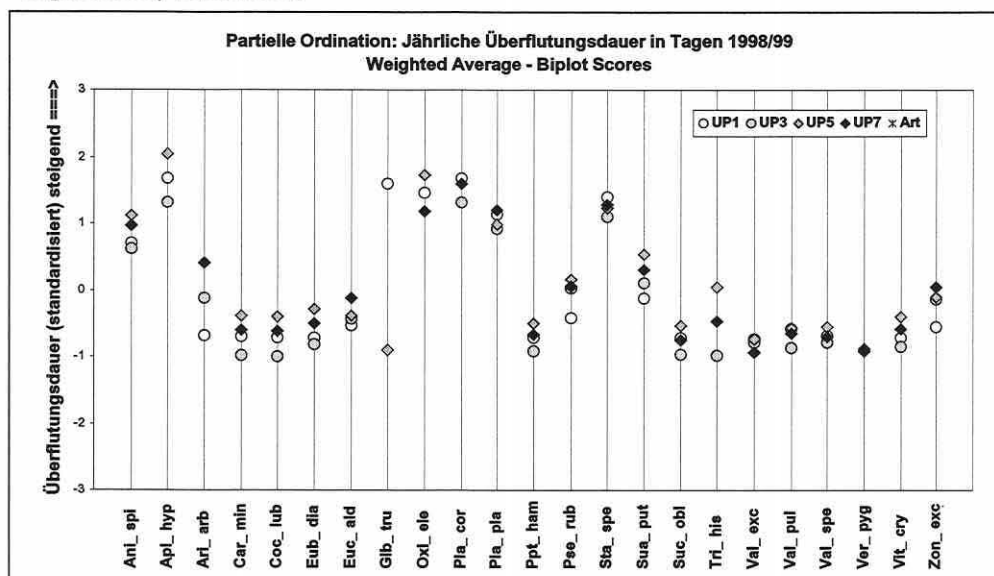


Abbildung 5.4-39: Berechnete Optima der realisierten univariaten Nische für die Molluskenarten über die vier Untersuchungsphasen für den Modellparameter "jährliche Überflutungsdauer in Tagen" (FDAJd98/99). Das Nischenoptimum der Arten ist dem File "Spec\_Env.tab" entnommen (unimodale Methoden, vgl. Abschnitt 3.6.1). Die Werte entsprechen den Positionskordinaten der Arten für die univariaten Analysen der Modellparameter.

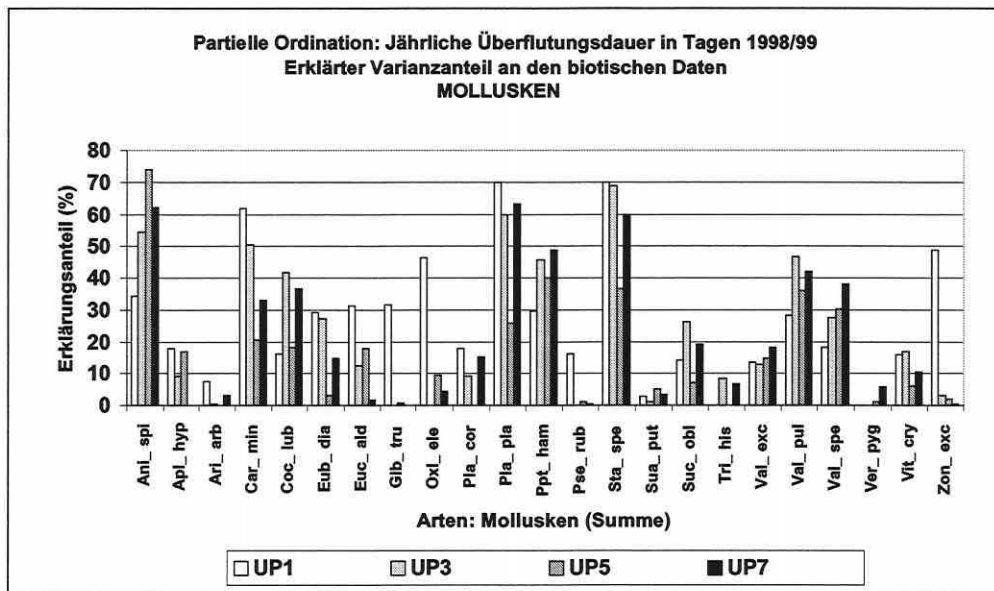


Abbildung 5.4-40: Univariate Erklärungsanteile des Modellparameters "jährlichen Überflutungsdauer in Tagen" (FDAJd98/99) an der Gesamtvariabilität der einzelnen Molluskenarten in den vier Untersuchungsphasen (UP1 bis UP7).

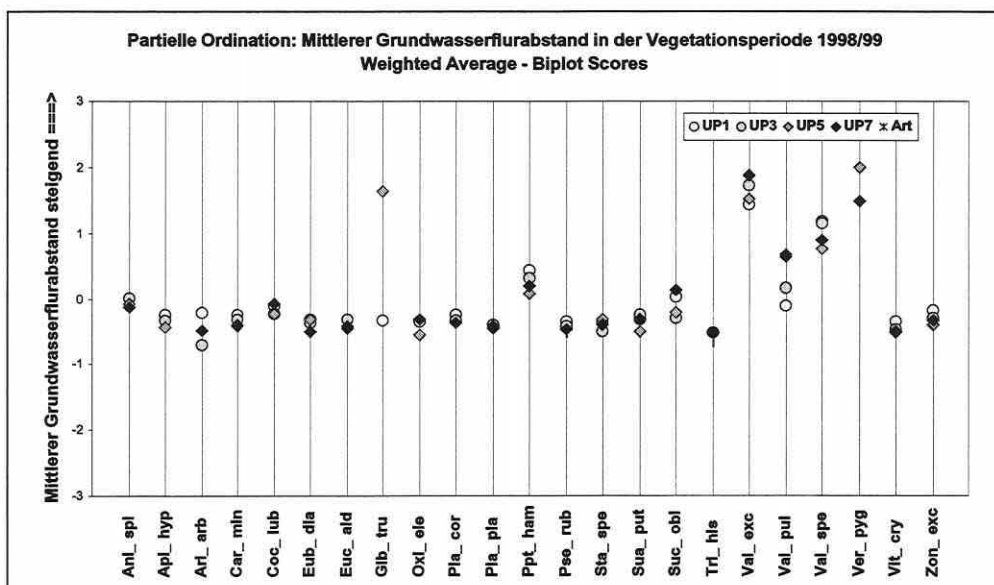
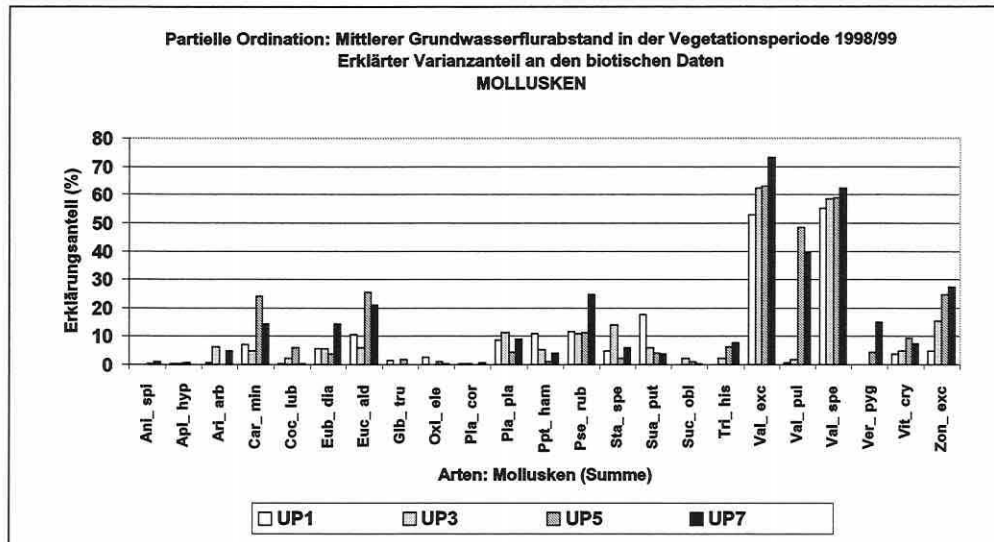


Abbildung 5.4-41: Berechnete Optima der realisierten univariaten Nische für die Molluskenarten über die vier Untersuchungsphasen für den Modellparameter "mittlerer Grundwasserflurabstand über die Monate April bis einschließlich September (MnGW498/499). Das Nischenoptimum der Arten ist dem File "Spec\_Env.tab" entnommen (unimodale Methoden, vgl. Abschnitt 3.6.1). Die Werte entsprechen den Positionskoordinaten der Arten für die univariaten Analysen der Modellparameter.



**Abbildung 5.4-42: Univariate Erklärungsanteile des Modellparameters "mittlerer Grundwasserflurabstand über die Monate April bis einschließlich September (MnGW498/499) an der Gesamtvariabilität der einzelnen Molluskenarten in den vier Untersuchungsphasen (UP1 bis UP 7).**

Diese Tendenz und die veränderlichen Erklärungsanteile der Parameter in den Untersuchungsphasen (vgl. Abbildung 5.4-40 und 5.4-42) bewirken, dass die Modelle für weitgehend ortsfeste Artengruppen zeitlich abhängig sind. Ein Beispiel kann die Aussage verdeutlichen helfen: Betrachtet man die Schirmart *Anisus spirorbis* (Ani\_spi), welche die erste Art von links in den Abbildungen ist, erkennt man den generell hohen Erklärungsanteil der Überflutungsdauer in Tagen (FDAJd; Abbildung 5.4-40, S. 207) und den sehr geringen des mittleren Grundwasserflurabstandes über die Monate April bis einschließlich September (MnGW498/499; Abbildung 5.4-42). Eine zuverlässige Aussage des MnGW ist bezogen auf das Optimum der Art daher nicht zu erwarten. Für die FDAJd zeigt sich die Verschiebung der Erklärungsanteile und der Optima speziell über die Untersuchungsjahre. Das stärker durch Überflutung geprägte Jahr 1999 besitzt höhere Erklärungsanteile an der Variabilität der Art und belegt eine Verlagerung des Existenzoptimums in intensiver überflutete Bereiche. Da beide angesprochenen Standortparameter über die Frühjahrs- und Herbstuntersuchungsphase integrieren, wird der Aspekt saisonal nicht so deutlich.

Weitere Auswertungen zeigen, dass beispielsweise die maximalen Grundwasserflurabstände in den Drei-Monatsphasen Februar bis Mai des logistischen Regressionsmodells für die Art *Arrhenatherum elatius* gleiches auch saisonal zum Ausdruck bringen (vgl. Abbildung 5.5-4, S. 223). Während die Art standorttreu über das Jahr bleibt, ist der Wert des Parameters bezogen auf eine bestimmte Vorkommenswahrscheinlichkeit im Herbst relativ niedriger als im Frühjahr, d.h. die Nische wird korrekt berechnet, ist aber nur auf die Jahreszeit bezogen gültig. Die Verschiebung der Optima ist stark parameterabhängig. Für die Carabiden gilt, dass durch die Wanderbewegungen der Effekt ausgeglichen wird, aber dafür eine Ortskonstanz nicht gegeben ist.

An der Abbildung 5.4-39 (S. 206) sowie der Abbildung 5.4-41 (S. 207) lässt sich weiterhin ablesen, welchen Einfluss Modellparameter auf den Gesamtdatensatz haben. Je stärker die Arten in ihren Optima zueinander streuen, desto besser kann der Parameter den Artenpool trennen. Der Verlust an Trennkraft des MnGW verglichen zu der FDAJd wird durch das ausgeglichene Muster der Verteilung klar belegt. Der erstgenannte Parameter trennt nur die Arten trockenerer Lebensräume, wie beispielsweise *Vallonia excentrica* (Val\_exc), gut vom restlichen Datensatz.

### Passive zeitliche Modellübertragung

Die Vorgehensweise in diesem Test ist, in bestehende Ordinationsmodelle die biotischen Daten anderer Untersuchungsphasen passiv einzuordnen, um die mit den Untersuchungsphasennummern codierten Arten in ihren gewichteten Schwerpunkten zueinander vergleichen zu können. Prinzipiell entspricht dies der univariaten Betrachtung aus Abbildung 5.4-39 bzw. 5.4-41 (S. 206f); Abbildung 5.4-43 zeigt aber die Verschiebung der Optima für den multivariaten Fall auf. Die Aussagen für die univariaten Betrachtungen bestätigen sich für diese Ebene: Die Optima streuen für die meisten Arten sowohl jahreszeitlich als auch über die Untersuchungsjahre hinweg bezogen auf die Modellparameter bestimmten Achsen. Dennoch bleibt das allgemeine Verhalten der Arten zu den Modellparametern ähnlich. Bei der Anwendung einer Szenariotechnik im GIS-Prognosemodell ist daher die Übertragbarkeit der Modelle innerhalb der vom Modell erfassten Wertespanssen realistisch.

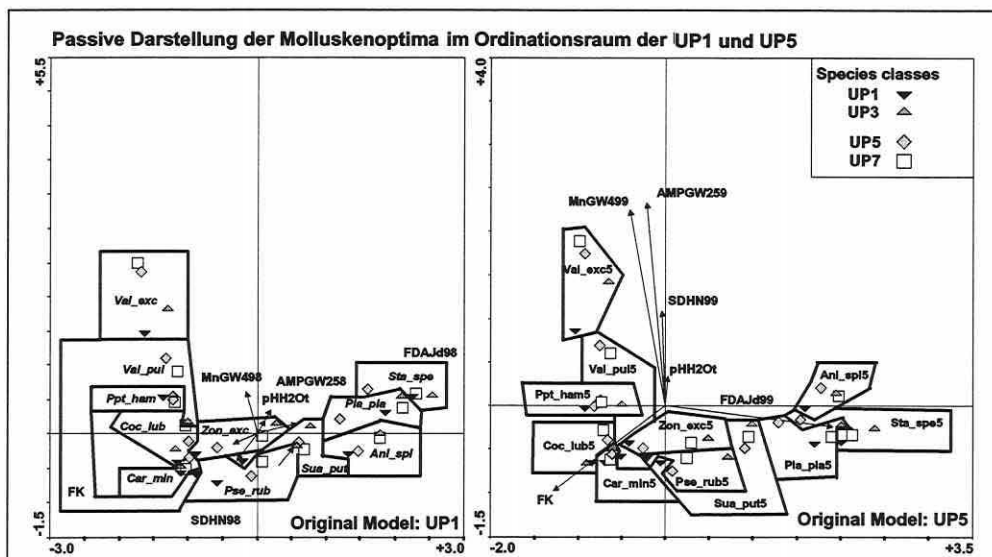


Abbildung 5.4-43: Zeitliche Modellübertragung der Molluskenmodelle (UP1 und UP5) für den multivariaten Fall. Die Optima der anderen Untersuchungsphasen wurden passiv in die Originalmodelle eingelesen.

### Optimalmodelle und zeitlich abgestimmte Modelle

Unter den Optimalmodellen werden die Ordinationen verstanden, deren Modellparameter nicht über die Untersuchungsphasen abgeglichen wurden (vgl. Abschnitt 5.4.1). Ihr Modellparametersatz ist nicht kongruent, sondern so gewählt, dass er die maximale Erklärung (Trennung) des biotischen Datensatzes ermöglicht.

Tabelle 5.4-26: Vergleich der gerundeten Eigenwerte (EV) und prozentualen Erklärungsanteile (%): Optimalmodelle (VSM) und angepasste Modelle (AM) der Mollusken über die Untersuchungsphasen (UP).

		Mollusken							
		UP1		UP3		UP5		UP7	
		VSM	AM	VSM	AM	VSM	AM	VSM	AM
EV	1. Achse	0.65	0.65	0.59	0.59	0.52	0.53	0.60	0.60
	2. Achse	0.29	0.29	0.30	0.29	0.32	0.31	0.34	0.33
	3. Achse	0.26	0.22	0.12	0.13	0.13	0.10	0.33	0.13
%	1. Achse	35.9	36.1	36.5	36.3	31.4	31.8	31.4	31.5
	2. Achse	16.1	15.9	18.2	17.8	19.2	18.8	17.9	17.2
	3. Achse	14.1	12.0	7.7	8.1	7.6	6.1	6.2	6.8
Σ% über 4 Achsen		69.0	65.4	65.3	64.4	63.0	59.5	59.2	58.8

Tabelle 5.4-26 (S. 209) belegt nur geringfügige Einbußen in der Modellgüte bei der Abstimmung auf einen zeitlich robusten Parametersatz zur Erklärung des biotischen Musters. Die modelltechnischen Vorteile (Variablenreduktion/Vergleichbarkeit/GIS-Umsetzung/Übertragbarkeit) der die vier Untersuchungsphasen abgestimmten Modelle wiegen die Erklärungsverluste auf. Als weiteres Beispiel dienen die Floramodelle, deren Auswertung die obigen Ergebnisse untermauern (Tabelle 5.4-27).

**Tabelle 5.4-27: Vergleich der gerundeten Eigenwerte (EV) und der prozentualen Erklärungsanteile (%): Optimalmodelle (VSM) und angepasste Modelle (AM) der Flora über die Untersuchungsphasen (UP).**

	Flora	UP1		UP3		UP5		UP7	
		VSM	AM	VSM	AM	VSM	AM	VSM	AM
EV	1. Achse	0.86	0.86	0.85	0.85	0.82	0.82	0.85	0.84
	2. Achse	0.39	0.37	0.37	0.35	0.36	0.36	0.36	0.35
	3. Achse	0.24	0.23	0.25	0.22	0.27	0.26	0.24	0.22
%	1. Achse	22.2	22.0	22.1	22.0	21.0	21.0	22.8	22.8
	2. Achse	9.9	9.4	9.6	9.2	9.3	9.1	9.8	9.4
	3. Achse	6.2	6.0	6.6	5.7	7.0	6.7	6.6	6.0
	$\Sigma\%$ über 4 Achsen	43,2	42,0	43,5	41,8	42,2	41,7	44,1	43,2

### Zeitliche Modellübertragung des Optimalmodells "Mollusken" aus UP1

Die Kombination der Erklärungsvariablen aus dem Mollusken-Optimalmodell der UP1 wurde ohne Plausibilitätsprüfung auf die übrigen Untersuchungsphasen übertragen. Dies dient dem Vergleich der resultierenden Modellergebnisse mit den separat gebildeten Optimalmodellen der jeweiligen Untersuchungsphase (Tabelle 5.4-28).

**Tabelle 5.4-28: Ergebnisvergleich: Übertragung des Optimalmodells Mollusken der UP1 (VS-Modell bzw. UP1) und separat aufgestellte Optimalmodelle der UP3 bis UP7 (dunkelgrau).**

**** Summary **** VS-Modell HUG UP1						
Axes		1	2	3	4	Total inertia
Eigenvalues	:	0,654	0,294	0,257	0,052	1,821
Sum of all unconstrained EV						1,821
Sum of all canonical EV						1,286

<b>**** Summary ****</b>	<b>UP3</b>	<b>UP1</b>	<b>UP3</b>	<b>UP1</b>	<b>UP3</b>	<b>UP1</b>	<b>UP3</b>	<b>UP1</b>	<b>UP3</b>	<b>UP1</b>
<b>Axes</b>	<b>1</b>		<b>2</b>		<b>3</b>		<b>4</b>		<b>Total inertia</b>	
<b>Eigenvalues :</b>	<b>0,59</b>	<b>0,583</b>	<b>0,296</b>	<b>0,298</b>	<b>0,124</b>	<b>0,111</b>	<b>0,047</b>	<b>0,049</b>	<b>1,618</b>	<b>1,618</b>
<b>Sum of all unconstrained EV</b>									<b>1,618</b>	<b>1,618</b>
<b>Sum of all canonical EV</b>									<b>1,073</b>	<b>1,061</b>

[illegible]

**** Summary ****		UP7	UP1	UP7	UP1	UP7	UP1	UP7	UP1	UP7	UP1
Axes		1		2		3		4		Total	Inertia
Eigenvalues	:	0,6	0,592	0,343	0,327	0,118	0,109	0,072	0,068	1,914	1,914
Sum of all unconstrained EV										1,914	1,914
Sum of all canonical EV										1,183	1,173

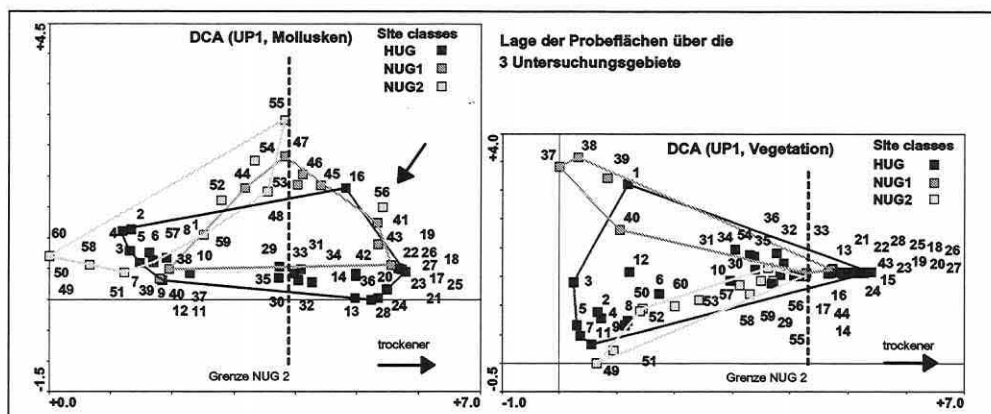


Das Optimalmodell der UP1 umfasst nachfolgende Erklärungsvariablen: "jährliche Überflutungsdauer in Tagen 1998", "maximaler Grundwasserflurabstand zwischen dem 01.02.-31.05.98", "Amplitude der Grundwasserstände zwischen dem 01.02.-31.05.98", "Wasserstoffkonzentration ( $H_2O$ ; multipliziert mit 1.000.000)", "Standardabweichung der Wasserstände in HN (Überflutung und Grundwasserstände zwischen dem 01.01.-31.12.98)", "maximale Überflutungshöhe 1998" und "Grundwasserstände niedriger als Deckschicht unter Flur (längste Periode in Tagen 1998)".

In Zusammenfassung der Tabelle 5.4-28 (S. 210) erkennt man, dass ein genereller Verlust in den Eigenwerten und der erklärten Gesamtvarianz zu verzeichnen ist. Setzt man diese Ergebnisse in Bezug zu denen der angepassten Modelle (Tabelle 5.4-26, S. 209), erkennt man deren Mittelstellung speziell bezogen auf die Eigenwerte der ersten bzw. bedeutendsten Ordinationsachse.

### Räumliche Modellübertragung

Die Ähnlichkeit der Untersuchungsgebiete wird qualitativ verdeutlicht, indem alle Probeflächen gleichzeitig aktiv ordiniert werden. Eine DCA zeigt die Situation in UP1 anhand der biotischen Daten für die Mollusken und die Flora (Abbildung 5.4-44). Es fällt auf, dass NUG2 (Dornwerder bei Sandau) von den übrigen Untersuchungsgebieten besonders im Molluskenmodell deutlich abweicht. In der Kenntnis, dass die Probeflächen 13 bis 28 des HUG die trockensten Standortbedingungen aufweisen, kann geschlussfolgert werden, dass das NUG2 innerhalb seiner Grenzen insgesamt feuchtere Standortverhältnisse beherbergt und vergleichbar trockene Bereiche des HUG und NUG1 vermissen lässt. Die einzige Probefläche des NUG2 (Mollusken), welche die in Abbildung 5.4-44 (links) eingezeichnete Grenzlinie überspringt, ist Nummer 56 (vgl. auch Abbildung 5.4-47, S. 213). Beim Floramodell gibt es keinen Ausreißer und es zeigt sich insgesamt ausgeglichener. In Folge der unterschiedlichen Feuchteverhältnisse ist die räumliche Übertragbarkeit der HUG-Modelle im nassen Bereich auf NUG2 gerade für das Molluskenbeispiel klar eingeschränkt, da die Wertespanne der Modellparameter überschritten wird und der Artenpool, je nach Artengruppe, entsprechend differiert.



**Abbildung 5.4-44:** Ähnlichkeit der Untersuchungsgebiete anhand einer DCA mit aktiver Probeflächenordination: Klassifizierte Probeflächenanordnung für die Gruppe der Mollusken und der Flora in UP1.

Die passive Einordnung der Probeflächen in die bestehenden CCA-Ordinationsdiagramme der Frühjahrsphasen (UP1 und UP5) im HUG bildet eine weitere Informationsquelle zur Einschätzung räumlicher Übertragbarkeit (Abbildungen 5.4-45 und 5.4-46, S. 212). In den Floramodellen wird speziell der sehr trockene Sektor im NUG2 nahezu nicht durch Probeflächen abgedeckt (Abbildung 5.4-44, Trennlinie Teilbild rechts). Die Lage und Gruppierung der Probeflächen lässt dennoch eine Übertragbarkeit der multivariaten Modelle mit unterschiedlichen

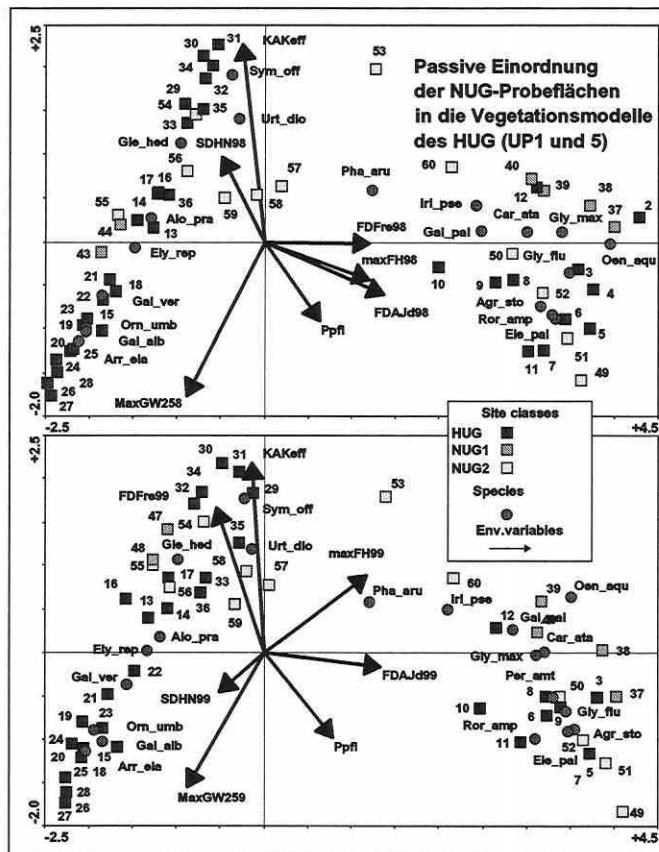


Abbildung 5.4-45: Räumliche Modellübertragung: CCA der Flora im HUG (UP1 oben und UP5 unten) mit passiven Probeflächen der NUGs.

Güteverlusten möglich erscheinen, da die biotische Ausstattung der Untersuchungsgebiete nicht zu stark voneinander abweicht und die meisten Arten, die als Stellvertreter für die ökologischen Gruppen selektiert wurden, über die Untersuchungsgebiete hinweg vertreten sind.

Die entsprechenden Darstellungen für die Frühjahrsuntersuchungsphasen Molluskenmodelle des HUG eröffnen Verzerrungen in der Anordnung der Probeflächen aus dem NUG2, während die Flächen aus NUG1 insgesamt positivere Ergebnisse zeigen (Abbildung 5.4-46). Die Ursache der Abweichungen wird anhand des Beispiels "Probefläche 50" des NUG2 erläutert. In der Stratifizierung zählen die Probeflächen 49 bis 52 zum Stratum "Flutrinne" (siehe Abbildung 5.4-47, S. 213, Teilbild rechts). In der Abbildung 5.4-46 (rechtes Teilbild) sollten diese Flächen

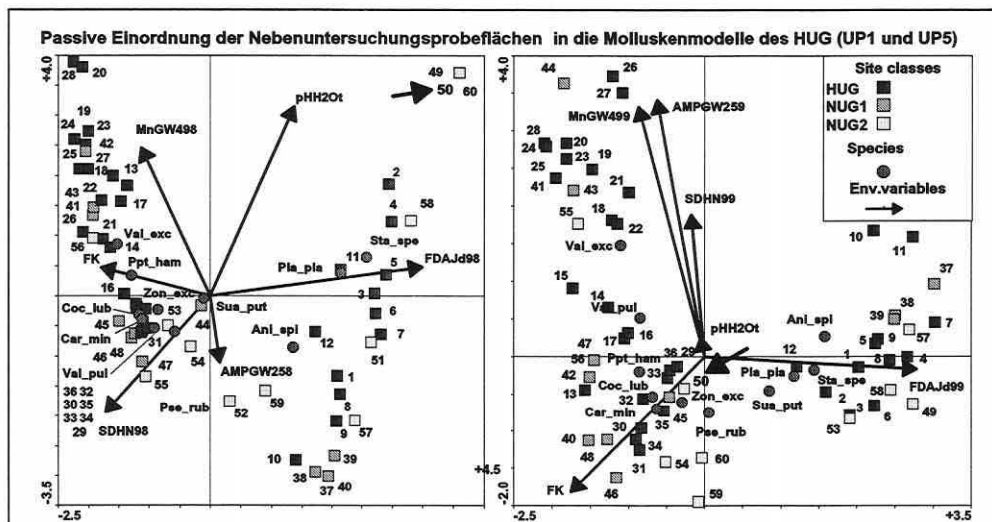
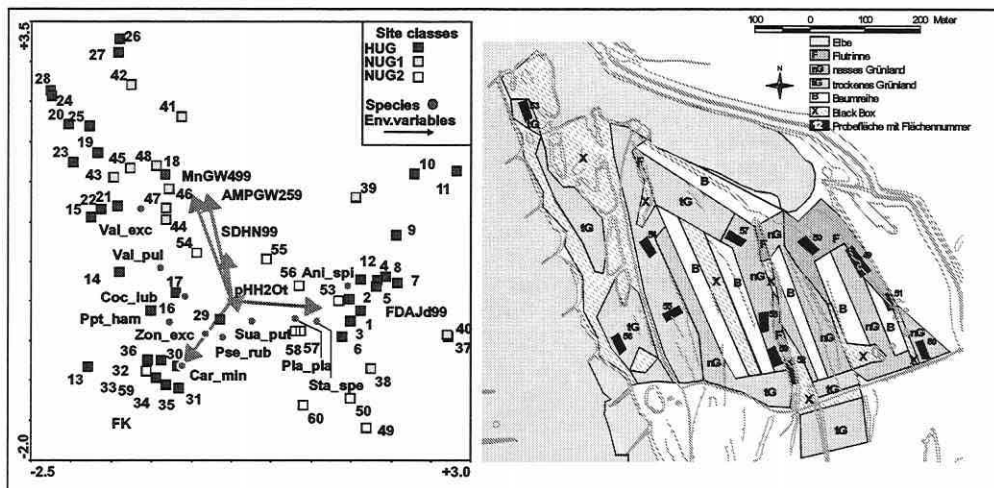


Abbildung 5.4-46: Räumliche Modellübertragung dargestellt anhand verschiedener CCA-Modelle der Mollusken im HUG (UP1 links und UP5 rechts) mit passiven Probeflächen der NUGs. Das Beispiel "Probefläche 50" ist durch Pfeile markiert (Erläuterungen im Text).

nahe der Spitze des Vektors für die Erklärungsvariable Überflutungsdauer (FDAJd99) erscheinen, da die Originaldaten zwischen 230 und 271 Tage Wasserbedeckung ausweisen. Fläche 49 bestätigt die Erwartungshaltung, während Fläche 50 weit abgeschlagen nahe dem Diagrammursprung in UP5 erscheint. Die Erklärung für dieses Phänomen geht auf die Mollusken-Fangergebnisse zurück. Die Passivsetzung der Probeflächen bewirkt, dass ihre Artenkomposition überwiegend die Lage im Ordinationsdiagramm bestimmt und nicht die Information der Abiotik. Fläche 50 weist die Arten *Stagnicola spec.* (Sta\_spe), *Zonitoides nitidus* (Zon\_exc), *Perpolita hammonis* (Per\_ham) und *Vallonia pulchella* (Val\_pul) mit jeweils einem Individuum auf. Diese Artenkomposition, welche hinsichtlich der Häufigkeiten nicht als verlässlich angesehen werden darf, bewirkt die Verzerrung.



**Abbildung 5.4-47: Räumliche Modellübertragung. CCA der Mollusken im HUG (UP5 links) mit aktiven Probeflächen der Nebenuntersuchungsgebiete und Stichprobenplan NUG2 (rechts).**

Werden die Probeflächen der Nebenuntersuchungsgebiete aktiv gesetzt (Abbildung 5.4-47, linkes Teilbild), erfolgt die "korrekte" Positionierung der Fläche 50 entlang des Überflutungsgradienten. Die Positionskoordinaten der Arten werden nur geringfügig beeinflusst, da sie gewichtete Schwerpunkte der Arten über die Probeflächen darstellen. Ein Individuum auf Fläche 50 ist daher relativ unbedeutend im Gesamtdatensatz. Der Argumentationsweg kann aber nicht zur Begutachtung einer Modellübertragbarkeit genutzt werden, da das Originalmodell durch die aktive Implementierung der NUG-Probeflächen verändert wird.

Die geringe Stichprobenzahl von 60 im Projekt kann in den Nebenuntersuchungsgebieten mit jeweils nur zwölf Probeflächen nicht die eigentlich gewünschte Ergebnissicherheit erbringen. Die Berechnungen sowie die Abbildungen 5.4-48 und 5.4-49 (S. 214) mit den Ordinationsmodellen für die Nebenuntersuchungsgebiete erhärten diese Hypothese eines grenzwertigen Stichprobenumfangs. Eine hohe Anfälligkeit gegen Ausreißer und damit Zufallspositionierungen von Arten in den Ordinationsdiagrammen sind die Folge. Nochmals betont werden muss dabei, dass die Modellübertragung in den Abbildungen 5.4-48 und 5.4-49 sich nur auf den Parametersatz bezieht, nicht aber auf eine Berechnung anhand der Modellgleichungen aus dem HUG.

Die konsequente Übertragung der vollständigen Erklärungsvariablenanzahl aus den HUG-Modellen lässt die NUG-Modelle zudem statistisch gesehen unzuverlässig werden. Die zu hohe Anzahl von Erklärungsvariablen verglichen mit dem Stichprobenumfang (Ausfälle von Probeflächen führten zu einer weiteren Verminderung) bewirkt instabile Regressionskoeffizienten der Modelle. Die Auswertungen sind letztlich nur graphisch-tendenziell zu verantworten.

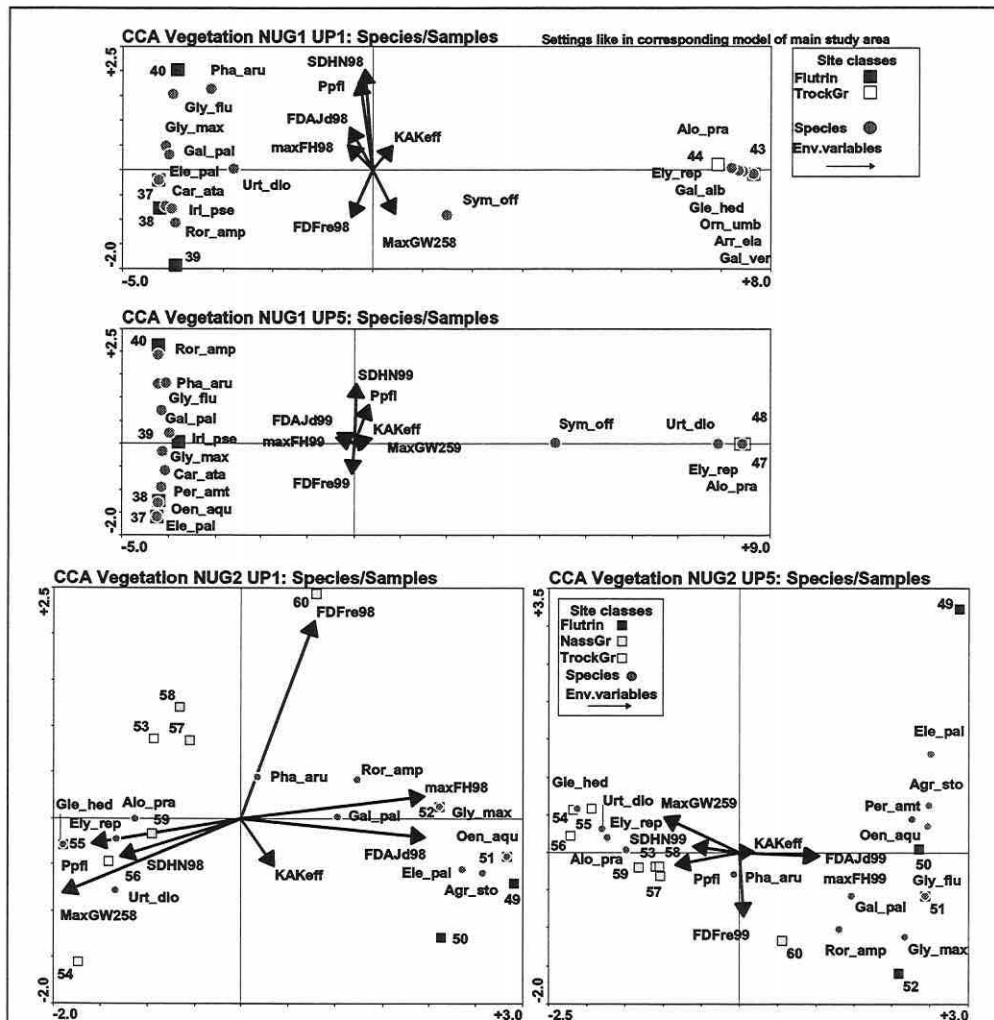


Abbildung 5.4-48: Räumliche Übertragung der HUG-Modellparameter auf die Nebenuntersuchungsgebiete am Beispiel der Flora (UP1 und UP5; NUG1 oben und NUG2 unten).

Die direkten Ordinationen zur Beschreibung des floristischen Raummusters (Abbildung 5.4-48) für die ausgewählten Arten sind für das NUG2 befriedigend. Es entsteht ein etwaiges Abbild der Ergebnisse aus dem HUG. Die Resultate für NUG1 sind durch das Ausfallen aller Probestellen des nassen Grünlandes und zweier weiterer Flächen des trockenen Grünlandes (nach Klassifikation aus der Stratifizierung) lückenhaft und wenig differenziert.

Die Molluskenmodelle (Abbildung 5.4-49, S. 215) sind tendenziell akzeptabel, aber durch die geringen Fangzahlen teils inkonsistent gegenüber den Aussagen aus dem HUG. So finden sich *Vallonia pulchella* (Val\_pul) und *Perpolita hammonis* (Ppt\_ham) als Schirmarten der gleichen ökologischen Gruppe (vgl. Tabelle 5.4-24, S. 199) oft weit voneinander entfernt in den Ordinationsdiagrammen wieder. Das zur UP5 im NUG2 korrespondierende Diagramm (Abbildung 5.4-49, S. 215, Teilbild rechts unten) zeigt das extremste Beispiel hierfür: *Perpolita hammonis* tritt mit einem und vier Individuen auf Fläche 50 bzw. 60 auf, *Vallonia pulchella* mit je einem Individuum auf Fläche 50 und 54 bis 56.

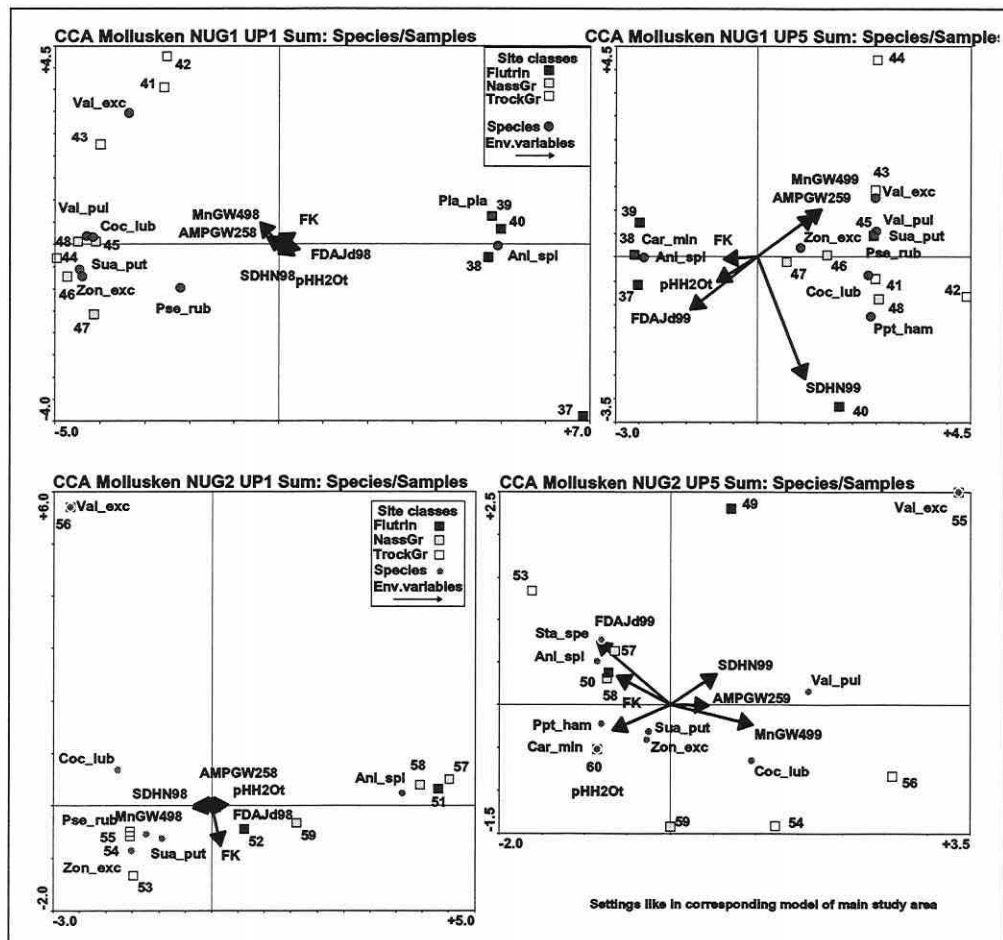


Abbildung 5.4-49: Räumliche Übertragung der Modellparameter des Hauptuntersuchungsgebietes auf die Nebenuntersuchungsgebiete (Mollusken: UP1 und UP5; NUG1 oben und NUG2 unten).

Weiterhin erwähnenswert ist Probefläche 53 in NUG2 (Abbildung 5.4-49, unten), da sie die unterschiedlichen Überflutungsverhältnisse 1998 und 1999 verdeutlicht. Obwohl nach der Stratifizierung als trockene Grünlandfläche eingeordnet, liegt sie auf einer relativ niedrigen Geländehöhe, d.h., die unterschiedliche Überflutungsdauer durch die Elbhochwässer wird durch ihre drastische Lageverschiebung in den Diagrammen direkt sichtbar. Ihre Artenkomposition bringt sie ebenfalls 1998 (*Succinea putris*: Sua\_put = ein Individuum und *Zonitoides nitidus*: Zon\_exc = ein Individuum) und 1999 (*Stagnicola spec.*: Sta\_spe = vier Individuen und andere Arten) in unterschiedliche Positionen. Die Fläche liegt dann letztlich extremer gegenüber der Überflutungsdauer-Achse als Fläche 50. Man erinnere sich, dass der Fokus der Analysen auf der möglichst korrekten Approximation der Artsignaturen im ökologischen Raum liegt (vgl. Abschnitt 3.5.1) und die Probeflächensignaturen um diese gestreut werden. Sie sind daher nicht exakt positioniert. Um dies zu erreichen, muss der Fokus "Intersample-Distances" gewählt werden.

## 5.5 Generalisierte lineare Modelle: Logistische Regression

Im Rahmen der vorliegenden Arbeit wurde die Möglichkeit einer Verknüpfung der Mustererkennung und Modellierung mit gradientenanalytischen Verfahren und der Modellierung von



Lebensraumeignungsmodellen mit logistischer Regression bereits angesprochen (vgl. Abschnitt 4.2). Es erscheint attraktiv, artspezifische Modelle zu erstellen, welche die Vorkommenswahrscheinlichkeit der Schirmarten eventuell exakter als die Ordinationsmodelle abzubilden vermögen. Die Modellanpassung wird bei der logistischen Regression nicht global (linear oder unimodal) vorgenommen, sondern kann auf die jeweilige Art genau abgestimmt werden. Außerdem ist eine quantitative Möglichkeit gegeben, die räumliche und zeitliche Übertragbarkeit der Modelle auch ohne Übertragung ins GIS zu testen. Gegenstand des Abschnittes ist es, Ergebnisse und Erkenntnisse aus der Generierung und Validierung von Lebensraumeignungsmodellen beispielhaft für die Arten *Anisus spirorbis* und *Arrhenatherum elatius* (Schirmarten der Gruppen 1 und 6; vgl. Tabelle 5.4-24, S. 199) darzulegen und mit den Ordinationsergebnissen zu vergleichen.

Ausgehend von einem Klassifikationsschwellenwert ( $P_{krit}$ ) mit dem Wert 0,5 wird eine erste Auswahl basierend auf den Modellvariablen der Ordinationsmodelle durch den vorwärtsschrittweisen Likelihood Ratio-Test vorgenommen. Die Wahrscheinlichkeit für den Einschluss wird auf 0,05 und für den Ausschluss auf 0,1 festgelegt. Über Kreuztabellen wird dann ein annähernd fehlergerechter Klassifikationsschwellenwert (Anteil der Sensitivität  $\approx$  Anteil der Spezifität) ermittelt (vgl. Abschnitt 3.7.1). Zur Modelldiagnostik werden folgende Maße herangezogen:

- (1) Bestimmtheitsmaß nach Nagelkerke ( $R^2_N$ )
- (2) Likelihood Ratio-Test zur Modellsignifikanz
- (3) Odd-Ratio (EXP B) und zugehöriges 95%-Konfidenzintervall
- (4) Wald-Statistik: Signifikanz der einzelnen Modellparameter
- (5) Prävalenz/Klassifikationsgüte des Nullmodells
- (6) Prozentsatz korrekter Prognosen
- (7) Spezifität
- (8) Sensitivität
- (9) Fläche (AUC) zwischen der ROC-Kurve und der Diagonale des Zufallsmodells mit 95%-Konfidenzintervall und Signifikanzwert

Ausgehend vom Nullmodell werden die univariaten sigmoidalen oder unimodalen Modelle tabellarisch aufgelistet (vgl. z.B. Tabelle 5.5-2, S. 217). Die Form der Antwortkurve richtet sich dabei danach, ob die Erklärungsvariablen linear oder zusätzlich quadratisch ins Modell eingehen. Lineare Terme bestimmen einen sigmoidalen Kurvenverlauf, während quadratische zu unimodalen Antwortkurven führen, welche Optimum und Toleranz (vgl. Abschnitt 2.3) der Art zum Gradienten beschreiben. Die Signifikanz der einzelnen Modellparameter (Wald-Statistik) belegt dann, dass deren Regressionskoeffizient von null verschieden ist und der Parameter relevant zur Erklärung des Vorkommens beiträgt. Kann eine Modellverbesserung durch die Hinzunahme weiterer Erklärungsparameter erreicht werden, führt dies zu bivariaten Antwortoberflächen in der Darstellung.

Eine andere Form "multivariater" Modelle wird durch die Generierung von Modellen basierend auf den kanonischen Variablen aus den Ordinationsverfahren, welche bereits eine Linearkombination mehrerer Parameter darstellen, getestet.

#### **Schirmart: *Anisus spirorbis***

Ausgehend von den sechs Modellparametern des CCA-Molluskenmodells und ihrer quadratischen Terme wird die Trennkraft der Einzelparameter gegenüber dem Nullmodell ausgegeben (Tabelle 5.5-1, S. 217). Es wird anschaulich, dass der mittlere Grundwasserflurabstand über die Monate April bis einschließlich September 1998 (MNGW498) für *Anisus spirorbis* die höchste Trennkraft aufweist. Verglichen mit dem multivariaten Modell der CCA wechselt die Rangfolge des Parameters vor die der Überflutungsdauer (vgl. Abbildung 5.4-4, S. 169). Die Einzelbetrachtung der Art belegt deren höhere Affinität zum mittleren Grundwasserflurabstand, während die Überflutungsdauer eine bedeutendere Stellung bei der Sammelbetrachtung des Vorkommens aller Molluskenarten im Untersuchungsgebiet einnimmt.

**Tabelle 5.5-1: Univariate Trennkraft (Wert) und Signifikanz (Sig.) der Modellparameter (basierend auf den Ordinationsmodellen) gegenüber dem Nullmodell in UP1: *Anisus spirorbis*.**

		Wert	Sig.
Feldkapazität	FK	,442	,506
pH-Wert H <sub>2</sub> O, transformiert (s.u.)	PHH2OT	,925	,336
Standardabweichung Grundwasserstand + Überflutung	SDHN98	,053	,818
Mittlerer Grundwasserflurabstand (Monate 4-9)	MNGW498	23,558	,000
Amplitude Grundwasserflurabstand (Monate 2-5 bzw. 6-9)	AMPGW258	3,610	,057
Überflutungsdauer in Tagen (1998)	FDAJD98	13,728	,000
Quadratische Terme der Parameter	FK2	,349	,555
	PHH2OT2	,002	,962
	SDHN982	,027	,869
	MNGW4982	22,175	,000
	APGW2582	3,268	,071
	FDAJD982	12,512	,000

Anmerkung: pH-Wert H<sub>2</sub>O transformiert bedeutet: pH-Wert delogarithmisiert (Wasserstoffionenkonzentration); Transformation: Multiplikation mit 1.000.000.

Die Modellgüte der Originalmodelle – erstellt für die jeweiligen Untersuchungsphasen – gibt das Bestimmtheitsmaß nach Nagelkerke ( $R^2_N$ ) mit 0,81 für UP1, 0,53 für UP3, 0,64 für UP5 und 0,81 für UP7 an. Die Werte können als Prozentanteil interpretiert werden, die das Modell an der Varianz der abhängigen Variable erklärt, d.h. 81% in UP1. Obwohl dieser Prozentsatz deutlichen Schwankungen unterworfen ist, trennen alle Originalmodelle nach dem Likelihood Ratio-Test (LR-Test) hochsignifikant ( $<0,001$ ).

Die vorwärtsschrittweise Modellbildung führt zu den für UP1 und UP5 zusammengefassten Modellen in Tabelle 5.5-2. Es wird immer (auch im Herbst, nicht dargestellt) der gleiche Modellparameter ausgewählt, was für dessen prädiktive Stabilität spricht.

**Tabelle 5.5-2: Lebensraumeignungsmodelle für *Anisus spirorbis* in UP1 (oben) und UP5 (unten) mit Signifikanzangaben zu den Modellparametern aus der Wald-Statistik.**

Variable	Regressionskoeffizient $\beta$	Standardfehler $\beta$	Signifikanz
<b>Schritt 1: Nullmodell UP1</b>			
Konstante	0,288	0,342	0,400
<b>Schritt 2: Sigmoidales logistisches Regressionsmodell UP1</b>			
Konstante	12,041	3,969	0,002
Mittlerer Grundwasserflurabstand 4-9 1998	-0,48	0,17	0,004
<b>Schritt 1: Nullmodell UP5</b>			
Konstante	0,452	0,342	0,186
<b>Schritt 2: Sigmoidales logistisches Regressionsmodell UP5</b>			
Konstante	6,849	1,952	0,000
Mittlerer Grundwasserflurabstand 4-9 1999	-0,30	0,009	0,001

Die niedrigen Werte des Regressionskoeffizienten für die Erklärungsvariable in Tabelle 5.5-2 beschreiben einen allmählichen Wahrscheinlichkeitsübergang des Vorkommens entlang der Kurve (vgl. Abbildung 5.5-1, S. 219). Das negative Vorzeichen des Beta-Wertes lässt bereits die Richtung des Einflusses der Erklärungsvariablen erkennen. Mit steigendem Parameterwert  $x$  (steigendem mittleren Grundwasserflurabstand) wird die Wahrscheinlichkeit für die Referenzprägung (hier: Gruppe 0 / Nichtvorkommen) höher, da folglich der Wert  $z$  für die Regressionsgleichung ( $z_i = \beta_0 + \beta_1 x_{i1} + \dots$ ) sinkt (Backhaus et al. 1996). Die Wald-Statistik testet, ob ein bestimmtes  $\beta_j = 0$  ist, d.h., die zugehörige Erklärungsvariable keinen Einfluss auf die Gruppentrennung hat. Der Signifikanzwert in UP1 von 0,004 bedeutet, dass mit einer Wahrscheinlichkeit von 99,6% die Nullhypothese verworfen werden darf ( $H_0$ : Der Parameter hat keinen Einfluss auf die Gruppentrennung).

Um Informationen über die Wirkungsstärke zu erhalten, kann die Odd-Ratio und deren Konfidenzintervall betrachtet werden (Tabelle 5.5-3, S. 218). Eine Odd-Ratio von 0,953 bedeutet,

dass das Wahrscheinlichkeitsverhältnis sich mit der Zunahme des x-Wertes der Erklärungsvariablen um eine Einheit, d.h. um das 0,953-fache zugunsten der Referenzgruppe Nichtvorkommen ändert. Ein sehr hohe Odd-Ratio belegt einen sprunghaften Anstieg der Wahrscheinlichkeit für einen bestimmten x-Wert. Liegen die Werte der Odd-Ratio und des zugehörigen Konfidenzintervalls über eins (unter eins) wird die Modellaussage bestätigt. Ansonsten kann nicht mit Signifikanz gesagt werden, ob die Modellaussage zutrifft und die Erklärungsvariable die Gruppen relevant trennt (Backhaus et al. 1996).

**Tabelle 5.5-3: Odd-Ratio (EXP ( $\beta$ )) und 95%-Konfidenzintervall (CI<sub>95%</sub>): Modelle der Art *Anisus spirorbis* für UP1 (oben) und UP5 (unten).**

Variable	EXP ( $\beta$ )	CI <sub>95%</sub> für EXP ( $\beta$ )	
		unterer Wert	oberer Wert
<b>Sigmoidales logistisches Regressionsmodell UP1</b>			
Konstante			
Mittlerer Grundwasserflurabstand 4-9 1998	0,953	0,922	0,985
Variable	EXP ( $\beta$ )	CI <sub>95%</sub> für EXP ( $\beta$ )	
		unterer Wert	oberer Wert
<b>Sigmoidales logistisches Regressionsmodell UP5</b>			
Konstante			
Mittlerer Grundwasserflurabstand 4-9 1999	0,97	0,953	0,988

Nachdem die grundlegenden Eigenschaften der Originalmodelle für UP1 und UP5 bekannt sind, wird deren klassifikatorische Güte bewertet. Dazu werden die Modelle mit einem anfänglichen Klassifikationsschwellenwert ( $P_{\text{krit}}$ ) von 0,5 und dem nahezu fehlergerechten ( $P_{\text{Fehler}}$ ) parallel analysiert.

**Tabelle 5.5-4: Parameter zur Klassifikationsgüte: Modelle für *Anisus spirorbis* in UP1 und UP5 (AUC: Fläche unter der ROC-Kurve; P: Klassifikationsschwellenwert; %korr: Gesamtprozentsatz der richtig klassifizierten Beobachtungen; Spe: Spezifität und Sen: Sensitivität).**

Modell	AUC mit [CI <sub>95%</sub> ]	P( $P_{\text{krit}}$ ) = 0,5			P( $P_{\text{Fehler}}$ )		
		%korr	Spe	Sen	%korr	Spe	Sen
UP1 mit P( $P_{\text{Fehler}}$ ) = 0,70	0,967 [0,917 1,016]***	88,9	86,7	90,0	91,4	93,3	90,0
UP5 mit P( $P_{\text{Fehler}}$ ) = 0,68	0,935 [0,846 1,024]***	88,9	78,6	95,5	86,1	85,7	86,4

Signifikanz des AUC mit [CI<sub>95%</sub>]: \*\*\* =  $\alpha < 0,001$ ; \*\* =  $\alpha < 0,01$ ; \* =  $\alpha < 0,05$  und ns = nicht signifikant

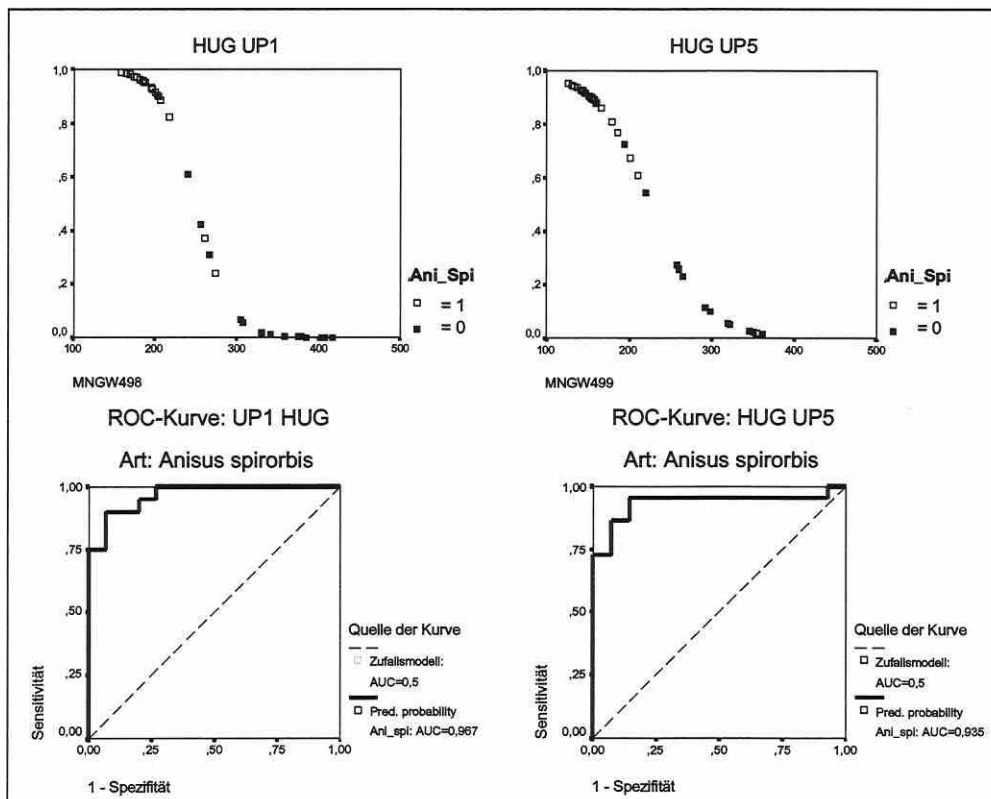
**Tabelle 5.5-5: Gesamtprozentsatz der richtig klassifizierten Beobachtungen (%korr) und Prävalenz: Nullmodelle für *Anisus spirorbis* in UP1 und UP5.**

Modell	%korr	Prävalenz (%)
UP1 mit P( $P_{\text{Fehler}}$ ) = 0,70	42,9	57,1
UP5 mit P( $P_{\text{Fehler}}$ ) = 0,68	38,9	61,1

Die beiden Frühjahrsmodelle zeigen eine sehr hohe diskriminatorische Kraft. Verglichen mit der Klassifikationsrate der Nullmodelle (Tabelle 5.5-5) liegt der Gesamtprozentsatz der richtig klassifizierten Beobachtungen deutlich höher. Der Wert des Gesamtprozentsatzes in UP1 erklärt, dass in 42,9% aller Fälle die Aussage "Die Art *Anisus spirorbis* kommt auf den Probeflächen nicht vor" korrekt ist, da umgekehrt die Prävalenz 57,1% beträgt. Die Nullmodelle sind unabhängig von Klassifikationsschwellenwerten und prognostizieren grundsätzlich 0% Spezifität und 100% Sensitivität oder exakt umgekehrt, d.h., sie liefern keine signifikante Klassifikationsmatrix bzw. keine verlässliche Prognose (Schröder 2000).

Der ebenfalls von einem Klassifikationsschwellenwert unabhängige AUC-Wert ist nahe seinem Optimum von eins (vgl. Tabelle 5.5-4 und Abbildung 5.5-1, S. 219 unten). Die Klassifikationsraten der Modelle mit einem fehlergerechten Klassifikationsschwellenwert liegen sowohl vom Gesamtprozentsatz der richtig klassifizierten Beobachtungen als auch bezüglich Spezifität und Sensitivität generell in Bereichen über 85 Prozent (Tabelle 5.5-4). Nach den

Kreuztabellen liegen die fehlergerechten Klassifikationsschwellenwerte ( $P_{\text{Fehler}}$ ) für UP1 zwischen einem Grundwasserflurabstand von 217 und 239cm und in UP5 zwischen 186 und 193cm. Dies zeigt, dass trotz hoher Trennkraft des Modells keine exakten Rückschlüsse auf einen kritischen Schwellenwert für den mittleren Grundwasserflurabstand in der Natur abgeleitet werden können. Es existieren keine echten Schwellenwerte zur Abgrenzung einer Nische, sondern es handelt sich um Übergangsbereiche. Zur verlässlichen Bestimmung dieser bedarf es eines hohen Stichprobenumfangs, falls es überhaupt möglich ist. Außerdem kann die beschriebene eindimensionale Nische generell die räumliche Variabilität der Art nicht genau festlegen (vgl. Abschnitt 2.3). Eine Tendenz ist aber ableitbar. Die Modelle für die Schirmart *Anisus spirorbis* sowie deren ROC-Kurven stellt die Abbildung 5.5-1 zusammen.



**Abbildung 5.5-1: Lebensraumeignungsmodelle und ROC-Kurven für *Anisus spirorbis* (UP1 und UP5; gefüllte Quadrate: Vorkommen).**

Eine Modellvalidierung ist essenziell zur Einschätzung der Modellrobustheit für prognostische Anwendungen. Grundsätzlich ergeben sich erst realistische Schätzungen der Klassifikationsgüte durch die Modellübertragung auf unabhängige Testdaten (Jørgensen 1994; Schröder 2000), welche aus anderen Zeitphasen oder Räumen stammen können (Fielding & Haworth 1995). Zum Test der **zeitlichen und räumlichen Übertragbarkeit der Modelle** für *Anisus spirorbis* wird daher unterschieden zwischen:

- (1) **Originalmodell:** Das für den Trainingsdatensatz optimal angepasste Modell.
- (2) **Übertragenes Modell:** Das Originalmodell wird auf andere Zeitphasen oder Räume übertragen und anhand eines Testdatensatzes seine Robustheit bestimmt.

Die Prognosegüte wird unter Beibehaltung des Klassifikationsschwellenwertes aus den Originalmodellen und dem schwellenwertunabhängigen AUC-Wert der ROC-Kurve einge-

schätzt. Die Ergebnisse der Modellübertragungen fasst Tabelle 5.5-6 und Abbildung 5.5-2 (S. 221) zusammen.

Die Auswertung der AUC-Werte für die zeitliche Übertragung überrascht, da die Werte der Originalmodelle identisch mit denen für die Übertragung der Originalmodelle auf UP1 und UP5 sind (vgl. Tabelle 5.5-6). Nach der Überprüfung der Kreuztabellen und der Diagramme in Anlage 7 wird offensichtlich, dass für nicht oder nahezu nicht mobile Arten eine sehr hohe zeitliche Autokorrelation über die Untersuchungsphasen hinweg herrscht. Daraus ergibt sich, dass für weitgehend standortgebundene, langsam reagierende Indikatorgruppen die zeitliche Übertragbarkeit nur langfristig zu beurteilen ist. Durch die Binärcodierung des Vorkommens werden geringe Populationsschwankungen durch geänderte Umweltbedingungen – im Gegensatz zu den Ordinationsmodellen – nicht erfasst. Dennoch lässt die Beibehaltung des Klassifikationsschwellenwertes Verschiebungen in der Antwort von Arten gegenüber dem Gradienten erkennen. Die übertragenen Modelle sind erwartungsgemäß, wie bereits für die multivariaten Ordinationsmodelle gezeigt, weniger erfolgreich bei der Klassifizierung der Testdaten (vgl. auch Tabelle 5.4-28, S. 210). Alle zeitlich übertragenen Modelle trennen aber hochsignifikant besser als die Nullmodelle.

**Tabelle 5.5-6: Modellübertragung der Modelle für *Anisus spirorbis* in UP1 und UP5 (AUC: Fläche unter der ROC-Kurve; P: Klassifikationsschwellenwert; %korr: Gesamtprozentsatz der richtig klassifizierten Beobachtungen; Spe: Spezifität und Sen: Sensitivität).**

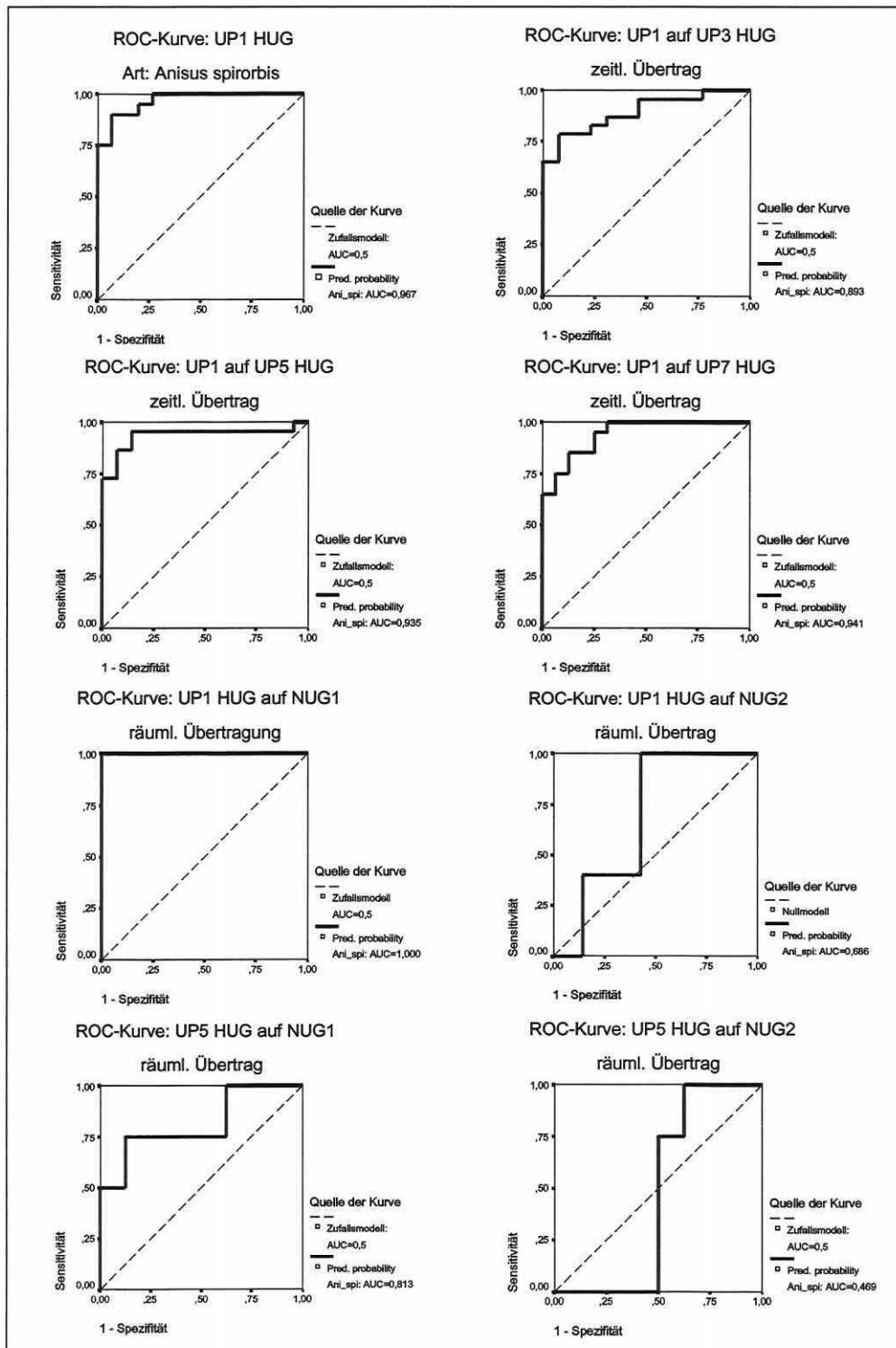
Modell	AUC mit [CI <sub>95%</sub> ]	P(Fehler) = 0,70			Nullmodell	
		%korr	Spe	Sen	%korr	Prävalenz
Originalmodell UP1	0,967 [0,917 1,016]***	91,4	93,3	90,0	42,9	57,1
zeitliche Übertragung						
übertragenes Modell auf UP3	0,893 [0,791 0,995]***	83,3	92,3	78,3	36,7	63,9
übertragenes Modell auf UP5	0,935 [0,846 1,024]***	88,9	78,6	95,5	38,9	61,1
übertragenes Modell auf UP7	0,941 [0,871 1,011]***	83,3	68,8	95,0	44,4	55,6
räumliche Übertragung						
übertragenes Modell auf NUG1 (UP1)	1,000 [1,000 1,000]**	91,7	100	75,0	66,6	33,3
übertragenes Modell auf NUG2 (UP1)	0,686 [0,368 1,004] ns	58,3	28,6	100,0	58,3	41,7

Signifikanz des AUC mit [CI<sub>95%</sub>]: \*\*\* =  $\alpha < 0,001$ ; \*\* =  $\alpha < 0,01$ ; \* =  $\alpha < 0,05$ ; <sup>0</sup> =  $\alpha < 0,1$  und ns = nicht signifikant

Modell	AUC mit [CI <sub>95%</sub> ]	P(Fehler) = 0,68			Nullmodell	
		%korr	Spe	Sen	%korr	Prävalenz
Originalmodell UP5	0,935 [0,846 1,024]***	86,1	85,7	86,4	38,9	63,9
zeitliche Übertragung						
übertragenes Modell auf UP1	0,967 [0,917 1,016]***	85,7	100	75,0	42,9	57,1
übertragenes Modell auf UP3	0,893 [0,791 0,995]***	77,8	100	65,2	36,7	63,9
übertragenes Modell auf UP7	0,941 [0,871 1,011]***	80,6	75,0	85,0	44,4	55,6
räumliche Übertragung						
übertragenes Modell auf NUG1 (UP5)	0,813 [0,523 1,102] <sup>0</sup>	75,0	87,5	50,0	66,6	33,3
übertragenes Modell auf NUG2 (UP5)	0,469 [0,132 0,806] ns	41,7	12,5	100	58,3	41,7

Signifikanz des AUC mit [CI<sub>95%</sub>]: \*\*\* =  $\alpha < 0,001$ ; \*\* =  $\alpha < 0,01$ ; \* =  $\alpha < 0,05$ ; <sup>0</sup> =  $\alpha < 0,1$  und ns = nicht signifikant



Abbildung 5.5-2: ROC-Kurven der zeitlichen/räumlichen Modellübertragungen für *Anisus spirorbis*.

Vergleicht man die Ergebnisse der Nischenstudien in Abschnitt 5.4.4, werden die dort getroffenen Aussagen prinzipiell bestätigt. Aus Abbildung 5.4-41 (S. 207) geht hervor, dass die Optima von *Anisus spirorbis* über die vier Untersuchungsphasen sehr ähnlich liegen. Dies ist wohl auch der Mittelung des Grundwasserflurabstandes über die Monate April bis September zuzusprechen. Abbildung 5.4-42 (S. 208) scheint dem Ergebnis der vorwärtsschrittweisen Modellbildung der Lebensraumeignungsmodelle zu widersprechen, da die partielle Ordination kaum einen Erklärungsanteil an der Art-Varianz zu erklären scheint. Die Einzelbetrachtung ergibt keinen unimodalen Zusammenhang der Antwort von *Anisus spirorbis* zu der Erklärungsvariablen mangels Länge des Gradienten, so dass die obige Aussage differenziert betrachtet werden muss. Hier zeigt sich der Einfluss des Modells auf die Ergebnisse. In den multivariaten Ordinationsmodellen gilt dies nicht in der direkten Übertragung, da hier die Antwort der Arten zur Linearkombination der Parameter approximiert wird. Ein Vergleich kann somit nicht ohne weiteres gezogen werden.

Betrachtet man die Resultate der räumlichen Übertragung (Tabelle 5.5-6, S. 220, und Abbildung 5.5-2, S. 221), bestätigen sich die Ergebnisse der Übertragbarkeitsprüfung für die Ordinationsmodelle (vgl. Abschnitt 5.4.4). NUG1 lässt eine mögliche Übertragbarkeit erkennen, während die Modelle im NUG2 gänzlich scheitern und nicht vom Nullmodell signifikant verschieden sind. Eine echte Zuverlässigkeit dieser Ergebnisse muss allerdings auch hier bezweifelt werden, da der Stichprobenumfang in den Nebenuntersuchungsgebieten mit zwölf extrem gering ist. Tendenzielle Aussagen mögen dennoch zu vertreten sein. Die visuelle Darstellung der sigmoidalen Modelle für *Anisus spirorbis* ist den Grafiken der Anlage 7 zu entnehmen.

#### Schirmart: *Arrhenatherum elatius*

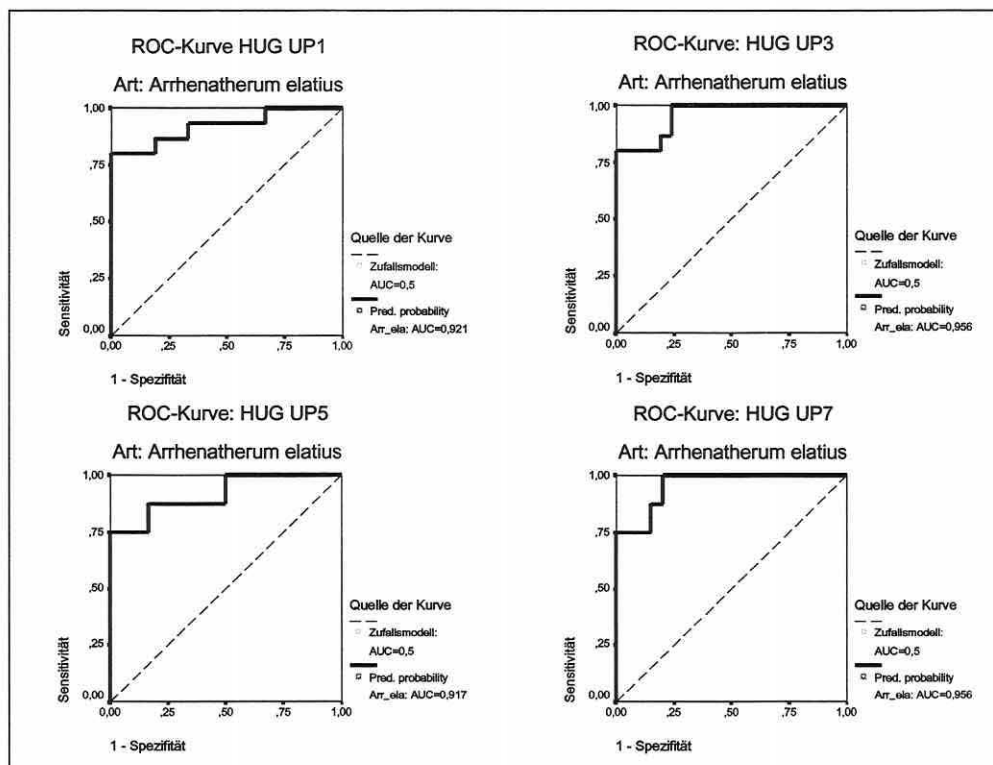
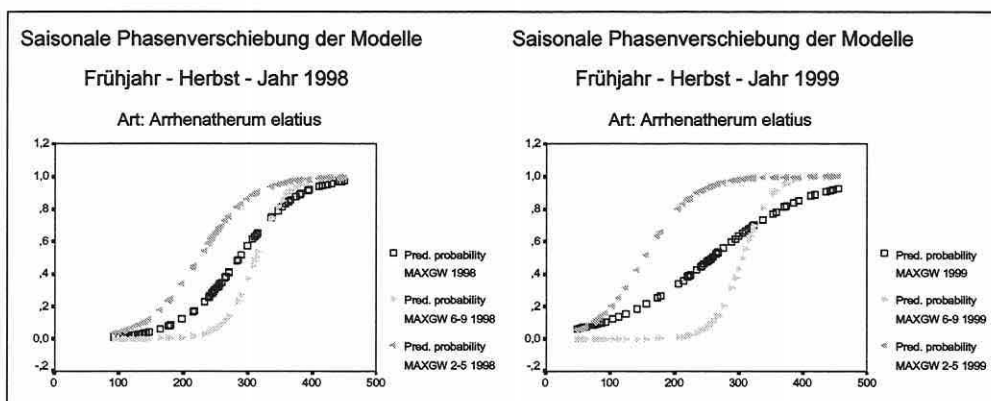


Abbildung 5.5-3: ROC-Kurven für *Arrhenatherum elatius* über die Untersuchungsphasen.

Ein interessantes weiteres Beispiel wird an *Arrhenatherum elatius* vor Augen geführt. Der Modellparameter maximaler Grundwasserflurabstand in den Monaten Februar bis Mai und Juni bis September zeigt eine robuste Trennkraft in allen Untersuchungsphasen, wie die ROC-Kurven der jeweiligen Originalmodelle beweisen (Abbildung 5.5-3, S. 222). Es wird anhand der Abbildung 5.5-4 eine relevante Zusatzinformation offen gelegt, die bei der Modellanwendung zu berücksichtigen ist. Sind Modellparameter auf saisonale Zeitphasen ausgelegt und bleibt die Art ortsfest, so ist ihre Nische nicht auf einen konkreten Wertebereich zu beziehen. Ihre Nische ist relativer Natur, d.h., es erfolgt eine Phasenverschiebung des Modells. Die Werte für den Modellparameter liegen im Herbst verständlicherweise sehr viel niedriger als im Frühjahr. Ein über das Jahr gemitteltes Modell fügt sich zwischen die saisonalen Modelle ein. Die Extrembedingungen sind daher für die Definition der Realnische von *Arrhenatherum elatius* verantwortlich. Eine Modellanwendung ist nicht ausgeschlossen, muss aber seasonspezifisch erfolgen, wenn die Modellparameter entsprechend erhoben wurden. Eine langfristige Veränderung der Lebensbedingungen durch einen anthropogenen Eingriff wirkt sich wiederum relativ auf die Grundwasserstände sowohl im Frühjahr als auch im Herbst aus.

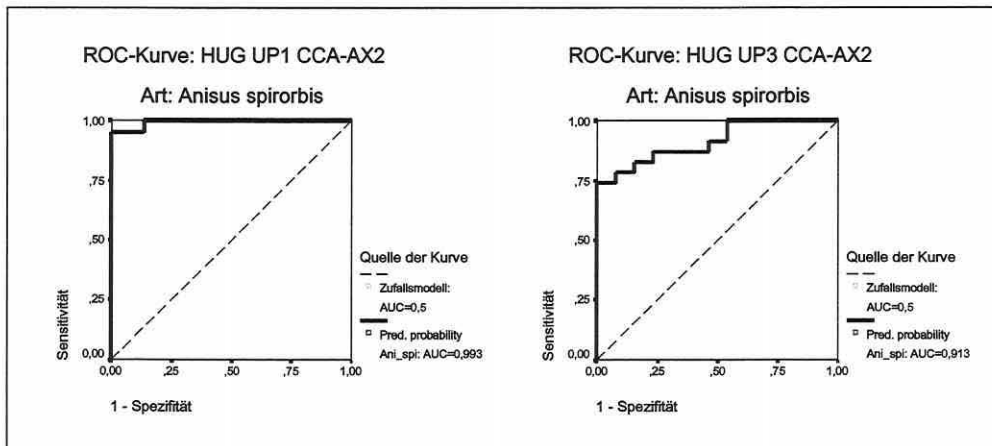


**Abbildung 5.5-4: Modelle für *Arrhenatherum elatius*: Frühjahr, Herbst und das gesamte Jahr 1998 und 1999. Als Modellparameter gingen der maximale Grundwasserflurabstand in cm über das Jahr (Quadrate), innerhalb der Monate Februar bis einschließlich Mai (dunkle Dreiecke) und innerhalb der Monate Juni bis einschließlich September (helle Dreiecke) ein.**

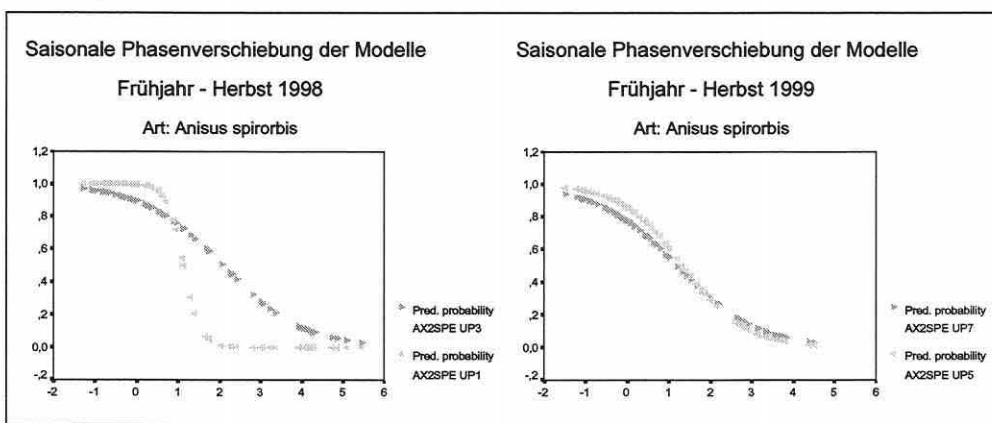
### Logistische Regression basierend auf kanonischen Erklärungsvariablen

Eine letzte Studie kombiniert die Ergebnisse der Ordinationsverfahren mit der Funktionalität der logistischen Regression. Es werden die kanonischen Variablen der Ordinationsachsen als linear kombinierte Erklärungsvariablen herangezogen. Die durch den mittleren Grundwasserflurabstand dominierte zweite Ordinationsachse des Molluskenmodells (vgl. Tabelle 5.4-12, S. 172) ist für die in diesem Abschnitt bereits besprochene Art *Anisus spirorbis* die trennkraftigste Komplexvariable. Die ROC-Kurven stellen die prädiktive Qualität der kanonischen Variable für 1998 dar (Abbildung 5.5-5, S. 224). Sie führt auch artenspezifisch zu sehr guten Modellergebnissen.

Abbildung 5.5-6 (S. 224) zeigt das abweichende Verhalten des Modells für UP1 gegenüber den Modellen der anderen Untersuchungsphasen auf. Bereits Tabelle 5.4-5 (S. 165) beschreibt die geringere Einflussnahme des mittleren Grundwasserflurabstandes bezogen auf UP1. Nahezu gleich hohe Erklärungsanteile in UP3 bis UP7 stehen geringeren in UP1 gegenüber. Eine Folge dieses Umstandes scheint der auf einen sehr viel engeren Bereich bezogene Wahrscheinlichkeitsübergang zu sein, da Grundwasserparameter gewöhnlich räumlich stärker autokorreliert sind (vgl. Abschnitt 5.4.1) und fließende Übergänge erzeugen.



**Abbildung 5.5-5: ROC-Kurven der Modelle für die kanonische Erklärungsvariable "zweite CCA-Achse": *Anisus spirorbis*.**



**Abbildung 5.5-6: Modelle für die kanonische Erklärungsvariable "zweite CCA-Achse" über die Untersuchungsphasen (UP1 und UP3 links, UP5 und UP7 rechts): *Anisus spirorbis*.**

Zusammenfassend muss betont werden, dass die Anwendung des Verfahrens "Logistische Regression" in RIVA grundsätzlich problembehaftet ist, da der Stichprobenumfang ( $n = 36$ ) die in der Literatur angesprochene Empfehlung von mindestens 50 Fällen unterschreitet (Backhaus et al. 1996). Trotz der guten Ergebnisse für einige Arten, hat dieser Aspekt bei der Modellerstellung bei vielen Arten zu mangelhaft trennenden Modellen (Datenrauschen) oder einer insgesamt nicht signifikanten Modelldiskriminierung geführt. Die schrittweise Implementierung von Erklärungsvariablen ist nicht oder nur selten über ein univariates sigmoideales Modell hinauszuführen gewesen. Unimodale Antworten gegenüber einem Parameter sind speziell bei der Überflutungsdauer auf einen so engen Bereich des Wahrscheinlichkeitswechsels zusammengedrängt, dass die Modelle nicht in einem eindeutigen 95%-Konfidenzintervall der Odd-Ratio ( $\text{EXP } \beta$ ) liegen.

## 5.6 GIS-Implementierung der statistischen Modelle zur Prognose

Der letzte Abschnitt dieses Kapitels beschäftigt sich mit der Implementierung der statistischen Modelle in das Geografische Informationssystem ARC/INFO®. Die zwei behandelten Modelltypen 'multivariate Ordinationsmodelle' und 'Lebensraumeignungsmodelle' basierend

auf logistischer Regression' finden auf verschiedenen Wegen Eingang in das Prognosesystem und werden deshalb getrennt behandelt.

#### **Multivariate Ordinationsmodelle:**

Die zum Transfer der CANOCO-Modellergebnisse eingesetzte ARC/INFO®-Applikation CANOGEN wurde von Andrew und Stuart Weiss an der Stanford University entwickelt (vgl. Guisan et al. 1999). Zur Modellierung des Artenvorkommens im realen Raum wird auf die multiplen Regressionsgleichungen der Ordinationen zurückgegriffen, die für jede Ordinationsachse berechnet werden (vgl. auch Abschnitt 4.2). Durch sie lassen sich die Koordinaten jeder Rasterzelle eines Grids entlang des Schlüsselkomplexgradienten, gebildet aus der gewichteten Linearkombination von Erklärungsvariablen im Modell, bestimmen. Da die Erklärungsvariablen eine Gewichtung und Standardisierung in der CCA erfahren, müssen die kanonischen Regressionskoeffizienten destandardisiert werden, bevor die Originalwerte der Erklärungsvariablen einer Flächeneinheit (Gridzelle) im Untersuchungsgebiet an die Modellgleichung übergeben werden können. Die Anwendung setzt voraus, dass regionalisierte Daten aller Modellparameter über das Untersuchungsgebiet hinweg vorliegen. Zur Destandardisierung werden das gewichtete Mittel und die Standardabweichung der Modellparameter, welche im log-File der CANOCO®-Ausgabefiles dokumentiert sind, benötigt. Die exakte Vorgehensweise bei der Destandardisierung kann unter anderem in ter Braak (1998) nachvollzogen werden.

Anhand der importierten regionalisierten Modellparameterwerte jeder Gridzelle ermittelt CANOGEN die Position der jeweiligen Zelle entlang der Ordinationsachse. Diese Werte legt das Modul in einem neuen Grid für die einzelnen Achsendimensionen ab. Die Verteilung kann nachfolgend visualisiert werden und zeigt farbkodiert die Verteilung der Flächeneinheiten entlang der Ordinationsachse an. Die Anlagen 8 und 9 zeigen je ein Beispiel für die erste und zweite Ordinationsachse des Floramodells. Im nächsten Schritt werden die Positionskoordinaten der jeweiligen Gridzelle entlang der Achsen erneut eingelesen und die euklidische Distanz zum Zentroid einer Art im kanonischen Raum errechnet. Klassifiziert nach der Toleranz dienen die Werte der Vorkommensprognose einer Art. Die Arttoleranzen bestimmen so letztlich ihre etwaige Realnische in der soziologischen Gemeinschaft gegenüber den ausgewählten Ordinationsachsenkombinationen (vgl. Abschnitt 2.3 und 3.6.1). Die Einteilung der Distanzen in zwei oder mehrere Klassen entspricht sinngemäß der Wahl eines Klassifikationsschwellenwertes in der logistischen Regression (vgl. Abschnitt 3.7.1) oder der von Guisan et al. (1999) angewendeten Klassifizierung der Vorkommenswahrscheinlichkeiten. Die klassifizierten Gridzellenwerte können schließlich für die räumliche Vorhersage der potenziellen Artenverteilung im GIS verwendet werden. Eine Karte der räumlichen Verteilung für jede Art im Datensatz bildet das Ergebnis (vgl. Anlagen 10 und 11).

Zur Prognose einer Veränderung des biotischen Musters nach einem anthropogenen Eingriff müssen als Basisinformation die Auswirkungen auf die Modellparameterwerte für verschiedene Szenarios festgelegt werden, um eine erste Prognose rechnen zu können. Optimal erscheint eine Modellkalibrierung auf das Gebiet noch vor dem eigentlichen Eingriff.

Klar definierte objektive Bewertungskriterien sind dann eine weitere Voraussetzung um den Vergleich der Prognoseergebnisse in der Planungspraxis tragfähig werden zu lassen.

#### **Lebensraumeignungsmodelle basierend auf logistischer Regression:**

Die Implementierung der Gleichungen ins GIS ist wenig problematisch. Die Werte der regionalisierten Modellparameter werden an die Modellgleichung übergeben und die Ergebnisse in einem neuen Datensatz abgelegt. Die resultierenden Vorkommenswahrscheinlichkeiten können dann beispielsweise klassifiziert in 0,2er-Schritten im GIS dargestellt werden und ergeben eine Vorkommensprognose für die jeweilige Art (vgl. Anlage 12). Durch die Klassifizierung der Vorkommenswahrscheinlichkeiten ist das Prognoseergebnis unabhängig von einem Klassifikationsschwellenwert, welcher nur zwei Gruppen zulässt.



Die Definition einer Schnittstelle zur Einbettung der statistischen Modelle in ein Geographisches Informationssystem ist essentiell, um einen Raumbezug herzustellen und die Funktionalität beider Komponenten voll für eine praxisrelevante Nutzung verfügbar zu machen (vgl. Richter et al. 1997). Speziell die Bewertung einer Prognose bei der praxisrelevanten Nutzung des Systems wird deutlich erleichtert durch die gleichzeitige Darstellung aller ökologischer Gruppen einer taxonomischen Artengruppe in einer Prognosekarte. Hierzu werden Einzelergebnisse im GIS verschnitten und Nischen-Überlappungsbereiche von ökologischen Gruppen nach der Nähe zu ihrem zugehörigen Schwerpunkt gewichtet, für nicht eindeutige Gridzellen priorisiert zugeordnet und abgebildet (Anlagen 13 und 14).

## 6 INTERPRETATION UND DISKUSSION DER ERGEBNISSE

Die Diskussion beschränkt sich auf Schwerpunktergebnisse aus Kapitel 5 und stellt diese in einen interpretierenden Kontext. Aus der kritischen Prüfung der Ergebnisse werden schlussfolgernd Erkenntnisgewinne für die übergreifende Fragestellung der Arbeit dargelegt und daneben Handlungsempfehlungen für Versuchsplanung, Datenanalyse und Modellierung abgeleitet. Die Schlussbetrachtung zieht ein Resümee und gibt einen Ausblick, wie in Folgeuntersuchungen auf die Erkenntnisse dieser Arbeit aufgebaut werden könnte.

### **Welche Erkenntnisse, Schlussfolgerungen und Empfehlungen liefert die vorliegende Arbeit?**

#### **Stichprobenplan, Indikatoren und Indikationsvermögen**

Die besondere Bedeutung des Stichprobenplans als Grundlage der Datenauswertung und Modellbildung wird immer wieder hervorgehoben (vgl. Jongman et al. 1987; Kent & Coker 1992; Köhler et al. 1996; Mühlenberg 1993; Økland 1990; Pickett et al. 1994; Richter & Söndgerath 1990; Rink 1997c, 1998; Rink et al. 2000 u.v.a.; s.a. Abschnitt 3.1). Die durch eine fachlich-statistische Zusammenführbarkeit der verschiedenen Datensätze erst realisierbaren Auswertungen und deren Ergebnisse bestätigen diese Aussage erneut.

Die speziell gewählte "stratifizierte Zufallsstichprobenerhebung" (Abschnitt 5.2) bei gewichtetem Stichprobenverhältnis über drei Untersuchungsgebiete (ein Haupt- und zwei Nebenuntersuchungsgebiete: HUG, NUG), vier Untersuchungsphasen (Frühjahr und Herbst 1998/1999: UP1, UP3, UP5 und UP7) und 14-tägigen Erhebungsintervallen erwies sich für zwei der drei bearbeiteten Indikatorgruppen (Abschnitt 5.2-3) hinsichtlich der Typisierung von Lebensgemeinschaften und der Auffindung von Koinzidenzen zwischen Lebensraumeigenschaften und Besiedlungsstruktur als effektiv.

Generell wird durch die Stratifizierung eine recht homogene Raumüberdeckung bei bestmöglicher Ausnutzung des vorgegebenen Stichprobenumfangs – auch bezogen auf physikalisch kleine, aber bedeutsame Raumeinheiten – erreicht. Die Integration des Zufallsaspektes sichert neben statistischen Ansprüchen auch die Einbeziehung von Ökotonen in die Erhebungen ab (Jongman et al. 1987; Kent & Coker 1992; Mühlenberg 1993 u.a.). Dies erweist sich als äußerst bedeutsam, um Schwellenbereiche des Artenübergangs entlang von Gradienten zu erfassen.

Die Gewichtung des Stichprobenverhältnisses über drei Untersuchungsgebiete im Verbundprojekt lässt erste Übertragbarkeitsstudien zu. 36 Probeflächen zur Modellbildung im Hauptuntersuchungsgebiet sind knapp bemessen (logistische Gründe), zeigen sich aber dennoch als ausreichend, um Ergebnisse zu erzielen, welche die Lebensraumbindungen verschiedener Arten (-gruppen) für die Flora und die Mollusken klar zu differenzieren vermögen. Der Stichprobenplantyp und die Erhebungsintervalle haben sich für die mobilere, wenig ortsgebundene Indikatorgruppe der Laufkäfer als suboptimal herausgestellt. Mangels Standortkonstanz in der Adultphase ist nur eine diffuse Zonation über das Hauptuntersuchungsgebiet hinweg zu erhalten gewesen. Das vergleichsweise raschere Reaktionsvermögen der Gruppe bewirkt speziell in der durch rezentes Überflutungsgeschehen geprägten Aue Wanderbewegungen mit der auflaufenden Welle oder zu noch bestehenden permanenten Wasserflächen in der trockeneren Herbstperiode. Somit sind die Tiere bei relativ gleichbleibenden Feuchteverhältnissen auf sehr unterschiedlichen Flächen über das Jahr anzutreffen. Als Folge lassen sich anhand der gewonnenen Laufkäferdaten nur sehr grobe ökologische Typisierungen vornehmen. Die Ordinationsdiagramme belegen, dass die Lebensgemeinschaften über die vier Untersuchungsphasen nicht statisch bleiben, sondern Durchmischungserscheinungen auftreten, die nur unklar abgrenzbare ökologische Typisierungen zulassen. Die feuchtebedingten Migrationsbewegungen blähen die eigentliche Nische bezogen auf die Modellpara-

meter stark auf, was an den starken Überlappungen der Antwortkurven deutlich wird (vgl. Abschnitt 5.4-2). Somit geben sich die Gradienten (approximierte DCA-Achsenlänge) im Vergleich zu den beiden anderen Indikatorgruppen stark eingekürzt.

Nur eine differenzierte Betrachtung hochdynamischer sowie der eher statischen Modellparameter zur Beschreibung der Artenverteilung vermag es, dass für bodenlebende mobile Artengruppen spezifische Problem der Nischenüberschätzung zu vermeiden. Die hierfür benötigte Informationsgrundlage ist nur durch hohen logistischen Aufwand bei der Datenerhebung zu realisieren. Durch den gerichteten Ablauf eines Überflutungsereignisses kann zur effektiven Datenerhebung nur eine Rasterlösung empfohlen werden. Transekte entsprechen nicht dem Anspruch bisher nicht berücksichtigte oder bekannte Gradienten neben der Überflutung aufzudecken. Speziell bei anstehenden auf- oder ablaufenden Flutwellen ist die zeitliche Dichte der Datenerhebung angemessen zu erhöhen, d.h. die Datenerhebung muss praktisch mit dem Vorkommensschwerpunkt der Arten wandern. Nur so kann die Migrationsbewegung der Arten in der Aue nachverfolgt werden, um differenzierte Aussagen über die aktuellen Feuchteverhältnisse zu ermöglichen.

In der Praxis wirft diese Empfehlung Probleme auf, da es die Logistik der Datenerhebung gerade bei der gleichzeitigen Studie verschieden mobiler Artengruppen erheblich kompliziert. Es müssten unterschiedliche Stichprobenpläne nebeneinander Anwendung finden, um den indikatorgruppenspezifischen Verhaltensweisen gerecht zu werden und gleichzeitig den Datenerhebungsaufwand noch realistisch erscheinen zu lassen. Als Handlungsempfehlung ist von einer Simultanbetrachtung verschieden mobiler Artengruppen innerhalb einer Studie abzusehen, da der Stichprobenumfang ohnehin meist durch finanzielle Rahmenbedingungen limitiert ist und somit verlässliche Ergebnisse nicht für alle Gruppen zu gewährleisten sind. Taxaübergreifende Datenerhebungen bleiben dennoch im engdefinierten Mobilitätsrahmen zu befürworten, da dann Erhebungs- und Bestimmungsaufwand mit dem zu erwartenden Erkenntnisgewinn im Einklang stehen und vergleichende Aussagen möglich werden.

Bei der separaten Betrachtung mobiler bodenlebender Artengruppen in der Aue kann auch bei angepassten Stichprobenplänen ein enger Raumbezug zu den Standortparametern durch die Wanderbewegungen nicht aufrecht erhalten werden. Das komplexe raum-zeitliche Zusammenwirken eines dominanten schnell veränderlichen Parameters wie der Überflutung und anderen standortdifferenzierenden Parametern ist daher modelltechnisch extrem schwierig zu beschreiben. Da Ordination und logistische Regression statische Modelle beruhend auf Koinzidenzen zwischen Standort (-parametern) und der biotischen Information ergeben, ist ihre Anwendung für mobile Arten bei der Integration von hochdynamischen Dominanzparametern – zu denen man die Überflutung zählen darf – auf niedriger Skalenebene in Frage gestellt. Der Prognosegedanke in Kombination mit der Visualisierung von Raumbindungen im Geographischen Informationssystem auf der Feldskala erfordert eine gewisse Standortkonstanz der Arten oder die weitgehend stabile Standortparameterausprägung über tolerable Zeitphasen.

Die unterschiedliche Mobilität von Artengruppen erfordert also eine klare Trennung bzw. die spezifische Anpassung der Stichprobenpläne und Auswertekonzepte auf die Verhaltensweisen der Taxa. Eine Globallösung, welche allen Ansprüchen gerecht wird und dabei vom finanziellen und logistischen Aufwand noch tragbar ist, erscheint nicht existent.

Weiterhin hat in der Studie der räumlich sehr zusammengedrückte rasche Übergang von hohen zu niedrigen Überflutungsdauern entlang der Flanken von Flutrinnen in der ansonsten nur schwach bewegten Aue zu disjunkten Gruppen bei den wenig mobilen Artengruppen geführt. Der Versuch einer Differenzierung der Arten innerhalb einer disjunkten Gruppe erbrachte auf der gegebenen Stichprobenbasis keinen zusätzlichen Erkenntnisgewinn. Eine Differenzierbarkeit bei höheren Stichprobenumfängen pro disjunkter Gruppe erscheint aber nicht ausgeschlossen.

Neben der unterschiedlichen Mobilität der Artengruppen ist für deren Auswahl in einer Studie bedeutsam, dass zwischen den faunistischen und den floristischen Indikatoren Unterschiede in ihrem Reaktionsvermögen auf Störungen im Standortsystem bestehen. Die Flora reagiert langsamer – teils extrem zeitverzögert (Abschnitt 5.5-2) – als die Mollusken und Carabiden. Die erhobenen floristischen Basisdatensätze belegen bei den Auswertungen über die zwei Jahre eine hohe zeitliche Autokorrelation, welche speziell in den Ergebnissen der Modellübertragungen mit Präsenz-Absenz-Daten zum Ausdruck kommt.

Reliktvegetation ist keine Seltenheit. Glavac (1996, S. 94) schreibt für Großseggenriede auf semiterrestrischen Böden: "Da viele hochwüchsige *Carex*-Arten auch nach einer Veränderung der Wasserhaushaltsbedingungen am gleichen Standort noch lange ausharren können, muss man auf die Lebensumstände bei ihrer Entstehung und ihrer Erhaltung achten, bzw. die Standortbedingungen ihrer Initial-, Optimal- und Terminalphasen mitberücksichtigen. Viele Großseggen-Bestände der Kulturlandschaft könnten demnach als 'Biomasse-Relikte' bezeichnet werden."

Das Zitat sensibilisiert, die meist schlaglichtartig erhobenen Daten und die daraus gezogenen Schlussfolgerungen kontextgemäß zu bewerten. Neben den langen Reaktionszeiten der Flora sollten auch Keimungs- und Neu-Etablierungsphasen Berücksichtigung finden, da diese andere Standortbedingungen erfordern können. Ist die Berücksichtigung nicht möglich, muss die Gültigkeit der Ergebnisse zumindest diskutiert werden.

Die Fauna lässt kurz- bis mittelfristig Veränderungen im Standortsystem erkennen, und Reliktpopulationen sind seltener zu erwarten, aber es gilt zu beachten, dass die Ansprüche an den Lebensraum ebenfalls von verschiedenen Entwicklungsstadien abhängen. Häufig unterscheiden sich die Standortansprüche der Individuen zwischen dem Larvalstadium und der adulten Phase. Da die Laufkäfer-Bodenfallen methodisch bedingt nur adulte Tiere erfassen, bleibt fraglich, ob die gefundenen Art-/Standort-Zusammenhänge wirklich allgemeingültig für den gesamten Lebenszyklus sind. Trotz erfüllten Standortansprüchen im Adultstadium sind die Arten vielleicht auf den Flächen nicht anzutreffen, wenn die entsprechenden Lebensraumforderungen für andere Entwicklungsstadien nicht adäquat gegeben waren.

Durch schnellere Reaktionszeiten gegenüber Lebensraumveränderungen, die eigentlich positiv aus dem Indikationsgedanken heraus zu bewerten sind, können sich aber auch Probleme bei der Ableitung von Nischeninformationen für die Erstellung von Modellen ergeben. Es kommt vor, dass Populationen einer Art durch ein kurz vor der Datenerhebung eingetretenes Extremereignis stark geschädigt werden und daher viele potenziell geeignete Lebensräume temporär unbesetzt sind ("Memory-Effekt"), da eine Neubesiedlung bzw. Wiederetablierung noch aussteht. Als Folge wird die Nischenbreite unterschätzt (vgl. Kleyer et al. 2000) und Modelle basierend auf Einmalhebungen sind ohne Kontrolle der Historie von Standortbedingungen unzuverlässig.

Das entgegengesetzte Phänomen – Überschätzung der Nischenbreite – entsteht gegebenenfalls durch die Berücksichtigung von Einzelfunden mobiler Taxa in der Analyse. Im Zuge von Wanderbewegungen werden zum Beispiel Arten vereinzelt beim Queren von ungeeigneten Lebensräumen gefangen. In der Folge kommt es zu Fehlinterpretationen der Nischenspanne und einer Erhöhung des Fehlers 1. Art bei der Modellbildung. Wird beim Versuch, diesem entgegenzuwirken, ein unteres Limit der zu berücksichtigenden Individuenzahl pro Probestrich zur Modellbildung eingeführt und dieses zu hoch angesetzt, d.h. zum Beispiel, alle Funde kleiner fünf Individuen bleiben als Zufallsfunde unberücksichtigt, entsteht bei der Modellbewertung ein erhöhter Fehler 2. Art. Diese Fehleinschätzung ist besonders kritisch, wenn dadurch anthropogene Lebensraumveränderungen als unbedenklich eingeschätzt werden und als ungeeignet prognostizierte Flächen einem Eingriff zum Opfer fallen. Da letzteres naturschutzfachlich schwerer wiegt, ist in der vorliegenden Arbeit von einem Grenzwert abgesehen worden (vgl. Morrison et al. 1998; Schröder 2000).

### Analyse-/Modellmethoden

Ein Nebeneinander von verschiedenen statistischen Analyse- und Modellierungsansätzen ist ratsam, damit kein wichtiger Informationsgehalt der Daten unentdeckt bleibt (Legendre & Legendre 1998; Orlóci 1978). Unterschiedliche Methoden erfassen oft verschiedene Aspekte eines Datensatzes und ermöglichen einen zusätzlichen Erkenntnisgewinn bei gleichen Datenressourcen als kostenintensivster Faktor einer Studie (Guisan et al. 1999). Die parallele Anwendung von direkter Ordination und logistischer Regression zeigt in dieser Arbeit die Bedeutungsschwankung von Modellparametern auf, wenn diese artspezifisch bzw. artenspezifisch eingesetzt werden. Grundsätzlich ist die Wahrscheinlichkeit, reale Eigenschaften entdeckt zu haben, deutlich gesteigert, wenn die Ergebnisse verschiedener Verfahren zu ähnlichen Schlussfolgerungen führen (Guisan et al. 1999).

Ein Nebeneinander von statistischen und konventionellen Methoden bei der Datenanalyse ist gleichermaßen vorteilhaft (vgl. Rink 1997c, unveröffentlicht), da die ausschließliche Anwendung konventioneller Methoden (z.B. Literaturstudie) zur Beurteilung der Biotik eines lokal begrenzten Raumes hinsichtlich ihrer Standortansprüche häufig irreführend ist, wie Unklarheiten und sogar Widersprüche verschiedener Autoren in der Literatur bestätigen (vgl. u.a. Rödel 1987 und Tabelle 6.0-1, S. 236). Die oft regional erstellten Beschreibungen können nicht problemlos auf andere Räume übertragen werden und sind im Falle von großräumigen Einstufungen so undifferenziert, dass sie gebietsspezifische Besonderheiten nicht erfassen (Skalenproblem). Ökosystemmodelle leiden unter dem gleichen Übertragbarkeitsproblem, lassen sich jedoch bei ihrem Einsatz in einem Gebiet kalibrieren, solange die Modellparameter ihre Relevanz beibehalten und die durch das Modell erfassten Beobachtungsobjekte vorkommen. Sind diese Voraussetzungen nicht gegeben, wird die mangelnde Aussagekraft schnell deutlich und der Bedarf an Grundlagenstudien aufgezeigt.

### Vergleich von Ordinationsverfahren

Auf der Basis von Abschnitt 4.1 werden die wichtigsten Erkenntnisse zu den Verfahren kurz zusammengefasst.

#### Indirekte Ordinationsverfahren:

Hauptkomponentenanalysen (PCA) erweisen sich für die Studie von Raummustern als ungeeignet, da durch die Annahme eines linearen Antwortmodells eine empfehlenswerte Anwendung erst bei Gradientenlängen unter zwei Standardabweichungen gegeben ist (vgl. Causton 1988; Gauch 1982a; Gauch et al. 1977; Jongman et al. 1987; Legendre & Legendre 1998; Økland 1996; Palmer 1993; ter Braak 1998). Dies impliziert, dass Arten über die Probeflächen verteilt weitgehend identisch sein müssen und sich insgesamt nur in ihren Häufigkeiten/Artmächtigkeiten unterscheiden dürfen (Legendre & Legendre 1998; Økland 1990). Zudem fehlt dem linearen Modellgedanken ein oberes und unteres Limit, wodurch eine Nischenabschätzung nicht möglich ist. Von allen Varianten der PCA kann unter gegebenen Randbedingungen nur die PCAc (zentriert) oder die PCAs (standardisiert) empfohlen werden, da die nicht zentrierte Variante der PCA ein unadäquates Modell für natürliche Situationen bildet und noch intensivere artifizielle bogenförmige Verzerrungen als die zentrierten Varianten zeigt (vgl. Causton 1988; Gauch et al. 1977; Jongman et al. 1987; Leps & Smilauer 1999; Minchin 1987; Økland 1990; Palmer 1998; ter Braak 1998 u.a.). Die PCA ist dennoch ein effektives Werkzeug, um Zusammenhänge zwischen Umweltparametern zu analysieren und sie zu Komplexgradienten zusammenzufassen (Økland 1990; Palmer 1998 u.v.a.).

Die Korrespondenzanalyse (CA) basiert auf einem unimodalen Antwortmodell und ist deshalb der PCA für Datensätze, welche Gradientenlängen von mehr als vier Standardabweichungen aufweisen immer vorzuziehen (Causton 1988; Gauch 1982a; Gauch et al. 1977; Kenkel & Orlóci 1986; Legendre & Legendre 1998; Økland 1990; Palmer 1998; ter Braak 1998 u.a.).



Die Korrespondenzanalyse mit Detrending-Prozess (DCA) gilt bei vielen Autoren als der PCA und CA überlegen (Gauch 1982a, b; Kent & Coker 1992; Økland 1986a - c, 1990; Palmer 1993; Peet et al. 1988; Pielou 1984). Betrachtet man jedoch die teils begründete Kritik von Austin (1985), Dargie (1986), Knox (1989) Kenkel & Orlóci (1986), Minchin (1987), Oksanen (1988) sowie Palmer (1993), muss man zu dem Schluss kommen, das Verfahren nur zur Approximation von Gradientenlängen einzusetzen. So kann über das geeignetste Modell bzw. Verfahren für die weitere Analyse der vorliegenden Daten entschieden werden (vgl. Jongman et al. 1987, Palmer 1998 und ter Braak 1998).

#### Direkte Ordinationsverfahren:

Die Redundanzanalyse (RDA) ist aus den gleichen Gründen wie die PCA (unterliegendes Modell) zur Mustererkennung nur bedingt geeignet (Palmer 1993; ter Braak 1987b). Das Gegenstück zur RDA mit unimodalem Modell bildet die kanonische Korrespondenzanalyse (CCA). Sie resultiert für die meisten Datensätze in realitätstreueren Ergebnissen und zeigt auch bei komplexen Versuchsanordnungen hervorragende Analyseergebnisse, wie die Tests von Palmer (1993) aufgezeigt haben. Dennoch bleibt auch die CCA anfällig für Ausreißer (Legendre & Legendre 1998; McCune 1997) und Einflüsse kleinräumiger Heterogenität (Palmer & Dixon 1990).

Als **Fazit** kann festgehalten werden, dass es keine generalisierte Empfehlung für den Einsatz eines Verfahrens gibt. Das "einzig wahre Ordinationsverfahren" ist nicht existent, so dass eine bewusste Methodenauswahl erfolgen und sich jeder Anwender über die Eigenschaften des angewandten Verfahrens informieren muss. Wird dies nicht berücksichtigt, kann es zu gravierenden Fehlern bei der Ergebnisfindung und -interpretation kommen, da in realen Datensätzen die Strukturen weitgehend unbekannt sind und ein fairer Vergleich der Analysekraft und Realitätstreue von Verfahren nur an synthetischen Datensätzen erfolgen kann (Causton 1988; Gauch et al. 1977; Økland 1990; Palmer 1993). Es empfiehlt sich bei jeder neuen Studie mit einer DCA zur Einschätzung der Gradientenlänge zu beginnen und anhand des Ergebnisses festzulegen, welches Verfahren für die weiteren Auswertungen geeignet ist (unimodales oder lineares Modell – Jongman et al. 1987; Palmer 1998; ter Braak 1998). Danach erfolgt die Ordination der Daten mit einem geeigneten indirekten Verfahren oder – je nach Datenlage – mit einem direkten Verfahren. Es ist grundsätzlich angeraten die Ergebnisse einer direkten Ordination mit denen aus einer indirekten zu vergleichen (Jongman et al. 1987 u.a.).

#### **Vergleich von Ordination und Clusteranalysen**

Speziell Ordinationsverfahren sind geeignet, dem Kontinuumgedanken des Artenübergangs entlang eines Gradienten gerecht zu werden, da sie im Gegensatz zu Clusterverfahren keine scharfen Gruppengrenzen erzeugen (vgl. Backhaus 1990, 1996; Legendre & Legendre 1998), sondern eher die Tendenz zur ökologischen Gruppierung bei fließenden Übergängen im Ordinationsraum zulassen (Begon 1998; Legendre & Legendre 1998). Ordination und Clusteranalysen können sich sinnvoll ergänzen (vgl. Legendre & Legendre 1998). Im reduzierten Ordinationsraum, in dem meist nur wenige Dimensionen berücksichtigt werden, kann es durch die Art der Projektion von Art- und Probeflächensignaturen zu Verzerrungen der Distanzen innerhalb einer Gruppe kommen. Fälschlicherweise können Arten zu einer Gruppe zusammengefasst werden, die in einer höheren Dimension durchaus noch eine Separierung erfahren. Als Folge können Zusammenhänge obskur erscheinen und bestehende Unterschiede übersehen werden. Clusterverfahren sind frei von dimensionsbezogenen Restriktionen. Wenn man beide Verfahren kombiniert und die Ergebnisse vergleicht, lassen sich daher die ökologischen Gruppen in ihrer Zusammensetzung über die ersten Dimensionen von Ordinationsverfahren prüfen und absichern. Die Clusteranalysen in dieser Arbeit belegen, dass die Gruppierung im reduzierten Ordinationsraum die Datenstrukturen weitgehend verzerrungsfrei wiedergibt und höherdimensionale Gruppentrennungen nicht zu erkennen sind. Die

ausgewiesen ökologischen Gruppen der Ordinationen können daher auf der vorliegenden Informationsbasis als verlässlich angesehen werden.

### **Vergleich von Ordination und logistischer Regression**

Einer der grundlegenden Unterschiede zwischen den beiden Verfahren liegt in der Art des resultierenden Modells. Die kanonische Korrespondenzanalyse (CCA) integriert alle zu betrachtenden Arten in ein Modell und sucht nach globalen Schlüsselgradienten, während die logistische Regression auf eine artspezifisch exakte Modellanpassung ausgelegt ist. Somit muss für jede Art ein neues Modell mit einer jeweils eigenständigen Modellparameterkombination angepasst werden, was sich bei vielen Arten als sehr zeitaufwendig erweist. Als Resultat ist dafür die Anpassungs- und Vorhersagegüte des Modells meist besser, da es nicht für alle Arten auf Globalannahmen (lineare/unimodale Antwort, Modellparametersatz) zurückgreift (Guisan et al. 1999). Nachteilig ist, dass in der logistischen Regression eine Binärikodierung der Daten vorausgesetzt wird, wodurch Populationsschwankungen außerhalb der Präsenz-Absenz-Information nicht erfasst werden. Manche Artengruppen weisen eine hohe zeitliche Autokorrelation auf, wodurch sich geringfügige Veränderungen schlecht bis nicht erkennen lassen.

Einige gewichtige Argumente sprechen für Ordinationsmethoden und deren globale Modellanpassung bei der Modellbildung und der prädiktiven Anwendung im Naturschutz. Durch die simultane multivariate Betrachtung aller Arten in nur einem Modell werden Muster in den biotischen Daten im ökologischen Ordinationsraum nachvollziehbar, wodurch wiederum ökologische Gruppen bestimmt werden können. Außerdem lassen sich im Gegensatz zur logistischen Regression auch seltene Arten mit wenigen Erscheinungen im Untersuchungsgebiet modellieren. Solche Arten haben oft ähnliche Lebensraumansprüche wie häufige Arten, die ihren natürlichen Schwerpunkt simulieren, und werden so ihren ökologischen Ansprüchen gerecht ins CCA-Modell eingefügt.

Ist der insgesamt verfügbare Stichprobenumfang in einer Untersuchung gering, treten bei der Modellerstellung mit der logistischen Regression verfahrensbedingte Probleme auf (Backhaus 1990). Oft kann entweder kein Modell oder nur ein Modell erstellt werden, welches die univariate Nische einer Art beschreibt. Diese Modelle sind trotz der Tatsache, dass der Modellparameter für die Art den höchsten Erklärungsanteil aufweist, sehr zweifelhaft und nur bedingt für den Prognoseeinsatz geeignet, da eindimensionale Nischen keine Spezies in ihrem wirklichen Raumverhalten gesichert zu beschreiben vermögen (Begon 1998). Tüxen (1954, S. 69) gibt sich sogar noch kritischer: "Wer glaubt, allein vom Boden, allein vom Grundwasser, allein vom Klima (oder auch von ihrem und anderer Faktoren angenommenen Zusammenwirken) her die Lebensmöglichkeit von Pflanzen verstehen zu können, hat sicher oft seine Rechnung ohne den Gast gemacht!"

Dieser Standpunkt ist nicht unbegründet und zeigt eines der grundlegenden Probleme von Ökosystemmodellen auf. Es können nicht alle Wirkfaktoren eines natürlichen Gefüges modelltechnisch erfasst werden, dennoch kann man sich an die Realität annähern. Die Kunst besteht darin, Modelle so einfach wie möglich zu halten, aber auch so verlässlich wie nötig zu gestalten (Hauhs & Lange 1996; Hauhs et al. 1996; Weiner 1995; Wissel 1989 u.v.a.).

Je nach Zielsetzung einer Studie kann die spezifische Überlegenheit der beiden Verfahren in Teilaspekten gewinnbringend genutzt oder eine Modellhierarchie erzeugt werden, wie die Ergebnisse dieser Arbeit zeigen. Die logistische Regression ergänzt die direkte Ordination optimal, da allgemeine Erkenntnisse aus der Ordination konkretisiert und für Schirmarten überprüft werden können.

Alternativ zu einzelnen Erklärungsparametern lassen sich die kanonischen Variablen aus der Ordination der logistischen Regression zuführen (Guisan et al. 1999; Hill 1991). Man arbeitet dann mit Faktorenkombinationen, den Schlüsselgradienten einer direkten Ordination, um exaktere Modellanpassungen für Einzelarten bei konstantem Modellparametersatz zu errei-

chen. Um erste räumliche oder zeitliche Übertragbarkeitstests der Modelle durchzuführen, bietet die quantitative Aussage der logistischen Regression verlässlichere Möglichkeiten. Als allgemeine Empfehlung bietet es sich an, die direkte Ordination zur Mustererkennung, zur Auffindung globaler Schlüsselgradienten und der Auswahl gut durch die Daten und das Modell repräsentierter Schirmarten heranzuziehen. Die Prognose für eine Vielzahl von Arten wird ebenfalls durch das Ordinationsmodell realisiert, während einzelne ausgewählte Arten durch die logistische Regression eine Überprüfung und gegebenenfalls eine exaktere Modellanpassung erfahren sollten, um bei Bedarf die Vorhersageergebnisse weiter zu optimieren.

#### **Standort- und Modellparameter**

Jede Studie zur Modellentwicklung erzwingt die fachliche Auswahl eines zu erhebenden Basisdatensatzes an abiotischen Parametern, welcher im zweiten Schritt zur Erklärung des biotischen Musters herangezogen wird und aus dem die eigentlichen Modellparameter selektiert werden.

Das solch ein Basisdatensatz geprägt ist von unterschiedlichen Expertenblickwinkeln, ist verständlich. Dieses durch Subjektivität im Sinne der objektiven Modellbildung negativ belastete Gedankengebäude erweist sich andererseits als vorteilhaft, da empirisches Wissen eingebracht und bewusst Spielräume für bisher nur imaginär angenommene Einflussfaktoren geschaffen werden können.

Letztlich dient jede Vorauswahl nur dem Versuch, ein komplexes kybernetisches System mit allen seinen Beziehungen zwischen den Faktoren auf wenige prägnante Einflussfaktoren zurückzuführen, um es einfachen Modellen zuzuführen und es letztlich unserem menschlichen Intellekt zugänglich zu machen. Diese Vorgehensweise ist durchaus gerechtfertigt (Hauhs & Lange 1996; Hauhs et al. 1996; Weiner 1995; Wissel 1989 u.v.a.).

Das Auffinden wichtiger bestandsdifferenzierender Faktoren für Modelle ist unter den angesprochenen Voraussetzungen generell relativer Natur (Reed et al. 1993; Richter & Söndgerath 1990 u.a.). Selbst innerhalb eines dieser relativen Basisdatensätze wird die Modellparameterauswahl selten absolut sein, da Umweltparameter fast ausnahmslos Abhängigkeiten untereinander zeigen und die Verfahren zur Auswahl diese korrelativen Strukturen zwischen Biotik und Abiotik prüfen. Die primär in dieser Arbeit eingesetzte Vorwärtsselektion wird somit zu einem streng hierarchischem Reduktionsverfahren. Als Folge ist die Reduktion auf einen einzigartigen und besterklärendsten Modellparametersatz unrealistisch, wenn man nicht alle erdenklichen Kombinationsmöglichkeiten des Basisdatensatzes auf ihren Erklärungsgehalt im Modell überprüfen möchte. Dennoch ist das Verfahren zur Parameterauswahl geeignet und führt zu einem der relevantesten Modellparametersätze innerhalb des Basisdatensatzes.

Allgemein muss bedacht werden, dass sich die Bedeutungsgehalte eines Erklärungsparameters nach der betrachteten Skala (hier: Feldskala), dem Ausschnitt im Lebenszyklus der Arten, dem Betrachtungswinkel (Sammelbetrachtung versus artspezifische Betrachtung) oder auch dem analysierten Zeitfenster bedingen. Das heißt, die prädiktive Stabilität von Modellparametern ist nur durch Vergleichsauswertungen ermittelbar. Konkret belegt wird die Aussage durch die separate Datenanalyse von vier Erhebungsphasen, welche zwar tendenziell ähnliche, aber keine identischen Ergebnisse zum Modellparametersatz geliefert hat. Somit ist in der Arbeit eine zeitlich robuste Abstimmung der Modellparameter präferiert und allen weiteren Bearbeitungen zu Grunde gelegt worden. Diese Abstimmung geht mit geringen Erklärungsverlusten bei der Artenseparierung einher, erzeugt dafür aber einen stabilen modelltechnisch erstrebenswerten Parametersatz, welcher klare logistische Vorteile bei der Modellübertragung in andere Gebiete durch die Beschränkung des Erhebungsaufwandes auf diese Parameter erbringt. Zusätzlich ist bei den Modellparametern eine Rangfolge des Bedeutungsgehaltes postulierbar, da die Parameter mit dem Auswahlverfahren vorwärts

schrittweise in das Modell überführt werden und so mit jedem Schritt ein Bedeutungsverlust zu dem nachfolgendem Parameter existiert (völlig unkorrelierte Parameter können theoretisch auch identische Bedeutungsgehalte besitzen).

Betrachtet man die Modellparameter als solches, wird schnell klar, dass manches empirische Wissen eine Bestätigung erfährt. Ellenberg stellte bereits 1952 fest, dass für Lebensgemeinschaften von Flussauen der Wasserfaktor das prägende Element ist. Andere Autoren wie beispielsweise Meyer (1957), Seibert (1975), Kiener (1984) oder Hügin (1990) kommen zu dem gleichen Schluss. So schreiben Hügin & Henrichfreise (1992, S. 10): "Das Wasser ist in Flusslandschaften entscheidend für die Abwandlung der Standorte. (...) Die Standortdifferenzierung einschließlich der Nährstoffverfügbarkeit wird überwiegend vom Wasser bestimmt." Diese Aussage spiegeln auch die Modellparameter für die bearbeiteten Indikatorgruppen deutlich wieder. Vorteilhaft ist jedoch, dass die statistischen Analysen diese bestehenden Kenntnisse zu parametrisieren vermochten und zu einem klar definierten Modellparametersatz geführt haben.

Für zwei von drei Indikatorgruppen hat sich die Überflutungsdauer als bedeutsamster Modellparameter zur Artendifferenzierung erwiesen. Dies vermutete bereits Dister (1983), als er den Parameter einsetzte, um Vegetationseinheiten der nördlichen Oberrheinaue in ihrer Verbreitung zu charakterisieren und die ökologischen Folgen von Wasserstandänderungen abzuschätzen.

Sind die Ergebnisse der Arbeit hinsichtlich der erhobenen Bodenparameter zu diskutieren und zu bewerten, muss in die Überlegungen einbezogen werden, dass indirekt auch die Nährstoffzufuhr – speziell bei Pflanzen – mit der Wasserversorgung gekoppelt ist: "Die Mobilität der Nährstoffe in Böden ist im Wesentlichen eine Funktion ihrer Konzentration in der Bodenlösung (Marschner 1985, S. 292)."

Somit ist es nicht verwunderlich, dass die vielen Bodenparameter im ersten Eindruck nur eine untergeordnete Rolle in den Modellen spielen. Speziell die Nährstoffversorgung stellt in Flussauen oft keinen limitierenden Faktor dar (Spink et al. 1998). Die Präsenz und der Bedeutungsgehalt von Bodenparametern nimmt vom Floramodell (effektive Kationenaustauschkapazität und pflanzenverfügbarer Phosphor als Nährstoffparameter) über das Molluskenmodell (Feldkapazität und Wasserstoffionenkonzentration:  $H_2O$ -Extraktionsmethode als Wassergehalts- und Lebensmilieuparameter des Oberbodens) bis zum Carabidenmodell (keine Bodenparameter) immer mehr ab. Fast alle aufgelisteten Parameter integrieren physiko-chemische Eigenschaften und stellen bereits einen komplexen Standortfaktor dar.

In das Carabiden-Modell geht kein Bodenparameter "direkt" ein. Bedenkt man die abnehmende Standortbindung der Artengruppen (zumindest in der Adultphase), ist der Bedeutungsverlust der Bodenparameter nachvollziehbar. Bemerkenswert ist, welche Parameter ersatzweise ins Modell eingeflossen sind: "Distanz zu permanenten Wasserflächen" und "Distanz zum Stratum Flutrinne". Hier wird der Einfluss unterschiedlicher Mobilität der Artengruppen nochmals unterstrichen. Durch die Möglichkeit eines Standortwechsels gewinnen Feuchtreifugien Einfluss auf die Verteilung der Arten. Vergleicht man die saisonalen Modelle miteinander, stellt man fest, dass die Stellung dieser beiden Parameter wechselt. Da viele Flutrinnen im Herbst trocken fallen, steigt die Relevanz der permanenten Wasserflächen und der Parameter wird in eine Position mit höherem Erklärungsanteil für die Artentrennung verschoben.

Eine weitere sehr interessante Erkenntnis ergibt sich aus dem Parameter "Standardabweichung des Wasserstandes im Kalenderjahr bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand". Er ist in allen Modellen enthalten und steht als Zeiger für hydrologische Dynamik in Auen. Sind häufiger starke Wasserstandsschwankungen an einem Standort gegeben, steigt der Wert des Parameters an und weist einen höher dynamischen Standort aus. Speziell die Dynamik eines Auestandorts



wird immer wieder als entscheidender Parameter von Ökologen angesprochen. Die konstante Bedeutung des Parameters über alle Modelle untermauert diese Vermutung.

Modelle, die den Wasserfaktor in Auen näher beschreiben und mit der biotischen Ausgestaltung eines Raumes in Beziehung setzen, sind von großer Bedeutung für Prognosen zur Vorhersage von Auswirkungen anthropogener Eingriffe in empfindliche natürliche Systeme. Tüxen (1954, S. 70) bemerkt: "(...) ohne sicheres Wissen von der Abhängigkeit der Pflanzengesellschaften überhaupt von der Höhe des Grundwassers (...) bleibt jede Melioration, die mit einer Senkung des Grundwasserspiegels verbunden ist, in der hergebrachten Form stecken und allen Fehlschlägen ausgesetzt, die mit ungenügenden Einblicken in gerade diese grundlegenden Beziehungen verbunden sind, um von der Unmöglichkeit der Beurteilung von Grundwasserschäden ohne Kenntnis dieser Beziehungen ganz zu schweigen!"

Ein Zitat von Kiener (1984, S. 104) geht dabei konkret auf wasserbauliche Maßnahmen und deren Auswirkungen ein: "Die Lage des Grundwasserspiegels, seine Bewegungen kennt man in vielen Fällen ziemlich genau, seine Veränderungen als Folge unterschiedlicher wasserbaulicher Maßnahmen und die Reaktion der von ihm mitbeeinflussten oder gar geprägten Vegetation wurde bislang nur in wenigen Fällen untersucht."

Mit der Anfertigung dieser Arbeit ist ein Schritt in Richtung einer Lösung der Problematik realisiert worden, da es ermöglicht wird, eine modelltechnische Vorhersage der Auswirkungen von hydrologischen Veränderungen auf das Raummuster der Biotik zu generieren.

### **Relativität der Nischen**

Aus der grundlegenden Tatsache, dass die Modellparameter eine Gruppierung der Arten zuließen, geht hervor, dass ausgedünnte Übergangsbereiche und Aggregationen von Arten entlang der Gradienten auftraten (vgl. Glavac 1996; Pfadenhauer 1993). Dieses Kontinuum mit Artensprüngen und die abzuleitende Nischeninformation bedürfen im Folgenden einer kritischen Betrachtung.

Generell kann keinesfalls von "Grenzwerten" für die Definition der Nischen von Arten oder Gruppen gesprochen werden, sondern nur von "Schwellenbereichen". Innerhalb dieser Zonen physiologischen Stresses kommt es zur Populationsausdünnung. Die Aussage lässt sich sehr gut in den Modellen der logistischen Regression nachvollziehen. In allen Modellen sind deutliche Überlappungsbereiche zwischen Probeflächen mit und ohne Vorkommen der Art zu erkennen, d.h., Probeflächen, auf denen die Art vorkommt, gehen entlang des Gradienten nicht abrupt über zu Probeflächen ohne Vorkommen. Im Übergangsbereich erfolgt eine Durchmischung von Präsenz und Absenz. Erst wenn die Zone der Intoleranz gegenüber einem limitierenden Modellparameter erreicht ist, tritt die Art nicht mehr auf. Um diese Übergangsbereiche zu erfassen und modelltechnisch zu charakterisieren, ist es von entscheidender Relevanz, dass solche Lebensräume (Ökotope) zwischen charakteristischen Standorten durch den Stichprobenplan erfasst worden sind.

Zudem ist für die Modellaussagen von besonderer Bedeutung, dass die Nischeninformationen relativer Art sind, da sowohl die zeitliche Anwendbarkeit der Modelle als auch deren räumliche Übertragbarkeit hiervon abhängig ist. Wichtig dabei ist, dass die räumliche Erhebungsintensität modifizierend auf die Nischenbreite wirkt. Gerade Ökotope sind hiervon betroffen, wie die Studie anhand der Mollusken erkennen lässt. Wertet man die Daten dahingehend aus, dass die Fangzahlen von fünf Stechrahmenproben auf einer Probefläche alternativ summiert bzw. nur von einer zufällig gezogenen Stechrahmenprobe stammen, ergibt sich, dass zwar eine identische Gesamtartenzahl im Gebiet ermittelt wird, aber das Vorkommen auf Probeflächen des Übergangsbereichs sich unterscheidet. Die Art fällt auf Probeflächen mit wenigen Erscheinungen bei fünf Stechrahmenproben bei einer geringeren Erhebungsintensität (nur eine Zufallsprobe) häufig ganz aus. Die Nische erscheint folglich enger als es der Realität entspricht.



Arbeitet man in Prognosen mit nur einer Schirmart, die stellvertretend für eine ökologische Gruppe steht, kommt es zu ähnlichen Modifikationen, da die artbezogenen Nischen verschiedener typischer Arten einer Gruppe zwar ähnlich, aber nicht identisch sind. Daher empfiehlt es sich, in Prognosen grundsätzlich zwei bis drei Schirmarten als Gruppenindikator einer Sammelbetrachtung zu unterziehen (vgl. Anlage 13 und 14).

Ein weiteres Argument zur Relativität von Nischen ergibt sich aus den Vergleichen der Nischenoptima über die vier Untersuchungsphasen. Sowohl in den Ordinationsergebnissen als auch in den Ergebnissen der logistischen Regression wird offensichtlich, dass gerade bei weitgehend ortsfesten Artengruppen trotz des "Memory-Effekts" die Optima und Nischen saisonal und somit auch über die Jahre hinweg verschoben sind. Erschwerend bei der Nischendefinition wirkt sich aus, dass vielfältige inter- und intraspezifische Verknüpfungen eine Biozönose nicht nur als Summe ihrer Mitgliedsarten erscheinen lassen, sondern dass sie durch die wechselseitigen Beziehungen neue zusätzliche Eigenschaften (emergente Merkmale) erhält, deren "Gesetzmäßigkeiten" nur selten exakt durch Felduntersuchungen messbar sind (Dierschke 1994 u.a.). Durch Konkurrenz kommt es zu einer Verschiebung des physiologischen Optimums von Arten gegenüber ihrem ökologischen Optimum. Selbst wenn dieser Einflussfaktor in den Erhebungsdaten des Modellerstellungsgebietes integriert ist, kann bei einer räumlichen oder zeitlichen Modellübertragung anhand von einzelnen stellvertretenden Arten eine differierende Artenzusammensetzung der ökologischen Gruppe Modifikationen der Nische bewirken.

Die Relativität von Nischen kommt auch in Literaturangaben qualitativ wie quantitativ zum Ausdruck (Rink 1997c, unveröffentlicht). Verschiedene Autoren treffen Aussagen zur ökologischen Kompensation, wobei der limitierende Einfluss eines Standortparameters auf die Existenz von Arten durch die spezifische Ausprägung eines anderen Parameters ausgeglichen wird (z.B. Ellenberg 1952, S. 28; Wilmanns 1989, S. 27). Weiterhin wirken Standortfaktoren untereinander beeinflussend und werden von der Lebensgemeinschaft rückwirkend wieder verändert (Marschner 1985, S. 291-309). Das sogenannte Gesetz der relativen Standortkonstanz kann ebenfalls von modelltechnischer Bedeutung sein. Es bedeutet, dass beispielsweise eine Gesellschaft humider Areale auch in ariden Gebieten auf sehr gut mit Wasser versorgten Sonderstandorten vorkommen kann (s.a. Walter & Breckle 1983 in Dierschke 1994, S. 33).

**Tabelle 6.0-1: Nische der Grundwasserflurabstände (cm) für das Rohrglanzgras-Röhrch (Phalaris arundinacea) nach verschiedenen Autoren.**

Optimale Verträglichkeit (cm)	Maximale Verträglichkeit (cm)	Minimale Verträglichkeit (cm)	Autor
+20 bis -100	+80	-200	DVWK 1996
keine Angaben	+130	-30	Meisel 1977
+40	+110	-30	Runge 1980
+10	+40	-5	Klemm 1976 in Rödel 1987
+50	+300	-50	Géhu 1961 in Rödel 1987
+10 bis +20	keine Angaben	keine Angaben	Vollrath 1965 in Rödel 1987
0 bis +40	keine Angaben	keine Angaben	Seibert 1962 in Rödel 1987
+30 bis -30 (Optimum Juli)	keine Angaben	keine Angaben	Ellenberg 1952a in Rödel 1987

**Tabelle 6.0-2: Lebensraumsansprüche der ökologischen Molluskengruppen G1 bis G6 bezogen auf die in der Untersuchung erfassten Modellparameterspannen (Angaben bezogen auf das HUG).**

Mollusken	G1	G2	G3	G4	G5	G6
Jährliche Überflutungsdauer in Tagen (Abflussjahr)	↑↑!	↑	↔!	↓	↓	↓↓!
Mittlerer Grundwasserflurabstand in der Vegetationsperiode (01.04.-30.09.) eines Kalenderjahres in cm	↔!	↔	↓	↓!	↔↑	↑!
Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand	↓!	↓	↔↓?	↔↑	↑	↑!
Amplitude der Grundwasserflurabstände zwischen dem 01.02.-31.05. bzw. 01.06.-30.09. eines Kalenderjahres in cm	↔	↔?	↓ F ↔↑ H	↓ F ↑ H	↔↑	↑! F ↓! H
Feldkapazität in mm; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994)	↓!	↓	↔↑	↑!	↔↑	↔?
pH-Wert H <sub>2</sub> O delogarithmisiert (Wasserstoffionenkonzentration in mol/l); Transformation: Multiplikation mit 1.000.000	↑↑!	↑	↔?	↓!	---	---

**Gruppenzuordnung:**

- G1: Tiefere Rinnen, meist länger mit Wasser bestanden  
 G2: Seichte, länger trocken liegende Rinnen  
 G3: Übergang von den Feuchtflächen zu den seichteren Rinnen  
 G4: Tieferliegende Feuchtflächen  
 G5: Übergang trockene/feuchte Flächen  
 G6: Höhergelegene, die meiste Zeit trocken liegende, genutzte Wiesenflächen

**Legende:**

↑↑	:	hohe Werte	—	:	nicht eindeutig
↓↓	:	niedrige Werte	!	:	besonders deutlich
↔	:	mittlere Werte	?	:	weniger deutlich
F	:	Frühjahr	H	:	Herbst

**Tabelle 6.0-3: Lebensraumsansprüche der ökologischen Carabidengruppen G1 bis G5 bezogen auf die in der Untersuchung erfassten Modellparameterspannen (Angaben bezogen auf das HUG).**

Carabiden	G1	G2	G3	G4	G5
Mittlerer Grundwasserflurabstand in der Vegetationsperiode (01.04.-30.09.) eines Kalenderjahres in cm	↓?	↓?	↔?	↔↑?	↑!
Jährliche Überflutungsdauer in Tagen (Abflussjahr)	↑↑?	↔?	↔?	↓↔?	↓↓!
Distanz zu permanenten Wasserflächen in m	↔↓ F? ↔ H?	↔?	↓?	↔?	↓ F? ↔ H?
Distanz zum Stratum Flutrinne in m	↔↓?	↔?	↑↔ F? ↔ H?	↑?	↑!
Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen max. Überflutung und max. Grundwasserflurabstand	↓?	↔↓?	↔?	↔?	↑

**Gruppenzuordnung:**

- G1: Tiefe, lange überschwemmte Flutrinnen  
 G2: Kürzere Zeit überschwemmte, flache Rinnen, Phalarisröhrichte und feuchte bis frische Mähwiesen  
 G3: Phalarisröhrichte  
 G4: Wechselfeuchte bis mäßig trockene Mähwiesen  
 G5: Trockene, allenfalls sehr kurz überschwemmte Mähwiesen

**Legende:** siehe Tabelle 6.0-2

**Tabelle 6.0-4: Lebensraumsprüche der ökologischen Floragruppen G1 bis G6 bezogen auf die in der Untersuchung erfassten Modellparameterspannen (Angaben bezogen auf das HUG).**

Flora	G1	G2	G3	G4	G5	G6
Jährliche Überflutungsdauer in Tagen (Abflussjahr)	↑↑!	↑↑	↔↑↑	↔↓↓	↓↓!	↓↓!
Maximum Grundwasserflurabstand zwischen dem 01.02.-31.05. bzw. 01.06.-30.09. eines Kalenderjahres in cm	↔!	↔↓	↔↓	↓↓!	↔!	↑↑!
Effektive Kationenaustauschkapazität in cmol/kg; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994)	↔	↔↑	↔↑	↑↑!	↔!	↓↓!
Maximale Überflutungshöhe in m (höchstes Überflutungsereignis im Abflussjahr)	↑↑!	↑↑	↔↑↑	↔↓	↓↓!	↓↓!
Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen max. Überflutung und max. Grundwasserflurabstand	↔↓	↓↓↔	---	---	↔?	---
Häufigkeit/Anzahl Überflutungsereignisse im Kalenderjahr	---	---	↔	---	↔↓	↓↓!
Pflanzenverfügbare Phosphor in mg/100 g	↑↑!	↑↑	↔	↓↓!	↔↓	↔↑↑

**Gruppenzuordnung:**

- G1: *Persicaria amphibia*-Gruppe: nasse, häufiger überschwemmte Standorte  
 G2: *Galium palustre*-Gruppe: nasse, etwas weniger häufig überschwemmte Standorte  
 G3: *Phalaris arundinacea*-Gruppe: feuchte bis nasse Standorte  
 G4: *Urtica dioica*-Gruppe: feuchte, nährstoffreiche, wenig genutzte Standorte  
 G5: *Alopecurus pratensis*-Gruppe: frische Standorte  
 G6: *Arrhenatherum elatius*-Gruppe: trockenere Standorte

**Legende:** siehe Tabelle 6.0-2

Einige Nischen-Kennwerte für das Rohrglanzgras-Röhrlicht sind in Tabelle 6.0-1 (S. 236) zusammengefasst worden und belegen die quantitativen Abweichungen von Literaturausagen. Sowohl das Nischenoptimum als auch speziell das Minimum und Maximum der Nische weisen drastische Differenzen auf. Die Isoliniendarstellungen zur Visualisierung von Nischenoptima für Arten im Ordinationsdiagramm in dieser Arbeit zeigen daher nur die lokale Ausprägung in den Untersuchungsgebieten. Hinzu kommt, dass die Optima und Antworten der Ordinationsmodelle sich auf die Linearkombinationen der Modellparameter beziehen, während die Isoliniendarstellungen nur jeweils einen Modellparameter berücksichtigen.

Die Tabellen 6.0-2 bis 6.0-4 (S. 237f) fassen die Ergebnisse der Mustererkennung dieser Arbeit für das Hauptuntersuchungsgebiet "Schöneberger Wiesen" zusammen und geben eine Übersicht der Lebensraumsprüche aller ökologischer Gruppen (vgl. auch Anlagen 13 und 14). Von numerischen Angaben zu Nischenspannen wird basierend auf der Diskussion zur Relativität von Nischen und Grenzwerten bewusst Abstand genommen. Als Folge der Unsicherheit in der Bestimmung einer wahren Nischenbreite lassen sich klare allgemeingültige Grenzwerte für die ökologische Existenz einer Art im natürlichen Gefüge für Standortparameter nicht festlegen. Daraus ergeben sich Auswirkungen für die Modellanwendung. Die Relativität von Nischeninformationen bewirkt Einschränkungen, die bei der Prognoserechnung unbedingt beachtet werden müssen:

Saisonal bezogene Modellparameter führen im Gegensatz zu Kalenderjahr bezogenen Modellparametern zu einer exakteren Vorhersagegüte des Modells, wenn die Modellanwendung zeitphasengerecht erfolgt. Gleichzeitig wird die zeitliche Übertragbarkeit solcher Modelle aber problematischer. Modellparameter, die über das Jahr oder lange Zeitphasen integrieren, bilden die gemittelte Nische einer Art ab. Dafür ist das Modell in seiner Anwendung nicht jahreszeitlich fixiert. Überträgt man die Zusammenhänge auf mehrere Untersuchungsjahre, wird verständlich, dass sich auch hierbei Verschiebungen speziell der Optima von Nischen ergeben können.

Um die Relativität der Nischen in eine Prognose einzubeziehen, muss entweder eine Modellkalibrierung bei jeder räumlichen Übertragung erfolgen oder man arbeitet mit der Szenariotechnik (klassifizierte Systemzustände), welche die Unschärfen des Modells soweit abfängt, dass vertretbare Aussagen und Prognosen innerhalb der definierten Grenzen von Modellen zu erwarten sind.

Trotz aller Unwegsamkeiten durch die Relativität von Nischen kann eine gebietsspezifische Annäherung an die realen Standortansprüche der Arten bzw. deren Existenzbereiche mit den eingesetzten Analyseverfahren erreicht werden. Dieses Wissen lässt sich allgemein als entscheidende Planungshilfe heranziehen und Veränderungen im Wirkungsgefüge der Arten durch anthropogene Eingriffe etwas vorhersehbarer machen.

#### **Zeitliche Übertragbarkeit:**

Eine zeitliche Übertragbarkeit von Modellen unter ähnlichen ökosystemaren Randbedingungen ist meist gegeben (vgl. Kleyer et al. 2000, Schröder 2000). Diese These unterstützen auch die Auswertungen dieser Arbeit. Zu bedenken ist dabei aber einerseits die Reaktionszeit der Arten, wodurch die Ergebnisse eng aufeinander folgender Zeitfenster bei wenig mobilen, langsam reagierenden Artengruppen von hoher zeitlicher Autokorrelation beeinflusst werden und bereits eingeleitete Veränderungen noch nicht sichtbar sind, und andererseits, dass die Modelle nur für die Umweltbedingungen während der Datenerhebungsphasen bei der Modellerstellung prognosefähig sind. Das heißt, ein sich nach extremen Veränderungen des Standortsystems einstellendes Artenspektrum kann außerhalb des durch die Modelle abbildbaren Gradienten liegen. Es entstehen "Platzhalter für unbekannte Folgebesiedler". Erst eine Modellentwicklung für verschiedene Systemzustände lässt eine Modellpalette entstehen, welche flexibel genug ist, die Folgen intensiver Standortveränderungen artenbezogen prognostizieren zu können.

Allgemein kann man auch nicht davon ausgehen, dass durch statische biozönotische Modelle vorhergesagte Artenvorkommen nach Ausgleichs- und Ersatzmaßnahmen oder einer Renaturierung kurzfristig im Untersuchungsgebiet anzutreffen sind. Dies ist eine Folge von Modellentwicklungen, welche die biologischen Parameter wie Artenpool, Ausbreitungsstrategie und -dauer, Sukzession, Biotopvernetzung und anderes nicht berücksichtigen (vgl. auch Schröder 2000).

#### **Räumliche Übertragbarkeit:**

Eine räumliche Übertragbarkeit der Modelle ist insgesamt nur begrenzt möglich. Entscheidend ist eine annähernde Analogie der Standortbedingungen und Artenspektren zwischen dem Herkunftsgebiet des Modells und dem Zielgebiet der räumlichen Modellübertragung. Nur so ist eine prinzipielle Anwendbarkeit gegeben.

Die Modellgültigkeit beschränkt sich zeitlich wie räumlich auf die im Modell integrierten und gut repräsentierten Beobachtungsobjekte und bewegt sich für die in dieser Arbeit erstellten Modelle innerhalb der Modellparameterspannen, welche Tabelle 6.0-5 (S. 240) anführt. Das heißt, eine Prognose des Artenwechsels ist nur innerhalb dieser Spannen möglich. Werden die Wertebereiche entlang des Gradienten überschritten entstehen die bereits erwähnten nicht definierbaren "Platzhalter" für die sich unter den neuen Bedingungen etablierenden Arten. Ohne die Kenntnis der in einem Gebiet zu erwartenden Gesamtlängen relevanter Gradienten sind im Extremfall nur Spekulationen möglich.

Speziell bei der räumlichen Übertragung sind daher erst Modelle mit entsprechend weiten Wertebereichen für die Modellparameter im Herkunftsgebiet zielführend, um den Istzustand eines Zielgebietes abbilden und darüber hinaus Prognoseszenarien approximieren zu können.



Tabelle 6.0-5: Spannen der Modellparameter über die Teilmodelle der Untersuchungsjahre 1998 und 1999 im HUG.

Modellparameter	Minimalwert	Maximalwert
Jährliche Überflutungsdauer in Tagen (Abflussjahr)	0	217
Maximale Überflutungshöhe in m (höchstes Überflutungsereignis im Abflussjahr)	0	4,10
Häufigkeit/Anzahl Überflutungsereignisse im Kalenderjahr	0	8
Mittlerer Grundwasserflurabstand (m) in der Vegetationsperiode (01.04.-30.09.) eines Kalenderjahres	1,26	4,17
Amplitude der Grundwasserflurabstände (m) im Zeitraum 01.02.-31.05. (Wert 1) bzw. 01.06.-30.09. (Wert 2) eines Kalenderjahres	Wert1: 0,34 Wert2: 0,69	2,63 1,52
Maximum Grundwasserflurabstand (m) im Zeitraum 01.02.-31.05. (Wert 1) bzw. 01.06.-30.09. (Wert 2) eines Kalenderjahres	Wert1: 0,47 Wert2: 2,13	4,23 4,55
Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen max. Überflutung und max. Grundwasserflurabstand	0,92	1,47
Feldkapazität [mm]; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994)	45,0	113,5
Effektive Kationenaustauschkapazität (cmolc/kg); abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994).	3,65	31,70
Pflanzenverfügbarer Phosphor (mg/100 g)	1,3	14,2
pH-Wert H <sub>2</sub> O delogarithmisiert (Wasserstoffionenkonzentration in mol/l); Transformation: Multiplikation mit 1.000.000	0,06	13,80
Distanz zu permanenten Wasserflächen (m)	15,38	613,47
Distanz zum Stratum Flutrinne (m)	0	574,83

### Schlussbetrachtung

Die Ergebnisse und ihre Diskussion lassen die Vielschichtigkeit an Aspekten bei der Datenanalyse zur Ökosystemmodellierung erkennen. Möchte man ökologische Strukturen/Muster aufdecken und vereinfachend beschreiben, ist ein ebenso komplizierter wie langer Weg, angefangen bei einer bewussten Versuchsplanung, der Auswahl geeigneter Auswertemethoden bis hin zur Analyse der Daten und der Modellerstellung selbst zu gehen. Gerade auf die Beschreibung eines solchen Gesamtkonzeptes und das Verständnis von Analyseverfahren ist in dieser Arbeit besonderer Wert gelegt worden, um zukünftigen Projekten ähnlicher Zielsetzung als Vorschlag und Orientierungshilfe dienen zu können.

Die vielen kritischen Gesichtspunkte bei der Modellerstellung rufen eine gewisse Skepsis gegenüber ökologischen Modellen hervor. Dennoch darf man sagen, dass die erstellten Modelle durchweg akzeptable bis gute Ergebnisse liefern. Erklärungsanteile von um die 60% bei den Ordinationsmodellen der Mollusken, 43% bei den Carabiden (trotz erwähnter Probleme) und 49% bei der sehr heterogenen Flora sind bemerkenswert.

Auch die Modelle der logistischen Regression und die Erkenntnisse zur Übertragbarkeit bestätigen, dass sich die mathematisch-statistische Ökosystemforschung als zukunftsträchtiges Arbeitsfeld erweist und die ökologische Forschung sowie das Ökosystemmanagement zu unterstützen vermag. Speziell die Koppelung der statistischen Modelle mit Geographischen Informationssystemen bietet ein enormes Potenzial, um für ökologische Prognosen Licht in manches Dunkel zu bringen (vgl. Anlagen 10 bis 14).

### Ausblick:

Die Einflüsse kleinräumiger (innerhalb der Probeflächen) im Vergleich zur großräumigen Heterogenität im Untersuchungsgebiet sowie die der Nutzung (Dominanz über Modellparameter) auf die Analyseergebnisse konnten nur benannt, aber in dieser Arbeit nicht analysiert werden. Weiterhin sind die erstellten biozönotischen Modelle bisher nicht dafür ausgelegt, die Populationsdynamik zu erfassen oder zusätzlich autökologische Parameter stärker zu berücksichtigen. Es ist wünschenswert, sowohl diese Themen aufzugreifen als auch der Frage nach der Robustheit der Modellparameter bei einer Übertragung auf andere Flussau-



en nachzugehen. Weiterführende Untersuchungen, welche Arten sich außerhalb der beobachteten Parameterspannen einnischen, sind ebenfalls von sehr großer Bedeutung für den Praxiseinsatz der Modelle.

Die vorgestellten Distanzparametermodelle eröffnen eine Perspektive, wie basierend auf der räumlichen Verteilung von Vegetationseinheiten, die implizit einen Lebensraum mit seinen abiotischen Bedingungen definieren, auch die Fauna des Raumes in ihrem Muster erklärt werden kann. Erste Ergebnisse hierzu sind vielversprechend und ausbaubar.

Schließen möchte ich mit einer Botschaft, welche trotz aller Schwierigkeiten bei der Mustererkennung und der Ökosystemmodellierung nicht ungehört bleiben sollte, denn zur Kunst des Ökologen gehört das Streben nach neuen Erkenntnissen und der Mut sich auf den Weg zu begeben:

***"It is only by them (statistical analyses like ordination)  
that a reasonable summary of our field data can be achieved;  
only by using these methods can we find some sort of order  
in what is otherwise an indigestible mass of data."***

***(Causton 1988, S. 324)***

## 7 ZUSAMMENFASSUNG

Die Aufarbeitung von Wissensdefiziten bei der Erforschung von Ökosystemen und die modelltechnische Formalisierung und Operationalisierung des Erfahrungsschatzes kann einen Beitrag zur integrativen Kommunikation zwischen Wissenschaft und Planung leisten. Der Wissenstransfer aus der Biozönoseforschung in die praktische Anwendbarkeit durch Modelltechniken soll Entscheidungsträgern helfen, Fehlentscheidungen der Vergangenheit zu vermeiden und allgemein zur Entwicklungslenkung beizutragen.

Ziele der Arbeit sind, Wissenslücken in Strukturzusammenhängen von Aueökosystemen durch statistische Datenanalyseverfahren und Modelltechniken zu schließen, bestehendes Wissen zu aktualisieren und zu validieren sowie erarbeitetes Wissen in einem Modellsystem zu aggregieren. Die Grundlage dafür bietet ein vom Bundesministerium für Bildung und Forschung gefördertes Verbundprojekt mit dem Titel "Übertragung und Weiterentwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen" (RIVA).

Als theoretischer Schwerpunkt der Dissertation wird eine wissenschaftlich fundierte, ganzheitliche Analysestrategie angefangen von der Datenerhebung bis hin zur Modellübergabe an ein Geographisches Informationssystem (GIS) erörtert und dokumentiert (Kapitel 2 bis 4). Dabei wird besonderer Wert gelegt auf Fragen der Stichprobenplanung und die umfassende Einführung in die Theorie und Leistungsfähigkeit von Ordinationsverfahren sowie der logistischen Regression als Analyse- und Modelltechniken. Zusätzlich wird dahingehend sensibilisiert, Ergebnisse durch Diagnostik und graphische Aufbereitung zu prüfen und ihre Repräsentativität detailliert zu hinterfragen.

Die Anwendung statistischer Analyseverfahren in der terrestrischen Ökosystemforschung zur Mustererkennung und Auffindung von Koinzidenzen zwischen Biotik und Abiotik bildet die Grundlage, auf der Modelle erstellt und letztlich in ein Geographisches Informationssystem zur Prognose eingebettet werden können. Die vergleichende Bewertung der Leistungsfähigkeit von Ordinationsverfahren und die Ergänzung dieser Verfahren um die logistische Regression zum Aufbau eines hierarchischen Modellsystems runden die Dokumentation ab. Durch die Wissensakquisition soll eine Übertragung der Analysestrategie, der Datenanalysemethoden und Modelltechniken auf ähnlich gelagerte Projektfragestellungen gefördert werden.

Der projektbezogene Teil der Arbeit zeigt die Anwendung der ganzheitlichen Analysestrategie am Beispiel eines Verbundprojektes auf (Kapitel 5). Dieses Ziel geht einher mit der Erarbeitung des Stichprobenplans, der Parametrisierung von Lebensraumbindungen (Nischenbestimmung) für die Artengruppen Flora, Mollusken und Carabiden, deren ökologische Typisierung und die Ableitung typischer Arten, welche stellvertretend für die jeweilige ökologische Lebensgemeinschaft in den Prognosemodellen bewertet werden können.

Anhand der Strukturanalyse der umfangreichen Felddaten mit multivariaten statistischen Methoden, werden biotische Muster beschrieben, die bedingenden Standortfaktoren ermittelt und Modelle abgeleitet, die das biotische Lebensraumgefüge soweit vereinfacht darstellen, dass es kommunizierbar und in Grenzen prognostizierbar wird.

Detailergebnisse theoretischer Teil: Die vergleichende Diskussion der multivariaten Analyseverfahren Hauptkomponentenanalyse (PCA), Korrespondenzanalyse (CA), Korrespondenzanalyse mit Detrending-Prozess (DCA), Redundanzanalyse (RDA) und kanonische Korrespondenzanalyse (CCA) zeigt, dass jede Ordinationsmethode ihre spezifischen Stärken und Schwächen besitzt, die eine generalisierte Anwendung von einem Verfahren verneinen.

Die PCA ist aufgrund ihrer linearen Modellannahme der CA meist deutlich unterlegen. Die reale Datenstruktur wird nicht aufgedeckt und für Gradientenlängen von mehr als vier Stan-

dardabweichungen ist eine bogenförmige Verzerrung der Datenstruktur sowie die Stauchung der Achsenenden das unvermeidliche Resultat. Die PCA kann dennoch für sehr kurze Gradienten Anwendung finden ( $<2$  Standardabweichungen), da somit ein größerer Anteil der Antworten linear ist. In diesem Fall sind die Arten auf den Probeflächen weitgehend identisch und unterscheiden sich nur in ihren Artmächtigkeiten/Häufigkeiten, um die Flächen zu differenzieren. Daraus erklärt sich, weshalb die PCA keine effektive Methode für die Studie von Räummustern in der Artenkomposition ist.

Im Vergleich zur PCA liefert die CA aufgrund des unterliegenden unimodalen Modells in den meisten Studien bessere Ergebnisse als die PCA. Für Gradientenlängen von mehr als vier Standardabweichungen ist sie der PCA immer vorzuziehen. Auf den ersten Blick vermag die DCA es, die Schwächen von CA und speziell der PCA zu vermeiden. Dennoch ist das Verfahren problematisch, da es durch den Detrending-Prozess die Daten systematisch manipuliert. Die Existenz realer Datenstrukturen kann verzerrt werden, wodurch das Verständnis für die Daten eingeschränkt wird. Neben der mangelnden Eleganz des Detrending-Algorithmus stellt sich eine geringe Analysekraft bei schiefen Verteilungen, ein Fehlschlagen der Analyse bei komplexen Stichprobenanordnungen und das gelegentliche Vorkommen einer Kompression des Gradientenendes in eine Zungenform ein. Die DCA bietet aber eine zuverlässige Möglichkeit, Gradientenlängen zu schätzen und so einen Hinweis zu geben, welches Modell bei der weiteren Datenanalyse zur Anwendung kommen sollte.

Bei den direkten Ordinationen kann der CCA eine hervorragende Analysekraft unter verschiedensten Bedingungen bescheinigt werden, wobei das unimodale Antwortmodell sie der RDA aus den gleichen Gründen überlegen werden lässt, wie die CA der PCA. Ein artifizieller Bogeneffekt ist nicht zu erwarten, solange er nicht real in den Daten vorhanden ist. Dieses direkte Ordinationsverfahren bleibt unbeeinflusst von vielen Problemen der CA, DCA und PCA.

Aus dem methodenkritischen Vergleich leitet sich die Notwendigkeit ab, differenziert vorzugehen und sowohl die grundlegenden Erfordernisse für die Anwendung von Ordinationsverfahren schon bei der Stichprobenplanung zu berücksichtigen als auch die Auswahl des Verfahrens den gegebenen Datensatzeigenschaften anzupassen.

Bei jeder neuen Analyse empfiehlt es sich mit einer DCA zur Einschätzung der Gradientenlänge zu beginnen und anhand des Ergebnisses festzulegen, welches Verfahren (unimodales oder lineares Modell) für die weiteren Auswertungen geeignet erscheint. Danach erfolgt die Ordination der Daten mit einem indirekten Verfahren oder – je nach Datenlage – mit einem direkten Verfahren. Es ist angeraten, die Ergebnisse einer direkten mit denen einer indirekten Ordination zu vergleichen. So lässt sich prüfen, ob die wichtigsten strukturdifferenzierenden Erklärungsparameter erfasst wurden.

In Zusammenhang einer bestmöglichen Datenexploration und der Ergebniskontrolle ist auch die Gegenüberstellung von Ordination und logistischer Regression (LR) zu sehen, da die parallele Anwendung verschiedener Analyseverfahren oft wichtige Informationsgehalte und Aspekte in den Daten erhellt, die bei der Beschränkung auf eine Methode verloren gehen. Analoge Ergebnisse sichern außerdem die Wirklichkeitsäquivalenz ab.

Die Frage, inwieweit sich die Modelltechniken zu einer Modellhierarchie ergänzen können, beantwortet sich aus den spezifischen Vor- und Nachteilen der direkten Ordination und der logistischen Regression. Durch die Möglichkeit einer Simultanauswertung vieler Arten und das generalisierte Modell einer CCA verglichen zur LR ist das Streben nach vielseitigen und kostengünstigen Modellen eher zu erfüllen. Das aus einer CCA resultierende übergreifende Modell minimiert den Aufwand der Modellerstellung durch wenige allgemeine Schlüsselparameter für alle Arten, bei gleichzeitiger Maximierung der Aussagekraft und Zuverlässigkeit. Das Verfahren eröffnet außerdem durch die Simultanbetrachtung der Arten eine ökologische Typisierung anhand deren Verteilungsmuster im Ordinationsraum und ermöglicht in Folge

die Ableitung typischer auch modelltechnisch gut repräsentierter Schirmarten als Stellvertreter der ökologischen Gruppen.

Ist ein möglichst exaktes Modell für eine spezifische Art gefordert, sollte die LR der CCA vorgezogen werden, da die Modellanpassung genauer vorgenommen werden kann und damit die Vorhersagegüte eventuell zu steigern ist. Die Verfahren ergänzen sich aufgrund ihrer spezifischen erörterten Eigenschaften sehr sinnvoll, indem die LR zur Ergebniskontrolle und für artspezifische Modellierungen aufbauend auf den CCA-Ergebnissen eingesetzt wird.

**Detaillierergebnisse projektbezogener Teil:** In diesem Abschnitt der Arbeit wird aufbauend auf den theoretischen Vorüberlegungen ein Modellsystem für den Themenbereich 'Auegrünland an der Elbe' erarbeitet. Basierend auf dem vergleichend ausgewählten Stichprobenplan stratifizierte Zufallsverteilung sind hierzu in drei Untersuchungsgebieten insgesamt 60 Probeflächen verortet und Erhebungen in den Jahren 1998 und 1999 durchgeführt worden. Die Daten umfassen Informationen zu den Artengruppen Mollusken, Carabiden und Flora sowie abiotischer Umweltbedingungen (238 Standortparameter), erhoben in vier separaten Untersuchungsphasen (Frühjahr und Herbst) innerhalb des zweijährigen Zeitfensters.

Die Parameter, welche sich primär für die Einnischung der Taxa verantwortlich zeigen, werden in einem ersten Schritt der Datenanalyse über Auswahlprozesse ermittelt. Die Selektion erfolgt vorrangig auf statistischer Basis. Die zeitliche Robustheit der Modellparameter wird durch die separate Auswertung der vier Erhebungsphasen optimiert. Als Resultat lassen sich so durchschnittlich sechs Modellparameter in nach Erklärungsanteilen gewichteter Rangfolge für die jeweiligen Taxa für die direkten Ordination ausweisen.

**Molluskenmodell:**

1. **Jährliche Überflutungsdauer in Tagen (Abflussjahr)**
2. **Mittlerer Grundwasserflurabstand in der Vegetationsperiode (01.04.-30.09.) eines Kalenderjahres in cm**
3. **Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand**
4. **Amplitude der Grundwasserflurabstände zwischen dem 01.02.-31.05. bzw. 01.06.-30.09. eines Kalenderjahres in cm**
5. **Feldkapazität in mm; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994)**
6. **pH-Wert H<sub>2</sub>O delogarithmisiert (Wasserstoffionenkonzentration in mol/l); Transformation: Multiplikation mit 1.000.000**

**Carabidenmodell:**

1. **Mittlerer Grundwasserflurabstand in der Vegetationsperiode (01.04.-30.09.) eines Kalenderjahres in cm**
2. **Jährliche Überflutungsdauer in Tagen (Abflussjahr)**
3. **Distanz zu permanenten Wasserflächen in m**
4. **Distanz zum Stratum Flutrinne in m**
5. **Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand**

**Floramodell:**

1. **Jährliche Überflutungsdauer in Tagen (Abflussjahr)**
2. **Maximum Grundwasserflurabstand zwischen dem 01.02.-31.05. bzw. 01.06.-30.09. eines Kalenderjahres in cm**
3. **Effektive Kationenaustauschkapazität in cmolc/kg; abgeleitet nach Bundesanstalt für Geowissenschaften und Rohstoffe (1994)**
4. **Maximale Überflutungshöhe in m (höchstes Überflutungsereignis im Abflussjahr)**
5. **Standardabweichung Wasserstand im Kalenderjahr, d.h. bezogen auf die Gesamtamplitude zwischen maximaler Überflutung und maximalem Grundwasserflurabstand**
6. **Häufigkeit/Anzahl Überflutungsereignisse im Kalenderjahr**
7. **Pflanzenverfügbare Phosphor in mg/100 g**

Durch die Mustererkennung, die ökologische Typisierung von Lebensgemeinschaften und die fachlich-statistisch gezielte Selektion von Schirmarten pro Gruppe wird das aufzustellende Modellsystem geschaffen. Es besteht aus den direkten Ordinationsmodellen, welche alle

Arten integrieren, und den Submodellen der logistischen Regression für die Schirmarten der ökologischen Gruppen. Als ökologische Gruppen und Schirmarten lassen sich ableiten:

#### Molluskenmodell:

- Gruppe 1: Tiefere Rinnen, meist länger mit Wasser bestanden  
 Typische Arten: *Planorbis planorbis*, *Stagnicola spec.*  
 Gruppe 2: Seichte, länger trocken liegende Rinnen  
 Typische Arten: *Anisus spirorbis*  
 Gruppe 3: Übergang von den Feuchtf Flächen zu den seichteren Rinnen  
 Typische Arten: *Succinea putris*, *Pseudotrichia rubiginosa*  
 Gruppe 4: Tieferliegende Feuchtf Flächen  
 Typische Arten: *Cochlicopa lubrica*, *Carychium minimum*, *Zonitoides nitidus*  
 Gruppe 5: Übergang trockene/feuchte Flächen  
 Typische Arten: *Vallonia pulchella*, *Perpolita hammonis*  
 Gruppe 6: Höhergelegene, die meiste Zeit trocken liegende, genutzte Wiesenflächen  
 Typische Arten: *Vallonia excentrica*

#### Carabidenmodell:

- Gruppe 1: Tiefe, lange überschwemmte Flutrinnen  
 Typische Arten: *Stenolophus skrimshirani*, *Bembidion biguttatum*, *Agonum versutum*, *Agonum duftschmidi*, *Pterostichus gracilis*, *Agonum fuliginosum*  
 Gruppe 2: Kürzere Zeit überschwemmte, flache Rinnen, Phalarisröhrichte und feuchte bis frische Mähwiesen  
 Typische Arten: *Carabus granulatus*, *Bembidion gilvipes*  
 Gruppe 3: Phalarisröhrichte  
 Typische Arten: *Amara communis*, *Epaphius secalis*  
 Gruppe 4: Wechselfeuchte bis mäßig trockene Mähwiesen  
 Typische Arten: *Syntomus truncatellus*, *Amara lunicollis*, *Calathus melanocephalus*, (*Poecilus versicolor*), *Amara strenua*  
 Gruppe 5: Trockene, allenfalls sehr kurz überschwemmte Mähwiesen  
 Typische Arten: *Carabus auratus*, *Amara equestris*, *Calathus fuscipes*

#### Floramodell:

- Gruppe 1: *Persicaria amphibia*-Gruppe: Nasse, häufiger überschwemmte Standorte  
 Typische Arten: *Oenanthe aquatica*, *Agrostis stolonifera*, *Eleocharis palustris*, *Persicaria amphibia*, *Rorippa amphibia*, *Glyceria maxima*, *Glyceria fluitans*  
 Gruppe 2: *Galium palustre*-Gruppe: Nasse, etwas weniger häufig überschwemmte Standorte  
 Typische Arten: *Carex acuta*, *Galium palustre*, *Iris pseudacorus*  
 Gruppe 3: *Phalaris arundinacea*-Gruppe: Feuchte bis nasse Standorte  
 Typische Arten: *Phalaris arundinacea*  
 Gruppe 4: *Urtica dioica*-Gruppe: Feuchte, nährstoffreiche, wenig genutzte Standorte  
 Typische Arten: *Glechoma hederacea*, *Symphytum officinale*, *Urtica dioica*  
 Gruppe 5: *Alopecurus pratensis*-Gruppe: Frische Standorte  
 Typische Arten: *Alopecurus pratensis*, *Elymus repens*  
 Gruppe 6: *Arrhenatherum elatius*-Gruppe: Trockenere Standorte  
 Typische Arten: *Ornithogalum umbellatum*, *Arrhenatherum elatius*, *Galium verum*, *Galium album*

Alle Modelle zeigen Lebensraumbindungen von Arten auf und ermöglichen das Extrahieren von Nischeninformation durch sorgfältige Prüfung der Einzelergebnisse. Erklärungsanteile von um die 60% bei den direkten Ordinationsmodellen der Mollusken, 43% bei den Carabiden und 49% bei der sehr heterogenen Flora sind bemerkenswert.

Die artspezifischen Modelle der logistischen Regression sind vergleichsweise kritisch zu sehen, da der verfügbare Stichprobenumfang rein verfahrenstechnisch limitierend wirkt. Trotz dieses Problems ergeben sich einige aussagekräftige Modelle aus denen weiterführende Erkenntnisse zur Relativität von Nischen gewonnen werden.

Durch den Test der Modelle auf ihre zeitliche und räumliche Übertragbarkeit, sind erste Erkenntnisse zur Validierung zu erhalten. Die Ergebnisse bestätigen die prinzipielle Übertragbarkeit und helfen, deren Grenzen zu definieren. Die zeitliche Übertragbarkeit stellt sich im Allgemeinen besser dar als die räumliche. Unbedingt bei der Modellübertragung zu beachten



ist die raum-zeitliche Relativität der Nischeninformationen. Sowohl saisonale und erhebungsjahrspezifische Unterschiede als auch Modifikationen durch differierende Erhebungsintensitäten werden beschrieben.

Die Gültigkeit der Modelle bleibt auf die Beobachtungsobjekte, eine Ähnlichkeit der abiotischen Bedingungen im Herkunfts- und Zielgebiet des Modelleinsatzes sowie die Skalenebene der Modellerstellung beschränkt.

Für den Transfer von einer segregativen, punktbezogenen Erhebung zu einer integrativen flächenbezogenen Aussage und Prognose ist die Anwendung mathematisch-statistischer Instrumentarien und die Verbindung mit den Funktionalitäten eines Geographischen Informationssystems (GIS) von entscheidender Bedeutung. Erst unter dieser Voraussetzung entsteht ein konkret in der Landschaftsplanung einsetzbares Prognosesystem, welches eine Abwägung zwischen Schutzgedanken und anderen Nutzungsinteressen ermöglicht.

Als letzter Schritt auf dem Weg zum Prognosesystem werden daher Schnittstellen zwischen den Modelltechniken und dem Geographischen Informationssystem definiert und für die Implementierung der statistischen Modelle bereitgestellt. Die Arbeit zeigt zwei Alternativen zur Einbettung: Die Modellfunktion der logistischen Regression wird in die Sachdatenbank des Geographischen Informationssystems integriert, während die direkten Ordinationsmodelle über ein eigenständiges Programm-Modul, welches mit der Datenbank kommuniziert, in das Geographische Informationssystem Eingang finden.

Das Prognosesystem wird es zukünftig vermögen die Nischeninformationen zur Beurteilung einer potentiellen Lebensraumeignung eines Standortes unter Berücksichtigung der Saisonalität zu verarbeiten und innerhalb der Modellgrenzen über eine Szenariotechnik Aussagen zu Impaktreaktionen nach anthropogenen Eingriffen zu generieren. Damit sollen Artenverluste, Änderungen der Artmächtigkeit/Häufigkeit und räumliche Verschiebungen von Habitaten im Ökosystem prognostiziert werden. Planungsszenarien sind so einer ersten Bewertung zu unterziehen.

Die vorliegende Arbeit leistet Grundlagenarbeit auf diesem Weg der Erstellung eines aue-ökologischen Prognosesystems als hochgestecktes Forschungsziel im RIVA-Verbundprojekt.

## 8 QUELLENVERZEICHNIS

- Agresti, A. (1996): *An Introduction to Categorical Data Analysis*. 290 S., New York, John Wileys & Sons.
- Altmoos, M. (1999): *Systeme von Vorranggebieten für den Tierarten-, Biotop- und Prozessschutz: Auswahlmethoden unter Einbeziehung von Habitatmodellen für Zielarten am Beispiel der Bergbaufolgelandschaften im Südraum Leipzig*. Fakultät für Physik und Geowissenschaften. Leipzig, Universität Leipzig: 252 S.
- Amarell, U. (2000): *Methodische Grundlagen zur Entwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen*. In: ATV-DVWK-Schriftenreihe. Henle, K. et al. Bonn-Hennef, Verlag für Abwasser, Abfall und Gewässerschutz 21: 185-186.
- Amler, K. et al., Eds. (1999): *Populationsbiologie in der Naturschutzpraxis*. In: *Isolation, Flächenbedarf und Biotopansprüche von Pflanzen und Tieren*. 336 S., Stuttgart, Gustav Fischer.
- Asshoff, M. (1999): *Die Erschließung und Modellierung ökologischen Wissens für das Management von Feuchtwiesenvegetation: ein Beispiel für die Aufbereitung ökologischen Wissens und den Transfer mit hybriden Expertensystemen*. 216 S., Kiel, Verein zur Förderung der Ökosystemforschung zu Kiel e.V.
- Austin, M. P. (1976): "Performance of four ordination techniques assuming three different non-linear species response models." *Vegetatio* 33: 43-49.
- Austin, M. P. (1985): "Continuum concept, ordination methods, and niche theory." *Annotated Review of Ecological Systems* 16: 39-61.
- Austin, M. P. (1987): "Models for the analysis of species response to environmental gradients." *Vegetatio* 1987 (69): 35-45.
- Austin, M. P. et al. (1984): "New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures." *Vegetatio* 1984 (55): 11-27.
- Austin, M. P. et al. (1990): *Measurement of the realized qualitative niche: environmental niches of five Eucalyptus species*. In: *Ecological Monographs*. 60 (2): 161-177.
- Austin, M. P. & I. Noy-Meir (1971): "The problem of non-linearity in ordination: experiments with two-gradient models." *Journal of Ecology* 59: 763-773.
- Backhaus, K. et al. (1990): *Multivariate Analysemethoden*. 416 S., Berlin, Springer-Verlag.
- Backhaus, K. et al. (1996): *Multivariate Analysemethoden - Eine anwendungsorientierte Einführung*. 591 S., Berlin, Heidelberg, New York, Springer-Verlag.
- Beck, J. R. & E. K. Shultz (1986): "The use of ROC curves in test performance evaluation." *Archives of Pathology and Laboratory Medicine* 110: 13-20.
- Begon, M. E., J. L. Harper & C.R. Townsend (1998): *Ökologie*. 750 S., Heidelberg, Spectrum, Akademischer Verlag.
- Birks, H. J. B. et al. (1994): *An annotated bibliography of canonical correspondence analysis and related constrained ordination methods 1986-1993*. Botanical Institute - University of Bergen, Sweden.
- BMBF (1995): *Ökologische Forschung in der Stromlandschaft Elbe (Elbe-Ökologie)*. Projektgruppe Elbe-Ökologie in der Bundesanstalt für Gewässerkunde, Bonn.
- Böhnke, R. & S. Geyer (2000): *Grundwasserdynamik in Auen-sedimenten der Mittleren Elbe*. In: *Stoffhaushalt von Auenökosystemen: Böden und Hydrologie, Schadstoffe, Bewertungen*. Friese, K. et al. Springer, Berlin: 145-152.
- Borcard, D. (1992): "Partialling out the spatial component of ecological variation." *Ecology* 73 (3): 1045-1055.
- Braak, C. J. F. ter (1987a): "The analysis of vegetation-environment relationships by canonical correspondence analysis." *Vegetatio* 69: 69-77.

- Braak, C. J. F. ter (1987b): Unimodal models to relate species to environment. Landwirtschaftliches Institut. Wageningen: 1-152.
- Braak, C. J. F. ter (1998): CANOCO Reference Manual and User's Guide to Canoco for Windows: Software for Canonical Community Ordination (version 4). Microcomputer Power, Ithaka, USA: 352 S.
- Braak, C. J. F. ter & C. W. N. Looman (1987): "Weighted averaging, logistic regression, and the Gaussian response model." *Vegetatio* 65: 3-11.
- Braak, C. J. F. ter & I. C. Prentice (1988): "A theory of gradient analysis." *Advances in Ecological Research* 18: 271-317.
- Bundesanstalt für Geowissenschaften und Rohstoffe, Ed. (1994): *Bodenkundliche Kartieranleitung*. 392 S., Hannover/Stuttgart, Schweizerbart'sche Verlagsbuchhandlung.
- Burrows, C. J. (1990). *Processes of Vegetation Change*. 551 S., London, Thomson Learning.
- Castella, E. & M. C. D. Speight (1996): "Knowledge representation using fuzzy coded variables: an example based on the use of Syrphidae (Insecta, Diptera) in the assessment of riverine wetlands." *Ecological Modelling* 85: 13-25.
- Causton, D. R. (1988): *Introduction to Vegetation Analysis*. 342 S., London, Unwin Hyman.
- Clausen, S.-E. (1998): *Applied Correspondence Analysis*. In: *Quantitative Applications in the Social Sciences*. M. S. Lewis-Beck. Thousand Oaks, London, New Dehli, Sage Publications. 121: 1-67.
- Cleveland, W. S. (1979): "Robust locally-weighted regression and smoothing scatterplots." *Journal of the American Statistical Association* 83: 596-610.
- Crawley, M. J. (1993): *GLIM for Ecologists*. 379 S., Oxford, London, Blackwell Science.
- Daget, P. & M. Godron (1982): *Analyse fréquentielle de l'écologie des espèces dans les communautés*. 163 S., Paris, Masson.
- Dargie, T. C. D. (1986): "Species richness and distortion in reciprocal averaging and detrended correspondence analysis." *Vegetatio* 65: 95-98.
- Ständige Arbeitsgruppe der Biosphärenreservate in Deutschland (1995): *Biosphärenreservate in Deutschland*. 377 S., Berlin, Springer-Verlag.
- Dierschke, H. (1994): *Pflanzensoziologie*. 683 S., Stuttgart, Gustav Fischer.
- Dister, E. (1983): *Anthropogene Wasserstandsänderungen in Flußauen und ihre ökologischen Folgen*. In: *Verhandlungen der Gesellschaft für Ökologie* 11: 89-100.
- Dolédec, S. & D. Chessel (1994): "Co-inertia analysis: an alternative method for studying species-environment relationships." *Freshwater Biology* 31: 277-294.
- DVWK - Deutscher Verband für Wasserwirtschaft und Kulturbau e.V., Ed. (1996). *Klassifikation überwiegend grundwasserbeeinflusster Vegetationstypen*. 492 S., Bonn, Wirtschafts- und Verlagsgesellschaft Gas und Wasser mbH.
- Eilertsen, O. et al. (1990a): "Data manipulation and gradient length estimation in DCA ordination." *Journal of Vegetation Science* 1: 261-270.
- Eilertsen, O. et al. (1990b): "The effects of scale range, species removal and downweighting of rare species on eigenvalue and gradient length in DCA ordination." *Journal of Vegetation Science* 12: 64-78.
- Ellenberg, H. (1952): "Auswirkung der Grundwassersenkung auf die Wiesengesellschaften am Seitenkanal westlich Braunschweig." *Angewandte Pflanzensoziologie* 6: 1-46.
- Eyre, M. D. et al. (1986): "An initial classification of the habitats of aquatic Coleoptera in north-east England." *Journal of Applied Ecology* 23: 841-852.
- Faith, D. P. et al. (1987): "Compositional turnover as a robust measure of ecological distance." *Vegetatio* 69: 57-68.

- Fielding, A. H. & P. F. Haworth (1995): "Testing the generality of bird-habitat models." *Conservation Biology* 9: 1466-1481.
- Föckler, F. et al. (2000): Weichtiergemeinschaften als Indikatoren für Wiesen- und Rinnenstandorte der Elbauen. In: Stoffhaushalt von Auenökosystemen: Böden und Hydrologie, Schadstoffe, Bewertungen. Friese, K. et al. Springer, Berlin: 391-402.
- Föckler, F. & H. Bohle (1991): Fließgewässer und ihre Auen - prädestinierte Standorte ökologischer und naturschutzfachlicher Grundlagenforschung. In: Arten- und Biotopschutzforschung für Deutschland. Henle, K. & G. Kaule. Forschungszentrum Jülich 4: 236-266.
- Fox, J. (1991): Regression Diagnostics. In: Quantitative Applications in the Social Sciences. M. S. Lewis-Beck. Newbury Park, London, New Dehli, Sage Publications. 79: 1-92.
- Fröhlich, H. & S. Grimm (1995): Punktbestimmung mit GPS für Einsteiger. 61 S., Bonn, Dümmler.
- Gauch, H. G. (1982a): Multivariate analysis in community ecology. 298 S., Cambridge, Cambridge University Press.
- Gauch, H. G. (1982b): "Noise reduction by eigenvector ordination." *Ecology* 63: 1643-1649.
- Gauch, H. G. & T. R. Wentworth (1976): "Canonical correlation analysis as an ordination technique." *Vegetatio* 33: 17-22.
- Gauch, H. G. & R. H. Whittaker (1972): "Coenocline simulation." *Ecology* 69: 135-52.
- Gauch, H. G. et al. (1977): "A comparative study of reciprocal averaging and other ordination techniques." *Journal of Ecology* 65: 157-174.
- Gerken, B. (1988): Auen - verborgene Lebensadern der Natur. 131 S., Freiburg, Rombach.
- Gerken, B. (1989): "Zur Bedeutung der Laufkäfer (Coleoptera: Carabidae) als Bioindikatoren." *Egge-Weser* 6: 5-11.
- Gigon, A. (1983): Welches ist der wichtigste Standortsfaktor für die floristischen Unterschiede zwischen benachbarten Pflanzengesellschaften? In: Verhandlungen der Gesellschaft für Ökologie 11: 145-160.
- Gittins, R. (1979): Ecological applications of canonical analysis. In: Multivariate Methods in Ecological Work. L. Orlóci et al. Burtonsville, Intl. Cooperative Pub House: 309-535.
- Glavac, V. (1996): Vegetationsökologie. 358 S., Jena, Gustav Fischer.
- Goldschmidt, B. (2000): Naturschutzbezogenes Monitoring in grünlandgenutzten Niedermoorlandschaften Nordostdeutschlands am Beispiel des Gartzzer Bruchs und des Niederoderbruchs. In: Eberswalder wissenschaftliche Schriften. Eberswalde, Universität Eberswalde. 4: 1-382.
- Gower, J. (1987): Introduction to ordination techniques. In: Developments in Numerical Ecology. P. Legendre & L. Legendre. Berlin, Springer-Verlag: 3-64.
- Guisan, A. et al. (1999): "GLM versus CCA spatial modeling of plant species distribution." *Plant Ecology* 143: 107-122.
- Hastie, T. & R. Tibshirani (1990): Generalized Additive Models. 384 S., London, Chapman & Hall.
- Hauhs, M. & H. Lange (1996): "Mensch und Umwelt. Eine komplexe Beziehung als interdisziplinäre Herausforderung." *Forum für interdisziplinäre Forschung* 16: 45-64.
- Hauhs, M. et al. (1996): "Summary of a workshop on ecosystem modelling: The end of an era?" *The Science of the Total Environment* 183: 1-5.
- Hill, M. O. (1974): "Correspondence analysis: a neglected multivariate method." *Journal of Royal Statistical Society Series C* 23: 340-54.
- Hill, M. O. (1991): "Patterns of species distribution in Britain elucidated by canonical correspondence analysis." *Journal of Biogeography* 18: 247-255.
- Hill, M. O. et al. (1991): TWINSpan-Manual. Cornell University, Ithaka, USA: 49 S.

- Hill, M. O. & H. G. Gauch (1980): "Detrended correspondence analysis, an improved ordination technique." *Vegetatio* 42: 47-58.
- Hirschfeld, H. O. (1935): A connection between correlation and contingency. In: *Proceedings of the Cambridge Philosophical Society*. 31: 520-524.
- Hosmer, D. W. & S. Lemeshow (1989): *Applied Logistic Regression*. 307 S., New York, John Wileys & Sons.
- Hovestadt, T. et al. (1991): Flächenbedarf von Tierpopulationen. In: *Berichte aus der ökologischen Forschung, Jülich, Selbstverlag Forschungszentrum Jülich* 1: 1-279.
- Hügin, G. (1981): "Die Auenwälder des südlichen Oberrheintals - Ihre Veränderung und Gefährdung durch den Rheinausbau." *Landschaft und Stadt* 13: 78-91.
- Hügin, G. (1990): Die Mooswälder der Freiburger Bucht. In: *Beihefte zu den Veröffentlichungen für Naturschutz und Landschaftspflege Baden-Württemberg* 29: 1-88.
- Hügin, G. & A. Henrichfreise (1992): Vegetation und Wasserhaushalt des rheinnahen Waldes. In: *Schriftenreihe für Vegetationskunde der Bundesforschungsanstalt für Naturschutz und Landschaftsökologie* 24: 1-48.
- IKSE (1994): Ökologische Studie zum Schutz und zur Gestaltung der Gewässerstrukturen und der Uferlandregionen an der Elbe. 10 S., Magdeburg, Internationale Kommission zum Schutz der Elbe.
- Isaaks, E. H. & R. M. Srivastava (1989): *Applied Geostatistics*. 561 S., New York, Oxford University Press.
- Jaccard, J. et al. (1990): Interaction Effects in Multiple Regression. In: *Quantitative Applications in the Social Sciences*. M. S. Lewis-Beck. Newbury Park, London, New Dehli, Sage Publications. 72: 1-94.
- Jongman, R. H. G. et al., Eds. (1987): *Data Analysis in Community and Landscape Ecology*. 299 S., Pudoc, Centre for Agricultural Publishing and Documentation.
- Jørgensen, S. E. (1994): *Fundamentals of Ecological Modelling*. 628 S., Amsterdam, Elsevier.
- Jørgensen, S. E. & F. Müller, Eds. (2000): *Handbook of Ecosystem Theories and Management*. 584 S., Boca Raton London New York, Lewis Publishers.
- Jurisch, R. (1998): Untersuchungen der Genauigkeiten von Tachymeter- und DGPS-Ortungen zur Ermittlung hydraulischer und hydrologischer Daten in Flüssen. In: *Mitteilungen der Bundesanstalt für Wasserbau* 77: 43-53.
- Kahlenborn, W. & R. A. Kraemer (2000): "Prinzipien einer nachhaltigen Wasserwirtschaft und ihre Anwendung in Deutschland." *Wasser & Boden* 52 (1+2): 36-39.
- Kalton, G. (1983): Introduction to Survey Sampling. In: *Quantitative Applications in the Social Sciences*. M. S. Lewis-Beck. Newbury Park, London, New Dehli, Sage Publications. 35: 1-96.
- Keddy, P. (1993): "Do ecological communities exist? A reply to Bastow Wilson." *Journal of Vegetation Science* 4: 135-136.
- Kenkel, N. C. et al. (1989): "On Sampling procedures in population and community ecology." *Vegetatio* 83: 195-207.
- Kenkel, N. C. & L. Orlóci (1986): "Applying metric and non-metric multidimensional scaling to ecological studies: some new results." *Ecology* 67: 919-928.
- Kent, M. & P. Coker (1992): *Vegetation Description and Analysis*. 363 S., London, John Wileys & Sons.
- Kiener, J. (1984): Veränderung der Auenvegetation durch die Anhebung des Grundwasserspiegels im Bereich der Staustufe Ingolstadt. In: *Berichte der Akademie für Naturschutz und Landschaftspflege, Laufen/Salzach* 8: 104-129.
- Kleyer, M. et al. (2000): "Habitatmodelle für Tierarten: Entwicklung, Methoden und Perspektiven für die Anwendung." *Zeitschrift für Ökologie und Naturschutz* 8 (1999/2000): 177-194.
- Knoke, D. & P. J. Burke (1980): Log-Linear Models. In: *Quantitative Applications in the Social Sciences*. M.-S. Lewis-Beck. Newbury Park, London, New Dehli, Sage Publications 20: 1-80.



- Knox, R. G. (1989): "Effects of detrending and rescaling on correspondence analysis: solution stability and accuracy." *Vegetatio* 83: 129-136.
- Kobus, H. (1995): Prognoseinstrumente und Meßdatenrealität in der Grundwasserwirtschaft. Ergebnisse aus der 50. Sitzung der Kommission am 10. und 11. März 1994. In: *Perspektiven der Wasserforschung*, Weinheim, VCH Verlagsgesellschaft. 14: 133-149.
- Köhler, W. et al. (1996): *Biostatistik*. 258 S., Berlin/Heidelberg, Springer-Verlag.
- Kuntze, H. e. a. (1994): *Bodenkunde*. 424 S., Stuttgart, Gustav Fischer.
- Lahiti, T. & R. A. Väisänen (1987): "Ecological gradientsof boreal forestsin South Finland: an ordination test of Cajander's forest site type theory." *Vegetatio* 68: 145-156.
- Legendre, P. (1993): "Spatial autocorrelation: trouble or new paradigm?" *Ecology* 74: 1659-1673.
- Legendre, P. & M.-J. Fortin (1989): "Spatial pattern and ecological analysis." *Vegetatio* 80: 107-138.
- Legendre, P. & L. Legendre (1998): *Numerical Ecology*. 853 S., Amsterdam, Elsevier.
- Lek, S. & J. F. Guégan (1999): "Artificial neural networks as a tool in ecological modelling, an introduction." *Ecological Modelling* 120: 65-73.
- Leland, J. J. et al. (2000): "An Introduction to the Practice of Ecological Modelling." *BioScience* 50 (8): 694-706.
- Liebetrau, A. M. (1983): Measures of Assoziation. In: *Quantitative Applications in the Social Sciences*. M. S. Lewis-Beck. Newbury Park, London, New Dehli, Sage Publications 32: 1-94.
- Maarel van der, E. (1969): "On the use of ordination models in phytosociology." *Vegetatio* 19: 1-46.
- Maarel van der, E. (1979): "Transformation of cover-abundance values in phytosociology and its effects on community similarity." *Vegetatio* 39: 97-114.
- Manel, S. et al. (1999): "Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird." *Ecological Modelling* 120: 337-347.
- Marschner, H. (1985): Nährstoffdynamik in der Rhizosphäre: Eine Übersicht. In: *Berichte der Deutschen Botanischen Gesellschaft* 98: 291-309.
- McCullagh, P. & J. A. Nelder (1997): *Generalized linear models*. 511 S., London, Chapman & Hall.
- McCune, B. (1997): "Influence of noisy environmental data on canonical correspondence analysis." *Ecology* 78: 2617-2623.
- Meisel, K. (1977). Die Grünlandvegetation nordwestdeutscher Flußtäler und die Eignung der von ihr besiedelten Standorte für einige wesentliche Nutzungsansprüche. In: *Schriftenreihe Vegetationskunde* 11: 1-121.
- Menard, S. (1995): *Applied Logistic Regression Analysis*. In: *Quantitative Applications in the Social Sciences*. M. S. Lewis-Beck. Thousand Oaks London New Dehli, Sage Publications 106: 1-98.
- Méot, A. et al. (1998): "Partialling out the spatial component of ecological variation: questions and propositions in the linear modelling framework." *Environmental and Ecological Statistics* 5: 1-27.
- Mercier, P. et al. (1992): "Complete correspondence analysis of an ecological profile data table: a central ordination method." *Acta Ecologica* 13(1): 25-44.
- Meyenburg, G. (1999): Heterogenität von Auenböden: Vergleich zweier Transekte durch ein Untersuchungsgebiet im Biosphärenreservat Mittlere Elbe. In: *Hamburger Bodenkundliche Arbeiten* 44: 108-113.
- Meyer, F. H. (1957). Über Wasser- und Stickstoffhaushalt der Röhrichte und Wiesen im Elballuvium bei Hamburg. In: *Mitteilungen aus dem Staatsinstitut für allgemeine Botanik in Hamburg* 11: 137-203.
- Minchin, P. R. (1987): "An evaluation of the relative robustness of techniques for ecological ordination." *Vegetatio* 69: 89-107.

- Mooney, C. Z. (1997): Monte Carlo Simulation. In: Quantitative Applications in the Social Sciences. M. S. Lewis-Beck. Newbury Park London New Dehli, Sage Publications 116: 1-101.
- Morrison, M. L. et al. (1998): Wildlife-habitat Relationships - Concepts and Applications. 435 S., Madison, Wisconsin, The University of Wisconsin Press.
- Mühlenberg, M. (1993): Freilandökologie. 512 S., Heidelberg, Wiesbaden, Gustav Fischer.
- Müller, F. (1997): Typen, Aufbau und Bewertung von Umwelt-Systemmodellen. 51. Deutscher Geographentag, Bonn, Internet: <http://giub.uni-bonn.de/internetcafe/tagungsband/>.
- Müller-Westermeier, G. (1996): Klimadaten der Bundesrepublik Deutschland. Zeitraum 1961-1990. 289 S., Offenbach, Selbstverlag des Deutschen Wetterdienstes.
- Namboodiri, K. (1984): Matrixalgebra. In: Quantitative Applications in the Social Sciences. M. S. Lewis-Beck. Newbury Park, London, New Dehli, Sage Publications. 38: 1-96.
- Nienhuis, P. H. et al., Eds. (1998): New Concepts for Sustainable Management of River Basins. 374 S., Leiden, Backhuys.
- Økland, R. H. (1986a): "Rescaling of ecological gradients. I. Calculation of ecological distance between vegetation stands by means of their floristic composition." *Nordic Journal of Botany* 6 (5): 651-660.
- Økland, R. H. (1986b): "Rescaling of ecological gradients. II. The effect of scale on symmetry of species response curves." *Nordic Journal of Botany* 6 (5): 661-669.
- Økland, R. H. (1986c): "Rescaling of ecological gradients. III. The effect of scale and niche breadth measurements." *Nordic Journal of Botany* 6 (5): 671-677.
- Økland, R. H. (1990): *Vegetation Ecology: Theory, Methods, and Applications with Reference to Fennoscandia*. 233 S., Oslo, Botanical Garden and Museum, University of Oslo.
- Økland, R. H. (1996): "Are ordination and constrained ordination alternative or complementary strategies in general ecological studies?" *Journal of Vegetation Science* 7: 289-292.
- Økland, R. H. & O. Eilertsen (1994): "Canonical Correspondence Analysis with variation partitioning: some comments and an application." *Journal of Vegetation Science* 5: 117-126.
- Oksanen, J. (1983): "Ordination of boreal heath-like vegetation with principal component analysis, correspondence analysis and multidimensional scaling." *Vegetatio* 52: 181-189.
- Oksanen, J. (1988): "A note of the occasional instability of detrending in correspondence analysis." *Vegetatio* 74: 29-32.
- Orlói, L. (1978): *Multivariate Analysis in Vegetation Research*. 451 S., The Hague, Kluwer Academic Press.
- Palmer, M. (1998): Ordination methods for ecologists, Internet: <http://okstate.edu/artsci/botany/ordinate/>.
- Palmer, M. W. (1993): "Putting things in even better order: the advantages of canonical correspondence analysis." *Ecology* 8: 2215-2230.
- Palmer, M. W. (1994): "Scale dependence and the species area relationship." *The American Naturalist* 5: 717-740.
- Palmer, M. W. (1995): "How should one count species?" *Natural Areas Journal* 15: 124-135.
- Palmer, M. W. & P. M. Dixon (1990): "Small-scale environmental heterogeneity and the analysis of species distributions along gradients." *Journal of Vegetation Science* 1: 57-65.
- Palmer, M. W. & T. A. Maurer (1997): "Does diversity beget diversity? A case study of crops and weeds." *Journal for Vegetation Science* 8: 235-240.
- Palmer, M. W. & P. S. White (1994): "On the existence of ecological communities." *Journal of Vegetation Science* 5: 279-282.
- Pearce, J. & S. Ferrier (2000a): "Evaluating the predictive performance of habitat models developed using logistic regression." *Ecological Modelling* 133: 225-245.

- Pearce, J. & S. Ferrier (2000b): "An evaluation of alternative algorithms for fitting species distribution models using logistic regression." *Ecological Modelling* 128: 127-147.
- Peet, R. K. et al. (1988): "Putting things in order: the advantages of detrended correspondence analysis." *The American Naturalist* 131: 924-934.
- Peeters, E. T. H. M. & J. J. P. Gardeniers (1998): "Logistic Regression as a tool for defining habitat requirements of two common gammarids." *Freshwater Biology* 39: 605-615.
- Peter, W. et al. (1999): Dynamik natürlicher und anthropogener Systeme und ihre Wechselwirkungen - Überprüfung der Höhengenaugigkeit eines Digitalen Geländemodells im Deichvorland der Mittellage mit einem differentiellen globalen Positionierungssystem (DGPS). Tagungsband Geoökon 99, Braunschweig, Institut für Geographie und Geoökologie der TU Braunschweig: 53-57.
- Petermeier, A. et al. (1996): Die ökologische und biologische Entwicklung der deutschen Elbe. In: Lauterbornia. Dinkelscherben, Erik Mauch Verlag 24: 1-95.
- Peterson, J. (1998): Rote Liste der gefährdeten Biotoptypen des Landes Sachsen-Anhalt. In: Berichte des Landesamtes für Umweltschutz Sachsen-Anhalt 30: 1-76.
- Pfadenhauer, J. (1993): Vegetationsökologie - ein Skriptum. 301 S., Eching, IHW-Verlag.
- Pickett, S. T. et al. (1994): Ecological Understanding (The Nature of Theory and the Theory of Nature). 206 S., San Diego, Academic Press.
- Pielou, E. C. (1977): Mathematical Ecology. 385 S., New York, John Wiley & Sons.
- Pielou, E. C. (1984): The Interpretation of Ecological Data. A primer on classification and ordination. 263 S., New York, John Wileys & Sons.
- Press, W. H. et al. (1986): Numerical Recipes - The Art of Scientific Computing. 818 S., Cambridge, Cambridge University Press.
- Reed, R. A. et al. (1993): "Scale dependence of vegetation-environment correlations: A case study of North Carolina piedmont woodland." *Journal of Vegetation Science* 4: 329-340.
- Reiter, K. & K. Fussenegger (1997): Verortung von Vegetationsmonitoringflächen mit GPS und Laserfeldstecher. Tagungsband AGIT, Salzburg, Institut für Geographie der Universität Salzburg.
- Richter, O. (1985): Simulation des Verhaltens ökologischer Systeme - Mathematische Modelle und Methoden. 219 S., Weinheim, VCH Verlagsgesellschaft.
- Richter, O. et al. (1997): Koppelung Geographischer Informationssysteme (GIS) mit ökologischen Modellen im Naturschutzmanagement. In: Geographische Informationssysteme im Naturschutz: Forschung, Planung, Praxis. Suhling, F. & R. Kralz, Magdeburg, Westarp Verlag: 5-29.
- Richter, O. & D. Söndgerath (1990): Parameter Estimation in Ecology. 218 S., Weinheim, VCH Verlagsgesellschaft.
- Rijf van de, C. W. C. J. (1996): "Vegetation zonation in a former tidal area: a vegetation-type response model based on DCA and logistic regression using GIS." *Journal of Vegetation Science* 7: 505-518.
- Rink, M. (1997a): Multivariate statistische Auswertung und Klassifizierung von Bodendaten aus dem Projekt Vynen/Rees. Unveröffentlichtes Gutachten vom 30.06.1997: M3/435.062/4360, 24. S., Bundesanstalt für Gewässerkunde, Koblenz.
- Rink, M. (1997b): Prüfung der Robustheit und modellhafte Gesamtbetrachtung des Ökosystems Elbaue in den Untersuchungsgebieten - Aufbau einer geeigneten Struktur der benötigten Datensammlung, Zusammenstellung von Fremddaten, Sichtung und Aufbereitung der Daten für die vergleichende statistische Auswertung. Unveröffentlichtes Gutachten vom 04.03.1997: M3/435.062/1442, 76. S., Bundesanstalt für Gewässerkunde, Koblenz.
- Rink, M. (1997c): Synökologische Untersuchung ausgewählter Pflanzengesellschaften am Niederrhein mit der kanonischen Korrespondenzanalyse - eine methodenkritische Ausarbeitung. Unveröffentlichte Diplomarbeit am Geographisches Institut der Johannes Gutenberg-Universität Mainz: 179 S.

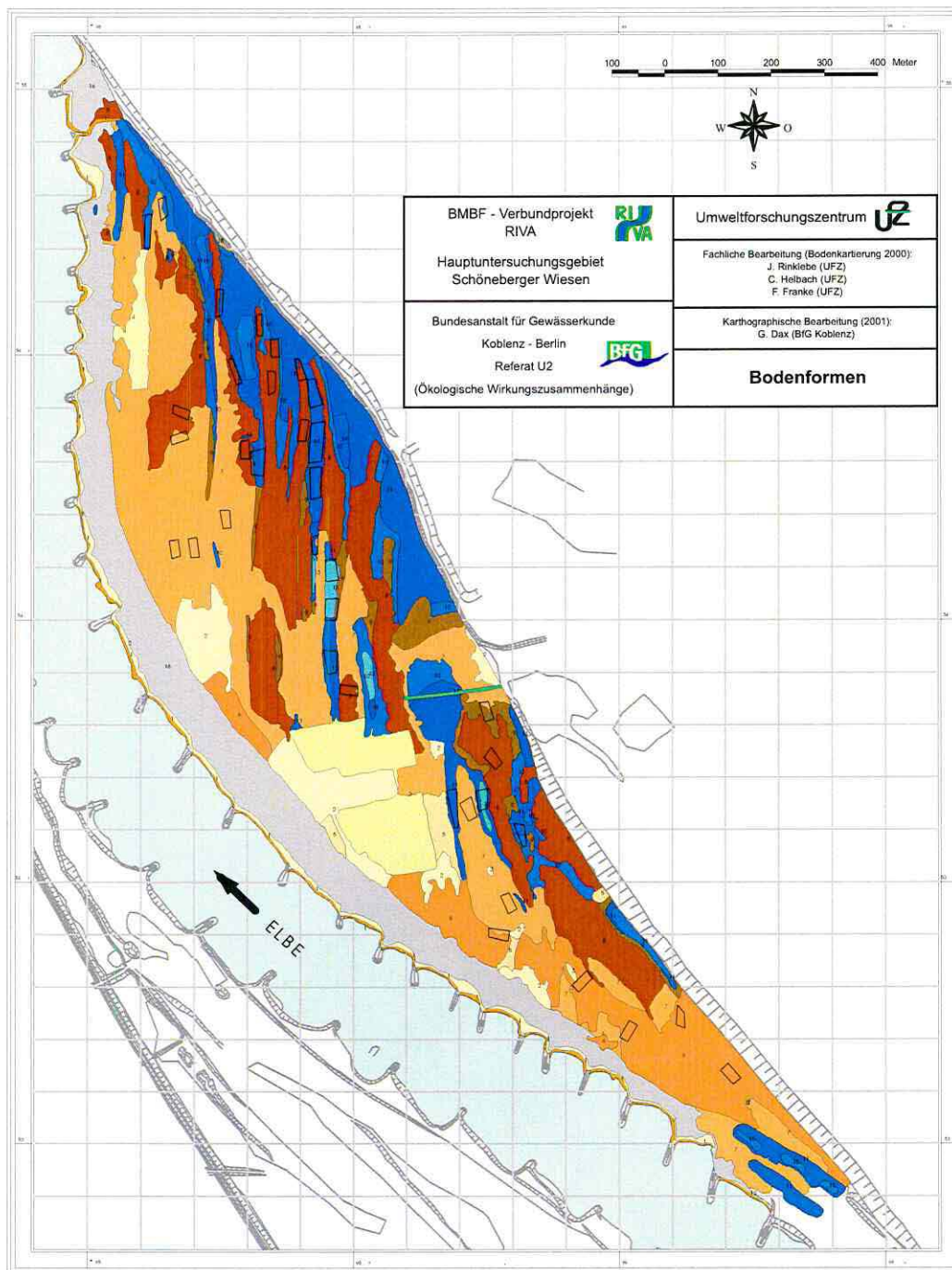
- Rink, M. (1998): Entwicklung eines Ökosystemmodells Elbaue zur Ableitung von Auenindikationsystemen durch den Einsatz multivariater statistischer Verfahren. Unveröffentlichter Bericht vom 30.04.1998: M3/323.04/2717, 66. S., Bundesanstalt für Gewässerkunde, Koblenz.
- Rink, M. (1999): Entwicklung eines Ökosystemmodells Elbaue zur Ableitung von Auenindikationsystemen durch den Einsatz multivariater statistischer Verfahren. Unveröffentlichter Abschlußbericht vom 08.06.1999: U2/323.04/4101, 26. S., Bundesanstalt für Gewässerkunde, Koblenz.
- Rink, M. et al. (2000): "Zur Erstellung einer fachlich-statistisch abgestimmten Datenerhebungsstrategie am Beispiel eines synökologisch orientierten Forschungsprojektes in den Elbauen." *Hydrologie und Wasserbewirtschaftung* 4: 184-190.
- Rinklebe, J. et al. (2000): Großmaßstäbige bodenkundliche Kartierung im Biosphärenreservat Mittlere Elbe. In: *Stoffhaushalt von Auenökosystemen: Böden und Hydrologie, Schadstoffe, Bewertungen*. Friese, K. et al. Springer, Berlin: 27-35.
- Roberts, D. W. (1986): "Ordination on the basis of fuzzy set theory." *Vegetatio* 66: 123-131.
- Robertson, G. P. (1998): *Manual GS<sup>+</sup>: Geostatistics for the Environmental Sciences*. 152. S., Plainwell USA, Gamma Design Software.
- Rödel, D. (1987). Vegetationsentwicklungen nach Grundwasserabsenkungen, dargestellt am Beispiel des Fuhrberger Feldes in Niedersachsen. In: *Landschaftsentwicklung und Umweltforschung, Schriftenreihe Fachbereich Landschaftsentwicklung der TU Berlin* S1: 1-245.
- Rönz, B., Ed. (1994): *Lexikon Statistik*. 438 S., Wiesbaden, Th. Gabler.
- Runge, F. (1980). *Die Pflanzengesellschaften Mitteleuropas*. 158 S., Münster, Aschendorff.
- Schanowski, A. & W. Figura (2000): Methodische Grundlagen zur Entwicklung eines robusten Indikationsystems für ökologische Veränderungen in Auen. In: *ATV-DVWK-Schriftenreihe*. Henle, K. et al. Bonn-Hennef, Verlag für Abwasser, Abfall und Gewässerschutz 21: 187-190.
- Scheiner, S. M. & J. Gurevitch, Eds. (1993): *Design and Analysis of Ecological Experiments*. 445 S., London, Chapman & Hall.
- Scholle, D. (1997): *GIS-gestützte Zusammenführung ökologischer Daten*. 201 S., Kiel, Verein zur Förderung der Ökosystemforschung zu Kiel e.V.
- Schröder, B. (2000): Zwischen Naturschutz und Theoretischer Ökologie: Modelle zur Habitateignung und räumlichen Populationsdynamik für Heuschrecken im Niedermoor. Institut für Geographie und Geoökologie. Braunschweig, Technische Universität Braunschweig: 1-202.
- Schröder, B. & O. Richter (2000): "Are habitat models transferable in space and time?" *Zeitschrift für Ökologie und Naturschutz* 8: 195-205.
- Schubert, R. & G. Wagner (1988): *Botanisches Wörterbuch*. 582 S., Stuttgart, Gustav Fischer.
- Schulze, H., Ed. (1982): *Alexander Weltatlas*. 223 S., Stuttgart, Klett.
- Sebald, O. et al., Eds. (1993): *Die Fam- und Blütenpflanzen Baden-Württembergs in acht Bänden*. Stuttgart, Ulmer.
- Seeber, G. (1996): "Grundprinzipien zur Vermessung mit GPS." *Der Vermessungsingenieur* 2: 53-64.
- Seibert, P. (1975). "Veränderung der Auevegetation nach Anhebung des Grundwasserspiegels in den Donauauen bei Offingen." In: *Beiträge zur Naturkundlichen Forschung Südwestdeutschlands* 34: 329-343.
- Siebeck, O. (1995): Zusammenführung von Hydrologie und Ökologie in der Umweltforschung. Welche Ökologie ist gefragt? Ergebnisse aus der 50. Sitzung der Kommission am 10. und 11. März 1994. In: *Perspektiven der Wasserforschung*. Weinheim, VCH Verlagsgesellschaft 14: 182-203.
- Sieber, H. (1982): *Mathematische Begriffe und Formeln*. 39 S., Stuttgart, Klett.
- Smilauer, P. (1992): *Cano Draw 3.00 Lite - User's Guide*. 118 S., University of Budevice, Tschechei.
- Sokal, R. R. & F. J. Rohlf (1995): *Biometry*. 887 S., New York, Freeman W. H. and Company.

- Sokal, R. R. & F. J. Rohlf (1996): *Introduction to Biostatistics*. 363 S., New York, Freeman W. H. and Company.
- Spang, W.D. (1995): Die Eignung von Regenwürmern (Lumbricidae), Schnecken (Gastropoda) und Laufkäfern (Carabidae) als Indikatoren für autotypische Standortbedingungen. In: *Heidelberger Geographische Arbeiten* 102: 1-236.
- Spink, A. et al. (1998). "Nutrient dynamics of Large River Floodplains." *Regulated Rivers* 14: 203-216.
- StatSoft (1997): *Statistica Online-Manual*. Tulsa, StatSoft Incorporation.
- Steinhardt, U. (1998): Dimensionen und Skalen in der Geoökologie - Theoretische Grundlagen und praktische Relevanz. *Forum der Geoökologie* 9 (3): 8-11.
- Stockwell, D. & D. Peters (1999): "The GARP modelling system: problems and solutions to automated spatial prediction." *International Journal for Geographical Information Science* 2: 143-158.
- Stohlgren, T. J. et al. (1995): "A Modified-Whittaker nested vegetation sampling method." *Vegetatio* 117: 113-121.
- Swan, J. M. A. (1970): "An examination of some ordination problems by use of simulated vegetation data." *Ecology* 51: 89-102.
- Thiele, H.-U. (1977): *Carabid Beetles and their Environment. A Study on Habitat Selection by Adaptions in Physiology and Behavior*. 369 S., Berlin, New York, Springer-Verlag.
- Thioulouse, J. et al. (1995): "Multivariate analysis of spatial patterns: a unified approach to local and global structures." *Environmental and Ecological Statistics* 2: 1-14.
- Thioulouse, J. et al. (1997): "ADE-4: a multivariate analysis and graphical display software." *Statistics and Computing* 7 (1): 75-83.
- Thioulouse, J. & J. R. Lobry (1994): Relationships between ferrisol properties and the structure of plant parasitic nematode communities on sugarcane in Martinique. *Acta Ecologica*. 15: 767-780.
- Thioulouse, J. & J. R. Lobry (1995): "Co-inertia analysis of amino-acid physico-chemical properties and protein composition with the ADE package." *Computer Applications in the Biosciences* 11: 321-329.
- Tüxen, R. (1954). "Pflanzengesellschaften und Grundwasserganglinien." *Angewandte Pflanzensoziologie* 8: 64-98.
- Tüxen, R. (1970): "Bibliographie zum Problem des Minimum-Areals und der Art-Arealkurve." *Excerpta Botanica* B10: 291-314.
- UFZ (1999): *RIVA-Zwischenbericht Übertragung und Weiterentwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen (RIVA) für den Projektzeitraum 9/97 bis 11/98*. Umweltforschungszentrum Leipzig/Halle: 166 S., unveröffentlicht.
- UFZ (2000): *Übertragung und Weiterentwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen (RIVA): Sachstandsbericht für den Projektzeitraum 2/99 bis 3/00*. Umweltforschungszentrum Leipzig/Halle: 117 S., unveröffentlicht.
- UNESCO (1998): *Das UNESCO-Programm: Der Mensch und die Biosphäre*. Jahrbuch 97/98 der deutschen UNESCO-Kommission, Bonn: 147 S.
- Vincent, P. J. & J. M. Haworth (1983): "Poisson regression models of species abundance." *Journal of Biogeography* 10: 153-160.
- Vogel, K. et al. (1996): "Einsatz von Zielarten im Naturschutz - Auswahl der Arten, Methode von Populationsgefährdungsanalyse und Schnellprognose, Umsetzung in der Praxis." *Natur- und Landschaftsplanung* 28: 179-184.
- Walter, H. (1986): *Allgemeine Geobotanik*. 279 S., Stuttgart, Ulmer.
- Walters, C. & J. Korman (1999): "Cross-scale modeling of riparian ecosystem responses to hydrologic management." *Ecosystems* 2: 411-421.
- Wartenberg, D. et al. (1987): "Putting things in order: a critique of detrended correspondence analysis." *The American Naturalist* 129: 434-448.



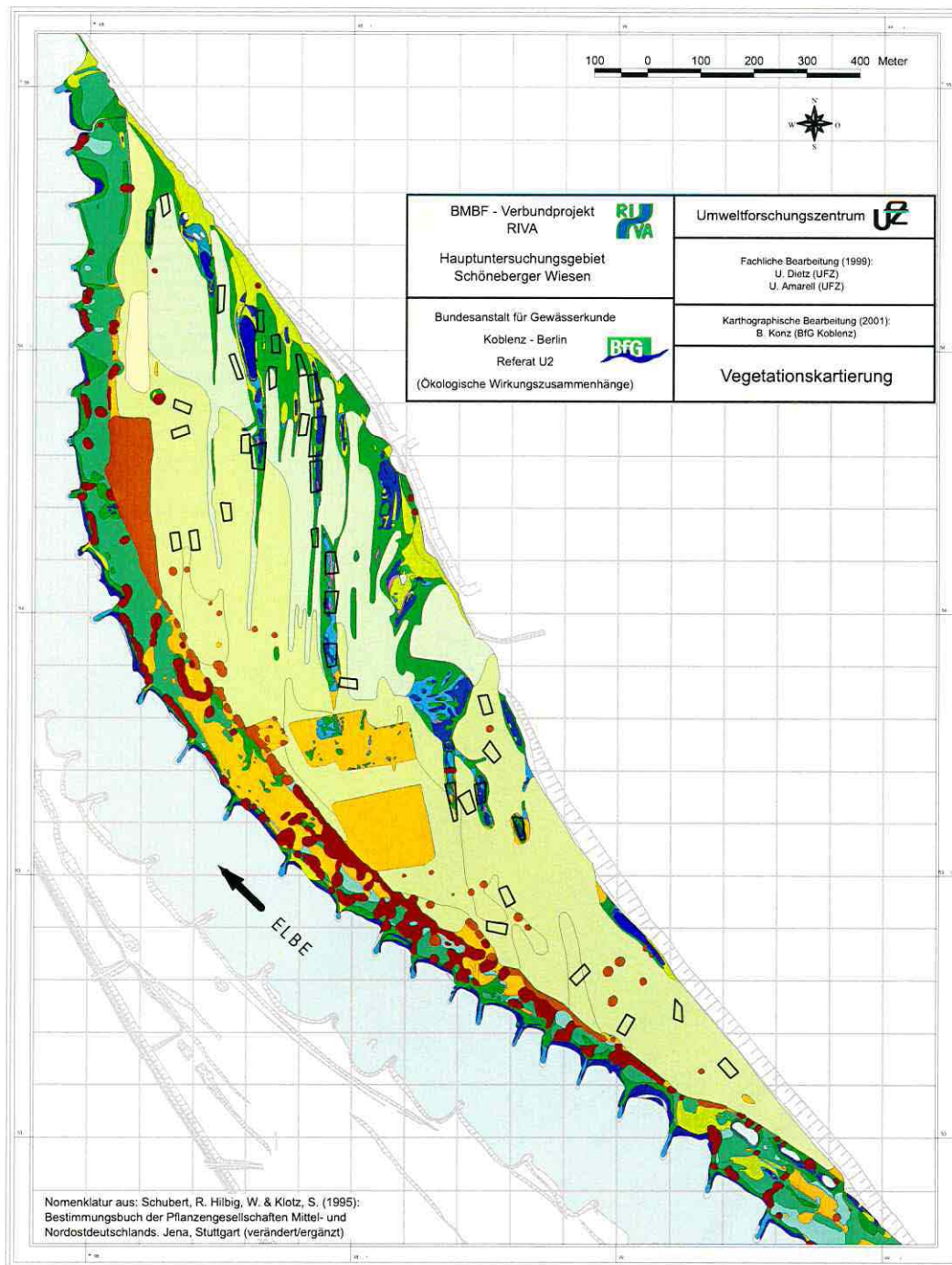
- Weiner, J. (1995): "On the practice of ecology." *Journal of Ecology* 83: 153-158.
- Wenkel, K.-O. (1997): Modellierung von Umweltveränderungen in unterschiedlichen räumlichen und zeitlichen Dimensionen. 51. Deutscher Geographentag, Bonn, Internet: <http://giub.uni-bonn.de/internetcafe/tagungsband/>.
- Whittaker, R. H. (1967): "Gradient analysis of vegetation." *Biological Reviews* 42: 207-264.
- Whittaker, R. H. (1973): Ordination and classification of communities. In: *Handbook of Vegetation Sciences*, The Hague, Junk 5: 1-737.
- Wildi, O. (1986): Analyse vegetationskundlicher Daten. Theorie und Einsatz statistischer Methoden. In: *Veröffentlichungen des Geobotanischen Institutes, ETH Stift, Rübel, Zürich* 90: 1-226.
- Wilmanns, O. (1989): *Ökologische Pflanzensoziologie*. 378 S., Heidelberg, Wiesbaden, Gustav Fischer.
- Wissel, C. (1989): *Theoretische Ökologie. Eine Einführung*. 299 S., Berlin, Heidelberg, New York, Springer-Verlag.
- Zar, J. H. (1996): *Biostatistical Analysis*. 662 S., Upper Saddle River, New Jersey, Simon&Schuster/A Viamcom Company.

# ANLAGEN



Anlage 1

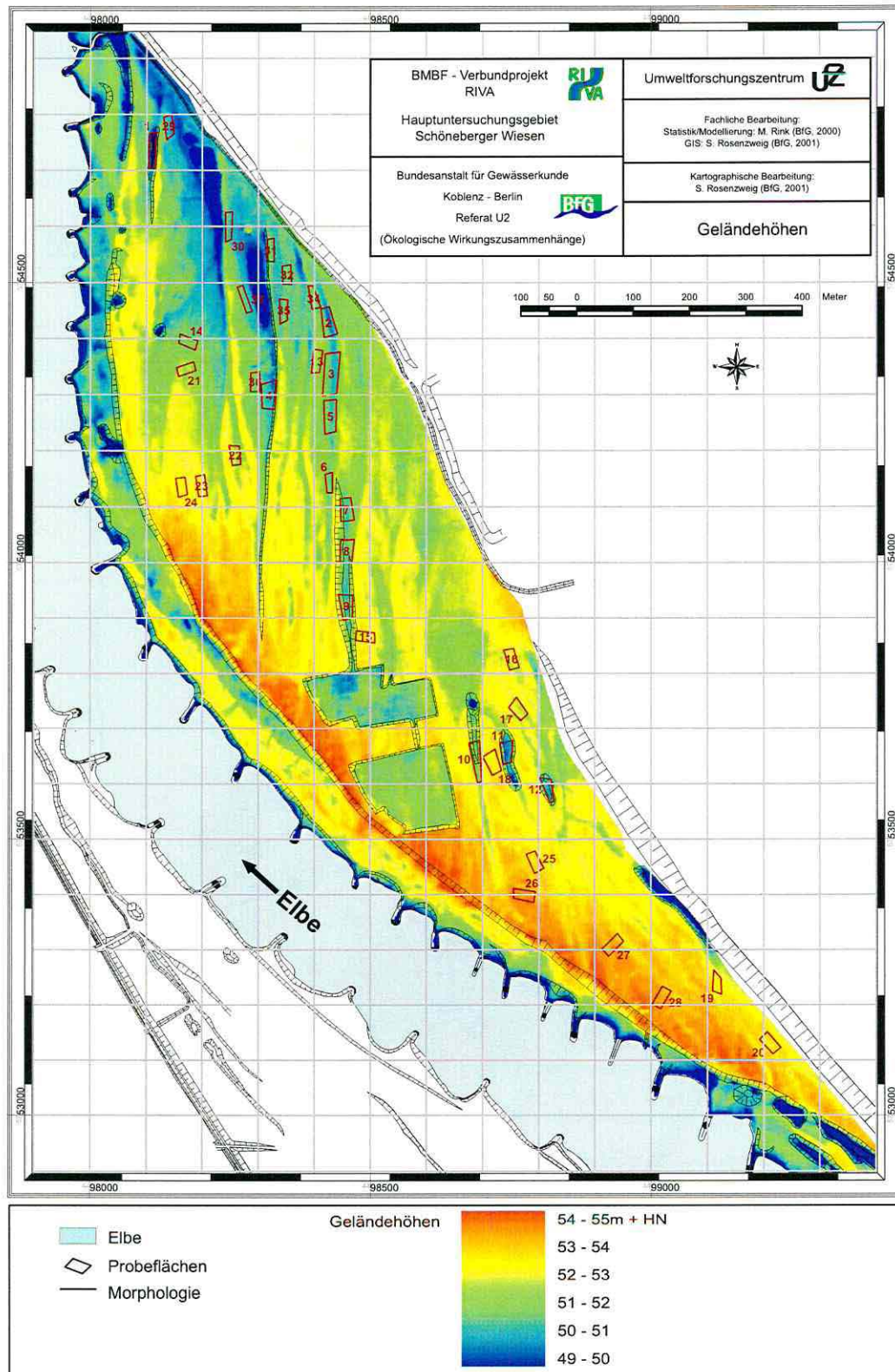
Elbe	6 Vega aus Auenlehmsand bis Auensandlehm über (tiefem) Auensand	13 Auengley aus flachem Auenlehm über Auensand über Auentonschluff
Probeflächen	7 Vega aus Auenormallehm über (tiefem) Auensand	14 Pelosol-Gley aus Auenton
Bodengesellschaft von Sapropel und Rambla aus Auensand in Wechsellagerung mit Organomudde	8 Vega aus Auentonschluff über (tiefem) Auensand	15 Sapropel
1 Paternia aus Auensand	9 Gley-Vega aus Auenlehm	16 (vergleyle) Tschernitza aus Auentonschluff über (tiefem) Auensand
2 Paternia aus Auesand bis Auenormallehm	10 Gley-Vega aus Auenlehm über (tiefem) Auensand	17 Weg mit Anthroskettlehm befestigt
3 Gley-Paternia aus Auesand bis Auenormallehm	11 (anmooriger) Auengley aus Auentonschluff bis Auenlehm über (tiefem) Auensand	
4 Gley-Paternia aus kiesführendem Auensand über geröllführendem Auenkiessand	12 Auengley aus Auensand über tiefem Auentonschluff	
5 Bodengesellschaft von Paternia und Vega aus Auenlehmsand bis Auensandlehm über (tiefem) Auensand		



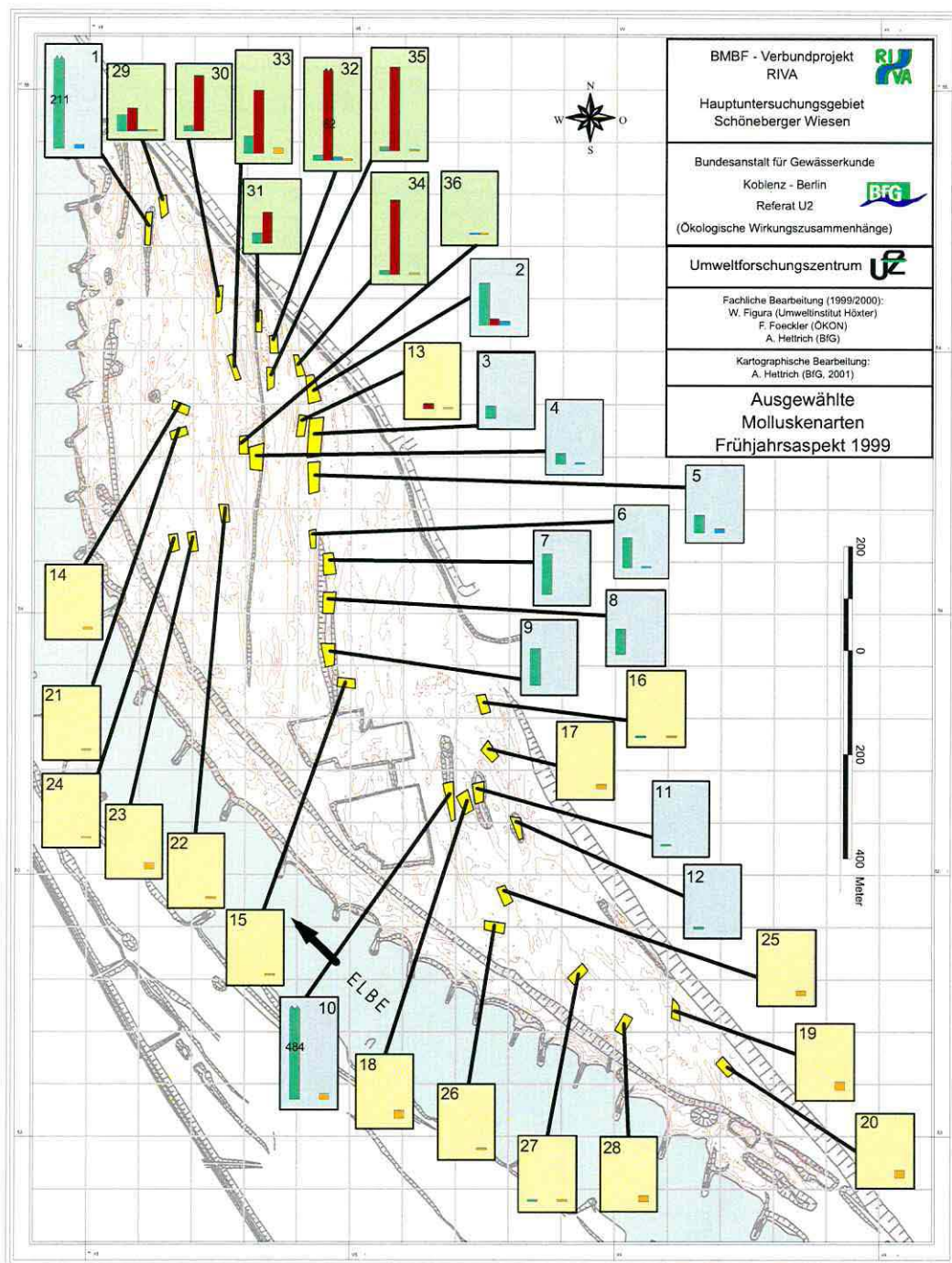
Anlage 2

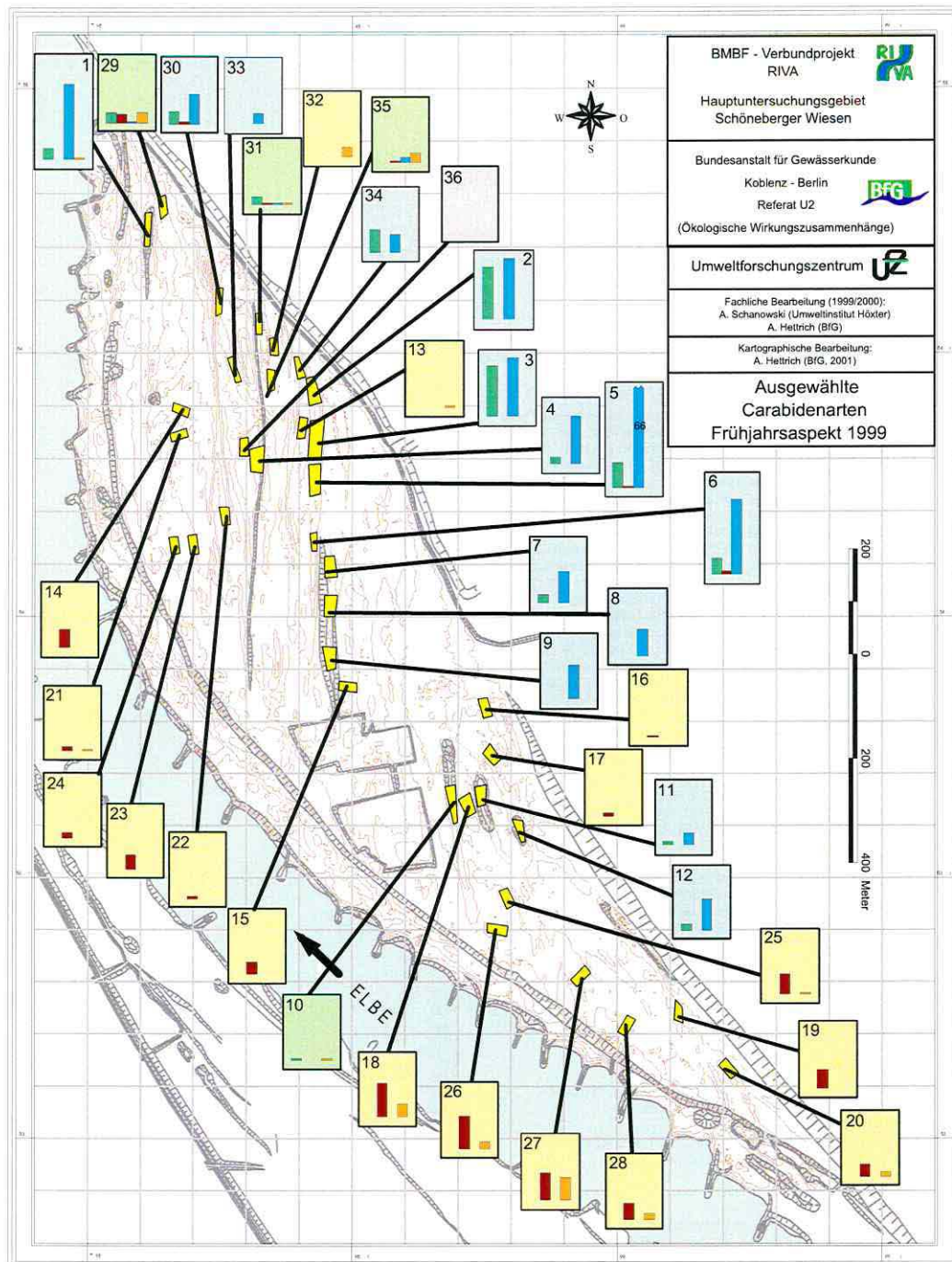
<ul style="list-style-type: none"> <li>Probeflächen</li> <li>Elbe</li> <li>Wasserfläche</li> <li>Temporäre Gewässer (z. Zt. der Erhebung überstaut)</li> <li>Buhnenköpfe (z. Zt. der Erhebung überflutet)</li> <li>z. Zt. der Erhebung vegetationslose Uferbereiche</li> <li>Alno-Ulmion; Quercus Ulmetum Einzelgehölze/Gehölzgruppen (Arten des Quercus-Ulmetum)</li> <li>Hybridpappel-Pflanzungen</li> <li>Salicion albae; Salicetum Einzelgehölze/Gehölzgruppen (Arten des Salicetum)</li> </ul>	<ul style="list-style-type: none"> <li>Convolvulion sepium</li> <li>Cuscuta europaea-Convolutum sepium</li> <li>Impatiens glandulifera-Convolutum sepium</li> <li>Ranuncion aquatilis; Ranunculetum aquatilis</li> <li>Phragmiton australis; Phragmitetum australis</li> <li>Bolboschoenetum maritimi</li> <li>Glycerietum maximae; Sparganietum erecti</li> <li>Rorippo-Oenantheum aquaticae</li> <li>Butometum umbellati; Acoretum calami</li> <li>Eleocharito-Sagittarion sagittifoliae</li> <li>Eleocharietum palustris</li> <li>Glycerio-Sparganion emersi</li> <li>Sparganio emersi-Glycerietum fluitantis</li> <li>Caricion elatae; Caricetum gracilis</li> <li>Phalaridetum arundinaceae</li> <li>Caricetum ripariae</li> <li>Cynosuion cristati; Lolietum perennis</li> </ul>	<ul style="list-style-type: none"> <li>Nanocyperion</li> <li>Cypero fusci-Limoselletum aquaticae</li> <li>Koelerion glaucae</li> <li>Diantho-Armerietum elongatae (verarmte, ruderalisiert Ausbildung)</li> <li>Arrhenatherion elatioris</li> <li>Galio molluginis-Alopecuretum pratensis</li> <li>Dauci carotae-Arrhenatheretum elatioris (verarmte Ausbildung)</li> <li>Deschampsion cespitosae</li> <li>Sanguisorbo officinalis-Silaetum silai</li> <li>Bidenton tripartitae; Rumicetum maritimi</li> <li>Rumici-Alopecuretum aequalis</li> <li>Bidenti-Polygonetum hydropiperis</li> <li>Xanthium albinum-Gemeinschaft</li> <li>Chenopodium glauci; Chenopodio polyspermi-Corrigioletum litoralis</li> <li>Xanthio albi-Chenopodietum rubri</li> <li>Echinochloo-Polygonetum lapatifolii</li> </ul>	<ul style="list-style-type: none"> <li>Potentillion anserinae</li> <li>Ranunculo repentis-Alopecuretum geniculati</li> <li>Rumici crisp-Agrostietum stoloniferae</li> <li>Potentilletum anserinae</li> <li>Inula britannica-Gesellschaft</li> <li>Lysimachia vulgaris-Gesellschaft</li> <li>Rorippa sylvestris-Gesellschaft</li> <li>Juncus compressus-Gesellschaft</li> <li>Polygonum amphibium-Gesellschaft</li> <li>Convolvulo-Agropyron repentis</li> <li>Agropyretum repentis</li> <li>trockenes Agropyretum repentis</li> <li>Calamagrostis epigejos-Gesellschaft</li> <li>Convolvulo-Brometum inermis</li> <li>Atriplicion nitentis</li> <li>Chenopodietum albi-suecici</li> <li>Arction lappae</li> <li>Lamio-Conietum maculati</li> </ul>
---	--	---	--











Anlage 5



# Gesamtliste der zur Modellbildung verfügbaren Erklärungsvariablen

pfnr	Probeflächennummer		
Rinne	Dummy Flutrinne		
NassGrue	Dummy Nasses Grünland		
Trogrue	Dummy Trockenes Grünland		
RWertCen	Rechtswert des Probeflächenzentroids (um die ersten drei Stellen gekürzt)		
HWertCen	Hochwert des Probeflächenzentroids (um die ersten drei Stellen gekürzt)		
MDeckS	Mächtigkeit der Auendeckschicht		
GOF	Geländeoberfläche am Pegel		
DeckSUn	Deckschichtunterkante in müHN		
gewmktDe	gewichteter mittlerer kF (cm/d) Deckschicht		
gewmktPr	gewichteter mittlerer kF (cm/d) Profil		
Ctotal	Kohlenstoff [%]; nach H3, DIN-EN 1484, Analyser		
Ntotal	Stickstoff [%]; nach H11, DIN 19684		
CtNtVerh	C/Nt Verhältnis		
Ppfl	pflanzenverfüg. P [mg/100 g]; nach Egner-Riehm (1960)		
Kpfl	pflanzenverfüg. K [mg/100 g]; nach Egner-Riehm (1960)		
Lf	elektr. Leitfähigkeit [µS/cm]		
mTon	mittlerer Tongehalt [%]		
mSand	mittlerer Sandgehalt [%]		
nFK	nutzbare Feldkapazität [mm]; abgeleitet nach KA4		
FK	Feldkapazität [mm]; abgeleitet nach KA4		
LK	Luftkapazität [mm]; abgeleitet nach KA4		
kcmd	hydraulische Leitfähigkeit [cm/d]; abgeleitet nach KA4		
KAKpot	potentielle Kationenaustauschkapazität [cmolc/kg]; abgeleitet nach KA4		
KAKeff	effektive Kationenaustauschkapazität [cmolc/kg]; abgeleitet nach KA4		
pHKCl	pH-Wert KCl delogarithmisiert (Wasserstoffionenkonzentration)		
pHCACl2d	pH-Wert CaCl2 delogarithmisiert (Wasserstoffionenkonzentration)		
pHH2Od	pH-Wert H2O delogarithmisiert (Wasserstoffionenkonzentration)		
pHKCl	pH-Wert KCl delogarithmisiert (Wasserstoffionenkonzentration)	Transformation * 1.000.000	
pHCACl2t	pH-Wert CaCl2 delogarithmisiert (Wasserstoffionenkonzentration)	Transformation * 1.000.000	
pHH2Ot	pH-Wert H2O delogarithmisiert (Wasserstoffionenkonzentration)	Transformation * 1.000.000	
Bodytp3	Fuzzy codierter Bodentyp 3 Gley-Paternia aus Auensand bis Auenormallehm		
Bodytp6	Fuzzy codierter Bodentyp 6 Vega aus Auenlehmsand bis Auensabdelehm über (tiefen) Auensand		
Bodytp7	Fuzzy codierter Bodentyp 7 Vega aus Auenlehmsand über (tiefen) Auensand		
Bodytp8	Fuzzy codierter Bodentyp 8 Vega aus Auentonschluff über (tiefen) Auensand		
Bodytp9	Fuzzy codierter Bodentyp 9 Gley-Vega aus Auenlehm		
Bodytp10	Fuzzy codierter Bodentyp 10 Vega-Gley aus Auenlehm über (tiefen) Auensand		
Bodytp11	Fuzzy codierter Bodentyp 11 (anmooriger) Auengley aus Auentonschluff bis Auenlehm über (tiefen) Auensand		
Bodytp12	Fuzzy codierter Bodentyp 12 Auengley aus Auensand über tiefem Auenschluffton		
Bodytp13	Fuzzy codierter Bodentyp 13 Auengley aus flachem Auenlehm über Auensand über Auentonschluff		
Bodytp14	Fuzzy codierter Bodentyp 14 Pelosol-Gley aus Auenton		
n verfügbar	Substrattyp		
n verfügbar	Gefügeform Oberboden		
MHoehCar	Mittlere Höhe der Carabidenerhebungen auf den Probeflächen		
MHoehMol	Mittlere Höhe der Molluskenerhebungen auf den Probeflächen		
MHoehVeg	Mittlere Höhe der Vegetationsfläche 01 auf den Probeflächen		
0streu98	1998: keine Steuauflage		
1streu98	1998: wenig Steuauflage		
2streu98	1998: mittel Steuauflage		
3streu98	1998: viel Steuauflage		
0streu99	1999: keine Steuauflage		
1streu99	1999: wenig Steuauflage		
2streu99	1999: mittel Steuauflage		
3streu99	1999: viel Steuauflage		
8ma7nab7	1: 1x gemäht nicht abgeräumt (ca 4.7.)	1998	
8mae6	2: 1x gemäht (ca 15.6.)	1998	
8ma6gs10	3: 1x gemäht (ca 15.6.) + 1x geschlegelt (ca 15.10.)	1998	
8flutr	4: Flutrinne	1998	
8ungenu	5: ungenutzt	1998	
8beweiX2	6: 2x beweidet	1998	kommt in HUG nicht vor
8poweima	7: Portionsweide + eine Mahd im Frühjahr	1998	kommt in HUG nicht vor
8poweifl	8: Portionsweide in Flutrinne	1998	kommt in HUG nicht vor
9ma6abg6	1: 1x gemäht (21.6.)	1999	
9ma7nab	2: 1x gemäht (Ende Juli.) nicht abgeräumt	1999	
9ma7abg8	3: 1x gemäht (Ende Juli.)	1999	
9flutr	4: Flutrinne	1999	
9ungenu	5: ungenutzt	1999	
9beweiX2	6: 2x beweidet	1999	kommt in HUG nicht vor
9poweima	7: Portionsweide + eine Mahd im Frühjahr	1999	kommt in HUG nicht vor
9poweifl	8: Portionsweide in Flutrinne	1999	kommt in HUG nicht vor
geminhoe	Index Bestandshöhe aus: Mittelwert([Deckung in %]*[Höhe minimal]) für alle Arten einer PF		
gemaxhoe	Index Bestandshöhe aus: Mittelwert([Deckung in %]*[Höhe maximal]) für alle Arten einer PF		
gkllhoe	Index Bestandshöhe Höhenklassen: Mittelwert([Deckung in %]*[Höhenklasse]) für alle Arten einer PF		
	Wuchshöhenklassen - Hkl 1 (< 100 mm)		
	Wuchshöhenklassen - Hkl 2 (100-299 mm)		
	Wuchshöhenklassen - Hkl 3 (300-599 mm)		
	Wuchshöhenklassen - Hkl 4 (600-999 mm)		
	Wuchshöhenklassen - Hkl 5 (1,0-1,5 m)		
	Wuchshöhenklassen - Hkl 6 (1,6-3,0 m)		
	Wuchshöhenklassen - Hkl 7 (3,1 - 6,0 m)		
	Wuchshöhenklassen - Hkl 8 (6,1- 15,0 m)		
	Wuchshöhenklassen - Hkl 9 (> 15,0 m)		

# Gesamtliste der zur Modellbildung verfügbaren Erklärungsvariablen (Fortsetzung)

WH2up1	Wuchshöhenklassen - Anteil von Arten der Hkl 2 (100-299 mm) (bezogen auf Gesamtdeckung)
WH3up1	Wuchshöhenklassen - Anteil von Arten der Hkl 3 (300-599 mm) (bezogen auf Gesamtdeckung)
WH4up1	Wuchshöhenklassen - Anteil von Arten der Hkl 4 (600-999 mm) (bezogen auf Gesamtdeckung)
WH5up1	Wuchshöhenklassen - Anteil von Arten der Hkl 5 (1,0-1,5 m) (bezogen auf Gesamtdeckung)
WH6up1	Wuchshöhenklassen - Anteil von Arten der Hkl 6 (1,6-3,0 m) (bezogen auf Gesamtdeckung)
ATrosup1	Architekturtypen - Anteil von Arten mit Grundrosetten und blattlosen Sproß (bezogen auf Gesamtdeckung)
ATrepup1	Architekturtypen - Anteil niederliegender bzw. kriechender Arten (bezogen auf Gesamtdeckung)
ATcaepup1	Architekturtypen - Anteil grundverzweigter oder horstiger, aufrechter Arten (bezogen auf Gesamtdeckung)
GW140m98	GW niedriger als 140cm unter Flur, längste Periode in Tagen 1998
GW140g98	GW niedriger als 140cm unter Flur, Gesamtzahl Tage 1998
GW140m99	GW niedriger als 140cm unter Flur, längste Periode in Tagen 1999
GW140g99	GW niedriger als 140cm unter Flur, Gesamtzahl Tage 1999
GW140v98	GW niedriger als 140cm unter Flur, längste Periode innerhalb Vegetationsperiode (01.04.-30.09.) in Tagen 1998
GW140v99	GW niedriger als 140cm unter Flur, längste Periode innerhalb Vegetationsperiode (01.04.-30.09.) in Tagen 1999
GW140258	GW niedriger als 140cm unter Flur für Monate Februar bis einschliesslich Mai 1998 in Tagen (UP1)
GW140698	GW niedriger als 140cm unter Flur für Monate Juni bis einschliesslich September 1998 in Tagen (UP3)
GW140259	GW niedriger als 140cm unter Flur für Monate Februar bis einschliesslich Mai 1999 in Tagen (UP5)
GW140699	GW niedriger als 140cm unter Flur für Monate Februar bis einschliesslich Mai 1999 in Tagen (UP7)
MinGW98	Minimum Grundwasserstand 01.01.-31.12.98
MaxGW98	Maximum Grundwasserstand 01.01.-31.12.98
AMPGW98	Amplitude der Grundwasserstände 01.01.-31.12.1998
MeanGW98	Mittlerer Grundwasserstand 01.01.-31.12.98
SDGW98	Standardabweichung Grundwasserstand 01.01.-31.12.98
AMPHN98	Amplitude Wasserstand in HN 01.01.-31.12.98
SDHN98	Standardabweichung Wasserstand in HN 01.01.-31.12.98
MinGW99	Minimum Grundwasserstand 01.01.-31.12.99
MaxGW99	Maximum Grundwasserstand 01.01.-31.12.99
AMPGW99	Amplitude der Grundwasserstände 01.01.-31.12.99
MeanGW99	Mittlerer Grundwasserstand 01.01.-31.12.99
MAXHN99	Maximum Wasserstand in HN 01.01.-31.12.99
AMPHN99	Amplitude Wasserstand in HN 01.01.-31.12.99
SDHN99	Standardabweichung Wasserstand in HN 01.01.-31.12.99
MinGW498	Minimum Grundwasserstand 01.04.-30.09.98 (Vegetationsperiode)
MaxGW498	Maximum Grundwasserstand 01.04.-30.09.98 (Vegetationsperiode)
AMPGW498	Amplitude der Grundwasserstände 01.04.-30.09.98 (Vegetationsperiode)
MnGW498	Mittlerer Grundwasserstand 01.04.-30.09.98 (Vegetationsperiode)
MINHN498	Minimum Wasserstand in HN 01.04.-30.09.98 (Vegetationsperiode)
MAXHN498	Maximum Wasserstand in HN 01.04.-30.09.98 (Vegetationsperiode)
AMPHN498	Mittlerer Wasserstand in HN 01.04.-30.09.98 (Vegetationsperiode)
SDHN498	Standardabweichung Wasserstand in HN 01.04.-30.09.98 (Vegetationsperiode)
MinGW499	Minimum Grundwasserstand 01.04.-30.09.99 (Vegetationsperiode)
MaxGW499	Maximum Grundwasserstand 01.04.-30.09.99 (Vegetationsperiode)
AMPGW499	Amplitude der Grundwasserstände 01.04.-30.09.99 (Vegetationsperiode)
MnGW499	Mittlerer Grundwasserstand 01.04.-30.09.99 (Vegetationsperiode)
SDGW499	Standardabweichung Grundwasserstand 01.04.-30.09.99 (Vegetationsperiode)
MaxOW499	Maximum Überflutungswasserstand 01.04.-30.09.99 (Vegetationsperiode)
MnOW499	Mittlerer Überflutungswasserstand 01.04.-30.09.99 (Vegetationsperiode)
SDOW499	Standardabweichung Überflutungswasserstand 01.04.-30.09.99 (Vegetationsperiode)
MINHN499	Minimum Wasserstand in HN 01.04.-30.09.99 (Vegetationsperiode)
MAXHN499	Maximum Wasserstand in HN 01.04.-30.09.99 (Vegetationsperiode)
AMPHN499	Mittlerer Wasserstand in HN 01.04.-30.09.99 (Vegetationsperiode)
MnGW499	Mittlerer Wasserstand in HN 01.04.-30.09.99 (Vegetationsperiode)
SDHN499	Standardabweichung Wasserstand in HN 01.04.-30.09.99 (Vegetationsperiode)
MinGW258	Minimum Grundwasserstand 01.02.-31.05.98 (für UP1: 24.03.-27.05)
MaxGW258	Maximum Grundwasserstand 01.02.-31.05.98 (für UP1: 24.03.-27.05)
AMPGW258	Amplitude der Grundwasserstände 01.02.-31.05.98 (für UP1: 24.03.-27.05)
MnGW258	Mittlerer Grundwasserstand 01.02.-31.05.98 (für UP1: 24.03.-27.05)
SDGW258	Standardabweichung Grundwasserstand 01.02.-31.05.98 (für UP1: 24.03.-27.05)
MaxOW258	Maximum Überflutungswasserstand 01.02.-31.05.98 (für UP1: 24.03.-27.05)
MnOW258	Mittlerer Überflutungswasserstand 01.02.-31.05.98 (für UP1: 24.03.-27.05)
SDOW258	Standardabweichung Überflutungswasserstand 01.02.-31.05.98 (für UP1: 24.03.-27.05)
MINHN258	Minimum Wasserstand in HN 01.02.-31.05.98 (für UP1: 24.03.-27.05)
MAXHN258	Maximum Wasserstand in HN 01.02.-31.05.98 (für UP1: 24.03.-27.05)
AMPHN258	Amplitude Wasserstand in HN 01.02.-31.05.98 (für UP1: 24.03.-27.05)
SDHN258	Standardabweichung Wasserstand in HN 01.02.-31.05.98 (für UP1: 24.03.-27.05)
MinGW698	Minimum Grundwasserstand 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
MaxGW698	Maximum Grundwasserstand 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
AMPGW698	Amplitude der Grundwasserstände 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
MnGW698	Mittlerer Grundwasserstand 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
SDGW698	Standardabweichung Grundwasserstand 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
MaxOW698	Maximum Überflutungswasserstand 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
MnOW698	Mittlerer Überflutungswasserstand 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
SDOW698	Standardabweichung Überflutungswasserstand 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
MINHN698	Minimum Wasserstand in HN 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
MAXHN698	Maximum Wasserstand in HN 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
AMPHN698	Amplitude Wasserstand in HN 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
SDHN698	Standardabweichung Wasserstand in HN 01.06.-30.09.98 (für UP3: 02.09.-07.10.)
MinGW259	Minimum Grundwasserstand 01.02.-31.05.99 (für UP5: 24.03.-27.05)
MaxGW259	Maximum Grundwasserstand 01.02.-31.05.99 (für UP5: 24.03.-27.05)
AMPGW259	Amplitude der Grundwasserstände 01.02.-31.05.99 (für UP5: 24.03.-27.05)
MnGW259	Mittlerer Grundwasserstand 01.02.-31.05.99 (für UP5: 24.03.-27.05)
SDGW259	Standardabweichung Grundwasserstand 01.02.-31.05.99 (für UP5: 24.03.-27.05)
MaxOW259	Maximum Überflutungswasserstand 01.02.-31.05.99 (für UP5: 24.03.-27.05)



# Gesamtliste der zur Modellbildung verfügbaren Erklärungsvariablen (Fortsetzung)

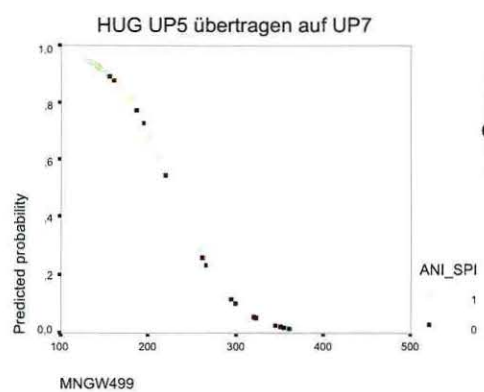
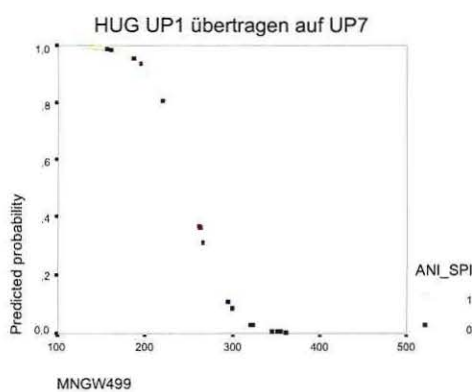
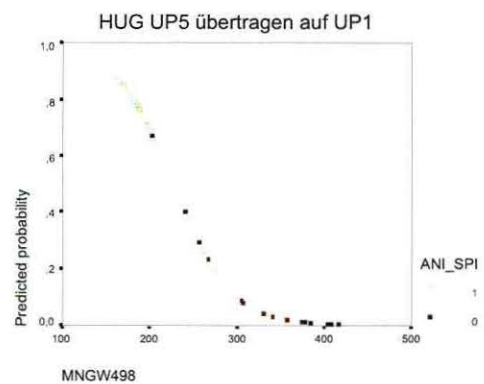
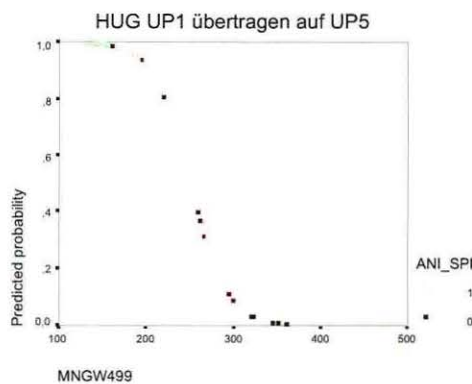
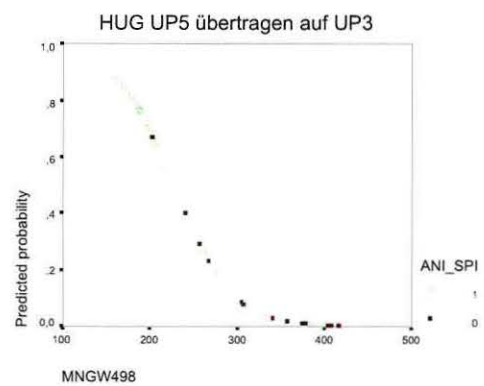
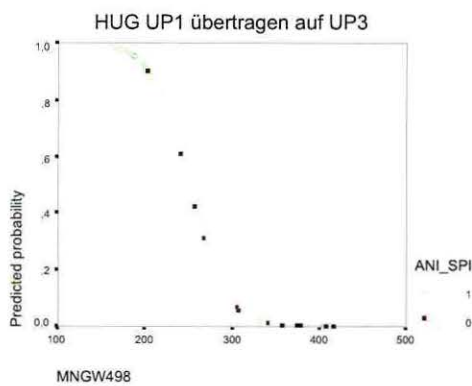
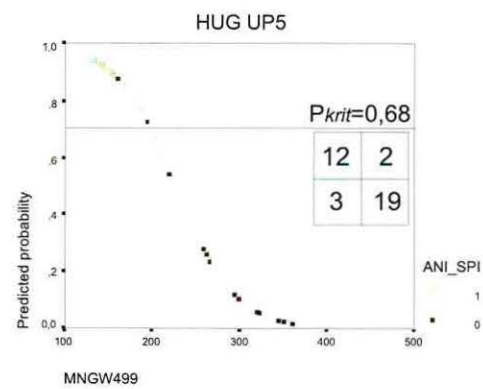
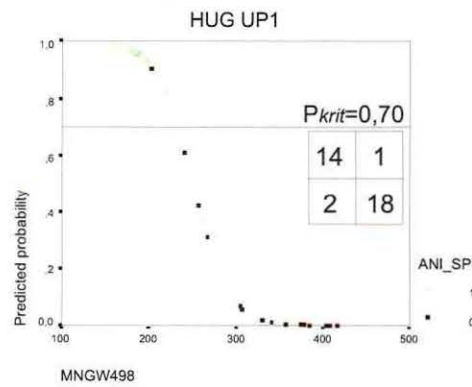
MnOW259	Mittlerer Überflutungswasserstand 01.02.-31.05.99 (für UP5: 24.03.-27.05)
SDOW259	Standardabweichung Überflutungswasserstand 01.02.-31.05.99 (für UP5: 24.03.-27.05)
MINHN259	Minimum Wasserstand in HN 01.02.-31.05.99 (für UP5: 24.03.-27.05)
MAXHN259	Maximum Wasserstand in HN 01.02.-31.05.99 (für UP5: 24.03.-27.05)
AMPHN259	Amplitude Wasserstand in HN 01.02.-31.05.99 (für UP5: 24.03.-27.05)
SDHN259	Standardabweichung Wasserstand in HN 01.02.-31.05.99 (für UP5: 24.03.-27.05)
MinGW699	Minimum Grundwasserstand 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
MaxGW699	Maximum Grundwasserstand 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
AMPGW699	Amplitude der Grundwasserstände 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
MnGW699	Mittlerer Grundwasserstand 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
SDGW699	Standardabweichung Grundwasserstand 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
MaxOW699	Maximum Überflutungswasserstand 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
MnOW699	Mittlerer Überflutungswasserstand 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
SDOW699	Standardabweichung Überflutungswasserstand 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
MINHN699	Minimum Wasserstand in HN 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
MAXHN699	Maximum Wasserstand in HN 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
AMPHN699	Amplitude Wasserstand in HN 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
SDHN699	Standardabweichung Wasserstand in HN 01.06.-30.09.99 (für UP7: 02.09.-07.10.)
FDAJd99	Jährliche Überflutungsdauer in Tagen (Abflußjahr 01.11.97-31.10.98) Pegelinterpolation, korrigiert auf DGM/GW
FDAJd99	Jährliche Überflutungsdauer in Tagen (Abflußjahr 01.11.98-31.10.99) Pegelinterpolation, korrigiert auf DGM/GW
FDUP1258	Dauer der Überflutung in Tagen UP1 Monate 2-5 98, korrigiert auf DGM/GW
FDUP3698	Dauer der Überflutung in Tagen UP3 Monate 6-9 98, korrigiert auf DGM/GW
FDUP5259	Dauer der Überflutung in Tagen UP5 Monate 2-5 99, korrigiert auf DGM/GW
FDUP7699	Dauer der Überflutung in Tagen UP7 Monate 6-9 99, korrigiert auf DGM/GW
PFDP99	Überflutungswahrscheinlichkeit (Abflußjahr 01.11.97-31.10.98) Pegelinterpolation, korrigiert auf DGM/GW
PFDP98	Überflutungswahrscheinlichkeit (Abflußjahr 01.11.98-31.10.99) Pegelinterpolation, korrigiert auf DGM/GW
PFDPUP-69	Überflutungswahrscheinlichkeit UP Monate 2-5 98, korrigiert auf DGM/GW
PFDPUP-25	Überflutungswahrscheinlichkeit UP Monate 6-9 98, korrigiert auf DGM/GW
FDPre99	Häufigkeit/Anzahl Überflutungsereignisse 1999
FDPre98	Häufigkeit/Anzahl Überflutungsereignisse 1998
MaxFD99	Maximale Überflutungsdauer in Tagen (längstes Überflutungsereignis im Abflussjahr 01.11.98-31.10.99, korrigiert auf DGM/GW
MaxFD98	Maximale Überflutungsdauer in Tagen (längstes Überflutungsereignis im Abflussjahr 01.11.97-31.10.98, korrigiert auf DGM/GW
maxFH98	Maximale Überflutungshöhe in m (höchstes Überflutungsereignis im Abflussjahr 01.11.97-31.10.98, korrigiert auf DGM/GW)
maxFH258	Maximale Überflutungshöhe in m (höchstes Überflutungsereignis in den Monaten 2-5.98, korrigiert auf DGM/GW)
maxFH99	Maximale Überflutungshöhe in m (höchstes Überflutungsereignis im Abflussjahr 01.11.97-31.10.99, korrigiert auf DGM/GW)
maxFH259	Maximale Überflutungshöhe in m (höchstes Überflutungsereignis in den Monaten 2-5.99, korrigiert auf DGM/GW)
IFHFD98	Index aus max Überflutungsdauer multipliziert mit max Überflutungshöhe für das Abflussjahr 98
IFHFD99	Index aus max Überflutungsdauer multipliziert mit max Überflutungshöhe für das Abflussjahr 99
GWDSm98	GW niedriger als Deckschicht unter Flur, längste Periode in Tagen 1998
GWDSg98	GW niedriger als Deckschicht unter Flur, Gesamtzahl Tage 1998
GWDSm99	GW niedriger als Deckschicht unter Flur, längste Periode in Tagen 1999
GWDSg99	GW niedriger als Deckschicht unter Flur, Gesamtzahl Tage 1999
GWDSv98	GW niedriger als Deckschicht unter Flur, längste Periode innerhalb Vegetationsperiode (01.04.-30.09.) in Tagen 1998
GWDSv99	GW niedriger als Deckschicht unter Flur, längste Periode innerhalb Vegetationsperiode (01.04.-30.09.) in Tagen 1999
GWDS258	GW niedriger als Deckschicht unter Flur für Monate Februar bis einschliesslich Mai 1998 in Tagen (UP1)
GWDS698	GW niedriger als Deckschicht unter Flur für Monate Juni bis einschliesslich September 1998 in Tagen (UP3)
GWDS259	GW niedriger als Deckschicht unter Flur für Monate Februar bis einschliesslich Mai 1999 in Tagen (UP5)
GWDS699	GW niedriger als Deckschicht unter Flur für Monate Februar bis einschliesslich Mai 1999 in Tagen (UP7)
CHCO3258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Hydrogenkarbonat-Gehalt in mg/l
CCl258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Chloridgehalt in mg/l
CSO4258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Sulfatgehalt in mg/l
CNH4258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Ammoniumgehalt in mg/l
CNO3258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Nitratgehalt in mg/l
CPO4258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Phosphatgehalt in mg/l
CNa258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Natriumgehalt in mg/l
CCa258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Calciumgehalt in mg/l
CK258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Kaliumgehalt in mg/l
CMg258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Magnesiumgehalt in mg/l
CSi258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Siliziumgehalt in mg/l
CP258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Phosphorgehalt in mg/l
CFe258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Eisengehalt in mg/l
CMn258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Mangangehalt in mg/l
CAI258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Aluminiumgehalt in Mikrogramm/l
CNi258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Nickelgehalt in Mikrogramm/l
CCu258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Kupfergehalt in Mikrogramm/l
CCr258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Chromgehalt in Mikrogramm/l
CCd258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Cadmiumgehalt in Mikrogramm/l
CPb258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Bleigehalt in Mikrogramm/l
CAs258	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP1 Feb.-Mai 1998: Arsengehalt in Mikrogramm/l
CHCO3698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Hydrogenkarbonat-Gehalt in mg/l
CCl698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Chloridgehalt in mg/l
CSO4698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Sulfatgehalt in mg/l
CNH4698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Ammoniumgehalt in mg/l
CNO3698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Nitratgehalt in mg/l
CPO4698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Phosphatgehalt in mg/l
CNa698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Natriumgehalt in mg/l
CCa698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Calciumgehalt in mg/l
CK698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Kaliumgehalt in mg/l
CMg698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Magnesiumgehalt in mg/l
CSi698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Siliziumgehalt in mg/l
CP698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Phosphorgehalt in mg/l
CFe698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Eisengehalt in mg/l
CMn698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998: Mangangehalt in mg/l

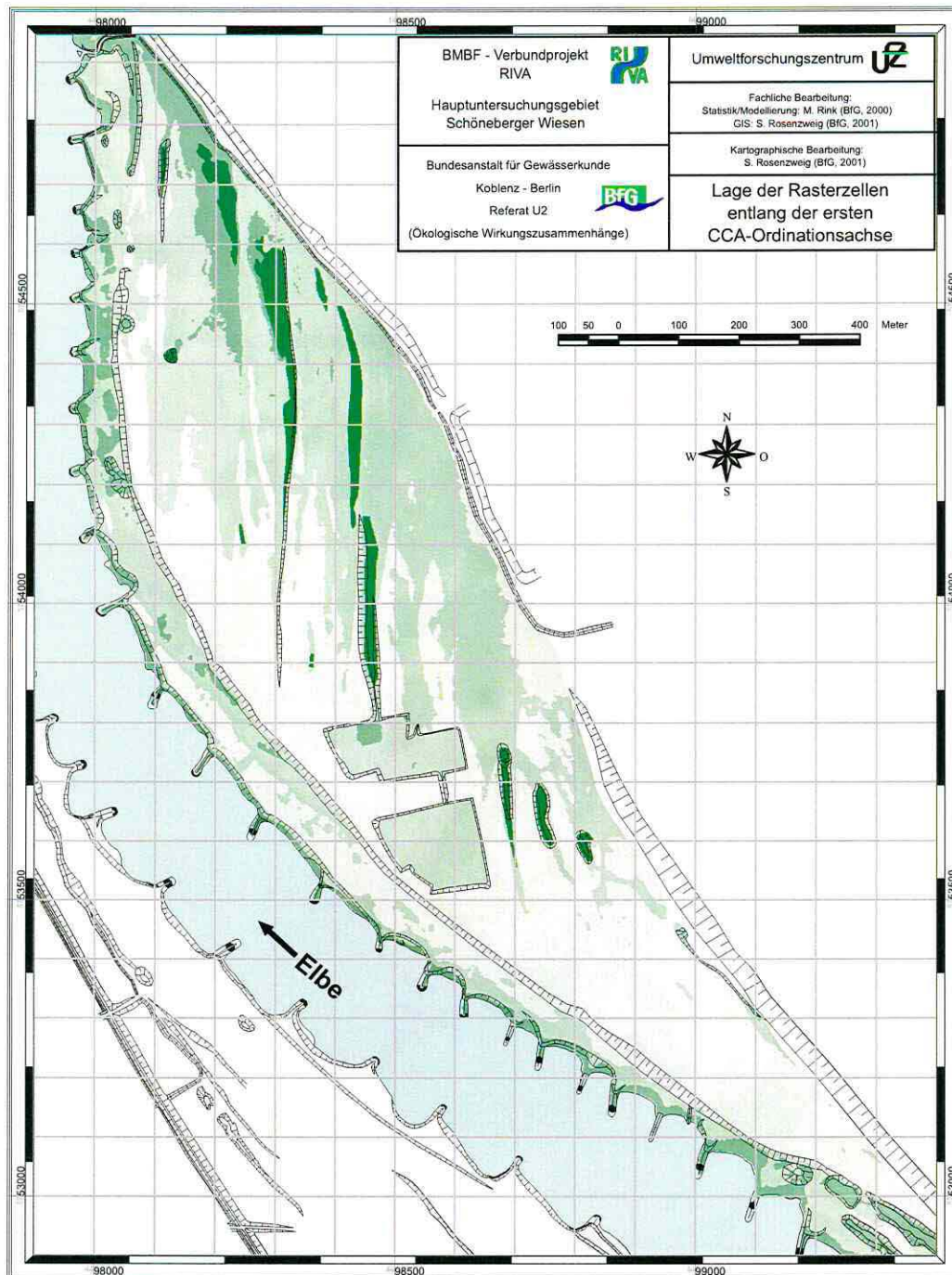


# Gesamtliste der zur Modellbildung verfügbaren Erklärungsvariablen (Fortsetzung)

CAI698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998;	Aluminiumgehalt in Mikrogramm/l
CNI698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998;	Nickelgehalt in Mikrogramm/l
CCu698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998;	Kupfergehalt in Mikrogramm/l
CCr698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998;	Cromgehalt in Mikrogramm/l
CCd698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998;	Cadmiumgehalt in Mikrogramm/l
CPb698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998;	Bleigehalt in Mikrogramm/l
CAs698	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP3 Jun.-Sep. 1998;	Arsengehalt in Mikrogramm/l
CHCO3259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Hydrogenkarbonat-Gehalt in mg/l
CCI259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Chloridgehalt in mg/l
CSO4259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Sulfatgehalt in mg/l
CNH4259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Ammoniumgehalt in mg/l
CNO3259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Nitratgehalt in mg/l
CPu4259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Phosphatgehalt in mg/l
CNa259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Natriumgehalt in mg/l
CCa259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Calciumgehalt in mg/l
CK259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Kaliumgehalt in mg/l
CMg259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Magnesiumgehalt in mg/l
CSi259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Siliziumgehalt in mg/l
CP259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Phosphorgehalt in mg/l
CFe259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Eisengehalt in mg/l
CMn259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Mangangehalt in mg/l
CAI259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Aluminiumgehalt in Mikrogramm/l
CNI259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Nickelgehalt in Mikrogramm/l
CCu259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Kupfergehalt in Mikrogramm/l
CCr259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Cromgehalt in Mikrogramm/l
CCd259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Cadmiumgehalt in Mikrogramm/l
CPb259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Bleigehalt in Mikrogramm/l
CAs259	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP5 Feb.-Mai 1999;	Arsengehalt in Mikrogramm/l
CHCO3699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Hydrogenkarbonat-Gehalt in mg/l
CCI699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Chloridgehalt in mg/l
CSO4699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Sulfatgehalt in mg/l
CNH4699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Ammoniumgehalt in mg/l
CNO3699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Nitratgehalt in mg/l
CPu4699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Phosphatgehalt in mg/l
CNa699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Natriumgehalt in mg/l
CCa699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Calciumgehalt in mg/l
CK699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Kaliumgehalt in mg/l
CMg699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Magnesiumgehalt in mg/l
CSi699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Siliziumgehalt in mg/l
CP699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Phosphorgehalt in mg/l
CFe699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Eisengehalt in mg/l
CMn699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Mangangehalt in mg/l
CAI699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Aluminiumgehalt in Mikrogramm/l
CNI699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Nickelgehalt in Mikrogramm/l
CCu699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Kupfergehalt in Mikrogramm/l
CCr699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Cromgehalt in Mikrogramm/l
CCd699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Cadmiumgehalt in Mikrogramm/l
CPb699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Bleigehalt in Mikrogramm/l
CAs699	Produktvariable GW-Chemie X Dauer Deckschichtkontakt UP7 Jun.-Sep. 1999;	Arsengehalt in Mikrogramm/l
DistPerm	Distanz zu permanenten Wasserflächen	
DistFluMi	Distanz zur Flussmitte (Strömung; mechan. Angriff Fluß; Lage in der Aue)	
D10	D10 (DXXXXX): Distanz zu Zieleinheit	X"Nummer des Biotopcodes nach Peterson"
A10	A10: Fläche der nächstgelegenen Zieleinheit	
B110	B110: Gesamtfläche der Zieleinheit in 100m Buffer	
DM10	DM10: Distanz zur Zieleinheit größer gewählte Minimumfläche (hier: 100m²)	
AM10	AM10: Fläche der Zieleinheit größer gewählte Minimumfläche (hier: 100m²)	
D20	s.D10	
A20	s.A10	
B120	s.B110	
DM20	s.DM10	
AM20	s.AM10	
D30	s.o.	
A30	s.o.	
B130	s.o.	
DM30	s.o.	
AM30	s.o.	
D3408	s.o.	
A3408	s.o.	
B13408	s.o.	
DM3408	s.o.	
AM3408	s.o.	
D3802	s.o.	
A3802	s.o.	
B13802	s.o.	
DM3802	B13806	A3901
AM3802	DM3806	B13901
D3805	AM3806	DM3901
A3805	D3807	AM3901
B13805	A3807	D3906
DM3805	B13807	A3906
AM3805	DM3807	B13906
D3806	AM3807	DM3906
A3806	D3901	AM3906
		D340403
		A340403
		D370202
		A370202
		D430404
		AM430405
		B1370202
		A430404
		B210
		B220
		B230
		B23408
		B23807
		B2350204
		B23901
		B2370202
		B23805
		B23906
		B2390702
		B2430404
		B2430405
		B210: Gesamtfläche der Zieleinheit
		in 200m Buffer

# Anisus spirorbis: sigmoidale Modelle/Streudiagramme aus der Logistischen Regression





Anlage 8

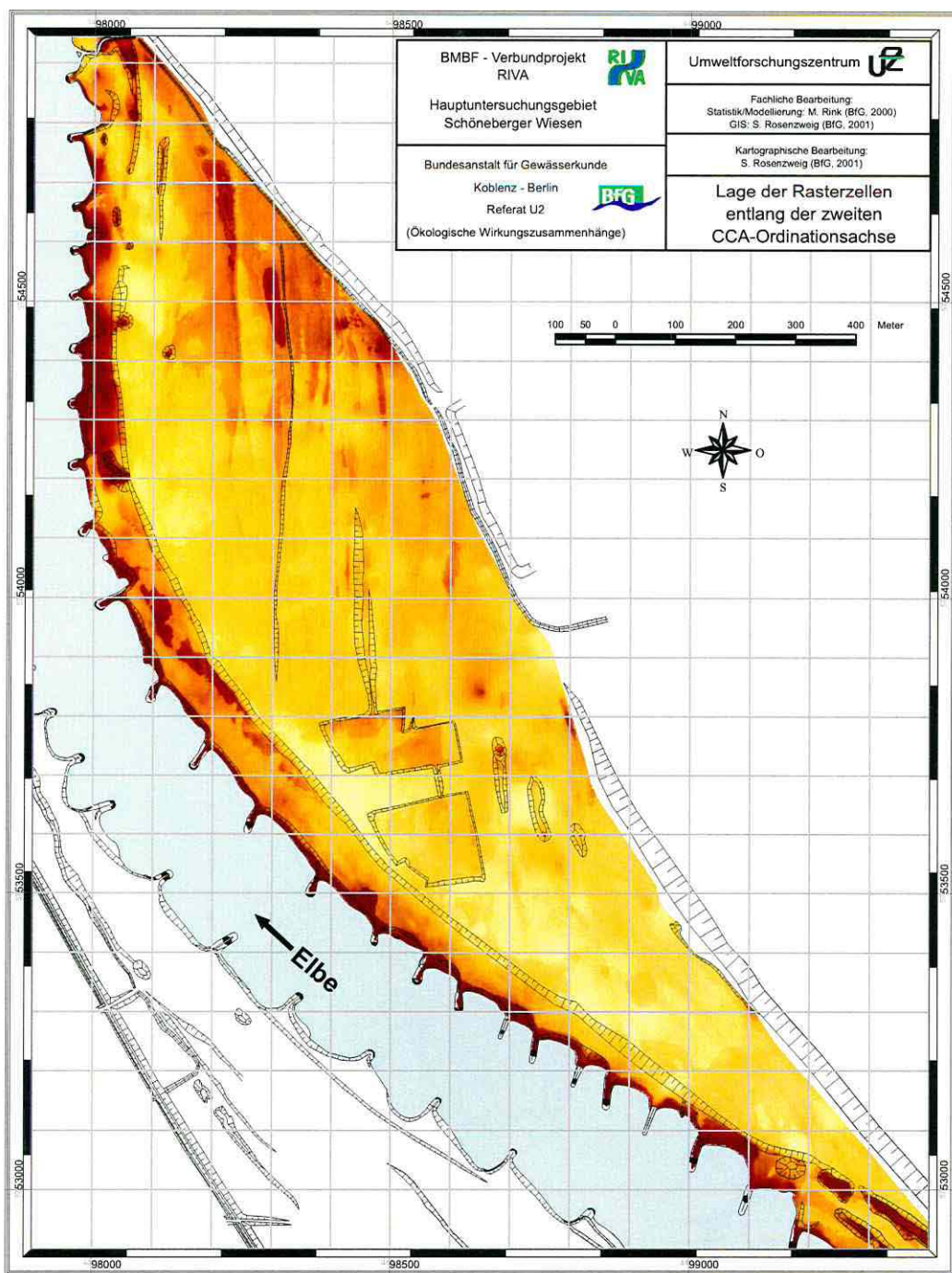
CCA-Modell der Flora  
 Untersuchungsphase  
 Frühjahr 1999 (UP5)

Elbe  
 Morphologie

Wert der Rasterzelle  
 entlang des ersten  
 Komplexgradienten





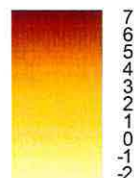


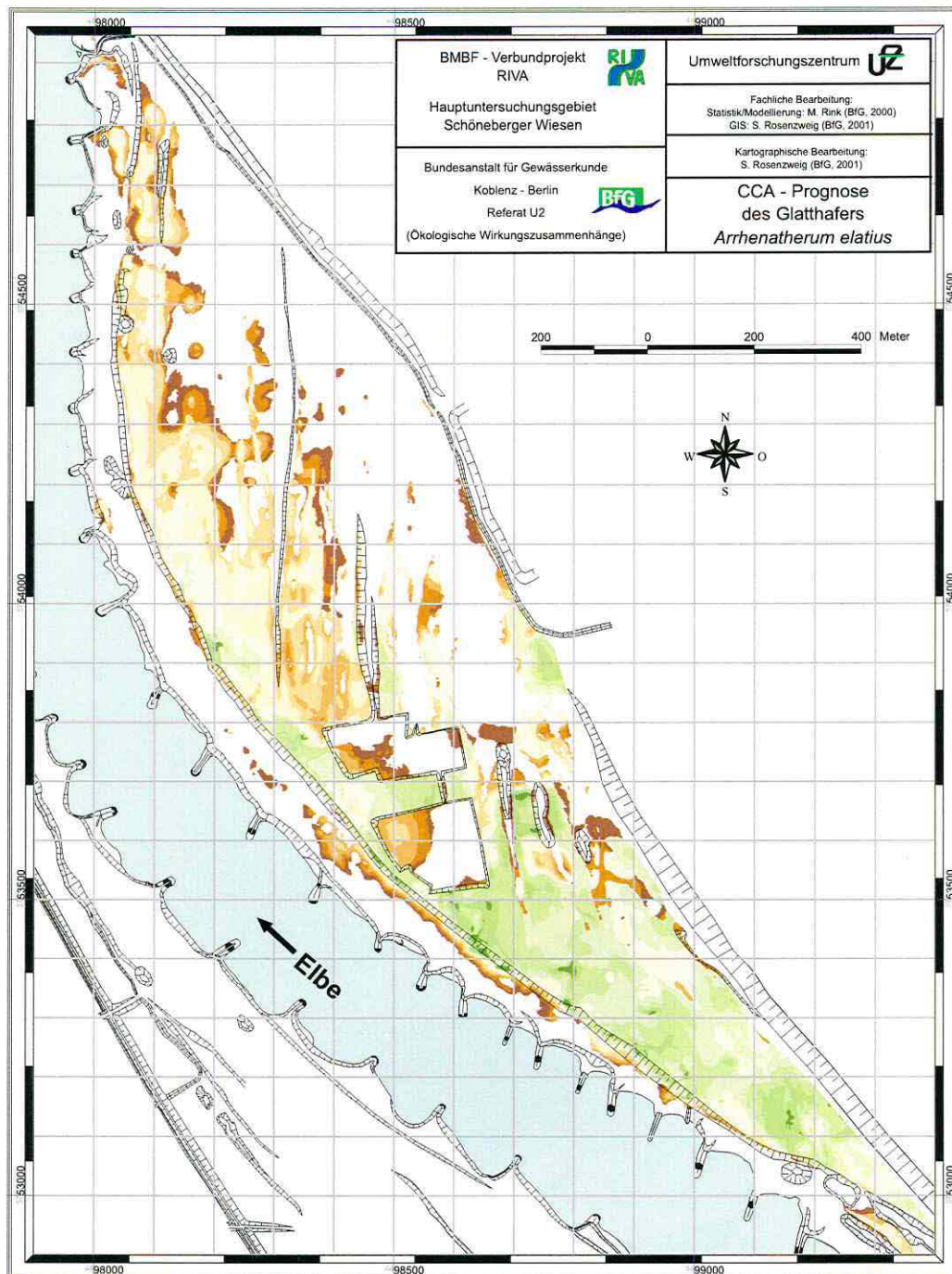
Anlage 9

CCA-Modell der Flora  
Untersuchungsphase  
Frühjahr 1999 (UP5)

Elbe  
Morphologie

Wert der Rasterzelle  
entlang des zweiten  
Komplexgradienten





Anlage 10

CCA-Modell des Glatthafer  
 Untersuchungsphase  
 Frühjahr 1999 (UP5)

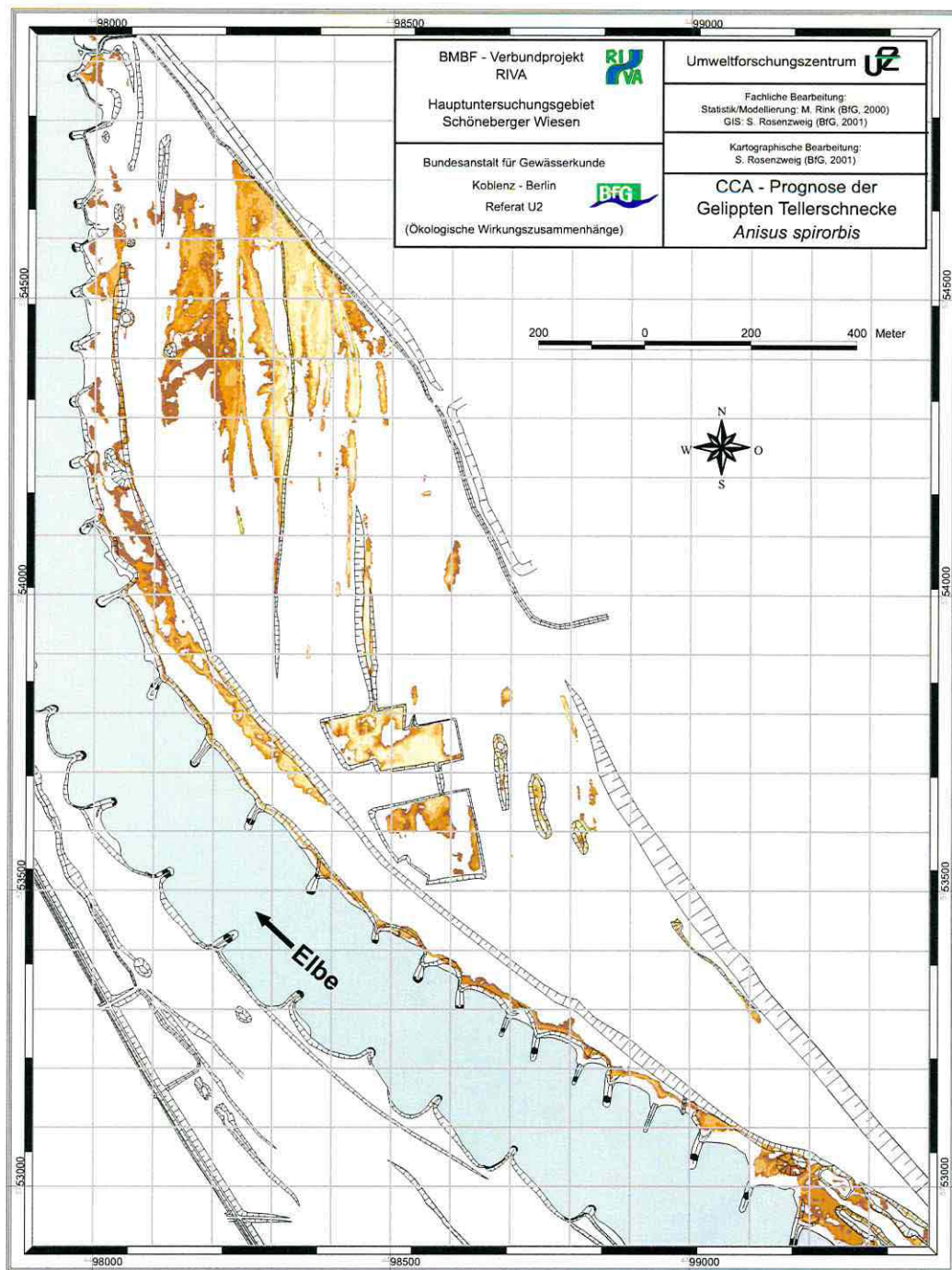
Elbe

Morphologie

Distanz der Rasterzelle  
 vom Optimum der Art im  
 Hauptuntersuchungsgebiet

0 - 10	50 - 60
10 - 20	60 - 70
20 - 30	70 - 80
30 - 40	80 - 90
40 - 50	90 - 100





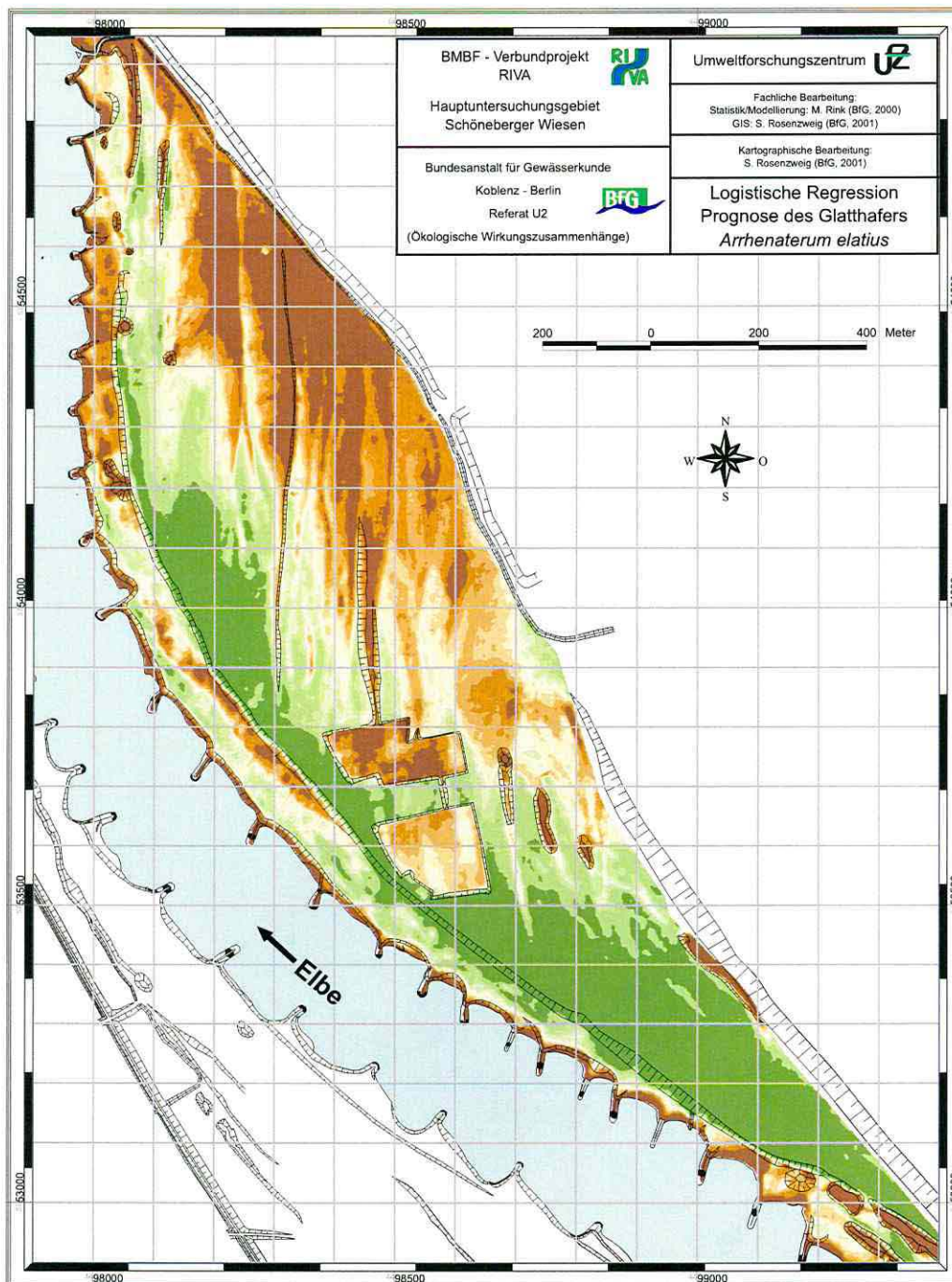
Anlage 11

CCA-Modell der  
 Gelippten Tellerschnecke  
 Untersuchungsphase  
 Frühjahr 1999 (UP5)

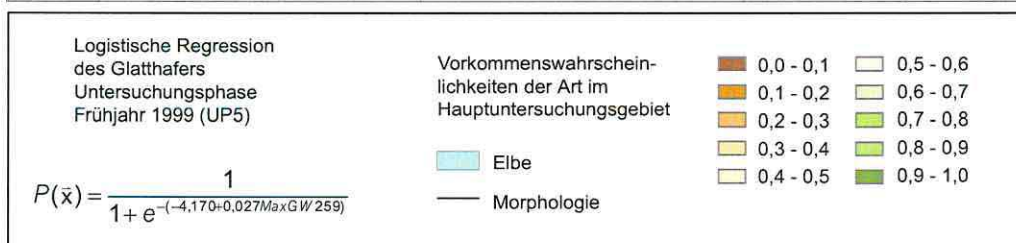
Elbe  
 Morphologie

Distanz der Rasterzelle  
 vom Optimum der Art im  
 Hauptuntersuchungsgebiet

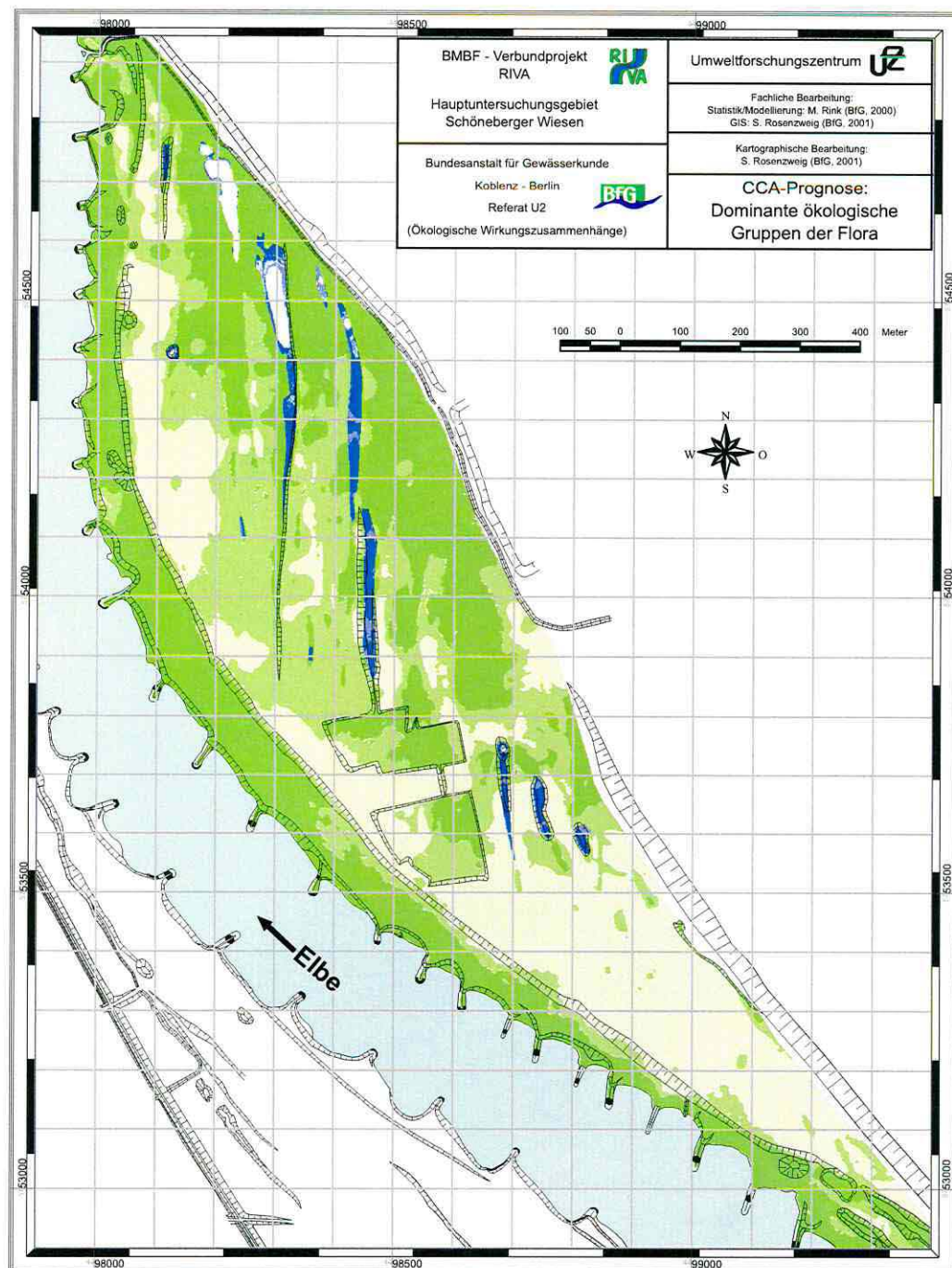
0 - 10	50 - 60
10 - 20	60 - 70
20 - 30	70 - 80
30 - 40	80 - 90
40 - 50	90 - 100



Anlage 12



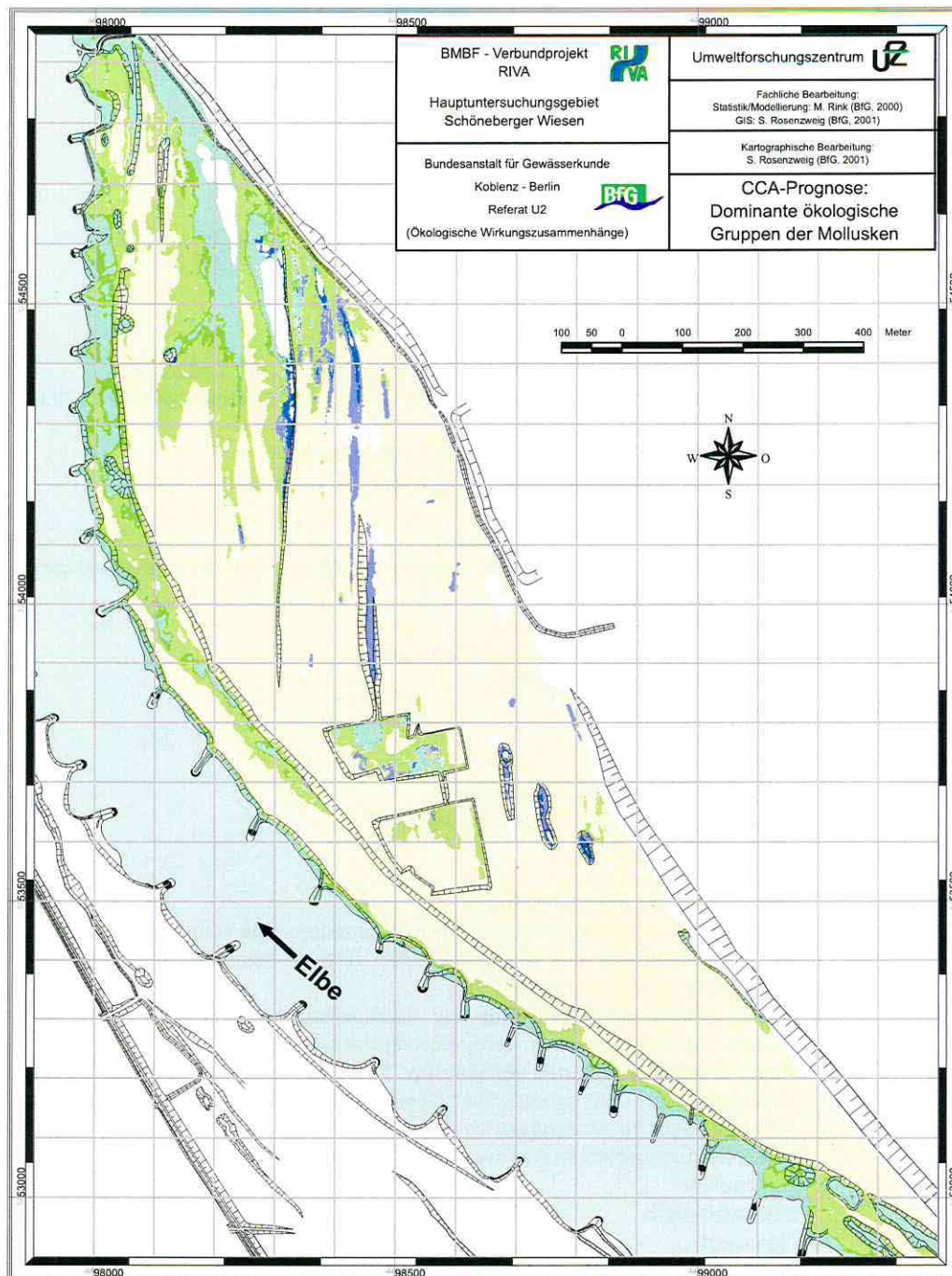




Dominante ökologische  
 Gruppen der Flora  
 Untersuchungsphase  
 Frühjahr 1999 (UP5)

Elbe  
 Morphologie

- Persicaria amphibia - Gruppe
- Nasse, häufiger überflutete Standorte
- Galium palustre - Gruppe
- Nasse, etwas weniger häufig überflutete Standorte
- Urtica dioica - Gruppe
- Feuchte, nährstoffreiche, wenig genutzte Standorte
- Alopecurus pratensis - Gruppe
- FrISChe Standorte
- Arrhenatherum elatius - Gruppe
- Trockene Standorte



Dominante ökologische  
 Gruppen der Flora  
 Untersuchungsphase  
 Frühjahr 1999 (UP5)

Elbe  
 Morphologie

- Tiefere Rinnen, meist länger mit Wasser bestanden  
 (Bsp. Schirmart: *Planorbis planorbis*)
- Seichte, länger trocken liegende Rinnen  
 (Bsp. Schirmart: *Anisus spirorbis*)
- Übergang von den seichteren Rinnen zu den Feuchtfächen  
 (Bsp. Schirmart: *Succinea putris*)
- Tieferliegende Feuchtfächen  
 (Bsp. Schirmart: *Cochlicopa lubrica*)
- Übergang von den feuchten zu den trockeneren Flächen  
 (Bsp. Schirmart: *Perpolita hammonis*)
- Höhergelegene, meist trocken liegende, genutzte Wiesenflächen  
 (Bsp. Schirmart: *Vallonia excentrica*)

Anlage 14



## Notationen

$Y = [y_{i,j}]$	Matrix mit Antwortvariablen (biotische Daten)
$\hat{Y}$	Matrix mit Modellwerten der Antwortvariablen (biotische Daten)
$X$	Matrix mit Erklärungsvariablen (abiotische Daten)
$y$	Originalantwortvariable
$\hat{y}$	Modellantwortvariable
$x$	Erklärungsvariable
$z$	Standardisierte Erklärungsvariable
$y_j$	Deskriptor Spalte $j$
$x_i$	Objekt Zeile $i$ Suffix $i$ nimmt Werte von 1 nach $n$ und $j$ von 1 nach $p$ ein
$y_{i,j}$	Zelle beschreibt dann den Zustand des Objektes $x_i$ in Abhängigkeit von Deskriptor $y_j$
$X^{-1}$	Inverse Matrix der Matrix mit Erklärungsvariablen $X$
$X'$	Transponierte Matrix der Matrix mit Erklärungsvariablen $X$
$A_{nn}$	Assoziationsmatrix der Objekte
$A_{pp}$	Assoziationsmatrix der Deskriptoren
$a_{1,2}$	Maß, welches die Intensität der Assoziation (Ähnlichkeit/Unähnlichkeit) zwischen Zellen anzeigt
$D(x)$	Diagonalmatrix des Vektors $[x_i]$
$D(\sigma)$	Diagonalmatrix der Standardabweichungen
$D(\lambda) = \Lambda$	Diagonalmatrix der Eigenwerte.
$D(1) = I$	Identitätsmatrix
$R$	Korrelationsmatrix
$S$	Streuungsmatrix
$S^{-1}$	Inverse Matrix der Streuungsmatrix $S$
$S'$	Transponierte Matrix der Streuungsmatrix $S$
$S_{xx}$	entspricht $S_{x1, x1} \rightarrow S_{xm, xm}$
$S_{xy}$	entspricht $S_{x1, y1} \rightarrow S_{xm, yp}$
$C$	Matrix der kanonischen Koeffizienten
$F$	Matrix, welche die Koordinaten der Punktsignaturen (site scores) aller Objekte im Ordinationsraum der Antwortvariablen (Biotik) beinhaltet (Hauptkomponenten sind skaliert zur Varianz $\lambda_k$ )
$Z$	Matrix, welche die Koordinaten der Punktsignaturen (sample scores) aller Objekte im Ordinationsraum der Erklärungsvariablen (Abiotik) beinhaltet
$G$	Matrix $G$ resultiert aus den Matrizen $F\Lambda^{-1/2}$ und beinhaltet die Koordinaten der Punktsignaturen (site scores) aller Objekte im Ordinationsraum (die Spalten der Matrix besitzen einheitliche Varianzen)
$U$	Matrix der Eigenvektoren (normiert zur Länge eins)
$u_k$	Eigenvektor
$\Lambda^{1/2}$	Eigenwertmatrix
$\lambda_k$	Eigenwert
$U\Lambda^{1/2}$	Matrix aus der Multiplikation der Eigenvektormatrix $U$ mit der Eigenwertmatrix $\Lambda^{1/2}$ , in der die Länge jedes Eigenvektors $u_i$ zur Quadratwurzel seines Eigenwertes ( $\sqrt{\lambda_i}$ ) skaliert ist
$r$	Korrelationskoeffizient
$R^2$	(multiples Bestimmtheitsmaß)
$B$	Matrix der Regressionskoeffizienten
$b_0, b_1, \dots, b_p$	Koeffizienten einer Modellgleichung
$O_{ij}$	Beobachtete Häufigkeiten



$E_{ij}$	Erwartete Häufigkeiten
$Q$	Matrix der relativen Häufigkeiten einer Kontingenztafel
$p_{ij}$	Relative Häufigkeit (absolute Häufigkeit der zugehörigen Zelle dividiert durch das Matrixtotal)
$f_{++}$	Matrixtotal einer Häufigkeitstabelle
$p_{i+}$	Zeilengewichte ( $f_{i+}$ dividiert durch $f_{++}$ )
$f_{i+}$	Summe aller Werte in Zeile $i$
$[p_{i+}]$	Vektor der Zeilengewichte
$p_{+j}$	Spaltengewichte ( $f_{+j}$ dividiert durch $f_{++}$ )
$f_{+j}$	Summe aller Werte in Spalte $j$
$[p_{+j}]$	Vektor der Spaltengewichte
$\chi^2_p$	Chi <sup>2</sup> -Statistik nach Pearson
$S_{\hat{Y}\hat{Y}}$	Matrix der Quadratsummen und Kreuzprodukte ohne eine Division durch die Freiheitsgrade in einer CCA
$I(d)$	Moran's $I$ für die Distanzklassen $d$
$y_h, y_i$	Moran's $I$ : beobachtete Werte auf den Flächen $h$ und $i$
$w_{hi}$	Moran's $I$ : Gewichte mit $w_{hi} = 1$ für $h$ und $i$ in der gleichen Distanzklasse und $w_{hi} = 0$ in anderen Fällen (es werden nur Probeflächenpaare in der gleichen Distanzklasse berücksichtigt)
$W$	Moran's $I$ : Summe der $w_{hi}$ für eine gegebene Distanzklasse

## Danksagung

An dieser Stelle möchte ich mich bei allen denjenigen Personen ganz herzlich bedanken, die mich bei der Erstellung meiner Doktorarbeit in verschiedenster Hinsicht unterstützt haben.

Die vorliegende Arbeit entstand als Dissertation an der Gemeinsamen Naturwissenschaftlichen Fakultät der Technischen Universität Carolo-Wilhelmina zu Braunschweig, eingebettet in das Forschungsprojekt "Übertragung und Weiterentwicklung eines robusten Indikationssystems für ökologische Veränderungen in Auen" (RIVA). So stellte sich mir die Möglichkeit, unzählige intensive, anregende und kontroverse Gespräche mit den beteiligten Fachwissenschaftlern zu führen und deren Wissensressourcen in die vorliegende Arbeit einzubringen. Speziell die Thematik meiner Doktorarbeit erforderte eine enge Zusammenarbeit mit allen Projektbeteiligten, die mir immer mit Rat und Tat zur Seite standen, und ohne deren Felderhebungen und Datenaufbereitungen diese Ausarbeitungen nicht vorstellbar gewesen wären.

Ganz besonderer Dank gebührt den Kollegen und Freunden am Ort meines Arbeitsplatzes, dem Referat für Ökologische Wirkungszusammenhänge der Bundesanstalt für Gewässerkunde Koblenz. Vor allem dem offenen Gedankenaustausch und den vielfältigen und klärenden Diskussionen verdanke ich weitreichende Einblicke in verschiedene fachliche Disziplinen. Vielen Dank für die stets enge Zusammenarbeit und die angenehme Arbeitsatmosphäre an Hr. Giebel, Hr. Dr. Fuchs, Hr. Hüsing, Hr. Peter, Fr. Konz und Fr. Hettrich. Speziellen Dank auch an Hr. Dax und Hr. Rosenzweig für die Betreuung bei computertechnischen Problemen und die Zusammenarbeit im GIS-Bereich.

Dank ebenfalls an Hr. Dr. Palmer und Hr. Dr. Bousquin für die angenehme sowie effektive Zusammenarbeit an der Oklahoma State University, Department of Botany, Stillwater/USA, die zahlreichen persönlichen Gespräche, für die sie sich Zeit nahmen und den herzlichen Empfang im Allgemeinen. Auch nach meiner Zeit am Institut beantworteten Sie mir neu entstandene Fragen zu Ordinationsverfahren immer in spontaner Weise. Gleiches gilt für Hr. Šmilauer und Hr. Lepš, die mir bei meinem lehrreichen Aufenthalt an der University of South Bohemia, Faculty of Biological Sciences, Ceske Budejovice/Czech Republic wertvolle gedankliche Anregungen für meine Auswertungen mit auf den Weg gaben und auch später noch beratend zur Seite standen.

Bedanken möchte ich mich weiterhin bei meinen Mitstreitern in der Geoökologie an der Universität Braunschweig, C. Vogel, B. Schröder und T. Rickfelder, die sich ebenfalls mit diversen Modellierungstechniken auseinander setzten und mir manchen Arbeitspunkt durch den Erfahrungsaustausch erleichterten.

Fr. Dr. Goldschmidt, Hr. Dr. Fuchs, Hr. Dr. Sundermeyer, Hr. Rosenzweig, Hr. von Landwüst sowie der als Lektorin tätigen Fr. Brigitte Rhein und Hr. Dr. Rhein möchte ich für die kritische Durchsicht der ersten schriftlichen Fassung des Manuskripts meinen Dank aussprechen.

Bei meinen Betreuern Hr. Prof. Dr. O. Richter an der Universität Braunschweig und Herr Dr. habil. C. Henle am Umweltforschungszentrum Leipzig/Halle bedanke ich mich besonders für die Übernahme des Referates und Korreferates, wie auch für die fachliche Unterstützung und die sachliche Hilfestellung während der Anfertigung dieser Dissertation.

Letztlich noch herzlichen Dank für die moralische Unterstützung und das erbrachte Verständnis für alle meine Gemütslagen über die Zeit der Ausarbeitung dieser Doktorarbeit an meine Familie, meine Pateneltern und meinen gesamten Freundeskreis.

Marcus Rink