

Springer-Lehrbuch

Multivariate Statistik in der Ökologie

Eine Einführung

Bearbeitet von
Ilona Leyer, Karsten Wesche

1. Aufl. 2007. Korr. Nachdruck 2008. Taschenbuch. X, 221 S. Paperback

ISBN 978 3 540 37705 4

Format (B x L): 15,5 x 23,5 cm

Gewicht: 730 g

[Weitere Fachgebiete > Geologie, Geographie, Klima, Umwelt > Umweltpolitik, Umwelttechnik > Umweltüberwachung, Umweltanalytik, Umweltinformatik](#)

Zu [Inhaltsverzeichnis](#)

schnell und portofrei erhältlich bei


DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.

5 Ordinationen – das Prinzip

5.1 Dimensionsreduktion als Analysestrategie

Als **Ordinationen** werden eine ganze Fülle von meist komplexen mathematischen Verfahren bezeichnet, denen gemeinsam ist, dass ein Hauptergebnis die grafische Darstellung der Daten in einem Koordinatensystem ist. Diese Darstellungen sind meist einfache Streudiagramme. Nehmen wir wieder den oben erwähnten multivariaten Datensatz, bestehend aus 3 Vegetationsaufnahmen im vordersten Bereich einer Salzwiese (Tabelle 4.3). Wir haben gesehen, dass sich die Ähnlichkeit der Aufnahmen untereinander sehr leicht grafisch darstellen lässt, indem wir sie in ein Koordinatenkreuz mit den Arten als Achsen eintragen (Abb. 4.1). Mit anderen Worten, wir bilden die Aufnahmen im **Artenraum** ab.

Genauso schlicht ist eine Grafik, die die Lage der Arten im **Aufnahme-raum** abbildet (Abb. 5.1). Dieser ist wegen der 3 vorhandenen Aufnahmen jetzt 3dimensional, also komplizierter, aber durchaus noch vorstellbar. Schwierig wird es erst, wenn es mehr als 3 Arten bzw. Aufnahmen sind. Für Mathematiker sind solche mehrdimensionalen Räume unproblematisch, für den Laien wird ihre Darstellung aber leicht zu abstrakt. Wichtig an diesem Beispiel ist aber, dass sich multivariate Datensätze als ***n*-dimensionale Hyperräume** verstehen lassen. Werden z. B. Barberfallenproben im Hinblick auf ihre Artenzusammensetzung analysiert, so hat dieser Hyperraum so viele Achsen, wie es Arten in den Fallen gibt. Wenn umgekehrt Arten in ihrer Ähnlichkeit betrachtet werden sollen, dann hat der Hyperraum so viele Achsen, wie es Objekte (also Fallen) gibt. Generell spricht man von **R-Analyse**, wenn Objekte (z. B. Barberfällen) im Raum der ihnen zugeordneten Variablen (z. B. Arten) dargestellt werden. Werden Variablen im Objektraum betrachtet, spricht man von **Q-Analyse**. Oft werden Q- und R-Analysen auch als 2 Betrachtungsweisen des gleichen Problems beschrieben. Dies stimmt aber nur bedingt, denn Variablen und Objekte sind insofern unterschiedlich, als dass in einem Datensatz zwar eine Art in einer Falle fehlen kann, in der Regel aber nicht umgekehrt eine Falle nur in einer Art fehlt (aus diesem Grunde sind für Q- und R-Analysen auch unterschiedliche Ähnlichkeitsmaße geeignet, s. Kap. 4).

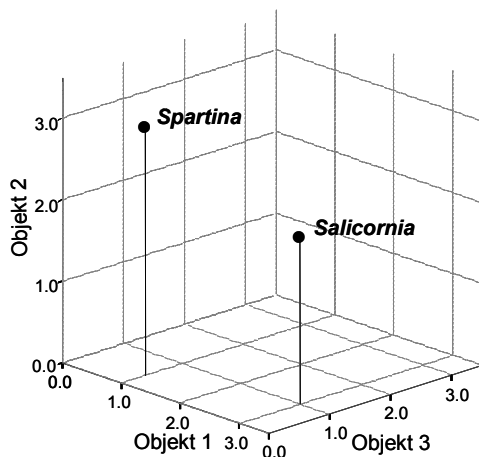


Abb. 5.1. Grafische Darstellung der Daten aus Tabelle 4.3; anders als in Abb. 4.1 wurden hier aber die Arten im Aufnahme-raum abgebildet

Ordinationen versuchen nun, n -dimensionale Hyperräume zu leichter verständlichen 2(-3) dimensionalen Darstellungen zu vereinfachen; sie sind Verfahren der **Dimensionsreduktion**. Ein Beispiel zur Veranschaulichung: Nehmen wir den Datensatz mit Vegetationsaufnahmen aus der Elbaue, der 33 Objekte und insgesamt 53 Arten umfasst (Tabelle 1.1). Theoretisch könnten wir jetzt Hunderte von einzelnen Grafiken zeichnen (Abb. 5.1), in denen wir jeweils die Arten im Hinblick auf 2 (oder 3) Aufnahmen untersuchen. Das wäre weder praktikabel noch anschaulich.

Auch können wir bei so einem Datensatz annehmen, dass die Artenzusammensetzung der Aufnahmen nicht vollkommen willkürlich variiert, sondern dass es einige wichtige Standorteigenschaften gibt, welche die Artenzusammensetzung bestimmen. Das könnte zum Beispiel der Grundwasserstand sein, entlang dessen sich die Artenzusammensetzung stark ändert. Wir könnten also die Aufnahmen viel effektiver entlang von einigen wenigen wichtigen Gradienten anordnen, d. h. ein Diagramm mit diesen wichtigen ökologischen Gradienten als Achsen zeichnen. Die Ordination versucht nun, die wesentlichen Gradienten herauszuarbeiten und sie in wenigen Dimensionen abzubilden. Aus diesem Grund wird für Ordination häufig auch der Begriff **Gradientenanalyse** benutzt. Die Objekte werden dabei so angeordnet, dass sich ähnelnde Aufnahmen auch nahe beieinander liegen.

Als Ökologen würden wir also Verfahren fordern, die uns die Beziehungen zwischen Aufnahmen so abbilden, dass wir die wichtigsten „Richtungen“ der floristischen oder faunistischen Änderungen erkennen können. Wir suchen also nach einer Anordnung (oder eben „Ordination“) der Aufnahmen, die möglichst sinnvoll die Ähnlichkeiten bzw. Unterschiede zwischen den Aufnahmen abbildet. Dies geschieht in der Regel in

einem Koordinatensystem. Etwas technischer ausgedrückt: Die Ordination sollte die Unterschiede zwischen den Aufnahmen so zusammenfassen, dass mit möglichst wenigen Dimensionen oder Achsen ein möglichst großer Anteil dieser Unterschiede abgebildet wird. Entsprechend sind die ersten Achsen dann Achsen maximaler Varianz im Datensatz.

Diese Achsen des Koordinatensystems beruhen bei den meisten einfachen Ordinationsverfahren auf den Unterschieden in der Artenzusammensetzung, und können dann in einem zweiten Schritt in Beziehung z. B. zu gemessenen Umweltgradienten gesetzt werden. So ein Vorgehen wäre eine **indirekte Gradientenanalyse**, bei der die ökologischen Zusammenhänge im ersten Schritt nur im Hinblick auf die Arten interpretiert werden. Das entspricht einer biologischen Sichtweise, bei der die Struktur der Biozöosen im Vordergrund steht (Gauch 1994). Eventuell verfügbare gemessene Umweltvariablen spielen für die Berechnung der eigentlichen Ordination keine Rolle und werden erst im zweiten Schritt zur Interpretation der Ordination genutzt.

Werden die Aufnahmen gleich in einen mit den Umweltvariablen aufgespannten ökologischen Raum (z. B. Meereshöhe, Niederschlag) eingetragen, spricht man von **direkter Gradientenanalyse**. Ein typisches Beispiel hierfür sind Ökogramme (Ellenberg 1996). Lassen wir Umweltvariablen direkt in unsere Ordinationen eingehen, so werden auch diese Verfahren als direkte Gradientenanalyse eingestuft; besser wäre es, in diesem Fall von multivariater direkter Gradientenanalyse zu sprechen. Häufig sind aber diese Umweltgradienten nicht bekannt, der erste Schritt ist also oft eine indirekte Ordination, welche die wesentlichen, z. B. floristischen, Zusammenhänge aufdecken soll.

Das leistet die Ordination auch für unser Beispiel. Abb. 5.2 ist eine simple (indirekte) Ordination, welche die Vegetationsaufnahmen aus der Elbaue entlang von 2 Achsen darstellt. Sehr deutlich zeichnet sich ein floristischer Hauptgradient ab, der die Aufnahmen entlang der ersten Achse differenziert. Die Aufnahmen sind dann noch entlang eines zweiten Gradienten angeordnet. In einem anschließenden Vergleich der Daten mit den Aufnahmelokalitäten zeigt sich, dass der erste Gradient weitgehend parallel zu den Unterschieden im mittleren Grundwasserstand (MWS) verläuft (s. Größe der Symbole in Abb. 5.2), der zweite Gradient hängt mit der Lage in der Aue zusammen (rezente Aue oder Altaue/Auenrand). Das Ordinationsverfahren hat also die insgesamt denkbaren 53 (Artenzahl!) Achsen auf die 2 ökologisch wichtigsten Achsen reduziert.

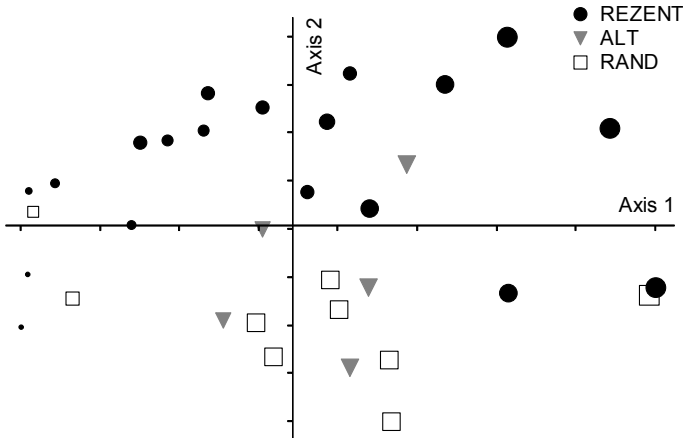


Abb. 5.2. Ordination von 33 Vegetationsaufnahmen mit insgesamt 53 Arten aus der Elbaue. Die erste Achse entspricht einem Feuchtegradienten. Dieser wird durch die Größe der Symbole angedeutet. Die Form der Symbole gibt die Lage der Aufnahmeflächen in der Aue an

Das Interpretationsprinzip ist bei allen diesen Grafiken ähnlich. Das Koordinatenkreuz entspricht Dimensionen floristischer oder faunistischer Ähnlichkeit, die Symbole sind Aufnahmepunkte/Objekte (R-Analyse) oder Arten bzw. andere Variablen (Q-Analyse). Die relative Entfernung der Objekte untereinander entspricht dann (näherungsweise) ihrer Unähnlichkeit; Objekte mit ähnlicher Artenzusammensetzung liegen also nahe beieinander, unähnliche Objekte sind dagegen weiter getrennt. Je nach Ordinationsverfahren lässt sich aber weitaus mehr aus diesen Grafiken lesen; davon mehr in den entsprechenden Kapiteln.

Wie wir gesehen haben, ist das grundsätzliche Ziel von Ordinationen das Herausarbeiten der wesentlichen Information. Wenn man Information als Position in einem mehrdimensionalen Raum betrachtet, ist das Ziel also Dimensionsreduktion. Damit ist hier gemeint, dass die ersten Achsen die wesentlichen ökologischen Zusammenhänge abbilden bzw. dass die ersten Achsen die wichtigsten Gradienten darstellen. Aus diesem Grund konzentriert sich die Interpretation einer Ordination auch meist auf die ersten Achsen, oft sogar nur auf die ersten beiden (streng genommen sollte man also eher von Dimensionskonzentration sprechen, denn die Achsen höherer Ordnung bleiben ja meist erhalten, haben allerdings geringe Bedeutung).

Ziel der Ordination ist es also, die wesentlichen Gradienten zu finden und zu visualisieren. In einer (sehr vereinfachten) wissenschaftstheoretischen Formulierung wäre dies die Suche nach Hypothesen (Gauch 1994), die dann, basierend auf dem Ordinationsdiagramm, so formuliert werden

könnten: „Der Wasserstand ist der wesentliche Faktor, der die Artenzusammensetzung des Elbauengrünlandes bestimmt“. Ordinationen sind also in der Regel Verfahren der **explorativen Datenanalyse**. Hier findet sich wieder ein großer Unterschied zu den meisten Anwendungen univariater Statistik, denn dort geht es in der Regel um das Testen bereits bestehender Hypothesen (z. B. „die Artenzahl nimmt mit dem Wasserstand ab“). Es handelt sich um **schließende Statistik**, hier sind Signifikanztests von fundamentaler Bedeutung.

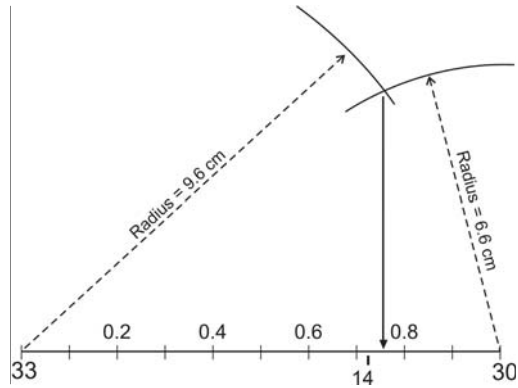
Schließende Statistik spielt in der multivariaten Analyse eine eher untergeordnete Rolle und wird deswegen am Ende des Buches behandelt (Kap. 17). Sie birgt einige Probleme und wird daher eher selten benutzt. Sehr häufig dagegen werden Ordinationen mit univariater schließender Statistik verknüpft. Ein Beispiel wäre die Verwendung von Ordinationen, um herauszuarbeiten, dass Landnutzung der wichtige Parameter für die Laufkäferfauna ist. Dann wird univariat nur mit der Landnutzung (z. B. in Beziehung zu einzelnen Arten) und nicht mit allen anderen Umweltvariablen weitergearbeitet (Vermeidung von *statistical fishing*, Kap. 1).

5.2 Polare Ordination

Zur Illustration des Gesagten wollen wir kurz ein sehr simples Ordinationsverfahren beschreiben, das heute aber fast nur noch historische Bedeutung hat. Es wurde in den 50er Jahren entwickelt (Bray u. Curtis 1957), also weit vor dem Einzug von Computern in ökologische Institute. Ausgangspunkt sind die in Kap. 4 beschriebenen Ähnlichkeits- oder Distanzmaße. Diese erlauben uns, die Unterschiede und Gemeinsamkeiten zwischen 2 Objekten in einer Zahl auszudrücken. Um diese räumlich abzubilden, könnten wir aus der Dreiecksmatrix die beiden unähnlichsten Aufnahmen auswählen und ihre Unähnlichkeit in Form einer Strecke symbolisieren. Wir nehmen die Bray-Curtis-Unähnlichkeit als Maß und berechnen sie mit einem geeigneten Programm, das Ergebnis ist eine diagonal strukturierte Dreiecksmatrix (analog Tabelle 14.1). Im typischen Fall beschreibt diese die Unähnlichkeit zwischen Aufnahmen, aber wir können je nach Fragestellung auch eine Ähnlichkeitsmatrix für Arten benutzen.

Im Fall des Elbauendatensatzes (Tabelle 1.1) haben z. B. Aufnahme 33 und Aufnahme 30 keine Arten gemeinsam; die Ähnlichkeit ist 0, die Unähnlichkeit oder Distanz 1. Wir können diese jetzt als Endpunkte (also als Pole) einer 10 Einheiten langen Achse auftragen; diese soll als erste Achse dienen, um unser Koordinatenkreuz aufzuspannen.

Abb. 5.3. Prinzip der polaren Ordination von Aufnahmen entlang einer Achse, die durch möglichst unähnliche Aufnahmen aufgespannt wird. In diesem Fall sind das 2 Aufnahmen mit der maximalen Unähnlichkeit 1 (ursprünglich als $1 = 10$ cm abgebildet, hier etwas verkleinert)



Jetzt können wir mit einem Zirkel alle anderen Aufnahmen in Beziehung setzen. Abb. 5.3 verdeutlicht das Prinzip. Aufnahme 14 hat eine Unähnlichkeit von 0.97 zu Aufnahme 33 und von 0.68 zu Aufnahme 30. Wir zeichnen entsprechend lange Bögen von den Endpunkten ausgehend; der Radius ist 9.7 Einheiten rund um Aufnahme 33 und 6.8 Einheiten rund um Aufnahme 30. Vom Schnittpunkt der beiden Kreise fallen wir ein Lot in Richtung der bereits gezeichneten Achse 1, auf der die Aufnahme 14 bei ca. 0.73 zu liegen kommt. Das können wir für alle Aufnahmen wiederholen; das Prinzip zeigt Abb. 5.4 a. Das Ergebnis ist eine Ordination aller Aufnahmen entlang der ersten (synthetischen) Achse.

Eine zweite Achse lässt sich aufspannen, in dem wir 2 weitere Aufnahmen wählen, die 1. einen großen Abstand zueinander haben und 2. möglichst im Zentrum der ersten Achse liegen, z. B. die Aufnahmen 18 und 31. Hier kommt eine weitere Anforderung an Ordinationen ins Spiel. Ziel ist, dass alle Achsen wirklich zusätzliche Information anzeigen, dass also die Positionen der Objekte entlang der zweiten Achse unabhängig von den Koordinaten der (selben) Objekte entlang der ersten Achse sind. Anders ausgedrückt, die Achsen sollen **unkorreliert** oder **orthogonal** sein. Für uns bedeutet das, dass die zweite Achse möglichst wenig mit der Varianz zu tun haben soll, die auf Achse 1 abgebildet ist. Das ist der Grund, warum wir für die zweite Achse 2 Aufnahmen auswählen, die entlang der ersten Achse möglichst kleine Unterschiede zeigen (also dicht beieinander liegen).

Auch für die zweite Achse lassen sich nun mit der Zirkelmethode alle Aufnahmen auf die Achse abtragen (5.4 b). Wenn wir nun die beiden Achsen senkrecht aufeinander stellen (sie sind ja weitgehend unkorreliert, s. oben!) und die jeweiligen Werte als x - bzw. y -Koordinaten abtragen, entsteht eine 2dimensionale Ordination der Aufnahmen.

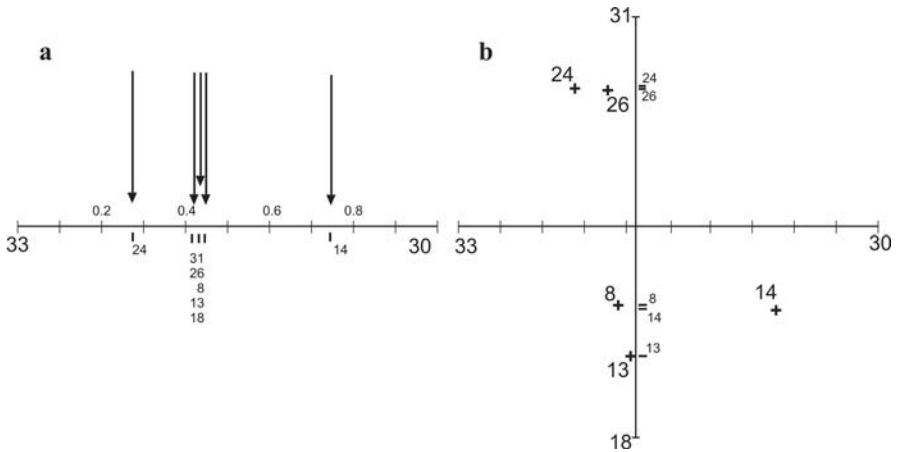


Abb. 5.4. **a** Ordination mehrerer Aufnahmen auf der ersten Achse durch Fällen der Lote. **b** Gleiches Vorgehen für eine zweite Achse, die von 2 Aufnahmen im mittleren Bereich der ersten Achse aufgespannt wird (Nr. 18 und 31). Die Kreuze geben die endgültige Lage der Objekte im 2dimensionalen Raum an

Abbildung 5.5 a zeigt das Ergebnis einer polaren Ordination für unseren Standarddatensatz. Sie zeigt die beiden wichtigsten Ordinationsachsen, also in diesem Fall die beiden wichtigsten Gradienten floristischer Ähnlichkeit. Um die Interpretation zu erleichtern, wurde in einem zweiten Schritt wieder der Grundwasserstand der Aufnahmen mit der Ordination verknüpft (Abb. 5.5 b), und in der Abbildung durch unterschiedlich große Symbole dargestellt. Es zeigt sich, dass auch hier die erste Achse einem Feuchtigkeitsgradienten entspricht, denn Flächen mit einem niedrigen Grundwasserstand liegen eher im rechten Bereich der Grafik, Aufnahmen aus der rezenten Aue liegen im oberen Bereich, Aufnahmen vom Auenrand eher im unteren Bereich.

Durch die polare Ordination wurden also die maximal möglichen 53 Achsen erfolgreich auf wenige wichtige reduziert; dennoch ist das Ergebnis dieser simplen Ordinationsmethode nicht ideal. Zwar werden die wichtigsten Gradienten erkannt, aber die Reihenfolge der Achsen ist gerade bei sehr heterogenen Datensätzen relativ zufällig, da viele Aufnahmen keine Arten gemeinsam haben und somit als Endpunkte der ersten Achsen in Frage kommen. Die ersten Achsen sind also nicht unbedingt Achsen absteigender Wichtigkeit.

Weil die ersten beiden Achsen letztlich nur von 4 Aufnahmen aufgespannt werden, ist auch die Dimensionsreduktion nicht immer günstig. Daher hat die polare Ordination heute kaum noch praktische Bedeutung.

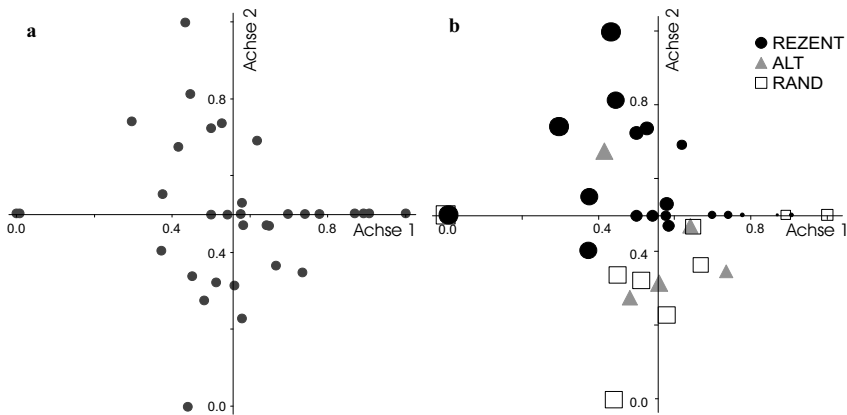


Abb. 5.5. **a** Polare Ordination der Elbauendaten. **b** Gleiche Ordination, aber zur Verdeutlichung der wichtigsten abiotischen Gradienten wurde der mittlere Grundwasserstand der Aufnahmelokalitäten durch unterschiedlich große Symbole angedeutet; die Form der Symbole gibt die Lage der Aufnahmen in der Aue an (Ähnlichkeitsmaß Bray-Curtis-Ähnlichkeit, Auswahl der Achsen nach der Standardmethode; zum Vergleich s. a. Abb. 5.2)

Allerdings gibt es inzwischen verbesserte Methoden (v. a. im Hinblick auf die Auswahl der Endpunkte), die in Vergleichsstudien gute Ergebnisse brachten (McCune et al. 2002). Polare Ordination sollte daher möglicherweise stärker zur Anwendung kommen. Je nach Fragestellung und Aufnahmedesign kann es auch gerechtfertigt sein, die Endpunkte bzw. Pole anhand externer Daten direkt auszuwählen. Dies gilt z. B. bei Transektstudien, wenn es darum geht, alle Aufnahmen in Bezug auf die Endpunkte des Transektes anzuordnen; auch bei Zeitreihenanalysen ergibt sich die Wahl der Endpunkte direkt aus dem Aufnahmedesign (z. B. Süß et al. 2004). Weitergehende Bemerkungen zu polarer Ordination finden sich bei Gauch (1994) und McCune et al. (2002).