

Diskussion und Schlussfolgerungen

Objekttyp: **Chapter**

Zeitschrift: **Veröffentlichungen des Geobotanischen Institutes der Eidg. Tech. Hochschule, Stiftung Rübél, in Zürich**

Band (Jahr): **90 (1986)**

PDF erstellt am: **25.08.2024**

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

10. Diskussion und Schlussfolgerungen

Vegetationskunde gehört zu den sogenannten "weichen" Wissenschaften. Sie beschäftigt sich mit Gegebenheiten, die sehr komplex, in der Regel schwer erkenn- und nachvollziehbar sind. Viele Störfaktoren verursachen Streuung in den Daten und erschweren das Erkennen einfacher Gesetzmässigkeiten. Ziel der statistischen Analyse ist es, interpretierbare Strukturen von blosser "Rauschen" zu trennen. Man kann nun davon ausgehen, dass die in den vorhergehenden Kapiteln dargestellten Lösungswege und Methoden bei vegetationskundlichen Daten eine solche Trennung herbeiführen können. Beweisbar ist diese Annahme nicht. Misserfolge sind zu erwarten. In der Literatur tauchen sie kaum auf, denn negative Ergebnisse werden selten publiziert. Die Vielfalt der statistischen Methoden lässt es als sehr wahrscheinlich erscheinen, dass Fehler im Auswertekonzept für viele Fehlversuche verantwortlich sind. Selbst bei Verwendung der besten Computerprogramme ist es möglich, klare Strukturen zu übersehen und richtige Ergebnisse als falsch zu betrachten. Erfolgsentscheidend ist insbesondere nicht nur die Auswertemethode, sondern der ganze Analyseweg (WILDI und ORLOCI 1983, S.10). Bei seiner Konzeption müssen die grundlegenden Probleme der Vegetationskunde berücksichtigt werden, wie zum Teil in den vorhergehenden Kapiteln angedeutet wurde. Wenn es schon schwierig ist, die Einsatzmöglichkeiten einzelner Methoden in verallgemeinerter Form darzustellen, so gilt dies noch viel mehr für den gesamten Auswertegang. Einige Grundsätze sollen gleichwohl nachfolgend diskutiert werden.

10.1 Grundgesamtheit und Stichprobe

Dem Problem der Stichprobenerhebung wurde in der Vegetationskunde bislang wenig Beachtung geschenkt (Kap. 2). Das liegt unter anderem daran, dass das Ziel vieler Untersuchungen in der Ermittlung qualitativer Zustände liegt. Im Gegensatz zu land- und forstwirtschaftlichen Fragestellungen, die oft dynamische Elemente enthalten, spielt die Interpolier- und Extrapolierbarkeit der Messungen, die sehr

stark auf dem Stichprobenkonzept beruht (ZÖHRER 1982), eine untergeordnete Rolle. Es ist klar, dass jede Stichprobe eine Schätzung erlaubt für die Grundgesamtheit, der sie entnommen wurde. Der Fehler dieser Schätzung ist jedoch nur quantifizierbar, wenn die Stichprobe mit Hilfe eines Zufallsstichprobenplanes gewonnen wurde. Solche Überlegungen sind bei explorativer Datenanalyse nicht zwingend. PIELOU (1984) betrachtet konsequenterweise ihre Datensätze als Gegenstand der Untersuchung. Ob es sich dabei um die eigentliche Grundgesamtheit oder um eine Stichprobe handelt, hinterfragt sie nicht. Und in der Tat bleibt die explorative Analyse der Daten in beiden Fällen genau dieselbe.

Wenn nun in Kapitel 2 die Ansicht vertreten wird, es sollte mit Zufallsstichproben gearbeitet werden, so liegt der Grund nicht in den Erfordernissen der Analyse, sondern in der besseren Interpretierbarkeit der Ergebnisse. Die Vegetationsdecke ist von einer enormen Variabilität gekennzeichnet. Bei subjektiver Auswahl der Stichprobeflächen wird meist versucht, ihr aus dem Wege zu gehen. Die Ergebnisse werden damit klarer, verlieren aber an Allgemeingültigkeit. In Zufallsstichproben erscheinen anstelle von Gruppen von Aufnahmen oft Gradienten. Letztere sind schwieriger zu beschreiben und weniger leicht umzusetzen für praktische Anwendungen. Besonders Ordinationen decken das Fehlen echter Unstetigkeiten unweigerlich auf. Die Erkenntnis, dass eine Stichprobe kaum eine Gruppenstruktur aufweist und mit viel Rauschen behaftet ist, kann aber nicht als negativ beurteilt werden. Es ist vielmehr zu prüfen, ob der Befund nicht den wirklichen Verhältnissen entspricht - ganz im Sinne von GREEN (1979), S. 63f, der feststellt, dass man einem sauber erarbeiteten Ergebnis Vertrauen entgegenbringen soll.

10.2 Methodische Entscheidungsschritte

Die explorative Analyse von Daten ist ein Experimentieren. Das Experiment soll aber nicht planlos verlaufen, sondern einer noch zu diskutierenden Strategie entsprechen. Von Experimenten im Labor und von Felduntersuchungen wissen wir,

dass sie auf zahlreichen, willkürlichen Entscheidungen beruhen. Entscheidungen schränken den Gültigkeitsbereich der Ergebnisse ein. Sie sind in jegliche Interpretation als zusätzliche Randbedingung mit einzubeziehen. Beispiele sind die Wahl des Untersuchungsgegenstandes, des Untersuchungszeitpunktes, der Messparameter und der Messmethode, usw. Bei der explorativen Datenanalyse geht dieser Prozess weiter, denn in der Regel gibt es kein ganz bestimmtes statistisches Modell, auf welches die Untersuchung zugeschnitten ist. Der Gliederung dieser Schrift entsprechend kann aber wiederum schrittweise vorgegangen werden. Es sind namentlich Entscheide zu fällen über die Skalierung, die Transformation, die Wahl des Aehnlichkeitsmasses, die Methode zur Analyse der Aehnlichkeitsstruktur, die Art der Weiterverarbeitung der Ergebnisse usw. Was resultiert, sind Hypothesen - zur Entstehung und Funktionsweise eines Vegetationsmusters, zur Ursache und zum Verlauf einer Veränderung, zur Möglichkeit der Steuerung usw. Die Methoden der explorativen Statistik helfen deshalb, Hypothesen zu formulieren. Liegen letztere vor, so kann eigentlich erst abgeklärt werden, ob sie mit Hilfe neuer Experimente oder Untersuchungen statistisch getestet werden können und sollen. Das wird in einigen Fällen sinnvoll sein, in anderen nicht. Ueberlegungen zum Sinn statistischer Tests in der Feldbiologie diskutiert GREEN (1979) eingehend. Wir gehen hier davon aus, dass der Engpass zum Verständnis ökologischer Prozesse beim Finden von Hypothesen hinreichenden Erklärungsgrades liegt.

10.3 Elemente einer Standardstrategie

Nehmen wir an, dass ein Datensatz zu analysieren ist, der eine Lebensgemeinschaft beschreibt. Sowohl in der Pflanzensoziologie, wie auch in der Faunistik fallen dabei zumeist Zähl- oder Deckungswerte an. Der erste Schritt besteht nun darin, eine geeignete Skalierung zu finden. Besonders Individuenzahlen schwanken im allgemeinen in einem so grossen Bereich, dass die Vergleichbarkeit verschiedener Arten mit einem der üblichen Standardisierungsverfahren nicht zu erreichen ist. Es ist deshalb eine geeignete (ordinale) Skalierung zu finden, die sogar von Art zu Art verschieden

sein kann. Die sich später ergebenden Resultate basieren natürlich vollständig auf der neuen Skala. Dieser Schritt ist sogar oft der kritischste in der ganzen Analyse. Besonders bei Deckungswerten ist es sinnvoll, den gesamten Auswertegang mit unterschiedlichen Skalen durchzuspielen. Im einen Fall soll vor allem die Gegenwart der Arten berücksichtigt werden (Präsenz-Absenz Skala), im andern Falle vor allem deren Quantitäten. Die sich ergebenden Unterschiede zeigen, welches Vorgehen sinnvoll ist und ob allenfalls eine Zwischenlösung gewählt werden muss.

Neu skalierte Daten weiter zu transformieren, ist durchaus denkbar. Es ist zu überlegen, ob sie als Zähldaten interpretiert werden können, so dass eine Weiterverarbeitung mit einem Kontingenzmass oder im Sinne der Korrespondenzanalyse möglich ist. Bereichsanpassung, Normalisierung und Standardisierung eignen sich allenfalls für den Vergleich von Aufnahmevektoren. Sie sind mit Bedacht zu verwenden für Artvektoren, die in der Regel viele Nullen aufweisen. Auch hier lohnt es sich, im Zweifelsfalle mit und ohne Standardisierung zu rechnen - womit bei Verwendung zwei verschiedener Skalierungen bereits vier Analysesequenzen durchzuspielen sind.

Ein dritter Schritt ist die Wahl des Vergleichsmasses. In vielen Fällen wird sie diktiert durch die später anzuwendende Analysemethode. Zu beachten ist, dass verschiedene Vergleichsmasse eine Transformation in sich bergen, was einen ungewollten Informationsverlust zur Folge haben kann. Beliebte sind einfache Masse, wie die Euklidische Distanz oder eine Absolutwertfunktion. Beide sind für die Analyse von Lebensgemeinschaften unbedenklich, sofern eine sinnvolle Skalierung gewählt wurde. Dagegen erscheint die Verwendung des Korrelationskoeffizienten meist ungerechtfertigt, denn er bringt eine oft ungewollte Angleichung der Streuungsverhältnisse mit sich.

Eine der charakteristischen Eigenschaften eines multidimensionalen Datensatz ist dessen Dimensionalität. Diese sollte in einem vierten Schritt geklärt werden. Die Haupt-

komponentenverfahren bringen zwar nur eine lineare Lösung des Problem, die aber meist gute Hinweise auf die Komplexität der Stichprobe liefert. Lebensgemeinschaften können im allgemeinen nur unter Berücksichtigung mehrerer Dimensionen befriedigend beschrieben werden. Typisch ist, dass auch sechs bis zehn Achsen oft nur 30 bis 40% der Gesamtvarianz erklären. Die Erfahrung lehrt, dass dieser an sich bescheidene Anteil genügt, um die wichtigsten Strukturen und Gesetzmässigkeiten erkennen zu können. Eine Ordination der Aufnahmen zeigt auch gleich, ob eine Gruppenstruktur vorliegt. Sofort erkennbar sind Ausreisser oder deutlich abgetrennte Teile der Stichprobe. Sie können die Analyse dominieren und alle andern Strukturen verwischen. Bei Eindeutigkeit sind sie aus der Stichprobe zu entfernen (im Sinne einer Stratifizierung) und separat auszuwerten.

Hauptkomponentenverfahren geben auch Auskunft über die Bedeutung der Arten in einer Ordination der Aufnahmen. Eine Alternative dazu stellen Gewichtungsverfahren (Rangierungen) dar, die meist leichter zu interpretieren sind. Eine sehr sorgfältige Interpretation der Ergebnisse einer Artenrangierung lohnt sich nicht zuletzt deshalb, weil hier noch am ehesten Fehlgriffe in der Methodenwahl erkannt werden können. Bei unzuweckmässig skalierten Daten können Arten ein hohes Gewicht erhalten, die ausnahmsweise massenhaftes Vorkommen aufweisen - eine in der Natur ganz normale Erscheinung. Gegebenenfalls ist das Auswertekonzept anzupassen.

Die Suche nach Gruppen hat als fünfter Schritt praktische Bedeutung, denn viele Anwendungen erfolgen auf der Basis von generalisierenden Typen von Lebensgemeinschaften. Der alleinige Einsatz von Gruppierungsverfahren ist möglich, die Ergebnisse lassen sich aber im allgemeinen nicht genügend gut interpretieren. Insbesondere sollte der Gruppierungsalgorithmus der aus der Ordination ersichtlichen Aehnlichkeitsstruktur angepasst sein. Liegt eine diffuse Punktwolke vor, so generiert ein Teilungsverfahren die gewünschten Gruppen - welche dann aber künstlicher Natur sind. Agglomerative Verfahren liefern immer ein Ergebnis. Es

bleibt aber meist einem weiteren Analyseschritt vorbehalten, herauszufinden, ob die Gruppenunterteilungen Unstetigkeiten in den Daten entsprechen. Methoden wie die GRID-Analyse lassen eine beliebige Teilung der Stichprobe nicht zu. Ihrem Verhalten kann entnommen werden, ob allenfalls echte Gruppen vorliegen. Zusammenfassend kann festgestellt werden, dass auch beim Einsatz von Gruppierungsverfahren oft zweckmässigerweise Versuche mit unterschiedlichen Methoden durchgeführt werden sollten.

Nach einer sorgfältigen Interpretation der soweit angefallenen Ergebnisse umfasst der sechste Schritt die Weiterverarbeitung der Resultate. Diese liegen ja in der Regel wieder als Daten vor, nämlich als Ordinationskoordinaten, Gruppenmittelwerte und Streuungen usw. Das Spektrum der in Frage kommenden Methoden erstreckt sich über die bereits dargelegten Möglichkeiten hinaus auf spezielle Methoden wie die Konzentrations-, die Diskriminanzanalyse und viele andere. In komplexen Fällen wird man nicht umhin kommen, die verfügbaren Hypothesen zum Entwurf eines mathematischen Simulationsmodells zu verwenden, um zu einer Ueberprüfung der Ergebnisse zu gelangen.

Die meisten Analysemethoden ergeben informativere Ergebnisse, wenn mit kleinen Datensätzen gearbeitet wird. Es ist deshalb wohl immer vorteilhaft, wenn auf Grund einer ersten Analyse eine begründbare Unterteilung des Datensatzes vorgenommen werden kann. Für jeden Datenteil beginnt dann die Analyse wieder von vorn. Viele strukturelle Details werden damit leichter erkennbar sein.