

Habitat-eignungsmodelle auf der Grundlage von Expertenwissen – ein Vergleich von Fuzzy-Logik Ansätzen am Beispiel westmediterraner Zwergbinsenrasen

Habitat suitability models based on expert knowledge – a comparison of fuzzy set approaches using the example of Western Mediterranean dwarf rush communities

Michael Rudner

Abstract

This study examines the ability of two fuzzy logic approaches to adequately describe the realised niche of a specific target taxon in Western Mediterranean dwarf rush communities (Isoeto-Nanojuncetea). In a strongly fuzzified approach, which required a large set of rules, all independent and dependent variables are coded as linguistic variables. This approach is compared to a different approach which defines habitat suitability through a membership function per taxon and independent variable. Both approaches yield acceptable results. Advantages and disadvantages of both fuzzy logic approaches are discussed. It is shown that this method is well suited for the description of realised niches and that the approach should be chosen depending on the question of the study. Knowledge based model building is a crucial step for the integration of rare taxa in landscape models.

Keywords: Fuzzy Logic, Isoeto-Nanojuncetea, Weighted kappa, Vegetation

Zusammenfassung

Ausgehend von der Frage, ob es möglich ist, für selten auftretende Pflanzenarten oder Vegetationstypen auf der Grundlage vorhandenen Expertenwissens ein Modell zu erstellen, das die Realniche des Zieltaxons zutreffend beschreibt, werden hier zwei Fuzzy-Logik-Ansätze verglichen. Am Beispiel westmediterraner Zwergbinsenrasen (Isoeto-Nanojuncetea) wird untersucht, ob sich ein weitgehend fuzzifizierter Ansatz, der alle Eingangs- und Ergebnisvariablen als linguistische Variablen behandelt und einen umfangreichen Regelsatz benötigt, sich besser eignet als ein Ansatz, bei dem für jedes Zieltaxon die Habitat-eignung über eine Zugehörigkeitsfunktion pro Umweltvariable formuliert wird. Die Modellierung gelingt mit beiden Verfahren ähnlich gut. Die Vor- und Nachteile der beiden Fuzzy-Logik-Ansätze werden abgewogen. Insgesamt wird klar, dass sich dieser Weg für die Beschreibung der Realniche gut eignet und die Wahl der Variante in Abhängigkeit von der Fragestellung erfolgen sollte. Die Erarbeitung wissenschaftlicher Habitat-eignungsmodelle stellt einen sinnvollen Schritt dar, um seltene Taxa in Landschaftsmodelle mit aufzunehmen.

Schlüsselwörter: Fuzzy-Logik, Isoeto-Nanojuncetea, Gewichtetes Kappa, Vegetation

1 Einleitung

In jüngerer Zeit machen der Aufbau, die Vereinheitlichung und die Nutzung von Datenbanken für Vegetationsdaten gro-

ße Fortschritte. Parallel werden auch Modellierungsverfahren weiterentwickelt, die gestützt auf diese Daten die ökologische Nische von Pflanzenarten und Vegetationstypen beschreiben und auch zur Vorhersage der Auswirkungen klimatisch oder durch Nutzung bedingter Veränderungen herangezogen werden können (vgl. GUI SAN & ZIMMERMANN 2000, AUSTIN et al. 2006, ELITH et al. 2006, 2009, SCHRÖDER et al. 2008).

Häufig werden statistische Habitatmodellierungsansätze, wie GLM (verallgemeinerte lineare Modelle), GAM (verallgemeinerte additive Modelle), MARS (multivariate adaptive Regression-Splines) oder Klassifikationsverfahren wie CART (Klassifikations- und Regressions-Entscheidungsbäume) verwendet (GUI SAN & ZIMMERMANN 2000, ELITH et al. 2006). Maschinelle Verfahren wie BRT (verstärkte Regressions-Entscheidungsbäume) oder MaxEnt (Maximum-Entropie) kommen in jüngerer Zeit zum Einsatz (ELITH et al. 2006). Liegen nur Daten zu Vorkommen vor, kommen Verfahren wie bioklimatische Umgrenzung (*envelope models*) oder ökologische Nischenanalyse (ENFA) in Frage (HIRZEL et al. 2002). Die statistischen Modellierungsverfahren können auch verwendet werden, wenn zusätzlich zu den Vorkommensdaten Proben zum standörtlichen Hintergrund (sog. Pseudo-Absenzen) einbezogen werden. Unter den Programmen, die für seltene Arten oder für Datensätze mit starkem Rauschen geeignet sind, führen diejenigen, die mit zusätzlichen Pseudo-Absenzen arbeiten, zu den besseren Ergebnissen. Insbesondere BRT, eine maschinelle Erweiterung von CART, zeigt vergleichsweise gute Ergebnisse. Die Regressionsansätze landen im Mittelfeld (GLM, GAM, MARS) während die Verfahren, die nur mit Präsenzen arbeiten, abgeschlagen sind (ELITH et al. 2006, 2009).

Für zuverlässige Modellierungen mit den genannten Ansätzen ist eine Mindestanzahl von beobachteten Vorkommen der jeweiligen Art oder des jeweiligen Vegetationstyps im Datensatz erforderlich (HARREL 2001, STOCKWELL & PETERSON 2002, COUDUN & GÉGOUT 2006). Für seltene Taxa oder Syntaxa, deren Erhaltung in der Landschaft für den Naturschutz von großer Bedeutung ist, ist die formale Beschreibung der Nische mit Habitat-eignungsmodellen auf diesem Wege daher schwierig – zuverlässige Vorhersagen sind unmöglich. In solchen Fällen ist es sinnvoll, über die Formalisierung von Expertenwissen in regelbasierten Modellen nachzudenken. Das Wissen von Experten birgt in der Regel gewisse Unschärfen, auch decken sich die Einschätzungen verschiedener Experten nicht exakt – zudem ist sowohl biotischen wie auch abiotischen Variablen eine gewisse Unschärfe inhärent (MOUTON et al. 2009, SVORAY et al. 2004). Diese Unschärfen sollten ebenso wie Übergänge entlang der Umweltgradienten beim

Übertragen des Wissens in Regeln abgebildet werden. Beides kann im Fuzzy-Logik-Ansatz umgesetzt werden. Damit werden handhabbare Habitatmodelle auch für seltene Arten oder Vegetationstypen möglich. Bei diesem Ansatz, auch bekannt als *Fuzzy Sets*, der in den 1960er Jahren von Zadeh in der Informatik eingeführt wurde (ZADEH 1965), werden diskrete Zuordnungen herkömmlicher Klassifikation in Anteile der Zugehörigkeit zu den verschiedenen Zuständen aufgelöst. In der Ökologie findet dieser Ansatz seit der Mitte der 1990er Jahre auch zunehmend Verwendung (von 1996 waren 40 Arbeiten, von 2008 über 140 Arbeiten bei *web of science* gelistet). Die räumlich explizite Modellierung des Vorkommens von Vegetationstypen mit funktionellem Hintergrund ist auf diesem Weg eher umzusetzen als mit statistischen Ansätzen (SVORAY et al. 2004, 2008). Auch dynamische Systeme werden modellierbar (RÜGER et al. 2005).

Ziele der Anwendung von *Fuzzy Sets*

- Formalisierung des vorhandenen Expertenwissens zur möglichst zutreffenden Beschreibung der Realnische, Überprüfen dieses Modells auf Vollständigkeit (d. h. auf die Berücksichtigung aller wesentlichen Standortfaktoren).
- Gültige Habitateignungsmodelle für die untersuchten Arten, die auch verwendet werden können, um a) weitere Vorkommen zu suchen (geeignete Räume), b) Auswirkungen von Veränderungen der Umwelt vorherzusagen – in die Regeln können auch funktionale Überlegungen einfließen, was einen Vorteil gegenüber statistischen Modellen darstellt (SVORAY et al. 2008).

Diese beiden Ziele machen Wissen über seltene Arten oder Zönosen nachvollziehbar und erlauben es im günstigen Falle, Prognosen auf einer formalen Basis zu treffen. Dies kann als Argumentationshilfe für Entscheidungen im Naturschutz herangezogen werden (z. B. ENGLER et al. 2004, BUCERIUS et al. 2006, HEATON et al. 2008).

Maschinenlernende Systeme wie z. B. Neuronale Netze werden nicht als Methoden in Betracht gezogen, da das Regelwerk selbst ein wertvolles Ergebnis ist. Es trägt maßgeblich zum Verständnis des Vorkommens der untersuchten Taxa bei. Maschinenlernende Systeme beinhalten i. d. R. eine *black box* und geben damit keine klar und einfach formulierten Regeln aus (KRUSE et al. 1995).

Ziel der vorliegenden Arbeit ist der Vergleich zweier *Fuzzy Set*-Methoden am Beispiel westmediterraner Zwergbinsenrasen, die in Portugal auf Gruskissen über Fels auftreten und im Wesentlichen durch ihren Wasserhaushalt bestimmt sind. Die Einnischung der unterschiedlichen Typen der Zwergbinsenrasen ist schwer datengestützt zu modellieren, lässt sich aber auf der Grundlage von Expertenwissen formalisiert beschreiben. Der Vergleich wird zeigen, ob eine parallele Modellierung verschiedener Vegetationstypen mit anschließender Auswahl des vorhergesagten Typs – eine Variante, die Entscheidungsbäumen wie z. B. CART ähnelt – besser geeignet ist als eine Reihung der Vegetationstypen entlang eines Gradienten mit direkter Vorhersage des Typs, was vom Vorgehen an die Regressionsverfahren wie GLM oder GAM erinnert.

2 Methoden

2.1 *Fuzzy Set*-Ansatz

Im betrachteten Anwendungsfall – quantitative Beschreibung der Nische von Pflanzenarten oder Vegetationstypen, die relativ selten auftreten – müssen die Regeln des Modells auf verfügbares Expertenwissen gegründet werden, d. h. auf schriftliche Quellen wie regionale Floren, Zeigerwerte oder das Wissen der Experten selbst.

Bei sogenannten *Fuzzy-Sets* wird die Zugehörigkeit durch einen Wert auf dem Einheitsintervall [0;1] ausgedrückt. Der Wert der Kurve kann als Habitateignungs-Wert im Gegensatz zur Auftretenswahrscheinlichkeit interpretiert werden. Bei Kurven der Wahrscheinlichkeitsdichte müsste die Fläche unter der Kurve den Wert 1 annehmen. Das ist bei *Fuzzy-Sets* nicht der Fall. Mit diesen Zugehörigkeitsfunktionen können sogenannte linguistische Variablen definiert werden (Abb. 1). Die Summe der Kurven muss ebenfalls nicht den Wert 1 ergeben (KRUSE et al. 1995). Die klassische Funktion ist die Trapezfunktion, die einen linearen Anstieg zwischen Minimum und Maximum beschreibt (LEHMANN et al. 1992). Um den Zusammenhang zu verstärken oder abzuschwächen, kann die Funktion quadriert oder radiziert werden. Auch ein sigmoidaler Kurvenverlauf ist denkbar. Linguistische Variablen sind sehr nützlich, wenn die Umweltvariablen nicht gemessen werden können, sondern geschätzt werden müssen.

Menschen können in der Regel maximal neun Zustände voneinander unterscheiden (vgl. Sherman-Kent-Skala in KRUSE et al. 1995). Deshalb und weil die Anzahl der Regeln mit der Anzahl der Zustände kardinal steigt, ist anzuraten, möglichst wenige Zustände pro Variable zu beschreiben. Dabei kann die Differenzierung entlang des Gradienten unterschiedlich fein sein. Eine Zugehörigkeitsfunktion kann für die Definition einer linguistischen Variablen aber auch als direkte Definition der Habitateignung für einen Typ in Abhängigkeit vom Wert der (gemessenen) unabhängigen Variablen eingesetzt werden. In diesem Artikel werden zwei *Fuzzy Set* Ansätze vorgestellt, die im Folgenden als Variante 1 bzw. Variante 2 bezeichnet werden.

2.2 Variante 1

Dieser Ansatz kann als MISO-Typ (*Multiple Input Single Output*) im Sinne von PEREIRA et al. (2009) betrachtet werden. Zunächst sind die im Modell zu verwendenden Umweltvariablen (unabhängige Variablen) als linguistische Variablen aufzustellen (Abb. 1). Dann müssen die Vegetationstypen (abhängige Variable) entlang eines Gradienten gereiht werden (linguistische Ergebnisvariable).

Sodann ist ein Regelwerk aus linguistischen Variablen, das pro Kombination der Zustände der Umweltvariablen einen Vegetationstyp voraussagt, zu erstellen. Dadurch entsteht eine große Anzahl an Regeln (vgl. Tab. 1) (PEREIRA et al. 2009).

Bei der Auswertung der einzelnen Regeln wird der Minimum-Operator (logisches UND) angewandt, d. h. das Ergebnis wird mit dem am wenigsten zutreffenden Zustand der Umweltvariablen gewichtet. Aus den Ergebnissen wird entlang der Outputvariablen jeweils der höchste Ergebniswert zurückgehalten (Maximum-Operator, logisches ODER). Aus der Ergebniskurve soll wieder ein diskretes Ergebnis ermittelt

Tab. 1: Ausschnitt eines Regelwerks zu einem Modellbeispiel (vgl. Abb. 1).

Tab. 1: Extract of a rule set for a modelling example (see Fig. 1).

WENN Hangneigung = flach	UND Einzugsgebiet = groß	DANN Veg.typ = C
WENN Hangneigung = flach	UND Einzugsgebiet = klein	DANN Veg.typ = B
WENN Hangneigung = steil	UND Einzugsgebiet = groß	DANN Veg.typ = B
WENN Hangneigung = steil	UND Einzugsgebiet = klein	DANN Veg.typ = A
WENN Hangneigung = sehr steil	UND Einzugsgebiet = groß	DANN Veg.typ = A
WENN Hangneigung = sehr steil	UND Einzugsgebiet = klein	DANN Veg.typ = zu trocken

werden. Dazu wird der Schwerpunkt ausgelesen (vgl. Abb. 1). Es sind auch andere Algorithmen, wie etwa das gewichtete Mittel oder das Mittel der Maxima, denkbar (Joss et al. 2008). Diese Verfahren können nur bei konvexem Verlauf der Ergebniskurve sinnvoll angewandt werden. Andernfalls ist eine Entscheidung vonnöten, welcher Teil der Kurve weiter ausgewertet wird oder ob mehrere Ergebnisse nebeneinander gestellt werden. Als Werkzeug wird hier das frei verfügbare Java-basierte Programm *RockOn Fuzzy* verwendet (LIND et al. 2008).

2.3 Variante 2

Dieser zweite Ansatz, der auch für räumlich explizite Modellierungen verwendet wird (SVORAY et al. 2008), entspricht dem SISO-Typ (*Single Input Single Output*) im Sinne von PEREIRA et al. (2009). Bei dieser Variante müssen die einzel-

nen Zugehörigkeitsfunktionen für jede Umweltvariable und jeden Vegetationstyp aufgestellt werden (vgl. Abb. 2).

Die Variablen fungieren in diesem Ansatz als limitierende Größen. Eine Codierung als linguistische Variablen ist in diesem Fall nicht erforderlich. Aus dem kleinsten Wert, den ein Vegetationstyp in seinen Zugehörigkeitsfunktionen erhält, wird der Index für die Habitataignung abgeleitet (Minimum-Operator). Eine andere Möglichkeit stellt das gewichtete Mittel der Werte der einzelnen Zugehörigkeitsfunktionen dar. Die Gewichte der Faktoren werden nach Expertenwissen vergeben (SVORAY et al. 2004). Der vorhergesagte Vegetationstyp ist der mit dem höchsten Indexwert. Gegebenfalls erfolgt hier eine Gewichtung mit dem Kehrwert der Prävalenz (PEPPLER-LISBACH 2003). Die Berechnung wird mit dem Statistikprogramm R (R DEVELOPMENT CORE TEAM 2008) mit Hilfe von ad-hoc-Skripten und eigens definierten Funktionen durchgeführt.

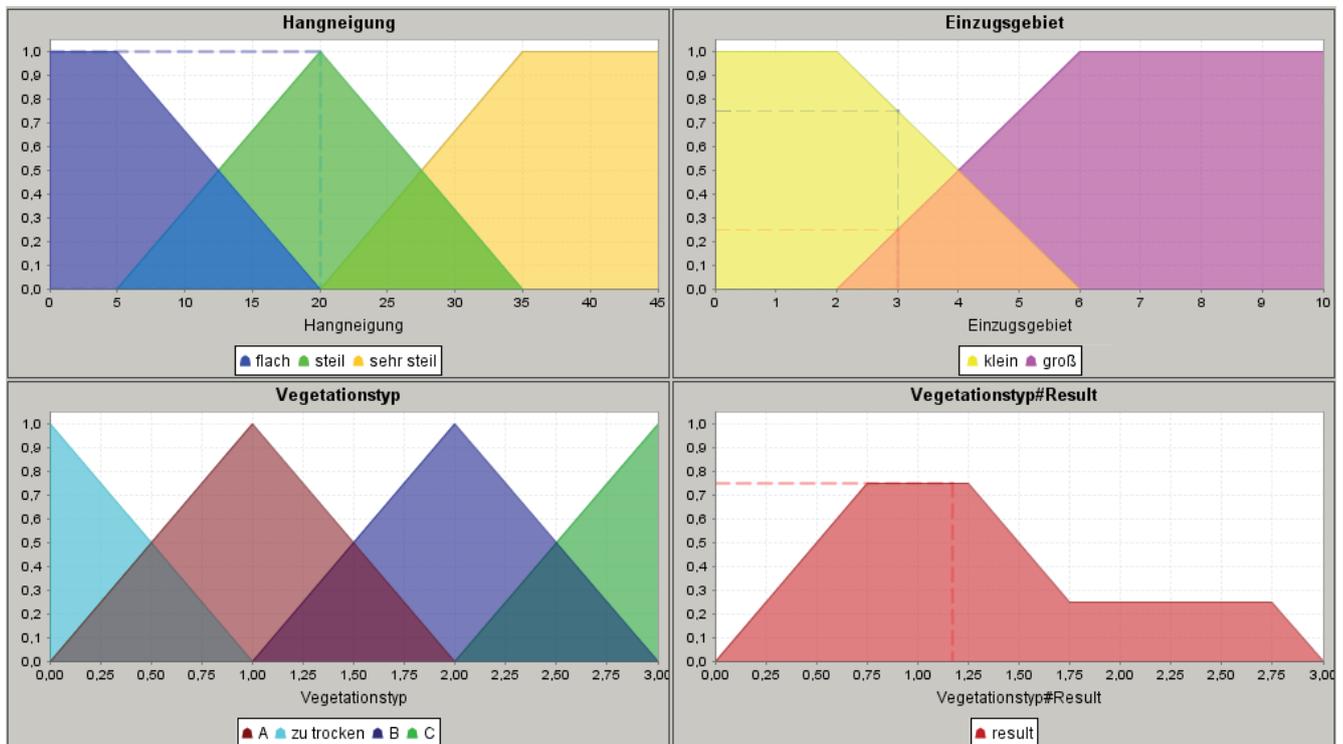


Abb. 1: Definition eines Modellbeispiels mit Hangneigung als 3-stufige linguistische Variable, Einzugsgebietsgröße als 2-stufige Variable und dem Vegetationstyp als 4-stufige linguistische Ergebnisvariable. Das Modell ergibt für eine Hangneigung von 20° und ein 3 Einheiten großes Einzugsgebiet defuzzifiziert einen Wert von 1,2 mit einer Habitataignung von 0,75. (Screenshot von *RockOn Fuzzy* vgl. LIND et al. 2008) Die Regeln des Modells sind in Tabelle 1 zusammengefasst.

Fig. 1: Exemplary definition of a model including slope inclination as a linguistic variable with 3 levels, specific catchment area as a 2 level variable and vegetation type as resulting linguistic variable with 4 levels. The model predicts a defuzzified value of 1.2 and the habitat suitability of 0.75 for a specific catchment area of 3 units with a slope inclination of 20°. (Screenshot of the *RockOn-Fuzzy* program by LIND et al. 2008) The rules of the model are presented in table 1.

2.4 Gütemaß

Das Erarbeiten eines gültigen Modells folgt einem heuristischen Ansatz. Die Güte des fertigen Modells wird durch den Vergleich mit beobachteten Daten ermittelt. Als Gütemaß wird das verteilungsfreie Konkordanzmaß Cohen's Kappa verwendet (BORTZ et al. 2000). Für Variante 1 kommt das Konkordanzmaß für kardinale Daten (*weighted Kappa*) zum Einsatz. Nach der Auswahl der Vegetationstypen wird für die Beurteilung des Gesamtmodells aus Variante 2 auch *weighted Kappa* verwendet. Die Berechnung erfolgt mit dem R-Paket *vcd* (MEYER et al. 2009).

2.5 Justieren des Modells

Steht ein, wenn auch kleiner, Datensatz zur Prüfung der Modellgüte zur Verfügung, so kann dieser auch dazu verwendet werden, das Modell zu justieren und dadurch zu verbessern. Der Datensatz spielt dann wie in statistischen Modellierungsverfahren die Rolle des Trainingsdatensatzes. Für die falsch klassifizierten Probenpunkte werden die unabhängigen Variablen nachgeschlagen. Daraus wird erkennbar, an welcher Stelle der Standortgradienten die Zugehörigkeitsfunktionen bzw. die Regeln nicht zum Datensatz passen. Eine Anpassung an alle falsch zugeordneten Probenpunkte ist sehr wahrscheinlich nicht möglich, da auf ökologisch interpretierbare Regeln bzw. Zugehörigkeitsfunktionen geachtet werden muss. Der Verlauf der Kurven muss i. d. R. konvex sein bzw. die Region, für die ein Taxon vorhergesagt wird, muss im Ökogramm eine zusammenhängende Fläche ergeben, um eine Überanpassung zu vermeiden. Nach dem Justieren werden erneut die Modellwerte und das Gütemaß berechnet. Für die Einschätzung der Übertragbarkeit des Modells auf andere Räume oder Zeiträume ist die Validierung mit einem Testdatensatz nötig. Dies wird bei wissensbasierten Modellen, die mangels verfügbarer Daten aufgestellt werden, nicht immer möglich sein.

2.6 Modellbeispiel Zwergbinsenrasen

Am Beispiel der westmediterranen Zwergbinsenrasen in der Serra de Monchique (Provinz Algarve, Portugal) werden beide *Fuzzy Set*-Varianten angewendet. Die Zwergbinsenrasen sind an diesem Standort primär. Die kleinräumige Differenzierung der Zwergbinsenrasen in verschiedene Typen wird im Wesentlichen durch den Bodenwasserhaushalt bestimmt. Entscheidende Größen sind die Menge des zur Verfügung stehenden Wassers, die Neigung des Standortes zu Austrocknung, die Dauer der Wasserzufuhr nach Ende der Niederschläge bzw. die Dauer der Vernässung (RUDNER 2005). Die Grundlage für die Modelle bildet das Expertenwissen des Autors. Die Zwergbinsenrasen werden entsprechend RUDNER (2005) typisiert und folgendermaßen von geringmächtigen Böden mit kurzfristiger Überstauung und rascher Austrocknung zu mächtigeren Böden und längerer Durchfeuchtung gereiht:

- *Crassula tillaea*-Gesellschaft (CT)
- *Junco capitati*-Isoetetum *histicis* (JI)
- *Loto subbiflori*-*Chaetopogonetum fasciculati* (LC)
- *Laurentio michelii*-*Juncetum tingitani* (LJ)
- *Laurentio michelii*-*Juncetum tingitani pinguiculetosum lusitanicae* (LJp)

Die Modellierung erfolgt bei beiden Varianten in zwei Schritten analog zum Vorgehen bei RÜGER et al. (2005) oder SVORAY et al. (2004). Im ersten Schritt wird aus den Variablen spezifisches Einzugsgebiet, Hangneigung und Dauer der Wasserzufuhr (Überrieselung) ein Feuchte-Index ermittelt. Dieser Index wurde in Anlehnung an den *wetness index* von KIRKBY & WEYMAN (1974) sowie BEVEN (1986) aufgestellt

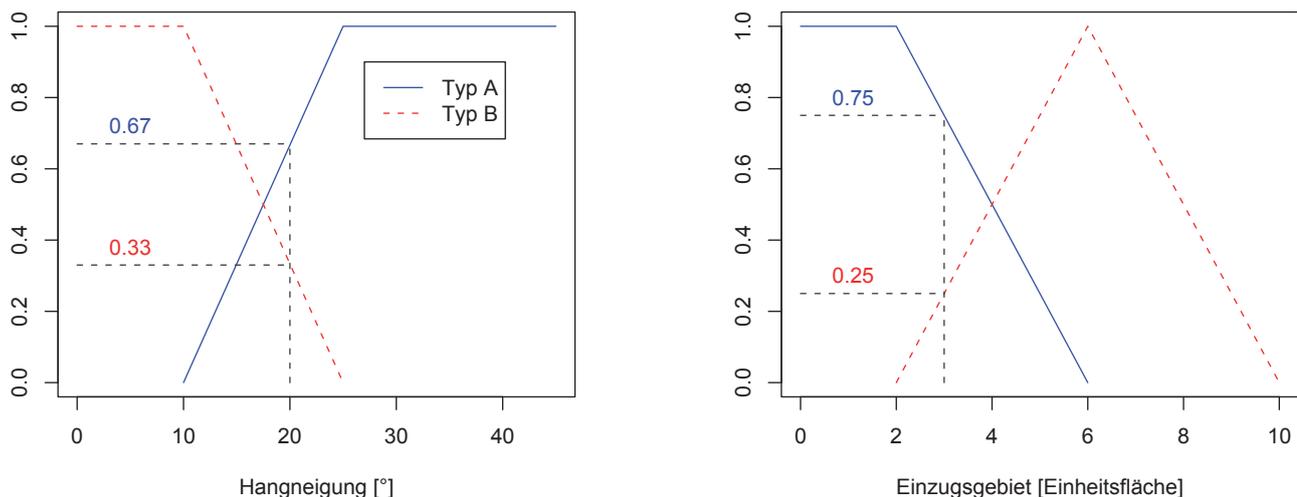


Abb. 2: Ableiten der Habitateignung aus den Zugehörigkeitsfunktionen von Vegetationstyp und Umweltvariablen. Für Typ A ergibt sich bei einer Hangneigung von 20° eine Habitateignung von 0,67, für Typ B von 0,33. Aus einer Einzugsgebietsgröße von 3 FE (Flächeneinheiten) leitet sich eine Habitateignung von 0,75 für Typ A und von 0,25 für Typ B ab. Unter Anwendung des Minimum-Operators ergibt sich für einen Standort mit 20° Hangneigung und einer Einzugsgebietsgröße von 3 FE eine Habitateignung von 0,67 für Typ A und 0,25 für Typ B.

Fig. 2: Deduction of the habitat suitability by use of the membership functions of vegetation type and environmental variables. The habitat suitability for type A at a slope inclination of 20° amounts to 0.67, for type B it is 0.33. The specific catchment area of 3 units leads to a habitat suitability of 0.75 for type A and of 0.25 for type B. Applying the minimum operator, a habitat suitability of 0.67 for type A and of 0.25 for type B result for a habitat with 20° inclination and a specific catchment area of 3 units size.

und steht für die zur Verfügung stehende Wassermenge und die Neigung des Standortes zu vernässen. In einem zweiten Schritt steht ein Regelwerk zum Auftreten der Vegetationstypen mit den Variablen Feuchte-Index, Textur (mittlere Korngröße, logarithmisch skaliert), Lage auf der Grusdecke und Bodenmächtigkeit. Die Lage auf der Grusdecke (oberer Rand, Mitte, unterer Rand) ist entscheidend für den Zeitpunkt des Einsetzens der Austrocknung nach dem Ende der Niederschläge (vgl. RUDNER 2004). Die Bodenmächtigkeit und mittlere Korngröße steuern die nutzbare Feldkapazität im effektiven Wurzelraum. Die Modelle werden im Ergebnisteil vorgestellt.

In einem kleinen Einzugsgebiet in der Serra de Monchique wurden entlang von mehreren Transekten über Grusdecken deren Hangneigung, Mächtigkeit und Vegetation vor Ort erfasst. Zusätzlich liegen Daten zur Textur der Grusdecken und zur Dauer der Wasserzufuhr vor. Die Vegetationstypen sind dort sehr kleinflächig ausgebildet. Die Mächtigkeit der Grusdecken über Fels, der als Stauer fungiert, variiert von 0 bis 40 cm. Dieser Datensatz wird für eine einmalige Justierung der Modelle und zur Bewertung der Güte der Modelle herangezogen.

3 Ergebnis

3.1 Variante 1 – Linguistische Variablen und Regeln

Der Regelsatz für Variante 1 ist in Abbildung 3 dargestellt. Die Diagramme, die für jede Korngrößengruppe dargestellt sind, können wie Ökogramme gelesen werden. Die Vorkommen der Zwergbinsenrasen werden auf der geringmächtigen, trockenen Seite des Standortgradienten begrenzt durch den Vegetationstyp „trocken“. Auf diesen Habitaten stehen meist annuelle mediterrane Sonnenröschenfluren (*Helianthemetea*). Die Ränder der Grusdecken neigen meist stärker zur Vernässung als die mittleren Bereiche (insbesondere bei grobkörnigem Substrat), weshalb hier häufiger Zwergbinsenrasen anzutreffen sind. Die Reihung der Zwergbinsenrasengesellschaften (*Isoto-Nanojuncetea*) von der sehr offenen *Crassula tillaea*-Gesellschaft, die ihre Entwicklung rasch abschließen kann und daher auch mit nur kurzzeitig gutem Wasserangebot auskommt, führt über das artenarme *Juncisoetum*, das noch relativ trocken stehen kann, und das *Loto-Chaetopogonietum*, das sich recht spät im Frühjahr entwickelt und seinen Schwerpunkt in besser und länger wasserversorgten durchaus auch

geringmächtigen Abschnitten der Grusdecken hat, zum *Laurentio-Juncetum*. Dieses ist auf eine gute Wasserversorgung und einen mächtigeren Bodenkörper angewiesen. Den Abschluss der Reihe bildet die Subassoziation *Laurentio-Juncetum pinguiculetosum*, die auf den unteren Rand von Grusdecken beschränkt ist, wo sehr lange Wasser austritt. Im Testdatensatz ist diese Gesellschaft unterrepräsentiert. Auf den tiefgründigen Böden schließen sich bei niedrigem Feuchte-Index ausdauernde Rasen, bei höherem Feuchte-Index ausdauernde Sumpfgesellschaften wie etwa *Juncus acutus*-Gesellschaften an.

3.2 Variante 2 – Zugehörigkeitsfunktionen

Diese Methode erlaubt eine relativ einfache Aufstellung von Regeln nach Expertenwissen und einmaliger Anpassung an einen Datensatz mit beobachteten Vorkommen und Nichtvorkommen. Sie ist flexibel bezüglich der Beschreibung des Vorkommens entlang der Gradienten der Umweltvariablen.

Aber jede Regel steht hier für sich, ist also nicht mit den anderen Regeln verknüpft. Die Zugehörigkeitsfunktionen sind in Abbildung 4 zusammengefasst. Die Eignung des Standorts ist für jeden Vegetationstyp und Standortfaktor in einer Kurve dargestellt. Die Regeln für die beiden Gesellschaften *Laurentio Juncetum* und *L.J. pinguiculetosum* unterscheiden sich bezüglich des Bereiches geeigneter Bodenmächtigkeit und in der Zuordnung der Position auf der Grusdecke, wobei die Fettkraut-Subassoziation auf den unteren Rand beschränkt ist. Das *Loto-Chaetopogonietum* ist auf flachgründigeren Böden eingemischt und dort auf die Randbereiche der Grusdecken beschränkt. Es kommt auf feinkörnigerem Substrat vor als das *Juncisoetum*. Dieses steht auf gering bis mittel mächtigen sandig-grusigen frischen Bereichen der Grusdecken. Die *Crassula tillaea*-Gesellschaft steht auf geringmächtigeren und etwas trockeneren Böden als die vorgenannte Gesellschaft.

3.3 Vergleich der Methoden

Beide Varianten eignen sich für die Beschreibung der Realnische für seltene Vegetationstypen. Es werden Kappa-Werte um 0,4 erreicht. Dabei ist bezüglich der Güte der Beschreibung die Variante 2 mit der Abbildung der Regeln in Zugehörigkeitsfunktionen etwas überlegen (vgl. Tab. 2). Die direkte Darstellung der Zugehörigkeitsfunktionen bietet einen guten Blick auf die Beziehung zwischen Vegetationstyp und

Tab. 2: Kontingenztafeln zu den vorhergesagten Vegetationstypen. Links Variante 1 ($\kappa = 0,38$), rechts Variante 2 ($\kappa = 0,41$), Beobachtungen in Spalten, Prognosen in Zeilen (---: kein Zwergbinsenrasen, CT: *Crassula tillaea*-Gesellschaft, JI: *Juncisoetum*, LC: *Loto-Chaetopogonietum*, LJ: *Laurentio-Juncetum*, LJp: *Laurentio-Juncetum pinguiculetosum*).

Tab. 2: Confusion matrices concerning the predicted vegetation types. Left: variant 1 ($\kappa = 0.38$), right: variant 2 ($\kappa = 0.41$), Observations in columns, predictions in rows (---: no dwarf rush communities, CT: *Crassula tillaea*-community, JI: *Juncisoetum*, LC: *Loto-Chaetopogonietum*, LJ: *Laurentio-Juncetum*, LJp: *Laurentio-Juncetum pinguiculetosum*).

	---	CT	JI	LC	LJ	LJp
---	158	7	3	1	3	0
CT	4	0	0	0	1	1
JI	2	0	1	0	0	0
LC	3	0	0	6	0	0
LJ	11	0	0	0	4	0
LJp	1	0	0	0	1	0

	---	CT	JI	LC	LJ	LJp
---	150	3	2	2	1	1
CT	1	3	0	0	0	0
JI	11	0	1	0	1	0
LC	10	1	1	5	0	0
LJ	5	0	0	0	5	0
LJp	2	0	0	0	2	0

Darstellung der Regeln offensichtlich und nicht erst im Modellergebnis zu erkennen. Bei der Modellierung von mehreren Typen als mögliches Ergebnis wird der prognostizierte Typ direkt ausgegeben und muss nicht über einen Vergleich der Eignung für verschiedene Typen nachträglich ermittelt werden.

Zieht man den Vergleich zu statistischen Habitatmodellierungsverfahren, so vereinigen z. B. GLM (verallgemeinerte lineare Modelle) einige Vorteile der beiden Varianten. Die Ermittlung des prognostizierten Typs bleibt jedoch einer nach-

träglichen Analyse überlassen. Eine integrierte Vorhersage des Vegetationstyps wäre mit Entscheidungsbaumverfahren z. B. CART möglich. Die Übergänge zu benachbarten Typen werden erst im Modellergebnis erkennbar. Wesentlich bleibt für GLM aber der Anspruch, einen Stichprobenumfang mit $n > 80$ und mehr als 10 Vorkommen zu haben, um verlässliche Modelle zu erzeugen.

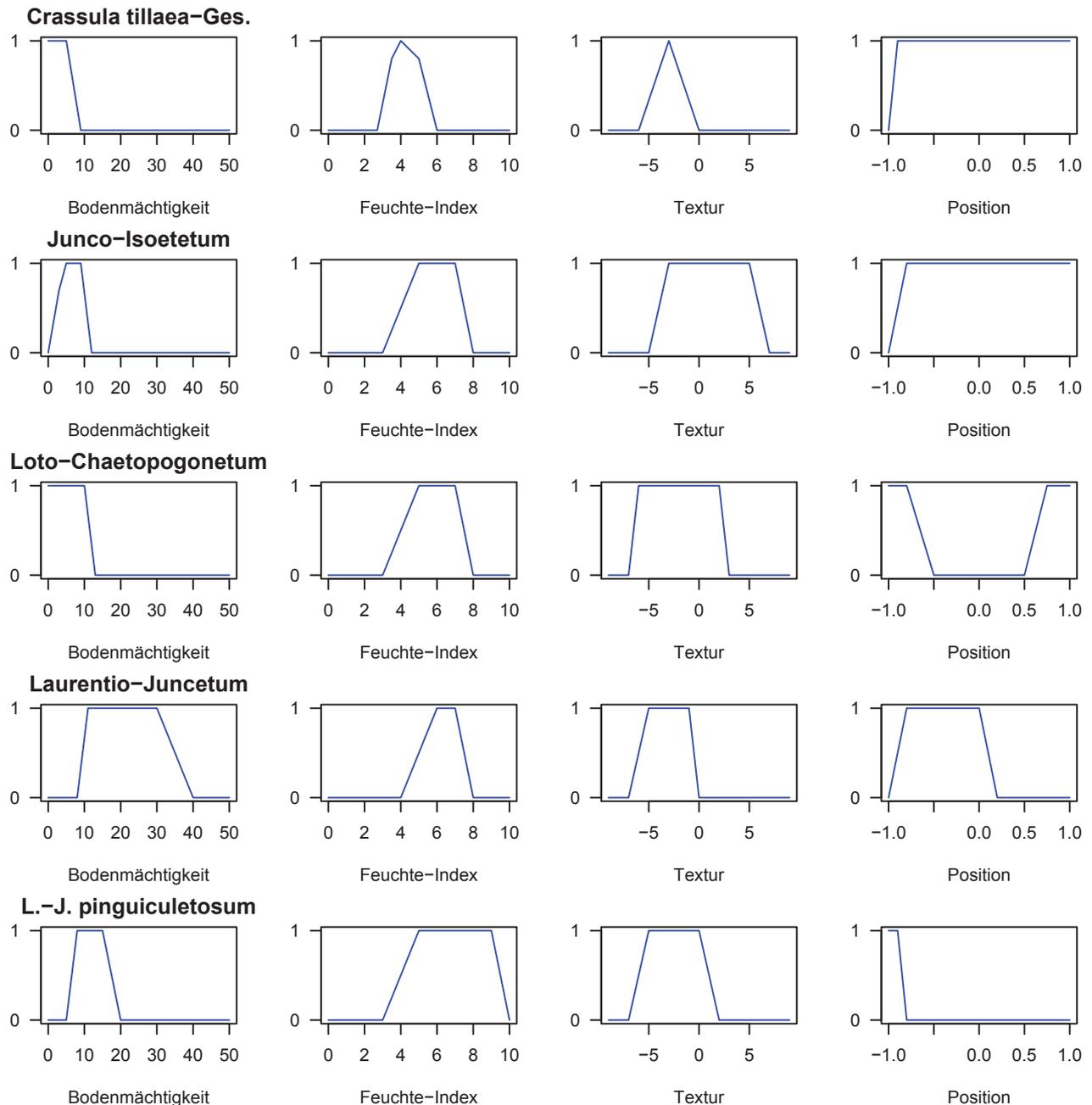


Abb. 4: Zugehörigkeitsfunktionen für Typen von Zwergbinsenrasen für die verwendeten Umweltvariablen (vgl. Abb. 2). Wesentliche Standortfaktoren: Bodenmächtigkeit [cm], ein topografischer Feuchte-Index (0: trocken, 10 stark zur Vernässung neigend), die Bodenart (Textur, 10er Logarithmus der mittleren Korngröße [cm]) und Position auf der Grusdecke (-1: unterer Rand, 0: Mitte, 1: oberer Rand).

Fig. 4: Membership functions for dwarf rush community types in dependence of the used environmental variables (see Fig. 2). Essential habitat factors: Soil depth [cm], a topographic wetness index (0: dry, 10 frequently saturated), soil texture (logarithm of the average grain size [cm]) and position on the gravel cushion (-1: lower margin, 0: center, 1: upper margin).

Tab. 3: Übersicht zum Vergleich der beiden *Fuzzy Set*-Varianten, als Referenz auf statistische Habitatmodellierungsansätze dienen verallgemeinerte lineare Modelle (GLM). (-: nicht möglich, +: zutreffend, mittlerer bis hoher Aufwand, ++: geringer Aufwand)

Tab. 3: Summary in order to compare the two fuzzy set variants, Generalized Linear Models (GLM) are included as reference for statistical habitat modelling approaches. (-: impossible, +: appropriate, middle to high effort, ++: low effort)

	Variante 1	Variante 2	GLM
<i>Modellerstellung</i>			
Abhängige Variable			
Wissensbasiert	+	+	
Präsenz/Absenz			+
Umweltvariablen			
Linguistische Variablen	+		
Metrische Variablen		+	+
Wechselwirkungen	++	-	++
Aufstellen der Regeln bzw. Parametrisieren des Modells	+	++	++
Übergang zu Nachbartypen			
Im Modell erkennbar	+		
Erst im Ergebnis ablesbar		+	+
<i>Feinjustierung</i>			
Anpassung an Datensatz	+	++	++
<i>Berechnung</i>			
Auswahl des prognostizierten Typs	+	++	++
Enthalten	+		
Post hoc		+	+
<i>Darstellung des Regelwerks</i>			
Ökogramm	+		
Responseoberfläche, bzw. -kurven		+	+

4 Diskussion

In zweierlei Hinsicht war zu prüfen, ob sich Habitatmodelle auf der Grundlage von Expertenwissen mit Hilfe eines *Fuzzy Set*-Ansatzes aufstellen lassen. Der erste Aspekt ist die Formalisierung des Expertenwissens, die dazu dienen soll, zum einen das vorhandene Wissen über die Habitatansprüche von Arten oder Zönosen unabhängig von der Verfügbarkeit des jeweiligen Experten nutzen zu können und zum anderen festzustellen, ob alle wesentlichen Standortfaktoren explizit benannt wurden. Der zweite Aspekt betrifft die Eignung derartiger Modelle für räumliche oder zeitliche Extrapolationen, also die Ermittlung geeigneter Habitate oder die Vorhersage von Vorkommen. Anschließend ist zu bewerten, in welchen Fällen ein wissensbasierter Ansatz Methoden der statistischen Habitatmodellierung vorzuziehen ist.

Die vorgestellten Ergebnisse zeigen, dass das Erarbeiten von Habitatmodellen auf der Basis von Expertenwissen möglich ist. Im vorgestellten Beispiel liegt das Gütemaß κ bei etwa 0,4, was auf mittelmäßige Modellgüte hinweist (LANDIS & KOCH 1977). Die formalisierte Darstellung der Modelle, ob in ökogrammartigen Tabellen (Abb. 3) oder in Zugehörigkeitsfunktionen (Abb. 4) macht die Habitatmodelle transparent und nachvollziehbar. Bei ähnlichen Schwächen bezüglich der Zuordnung des unterrepräsentierten *Laurentio-Juncetum pinguiculetosum* zeigt Variante 2 bessere Ergebnisse bei der Abbildung der anderen Zwergbinsenrasen-Typen. Bezüglich der richtigen Vorhersagen von Nichtvorkommen liegt dagegen Variante 1 um etwa 5 % besser. Am Beispiel der Zwergbinsenrasen ist gegenüber der einfachen Schematisierung, wie sie in RUDNER (2005) vorgenommen wurde, ein klarer Erkenntnisfortschritt zu konstatieren. Durch die klare Formalisierung des vorhandenen Wissens entsteht ein deutlich schärferes Bild, das eine bessere Verallgemeinerung erlaubt. Um jedoch die Übertragbarkeit zu beurteilen, ist die Validierung mit einem weiteren Datensatz erforderlich (BOCK & SALSKI 1998, PEREIRA et al 2009). Die Eignung der *Fuzzy Set*-Habitatmodelle für räumliche oder zeitliche Ex-

trapolation ist damit noch nicht nachgewiesen. Es ist davon auszugehen, dass sich die Güte bei der Übertragung verschlechtert, da der Anpassungsgüte an den Trainingsdatensatz immer ein gewisser Optimismus zu Grunde liegt, auch wenn bei der Justierung darauf geachtet wurde, keine Überanpassung vorzunehmen (GUISAN & ZIMMERMANN 2000).

Wird bei den statistischen Habitatmodellierungsansätzen immer angenommen, dass sich Habitat und Vorkommen im Gleichgewicht befinden (GUISAN & ZIMMERMANN 2000), so besteht bei den *Fuzzy Set*-Methoden die Möglichkeit auch dynamische Variablen einzubeziehen (SVORAY et al. 2004). Es sollte aber bedacht werden, ob die Gleichgewichtsannahme implizit im Modell steckt, sofern sich das Wissen im Wesentlichen auf beobachteten Vorkommen und Nichtvorkommen der Zielarten oder -gemeinschaften gründet (GUISAN & ZIMMERMANN 2000).

Wann ist nun der *Fuzzy Set*-Ansatz besser geeignet als statistische Habitatmodellierungsansätze? Die tabellarische Gegenüberstellung von den *Fuzzy Set*-Varianten und dem GLM-Ansatz als Vertreter der statistischen Habitatmodellierungsansätze zeigt, dass der GLM-Ansatz die Vorteile der beiden geprüften *Fuzzy Set*-Varianten vereint (Tab. 3). Bei entsprechenden Voraussetzungen sind geeignete statistische Habitatmodellierungsansätze klar vorzuziehen. Werden die grundsätzlichen Ansprüche insbesondere bezüglich einer Mindestanzahl von Vorkommen, die wiederum von der Ökologie der Arten abhängt (ELITH et al. 2006), an den Trainingsdatensatz nicht erfüllt, liegen die Vorteile gegenüber einer statistischen Habitatmodellierung auf der Hand. Als Grundlage für das Aufstellen eines Modells kann direkt Expertenwissen verwendet werden ohne einen explizit beobachteten Datensatz. Für die Justierung des Modells ist es dann notwendig, einen – möglicherweise auch kleinen – Datensatz heranzuziehen. Insbesondere ist auch daran zu denken, dass die Definitionen der linguistischen Variablen möglicherweise anzupassen sind.

Die Eignung der Variante der *Fuzzy Set*-Modellierung hängt sehr stark von der Fragestellung ab. Dreht es sich um die Beschreibung der Nische einer einzelnen Art bzw. eines einzelnen Vegetationstyps, ist die Variante mit den Zugehörigkeitsfunktionen vorzuziehen. Gleiches gilt, wenn beabsichtigt ist, das Modell zur Extrapolation zu nutzen (vgl. SVORAY et al. 2008). Wird beabsichtigt, die Einnischung standörtlich benachbarter Einheiten zu beschreiben, oder liegen keine Messwerte zu den Umweltvariablen vor, eignet sich die Variante 1 gut. Diese Variante bringt aufgrund der Vielzahl der notwendigen Regeln den Nachteil von vergleichsweise langen Rechenzeiten mit sich.

Die Aufstellung der Vegetationsreihe funktioniert gut bei einem starken, vorherrschenden Gradienten. Sind die Abhängigkeiten von den Umweltvariablen mehrdimensional gestaltet, entstehen Probleme mit der Reihung. So kann es bei nicht konvexen Ergebniskurven Schwierigkeiten bei der Defuzzifizierung geben.

Die Verknüpfung von mehreren Umweltvariablen zu einer Index-Variablen im Vorfeld des eigentlichen Modells erscheint sinnvoll, um eine übersichtliche Darstellung des Regelwerks erreichen zu können (vgl. Abb. 3, 4). Diese Aufteilung der Berechnung in zwei Schritte verkürzt auch die Rechenzeit erheblich, da deutlich weniger Regeln abgearbeitet werden müssen (RÜGER et al. 2005, SVORAY et al. 2004).

Der *Fuzzy Set*-Ansatz eignet sich in beiden untersuchten Varianten zur formalisierten Beschreibung der Realnische seltener Arten oder Vegetationstypen. Man sollte versuchen, für seltene Taxa derartige Modelle in Landschaftsmodelle (z. B. SCHRÖDER et al. 2008) zu integrieren. Gerade wenn Landschaftsmodelle im Naturschutz zur Entscheidungsunterstützung herangezogen werden und z. B. szenarienabhängig die Entwicklung geeigneter Habitate prognostiziert wird, ist es wesentlich auch für seltene Arten Modelle zur Verfügung zu haben, die auf GLM oder BRT-Basis nicht erarbeitet werden können (BUCERIUS et al. 2006, HEATON et al. 2008).

Abhängig von der Zielsetzung der Modellierung ist zu beachten, dass falsche positive und falsche negative Vorhersagen unterschiedlich zu werten sind. Entsprechend wären weiter entwickelte Gütemaße heranzuziehen (MOUTON et al. 2009).

Dank

Das Werkzeug *RockOn Fuzzy* (LIND et al. 2008) hat den Zugang zur *Fuzzy Set*-Modellierung geebnet. Ich bedanke mich bei den Entwicklern für die freundliche Berücksichtigung meiner Wünsche bei der Weiterentwicklung des Werkzeugs.

Die Kommentare zweier anonymer Gutachter haben wesentlich zur Verbesserung des Manuskripts beigetragen.

Literatur

AUSTIN, M.P., BELBIN, L., MEYERS, J.A., DOHERTY, M.D., LUOTO, M. (2006): Evaluation of statistical models used for predicting plant species distributions: Role of artificial data and theory. *Ecological Modelling* **199**: 197-216.

BEVEN, K.J. (1986): Runoff production and flood frequency in catchments of order n : an alternative approach. In: GUPTA, V.K., RODRIGUEZ-ITURBE, I., WOOD, E.F. (Eds.): *Scale problems in hydrology*. Dordrecht: 107-131.

BOCK, W., SALSKI, A. (1998): A fuzzy knowledge based model of population dynamics of the Yellow-necked mouse (*Apodemus flavicollis*) in a beech forest. *Ecological Modelling* **108**: 155-161.

BORTZ, J., LIENERT, G.A., BOEHNEKE, K. (2000): *Verteilungsfreie Methoden in der Biostatistik*. Springer, Berlin, Heidelberg: 939 S.

BUCERIUS, M., FUSS, R., STEINHOFF, J., HEIDLER, M., KRÄMER, P., HAUBENSACK, C. et al. (2006): Demonstrationsmodul für ein Landschaftsmodell im Landkreis Haßberge in Unterfranken – Methodenentwicklung und -erprobung zur Optimierung des Pflegemanagements. *Naturschutz und Landschaftsplanung* **38**: 276-281.

COUDUN, C., GÉGOUT, J.-C. (2006): The derivation of species response curves with Gaussian logistic regression is sensitive to sampling intensity and curve characteristics. *Ecological Modelling* **199**: 164-175.

ELITH, J., GRAHAM, C.H., ANDERSON, R.P., DUDIK, M., FERRIER, S., GUISAN, A. et al. (2006): Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**: 129-151.

ELITH, J., GRAHAM, C.H. (2009): Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography* **32**: 66-77.

ENGLER, R., GUISAN, A., RECHSTEINER, L. (2004): An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology* **41**: 263-274.

GUISAN, A., ZIMMERMANN, N.E. (2000): Predictive habitat distribution models in ecology. *Ecological Modelling* **135**: 147-186.

HARREL, F.E. (2001): *Regression modeling strategies. With applications to linear models, logistic regression, and survival analysis*. Springer, New York: 568 S.

HEATON, J.S., NUSSEAR, K.E., ESQUE, T.C., INMAN, R.D., DAVENPORT, F.M., LEUTERITZ T.E., MEDICA, P.A. et al. (2008): Spatially explicit decision support for selecting translocation areas for Mojave desert tortoises. *Biodiversity and Conservation* **17**: 575-590.

HIRZEL, A.H., HAUSSER, J., CHESSEL, D., PERRIN, N. (2002): Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology* **83**: 2027-2036.

JOSS, B.N., HALL, R.J., SIDDEERS, D.M., KEDDY, T.J. (2008): Fuzzy-logic modeling of land suitability for hybrid poplar across the Prairie Provinces of Canada. *Environmental Monitoring & Assessment* **141**: 79-96.

KRUSE, R., GEBHARDT, J., KLAUWONN, F. (1995): *Fuzzy-Systeme*. 2. Aufl. B.G. Teubner, Stuttgart: 273 S.

KIRKBY, M.J., WEYMAN, D.R. (1974): Measurement of contributing area in very small drainage basins. Seminar Series B3. Dept. of Geog., Univ. of Bristol.

LANDIS, J.R., KOCH, G.G. (1977): The Measurement of Observer Agreement for Categorical Data. *Biometrics* **33** (1): 159-174.

LEHMANN, I., WEBER, R., ZIMMERMANN, H.-J. (1992): *Fuzzy Set Theory – Die Theorie unscharfer Mengen*. OR Spektrum **14**: 1-9.

LIND, S., BACHMANN, F., STÜTZER, S. (2008): *Rock On Fuzzy. The best free fuzzy control and simulation tool*. URL: <http://www.timtomtam.de/rockonfuzzy/>.

MEYER, D., ZEILEIS, A., HORNIK, K. (2009): *vcd: Visualizing Categorical Data*. R package version 1.2-4. URL: <http://cran.r-project.org/>.

- MOUTON, A.M., DE BAETS, B., VAN BROEKHOVEN, E., GOETHALS, P.L.M. (2009): Prevalence-adjusted optimisation of fuzzy models for species distribution. *Ecological Modelling* **220**: 1776-1786.
- PEPPLER-LISBACH, C. (2003): Predictive modelling of recent and historical land-use patterns. *Phytocoenologia* **33** (4): 565-590.
- PEREIRA, G., EVSUKOFF, A., EBECKEN, N.F.F. (2009): Fuzzy modelling of chlorophyll production in a Brazilian upwelling system. *Ecological Modelling* **220**: 1506-1512.
- R DEVELOPMENT CORE TEAM (2008): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org>.
- RUDNER, M. (2004): Zwergbinsengesellschaften im Südwesten der Iberischen Halbinsel – eine Analyse räumlicher und zeitlicher Vegetationsmuster. Dissertation Universität Freiburg, Fakultät für Biologie. URL: <http://freidok.ub.uni-freiburg.de/volltexte/1413/> [PDF].
- RUDNER, M. (2005): Environmental patterns and plant communities of the ephemeral wetland vegetation in two areas of the Southwestern Iberian Peninsula. *Phytocoenologia* **35** (2/3): 231-265.
- RÜGER, N., SCHLÜTER, M., MATTHIES, M. (2005): A fuzzy habitat suitability index for *Populus euphratica* in the Northern Amudarya delta (Uzbekistan). *Ecological Modelling* **184**: 313-328.
- SCHRÖDER, B., RUDNER, M., BIEDERMANN, R., KÖGL, H., KLEYER, M. (2008): A landscape model for quantifying the trade-off between conservation needs and economic constraints in the management of a semi-natural grassland community. *Biological Conservation* **141**: 719-732.
- STOCKWELL, D.R.B., PETERSON, A.T. (2002): Effects of sample size on accuracy of species distribution models. *Ecological Modelling* **148**: 1-13.
- SVORAY, T., BAR-YAMIN, S., HENKIN, Z., GUTMAN, M. (2004): Assessment of herbaceous plant habitats in water constrained environments: predicting indirect effects with fuzzy logic. *Ecological Modelling* **180**: 537-556.
- SVORAY, T., SHAFRAN-NATHAN, R., HENKIN, Z., PEREVOLOTSKY, A. (2008): Spatially and temporally explicit modelling of conditions for primary production of annuals in dry environments. *Ecological Modelling* **218**: 339-353.
- ZADEH, L. (1965): Fuzzy Sets. *Information and Control* **8**: 338-353.

submitted: 01.07.2009

reviewed: 03.11.2009

accepted: 30.11.2009

Autorenanschrift:

Dr. Michael Rudner
 Albert-Ludwigs-Universität Freiburg
 Institut für Biologie II, Abteilung Geobotanik
 Schänzlestraße 1
 D-79104 Freiburg i. Br.

Tel.: +49 761 203-2822

Fax: +49 761 203-2696

Email: michael.rudner@biologie.uni-freiburg.de