



ÖKOLOGIE-WORKSHOP

Artenakkumulationskurven

Wie zählt man Arten, die man nicht sieht?

Gilles Colling

Division of BioInvasions, Global Change & Macroecology
Universität Wien

26.–27. Jänner 2026

INHALT

1 Tag 1 Theorie 3

Biodiversität & Artenakkumulationskurven

2 Tag 1 Übungen 12

R-Grundlagen bis zur ersten Datenanalyse

3 Tag 2 Theorie 25

Algorithmen & räumliche Ordnung

4 Tag 2 Übungen 38

Ein Mini-Forschungsprojekt

5 Nach dem Workshop Ergebnisse 56

Analyse-Ergebnisse

6 Nach dem Workshop Zusammenfassung 94

Workshop-Rückblick

Tag 1

Theorie

Biodiversität & Artenakkumulationskurven

Tag 1: Theorie

Biodiversität & Artenakkumulationskurven

Ökologie-Workshop

2026-01-25

Was ist Biodiversität?

Definition

Biodiversität = die Vielfalt des Lebens in einem Gebiet

Denk an zwei Wälder...

Eiche, Eiche, Eiche, Eiche, Eiche Eiche, Buche, Ahorn, Kiefer, Birke

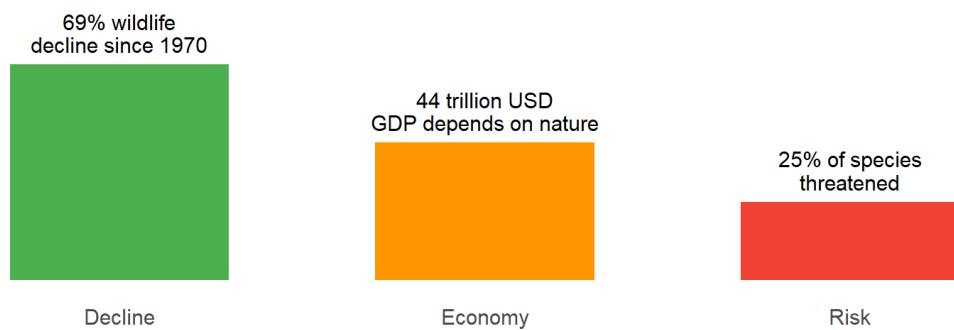
5 Bäume, **1 Art**

5 Bäume, **5 Arten**

Welcher Wald ist artenreicher? Wald B - obwohl beide gleich viele Bäume haben.

Warum ist Biodiversität wichtig?

Why Biodiversity Matters



- **Ökosystemleistungen:** Saubere Luft, Wasserfilterung, Bestäubung
- **Medizin:** 70% der Krebsmedikamente stammen aus natürlichen Quellen
- **Ernährungssicherheit:** Genetische Vielfalt schützt Pflanzen vor Krankheiten

Das Stichproben-Problem

Das Zählproblem

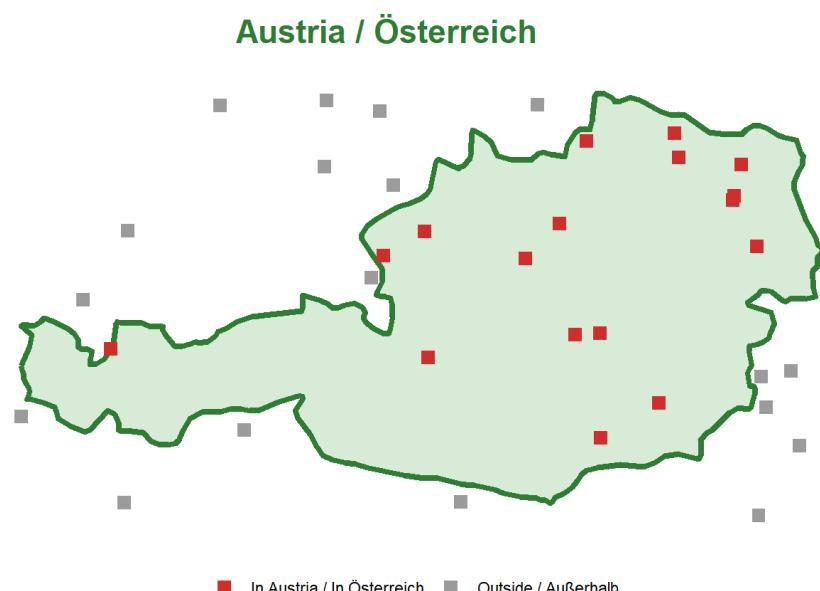
Wie viele Pflanzenarten gibt es in Österreich?

Können wir alle zählen? Bedenke:

- Österreich = 83.879 km²
- Manche Arten sind winzig (Moose, Flechten)
- Manche erscheinen nur zu bestimmten Jahreszeiten
- Manche sind extrem selten

Die Lösung: Stichproben!

Wir untersuchen kleine Flächen sorgfältig und nutzen Statistik zur Schätzung.



Kernidee: Wir untersuchen kleine Flächen sorgfältig und nutzen Muster, um das Ganze zu verstehen.

Diskussionsfrage

Stell dir vor, du willst wissen, wie viele verschiedene Süßigkeiten es in einem Süßwarenladen gibt.

Du hast **5 Minuten** und kannst in **10 Gläser** schauen (von insgesamt 100).

Welche Strategie ist besser?

- A. Die 10 Gläser direkt neben der Tür
- B. 10 Gläser verteilt im ganzen Laden

Warum ist das für Ökologie wichtig?

Vegetationsplots

Ein **Vegetationsplot** ist eine kleine Fläche (typisch 1-100 m²), auf der Ökologen jede Pflanzenart notieren.

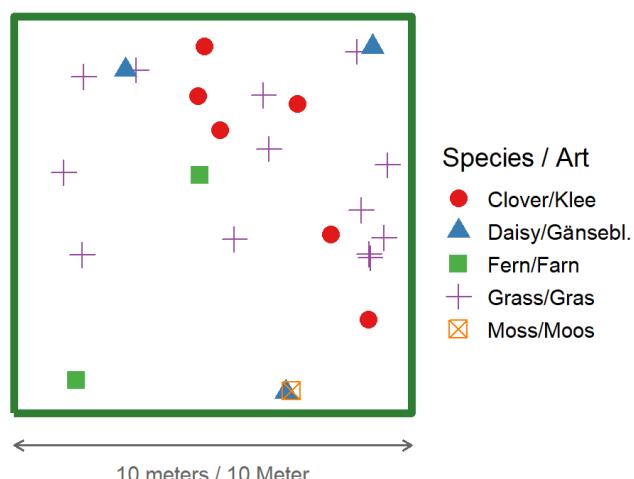
Standort (Koordinaten) 47.0707° N, 15.4395° E

Datum 2023-06-15

Alle vorhandenen Arten Festuca rubra, Trifolium repens, ...

Deckung (%) pro Art 25%, 10%, ...

10m × 10m Vegetation Plot



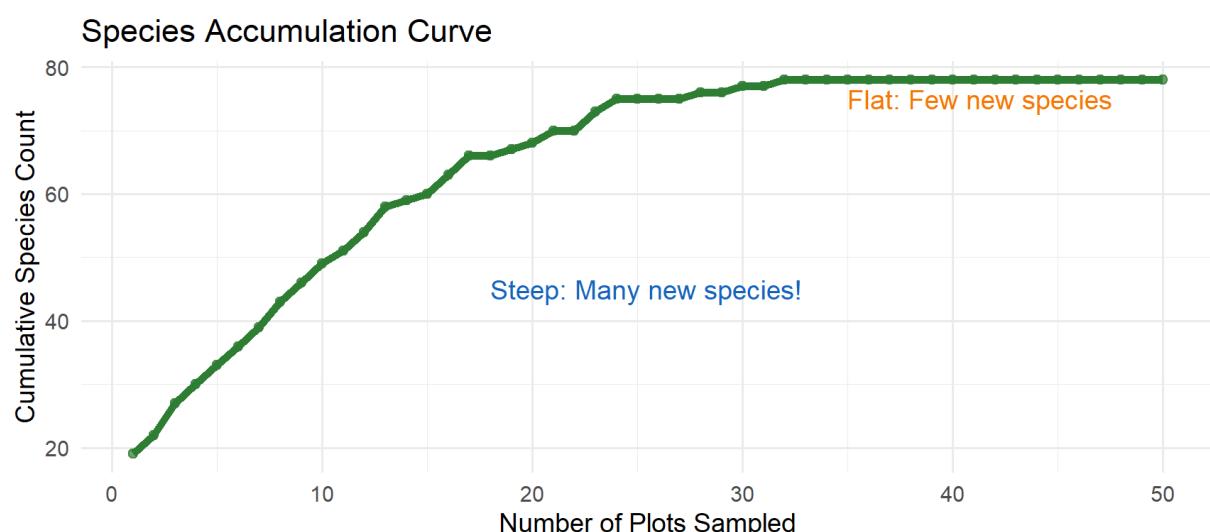
Artenakkumulationskurven

Die Kernidee

Je mehr Plots wir untersuchen, desto mehr Arten finden wir - aber mit **abnehmender Rate**.

Warum?

- Erste Plots: Hauptsächlich **häufige** Arten (leicht zu finden)
- Spätere Plots: Hauptsächlich **seltene** Arten (schwer zu finden)
- Schließlich: Sehr wenige neue Arten pro Plot



Warum flacht die Kurve ab?

Denk an ein **Kartenspiel**:

1. Karte Definitiv neu!
10. Karte Wahrscheinlich noch neu
30. Karte Vielleicht schon gesehen
50. Karte Sehr wahrscheinlich ein Duplikat

Je mehr Arten du schon gefunden hast, desto schwieriger wird es, neue zu finden.

Lasst uns eine Kurve bauen

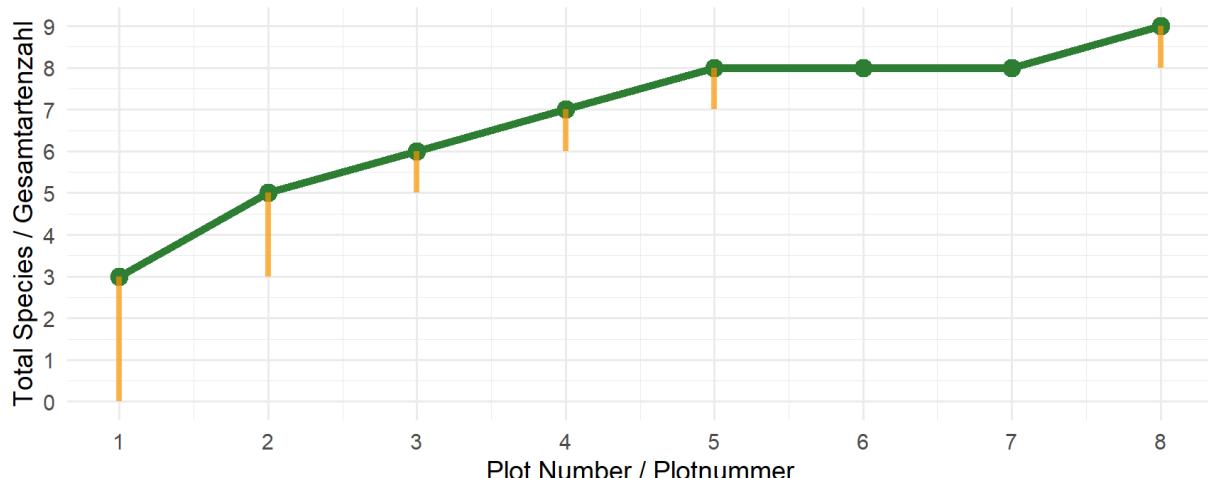
Stell dir vor, wir untersuchen eine Wiese. Jeder "Plot" zeigt einige Arten:

1	Gänseblümchen, Klee, Gras	3	3
2	Butterblume, Löwenzahn, Gras	2	5
3	Gänseblümchen, Distel, Klee	1	6
4	Gras, Spitzwegerich, Gänseblümchen	1	7
5	Klee, Butterblume, Schafgarbe	1	8
6	Gänseblümchen, Gras, Klee	0	8
7	Löwenzahn, Spitzwegerich, Gras	0	8
8	Klee, Gänseblümchen, Sauerampfer	1	9

Beachte: Bei Plot 6-7 finden wir nichts Neues mehr!

Unsere Beispielkurve

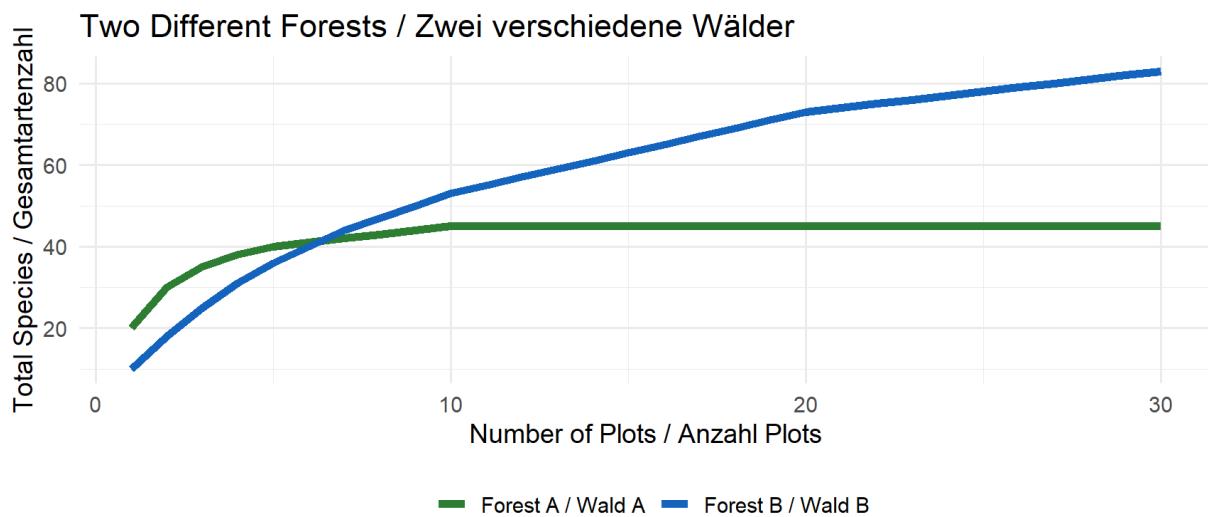
Meadow Accumulation Curve / Wiesen-Akkumulationskurve



Die **orangenen Segmente** zeigen, wie viele neue Arten jeder Plot hinzufügt. Beachte, wie sie kürzer werden!

Diskussionsfrage

Schau dir diese zwei Kurven an:



Fragen:

1. Welcher Wald hat wahrscheinlich mehr Arten insgesamt?
2. In welchem Wald sind die Arten gleichmäßiger verteilt?
3. In welchem Wald müsstest du MEHR Plots untersuchen, um alles zu finden?

Die Asymptote

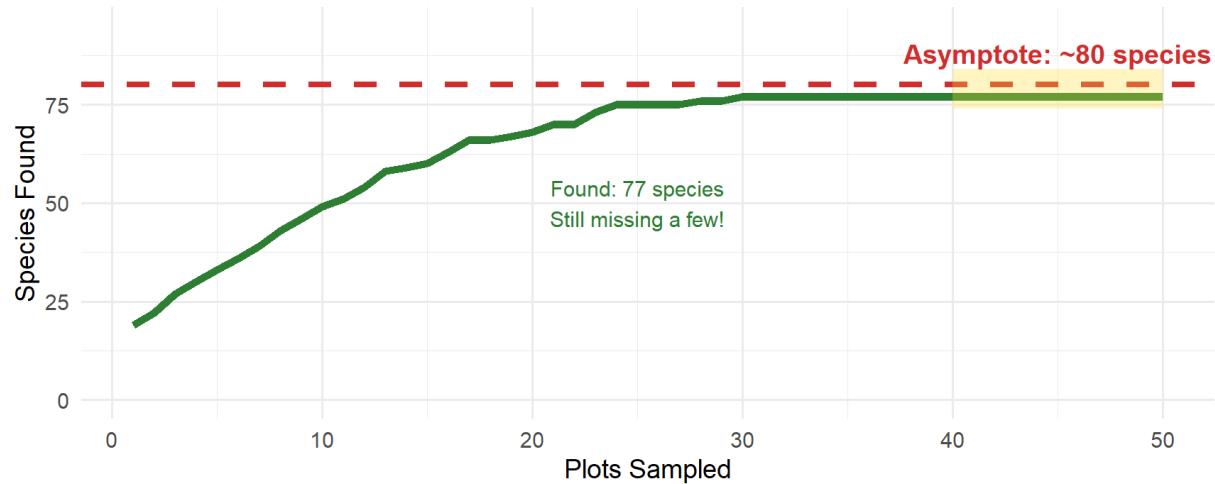
Das Unbekannte vorhersagen

Die **Asymptote** ist die Linie, der sich die Kurve nähert, aber nie ganz erreicht.

Sie repräsentiert: **die geschätzte Gesamtartenzahl**

Selbst ohne alles zu untersuchen, können wir schätzen, wie viel wir verpassen!

The Asymptote: Estimating Total Species



Das Europäische Vegetationsarchiv

EVA-Datenbank

EVA (European Vegetation Archive) ist die größte Vegetationsdatenbank Europas:

- **1,6+ Millionen** Plots
- **82** Länder
- Daten von **1900 bis heute**

Für Österreich haben wir Zugang zu ~52.000 Plots mit Standortdaten.

Heimische vs. Alien-Arten

Unser Datensatz unterscheidet zwischen:

Natürlich vorkommend	Vom Menschen eingeführt
Teil des ursprünglichen Ökosystems	Nach 1492 angekommen
Mit anderen Arten co-evolviert	Können Einheimische verdrängen

Forschungsfrage für Tag 2: Akkumulieren Alien-Arten anders als heimische Arten?

Warum nicht Excel?

Excel ist super für...

Du kennst wahrscheinlich Excel gut. Es ist perfekt für:

- Schnelle Berechnungen

- Einfache Tabellen und Diagramme
- Kleine Datensätze (Hunderte von Zeilen)

Könnten wir Akkumulationskurven in Excel bauen? Ja! Für ein kleines Beispiel:

Excel: Manual formulas for each row

A Plot	B Species	C New	D Total
1	Rose, Daisy...	3	=C2
2	Clover, Grass...	2	=D2+C3
3	Rose, Clover...	1	=D3+C4
...

Das funktioniert für 5 Plots. Aber was ist mit **52.000 Plots**?

Aber echte Forschung braucht mehr

Unser österreichischer Datensatz hat **52.000 Plots** und **850.000 Artenbeobachtungen**.

850.000 Zeilen laden	Stürzt ab oder friert ein	2 Sekunden
Analyse 100x wiederholen	100x klicken	Eine Schleife
Exakte Methode teilen	"Ich habe hier geklickt, dann dort..."	Code teilen
Fehler finden	Wo habe ich falsch geklickt?	Zeile 47 lesen
Mit neuen Daten aktualisieren	Von vorne anfangen	Skript erneut ausführen

Reproduzierbarkeit: Der Schlüsselunterschied

In der Wissenschaft müssen andere deine Ergebnisse **reproduzieren** können.

Mit Excel: > "Ich habe Spalte B sortiert, dann nach Werten > 10 gefiltert, dann ein Diagramm erstellt, dann... ich glaube, ich habe auch einige Ausreißer entfernt?"

Mit R:

```
data %>%
  filter(value > 10) %>%
  ggplot(aes(x, y)) +
  geom_line()
```

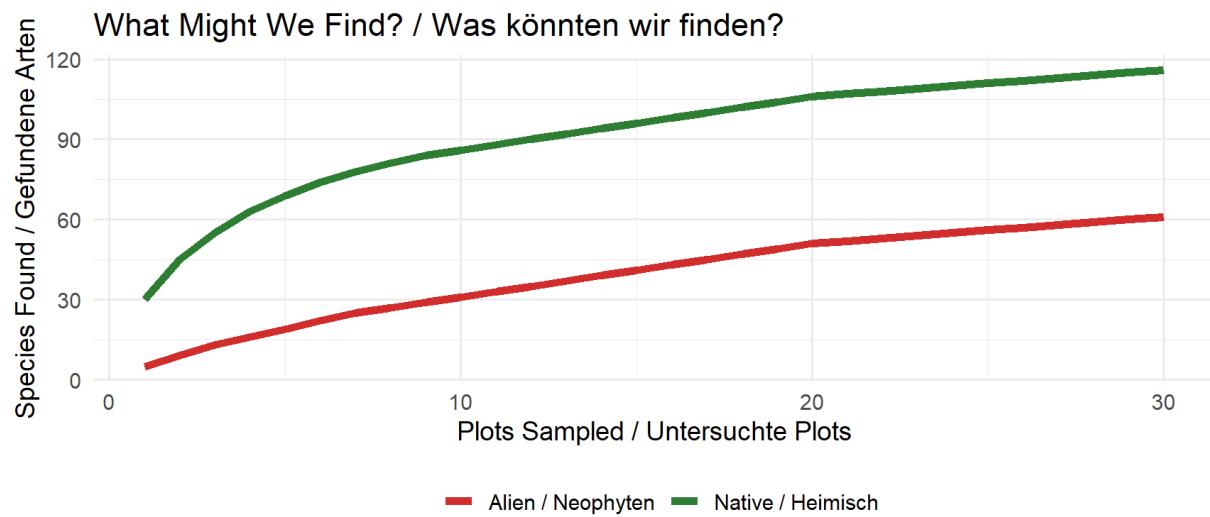
Der Code IST die Methode. Jeder kann ihn ausführen und das gleiche Ergebnis erhalten.

Zusammenfassung

Schlüsselkonzepte

Biodiversität	Artenvielfalt in einem Gebiet
Vegetationsplot	Kleine Fläche, auf der alle Pflanzen erfasst werden
Akkumulationskurve	Graph der Artenentdeckung über Stichprobenaufwand
Asymptote	Geschätzte Gesamtartenzahl (Kurve nähert sich, erreicht sie aber nie)
Heimische Arten	Natürlich in der Region vorkommend
Alien-Arten	Vom Menschen eingeführt (nach 1492)

Vorschau: Tag 2



Morgen werden wir:

1. Den **Nearest-Neighbor-Algorithmus** für räumliche Kurven lernen
2. Verstehen, warum die **Plot-Reihenfolge wichtig** ist
3. Kurven mit **mehreren Startpunkten** bauen
4. **Heimische vs. Alien**-Akkumulationsmuster vergleichen
5. Ergebnisse interpretieren und **biologische Mechanismen** diskutieren

Zu testende Hypothese: Akkumulieren Alien-Arten anders als heimische Arten?

Tag 1

Übungen

R-Grundlagen bis zur ersten Datenanalyse

Teil 1: R Grundlagen - Los geht's

Übung 1: Deine ersten Variablen

Übung 2: Vektoren - Listen von Werten

Übung 3: Die Magie von unique()

Übung 4: Data Frames

Übung 5: Echte Daten laden

Übung 6: Daten filtern

Übung 7: Die Akkumulationskurve bauen

Übung 8: Zeichne deine Kurve!

Bonus: Vergleiche Heimisch vs Alien!

Tag 1: Übungen

Deine erste Akkumulationskurve!

Fülle die Lücken aus!

2026-01-26

Teil 1: R Grundlagen - Los geht's

Willkommen bei R! In diesem Abschnitt lernst du die grundlegenden Bausteine der R-Programmierung. Nimm dir Zeit für jede Übung - wenn du diese Grundlagen gut verstehst, wird alles andere einfacher.

So verwendest du diese Übungen:

- Ersetze jedes _____ mit dem richtigen Code
- Führe jeden Chunk aus um zu sehen ob er funktioniert
- Lies die Hinweise wenn du nicht weiterkommst

Übung 1: Deine ersten Variablen

Konzept: Variablen sind Behälter, die Werte speichern. In R benutzen wir `<-` um Werte zuzuweisen.

1a: Einfache Variablen erstellen

```
# Store the number 42 in a variable called "answer"
# Speichere die Zahl 42 in einer Variable namens "answer"
answer <- ____
print(answer)

# Store your age in a variable
# Speichere dein Alter in einer Variable
my_age <- ____
print(my_age)

# Store a decimal number (height in meters)
# Speichere eine Dezimalzahl (Größe in Metern)
my_height <- ____
print(my_height)
```

1b: Text-Variablen (Strings)

```
# Text variables - use quotes!
# Text-Variablen - benutze Anführungszeichen!

# Store your name (text needs quotes)
# Speichere deinen Namen (Text braucht Anführungszeichen)
my_name <- "____"
print(my_name)

# Store a species name
# Speichere einen Artnamen
species <- "____"
print(species)
```

1c: Logische (TRUE/FALSE) Variablen

```
# Logical variables - TRUE or FALSE
# Logische Variablen - TRUE oder FALSE

# Is it raining? (answer TRUE or FALSE, no quotes!)
# Regnet es? (antworte TRUE oder FALSE, keine Anführungszeichen!)
is_raining <- ____

# Is 5 greater than 3? (R can calculate this!)
# Ist 5 größer als 3? (R kann das berechnen!)
five_greater_three <- 5 > 3
print(five_greater_three)

# Is 10 equal to 10? (use == for comparison)
# Ist 10 gleich 10? (benutze == für Vergleich)
ten_equals_ten <- 10 == 10
print(ten_equals_ten)

# Is "oak" equal to "Oak"? (R is case-sensitive!)
# Ist "oak" gleich "Oak"? (R unterscheidet Groß/Kleinschreibung!)
oak_equals_Oak <- "oak" == "Oak"
print(oak_equals_Oak)
```

1d: Grundrechenarten mit Variablen

```
# Create two number variables
# Erstelle zwei Zahlen-Variablen
a <- 10
b <- 3

# Addition
sum_ab <- a + b
print(sum_ab) # Should be 13 / Sollte 13 sein

# Subtraction / Subtraktion
diff_ab <- a - b
print(diff_ab) # Should be 7 / Sollte 7 sein

# Multiplication / Multiplikation
prod_ab <- a * b
print(prod_ab) # Should be 30 / Sollte 30 sein

# Division
div_ab <- a / b
print(div_ab) # Should be 3.333... / Sollte 3.333... sein

# Power (10 to the power of 3) / Potenz (10 hoch 3)
power_ab <- a ^ b
print(power_ab) # Should be 1000 / Sollte 1000 sein
```

1e: Variablen aktualisieren

```
# Variables can change!
# Variablen können sich ändern!

# Start with a count of species
# Starte mit einer Artenanzahl
species_count <- 0
print(species_count)

# Found 5 species / 5 Arten gefunden
species_count <- species_count + __
print(species_count) # Should be 5 / Sollte 5 sein

# Found 3 more / 3 weitere gefunden
species_count <- species_count + __
print(species_count) # Should be 8 / Sollte 8 sein

# Oops, 2 were misidentified (subtract)
# Ups, 2 waren falsch bestimmt (subtrahieren)
species_count <- species_count - __
print(species_count) # Should be 6 / Sollte 6 sein
```



Hinweis: Rechenoperatoren sind `+` `-` `*` `/` und `^` für Potenz. Benutze `==` für "gleich" Vergleich!

Übung 2: Vektoren - Listen von Werten

Konzept: Ein **Vektor** ist eine Liste von Werten des gleichen Typs. Wir benutzen `c()` um Werte zu kombinieren.

2a: Vektoren erstellen

```
# Create a vector of numbers
# Erstelle einen Vektor von Zahlen
species_counts <- c(5, 3, 7, 2, 4)
print(species_counts)

# Your turn - create a vector of 4 tree heights
# Du bist dran - erstelle einen Vektor mit 4 Baumhöhen
tree_heights <- c(__, __, __, __)
print(tree_heights)

# Create a vector of species names
# Erstelle einen Vektor von Artnamen
tree_species <- c("Oak", "Beech", "Pine", "Maple")
print(tree_species)
```

2b: Vektor-Funktionen

```
# How many elements in the vector?  
# Wie viele Elemente im Vektor?  
length(species_counts)  
  
# Calculate the sum  
# Berechne die Summe  
sum(species_counts)  
  
# Calculate the mean (average)  
# Berechne den Mittelwert (Durchschnitt)  
mean(species_counts)  
  
# Your turn - calculate mean of your tree heights  
# Du bist dran - berechne den Mittelwert deiner Baumhöhen  
mean(____)  
  
# Find min and max  
# Finde Minimum und Maximum  
min(species_counts)  
max(species_counts)
```

2c: Vektoren kombinieren

```
# Species from different plots  
# Arten von verschiedenen Plots  
plot1 <- c("Oak", "Beech", "Pine")  
plot2 <- c("Beech", "Maple", "Oak")  
plot3 <- c("Birch", "Oak", "Ash")  
  
# Combine all into one vector  
# Kombiniere alle in einen Vektor  
all_species <- c(plot1, plot2, plot3)  
print(all_species)  
  
# How many total observations?  
# Wie viele Beobachtungen insgesamt?  
length(all_species)
```



Hinweis: `c()` kombiniert Werte. `length()` zählt Elemente. `mean()` berechnet Durchschnitt.

Übung 3: Die Magie von `unique()`

Konzept: `unique()` ist die **Schlüsselfunktion** für Ökologie! Sie sagt uns, wie viele *verschiedene* Arten wir gefunden haben (ohne Duplikate).

```

# 3a: We observed these species (some repeats!)
# 3a: Wir haben diese Arten beobachtet (einige Wiederholungen!)

observations <- c("Oak", "Beech", "Oak", "Pine", "Beech", "Oak", "Maple")

# How many observations total?
# Wie viele Beobachtungen insgesamt?
length(observations)

# 3b: How many UNIQUE species?
# 3b: Wie viele EINZIGARTIGE Arten?
unique(observations)
length(unique(observations))

# 3c: Using our plots from before
# 3c: Benutze unsere Plots von vorher
plot1 <- c("Oak", "Beech", "Pine")
plot2 <- c("Beech", "Maple", "Oak")
plot3 <- c("Birch", "Oak", "Ash")

all_species <- c(plot1, plot2, plot3)

# Total observations vs unique species
# Gesamte Beobachtungen vs einzigartige Arten
length(all_species)      # Total / Gesamt
length(unique(all_species)) # Unique / Einzigartig

# 3d: Your turn - how many unique species after plot1 + plot2?
# 3d: Du bist dran - wie viele einzigartige Arten nach plot1 + plot2?
after_two_plots <- c(plot1, ____)
length(unique(after_two_plots))

```



Hinweis: Artenreichtum = `length(unique(artenliste))`. Das ist DIE Schlüsselformel!

Übung 4: Data Frames

Ein **Data Frame** ist wie eine Tabelle - er hat Zeilen und Spalten. Üben wir mit einem kleinen Beispiel, bevor wir echte Daten laden!

```

# 4a: Create a simple data frame
# 4a: Erstelle einen einfachen Data Frame

my_data <- data.frame(
  species = c("Oak", "Beech", "Pine", "Maple"),
  height = c(20, 18, 25, 15),
  native = c(TRUE, TRUE, TRUE, FALSE)
)

my_data

# 4b: How many rows and columns?
# 4b: Wie viele Zeilen und Spalten?

nrow(my_data)
ncol(my_data)

# 4c: Access a column with $
# 4c: Greife auf eine Spalte mit $ zu

my_data$species
my_data$height

# 4d: Your turn - get the mean height
# 4d: Du bist dran - berechne die mittlere Höhe

mean(my_data$__)

# 4e: Access a specific row (row 2)
# 4e: Greife auf eine bestimmte Zeile zu (Zeile 2)

my_data[2, ]

# 4f: Access a specific cell (row 2, column "height")
# 4f: Greife auf eine bestimmte Zelle zu (Zeile 2, Spalte "height")

my_data[2, "height"]

```



Hinweis: `$` greift auf Spalten per Name zu. Eckige Klammern `[zeile, spalte]` greifen auf bestimmte Zellen zu.

Übung 5: Echte Daten laden

Jetzt laden wir die **österreichischen Vegetationsdaten!**

```

# 5a: Load the tidyverse package
# 5a: Lade das tidyverse Paket

library(tidyverse)

# 5b: Load the species data
# 5b: Lade die Artdaten

species_data <- read_csv("../data/austria_species.csv")

# 5c: Look at the first few rows
# 5c: Schau dir die ersten Zeilen an

head(species_data)

# 5d: How many rows (observations)?
# 5d: Wie viele Zeilen (Beobachtungen)?

nrow(species_data)

# 5e: What columns do we have?
# 5e: Welche Spalten haben wir?

colnames(species_data)

# 5f: How many unique plots?
# 5f: Wie viele einzigartige Plots?

length(unique(species_data$PlotObservationID))

# 5g: How many unique species in the whole dataset?
# 5g: Wie viele einzigartige Arten im gesamten Datensatz?

length(unique(species_data$__))

```



Hinweis: Benutze `$` um auf eine Spalte zuzugreifen, z.B. `data$spaltenname`

Übung 6: Daten filtern

Oft wollen wir nur einen **Teil** unserer Daten betrachten. Die `filter()` Funktion hilft uns, bestimmte Zeilen auszuwählen!

```

# 6a: Look at the STATUS column - what values are there?
# 6a: Schau dir die STATUS-Spalte an - welche Werte gibt es?

unique(species_data$STATUS)

# 6b: Filter for native species only (STATUS == "nat")
# 6b: Filtere nur heimische Arten (STATUS == "nat")

native_only <- species_data %>%
  filter(STATUS == "nat")

nrow(native_only)

# 6c: Filter for alien/invasive species (STATUS == "neo")
# 6c: Filtere Alien/invasive Arten (STATUS == "neo")

alien_only <- species_data %>%
  filter(STATUS == "neo")

nrow(alien_only)

# 6d: How many unique native species?
# 6d: Wie viele einzigartige heimische Arten?

length(unique(native_only$WFO_TAXON))

# 6e: Your turn - how many unique alien species?
# 6e: Du bist dran - wie viele einzigartige Alien-Arten?

length(unique(__$WFO_TAXON))

# 6f: Filter for a specific plot
# 6f: Filtere einen bestimmten Plot

first_plot <- unique(species_data$PlotObservationID)[1]
first_plot

one_plot <- species_data %>%
  filter(PlotObservationID == first_plot)

# How many species in this one plot?
# Wie viele Arten in diesem einen Plot?

length(unique(one_plot$WFO_TAXON))

```



Hinweis: `filter()` behält Zeilen, wo die Bedingung TRUE ist. Benutze `==` für "gleich".

Übung 7: Die Akkumulationskurve bauen

Jetzt der spannende Teil - bauen wir eine Artenakkumulationskurve Schritt für Schritt!

```

# 7a: Get a list of all unique plots
# 7a: Hole eine Liste aller einzigartigen Plots

all_plots <- unique(species_data$PlotObservationID)
head(all_plots)

# 7b: Take a small sample (first 50 plots for speed)
# 7b: Nimm eine kleine Stichprobe (erste 50 Plots für Geschwindigkeit)

sample_plots <- all_plots[1:50]

# 7c: Initialize our accumulation tracking
# 7c: Initialisiere unsere Akkumulations-Verfolgung

species_found <- c()          # Empty vector to collect species
                                # Leerer Vektor um Arten zu sammeln
accumulation <- c()           # Will store species count after each plot
                                # Speichert Artenzahl nach jedem Plot

# 7d: Loop through each plot and count species!
# 7d: Schleife durch jeden Plot und zähle Arten!

for (i in 1:length(sample_plots)) {

  # Get current plot ID
  # Hole aktuelle Plot-ID
  current_plot <- sample_plots[i]

  # Get species in this plot
  # Hole Arten in diesem Plot
  plot_species <- species_data$WFO_TAXON[species_data$PlotObservationID == current_plot]

  # Add to our collection
  # Füge zu unserer Sammlung hinzu
  species_found <- c(species_found, plot_species)

  # Count unique species so far
  # Zähle einzigartige Arten bisher
  total_species <- length(unique(species_found))

  # Store in accumulation vector
  # Speichere im Akkumulationsvektor
  accumulation <- c(accumulation, total_species)
}

# 7e: Look at the result!
# 7e: Schau dir das Ergebnis an!

accumulation

```

Übung 8: Zeichne deine Kurve!

Der letzte Schritt - visualisieren wir unsere Akkumulationskurve!

```

# 8a: Create a data frame for plotting
# 8a: Erstelle einen Data Frame zum Plotten

curve_data <- data.frame(
  plots = 1:length(accumulation),
  species = accumulation
)

head(curve_data)

# 8b: Make the plot!
# 8b: Mache den Plot!

ggplot(curve_data, aes(x = plots, y = species)) +
  geom_line(color = "#2E7D32", linewidth = 2) +
  geom_point(color = "#2E7D32", size = 2) +
  labs(
    title = "My First Species Accumulation Curve!",
    x = "Number of Plots Sampled",
    y = "Cumulative Species Count"
  ) +
  theme_minimal(base_size = 14)

# 8c: Make it prettier with a filled area
# 8c: Mache es schöner mit gefüllter Fläche

ggplot(curve_data, aes(x = plots, y = species)) +
  geom_area(fill = "#C8E6C9", alpha = 0.5) +
  geom_line(color = "#2E7D32", linewidth = 2) +
  labs(
    title = "Species Accumulation Curve - Austria",
    subtitle = paste("Total species found:", max(accumulation)),
    x = "Plots Sampled",
    y = "Cumulative Species"
  ) +
  theme_minimal(base_size = 14)

```



Hinweis: Die Kurve flacht ab, weil wir weniger NEUE Arten finden, je mehr Plots wir untersuchen!

Bonus: Vergleiche Heimisch vs Alien!

Wenn du früh fertig bist, versuche heimische und Alien-Arten zu vergleichen!

```

# The species_data already has a STATUS column!
# Die species_data hat bereits eine STATUS-Spalte!

# Filter for native species only
# Filtere nur heimische Arten

native_species <- species_data %>%
  filter STATUS == "nat"

# How many native species?
# Wie viele heimische Arten?

length(unique(native_species$WFO_TAXON))

# How many alien species?
# Wie viele Alien-Arten?

alien_species <- species_data %>%
  filter STATUS == "neo"

length(unique(alien_species$WFO_TAXON))

# Now you can build separate curves for native vs alien!
# Jetzt kannst du separate Kurven für heimisch vs alien bauen!

```

Du hast es geschafft! Du hast deine erste Artenakkumulationskurve mit echten österreichischen Daten gebaut!

Gelernte Schlüsselkonzepte: - Variablen speichern Daten mit `<- - c()` erstellt Vektoren - `unique()` zählt verschiedene Werte - Data Frames organisieren Daten in Zeilen und Spalten - `filter()` wählt bestimmte Zeilen aus - Schleifen lassen uns Aktionen wiederholen - `ggplot()` erstellt schöne Plots

Tag 2

Theorie

Algorithmen & räumliche Ordnung

Tag 2: Theorie

Räumliche Algorithmen & Forschungsdesign

Ökologie-Workshop

2026-01-28

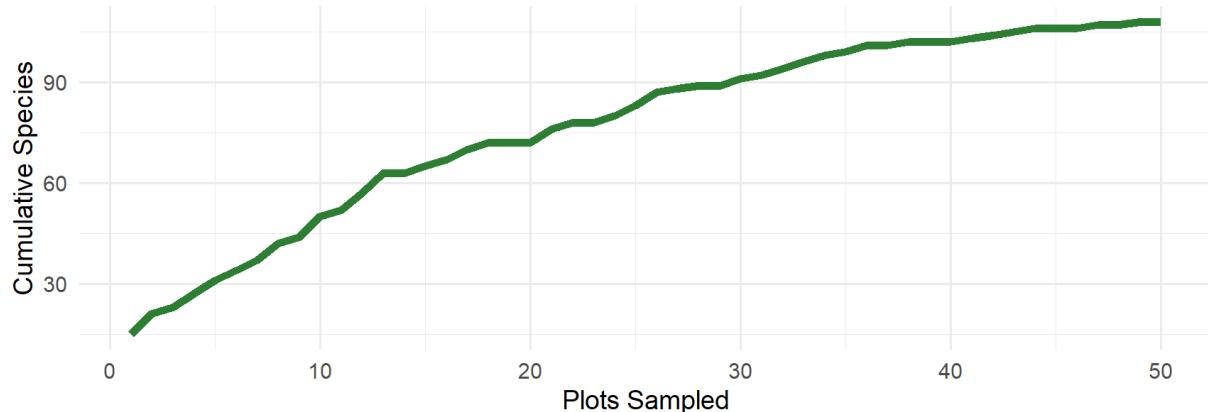
Wiederholung & Forschungsfrage

Kurze Wiederholung von Tag 1

Gestern haben wir gelernt:

- **Biodiversität** = Artenvielfalt in einem Gebiet
- **Akkumulationskurven** zeigen, wie die Artenentdeckung mit dem Stichprobenaufwand abnimmt
- **Asymptote** = geschätzte Gesamtartenzahl (was wir anstreben)
- Österreichische Daten haben **heimische** (natürliche) und **Alien-** (eingeführte) Arten

Review: Species Accumulation Curve



Heutige Forschungsfrage

Akkumulieren Alien-Arten anders als heimische Arten?

Hypothesen:

- H1: Alien-Arten akkumulieren **langsamer** (weniger Arten insgesamt)
- H2: Alien-Arten zeigen **andere räumliche Muster** als Einheimische
- H3: Die Kurven werden **unterschiedliche Formen** haben

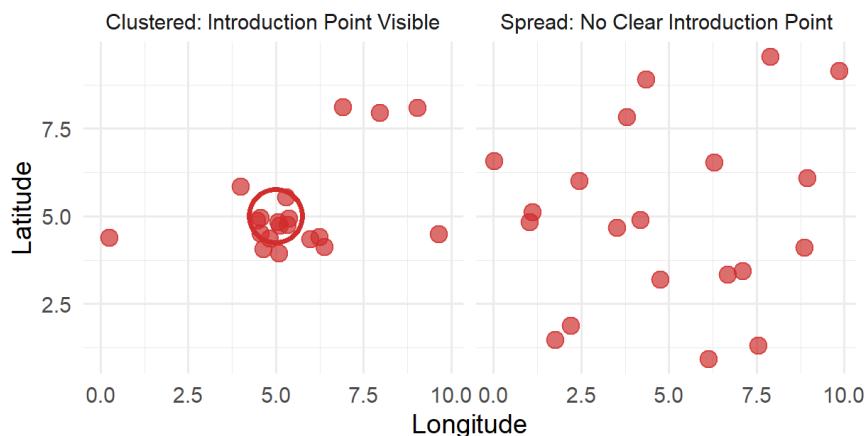
Warum räumliche Akkumulationskurven verwenden?

Die Schlüsselfrage: Warum nicht einfach Arten zählen? Warum brauchen wir Akkumulationskurven?

Antwort: Akkumulationskurven zeigen **wo** Arten sind, nicht nur **wie viele**.

Same Number of Alien Species, Different Stories

Left: We can identify WHERE aliens entered | Right: No spatial signal



Warum das für Invasionsbiologie wichtig ist:

Einführungspunkte

Wo kamen Aliens zuerst an? (Häfen, Städte, Gärten)

Ausbreitungsmuster

Wie verbreiten sie sich über die Landschaft?

Invasions-Hotspots

Welche Gebiete haben die meisten Alien-Arten?

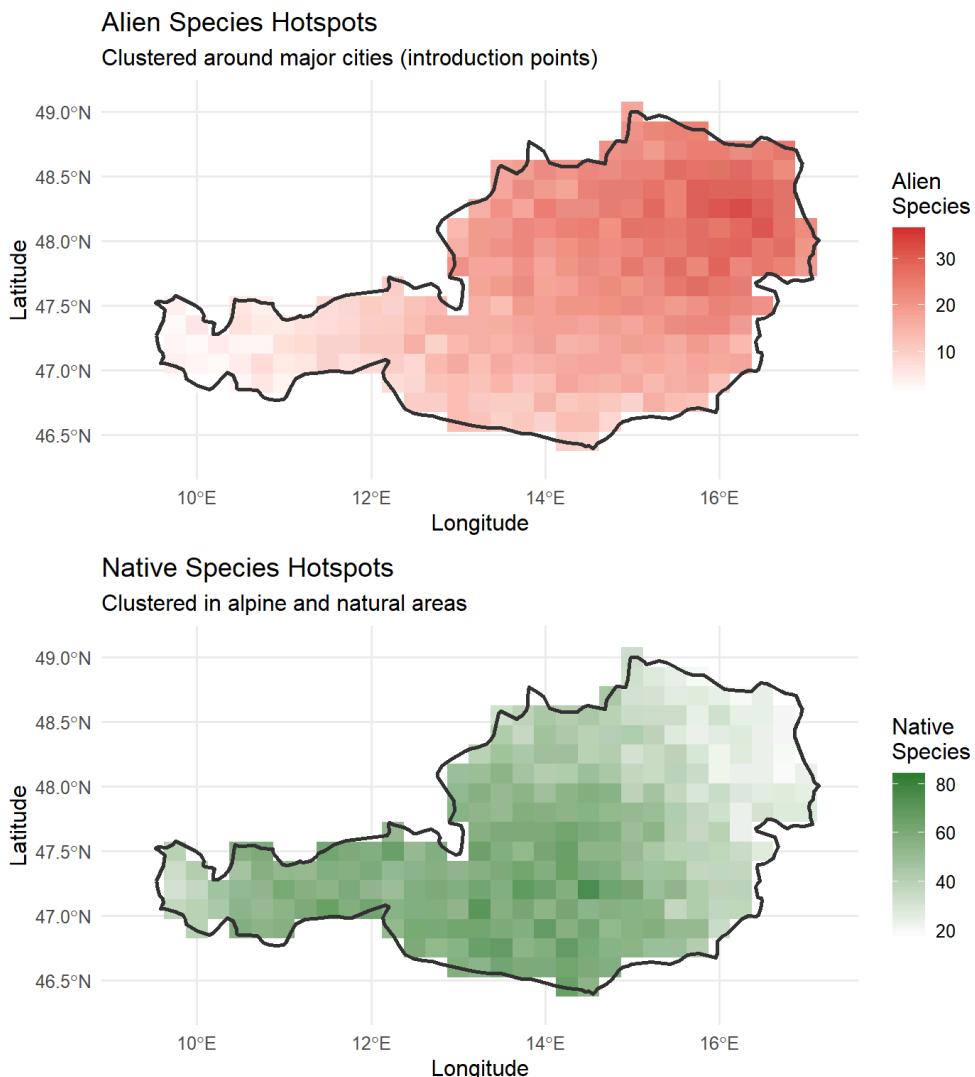
Management-Prioritäten

Wo sollten wir Kontrollmaßnahmen fokussieren?

Die Kurvenform erzählt die Geschichte: Wenn Aliens anfangs langsam akkumulieren und dann schneller werden, sind sie um Einführungspunkte geclustert. Wenn sie gleichmäßig akkumulieren, haben sie sich bereits überall verbreitet.

Hotspot-Karten: Österreich-Beispiel

Wie könnten Hotspot-Karten für Österreich aussehen? Hier ist ein hypothetisches Beispiel, das zeigt, wie Alien- und heimische Arten unterschiedlich über die Landschaft verteilt sein könnten.



Wichtige Beobachtungen (hypothetisch):

- **Alien-Arten** konzentrieren sich um **Wien, Linz, Salzburg** - Großstädte mit Häfen, Eisenbahnen und Gärten, wo Arten eingeführt werden
- **Heimische Arten** konzentrieren sich in **alpinen Regionen** - natürliche Lebensräume mit hoher endemischer Vielfalt
- Die **Muster sind fast gegensätzlich** - das können Akkumulationskurven aufdecken!

Der räumliche Algorithmus

Warum ist die Reihenfolge wichtig?

Gestern haben wir Kurven erstellt, indem wir Plots in **zufälliger Reihenfolge** untersucht haben.

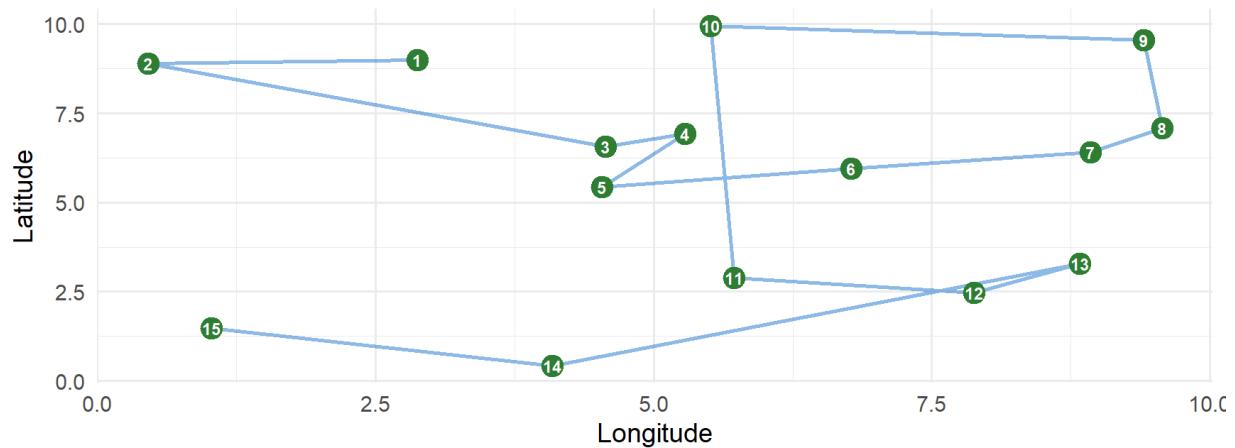
Aber in der realen Ökologie interessieren uns oft **räumliche Muster**:

- Sind Arten geografisch geclustert?
- Wie verändert sich die Diversität über eine Landschaft?

Lösung: Plots in **räumlicher Reihenfolge** untersuchen - immer zum nächsten unbesuchten Plot gehen!

Nearest-Neighbor Sampling Order

Numbers show visit order: always go to nearest unvisited plot



Der Nearest-Neighbor-Algorithmus

Schritt für Schritt:

1. Wähle einen **Startplot** (Seed)
2. Markiere ihn als **besucht**
3. Finde den **nächsten unbesuchten Plot**
4. Gehe dorthin, notiere neue Arten, markiere als besucht
5. **Wiederhole** bis alle Plots besucht sind

Die Entferungsformel

Entfernung zwischen zwei Punkten (x_1, y_1) und (x_2, y_2) :

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Dies ist die **Euklidische Distanz** - die Luftlinie zwischen zwei Punkten.

Mehrere Startpunkte

Warum mehrere Seeds?

Die Form der Kurve hängt davon ab, **wo wir starten!**

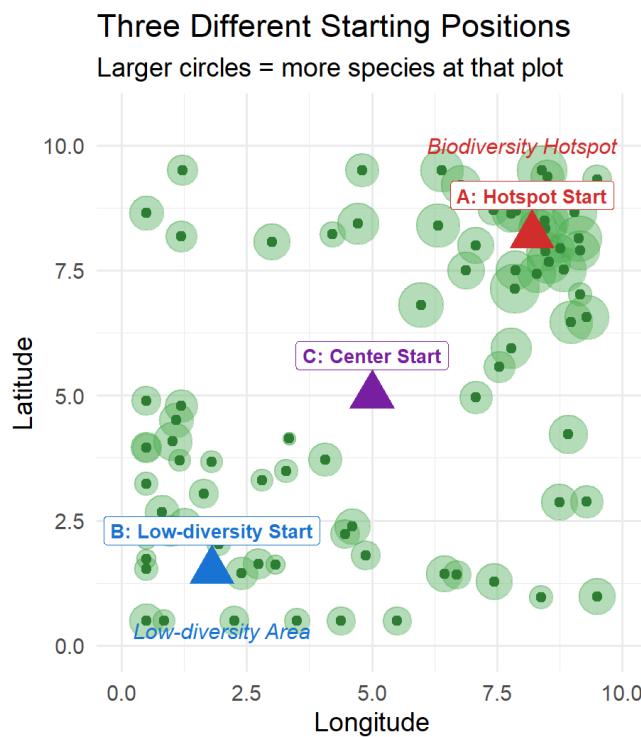
- Start in artenreichem Gebiet: schnelle anfängliche Akkumulation
- Start in artenarmem Gebiet: langsame anfängliche Akkumulation

Lösung: Den Algorithmus mehrmals mit verschiedenen Startpunkten ausführen, dann mitteln.

Startpositionen visualisieren

Stellen wir uns vor, wir haben Plots über eine Landschaft verteilt. Einige Gebiete sind **Biodiversitäts-Hotspots** (viele Arten), andere sind **artenarme Gebiete** (wenige Arten).

Wo wir unseren Probenweg beginnen, beeinflusst die Kurvenform dramatisch!



Was passiert bei jedem Start?

A: Hotspot Viele Arten schnell Steiler Anstieg, dann flach

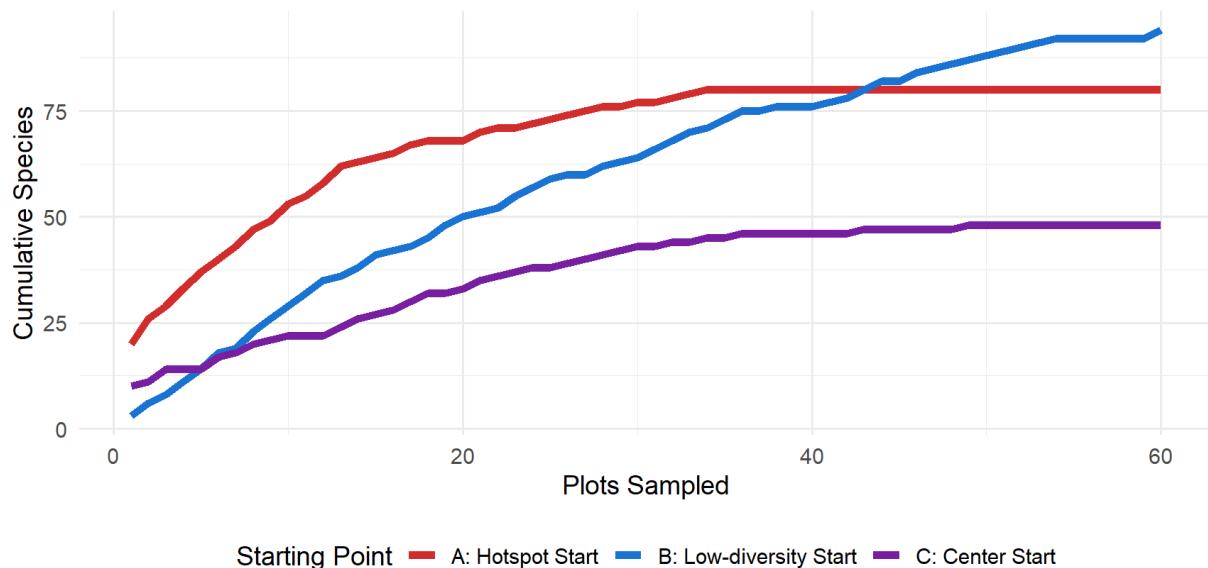
B: Artenarm Wenige Arten langsam Gradueller Anstieg

C: Zentrum Mittlerer Start Ausgewogene Kurve

Der Effekt der Startposition

Same Landscape, Different Starting Points

The curve shape depends heavily on where you begin!



Wichtige Beobachtung: Alle drei Kurven erreichen schließlich ungefähr die gleiche Artenzahl (sie konvergieren), aber der **Weg dorthin ist sehr unterschiedlich!**

Deshalb müssen wir mehrere Startpunkte in unserer Analyse berücksichtigen.

Die Nullhypothese

Was wenn die Startposition egal ist?

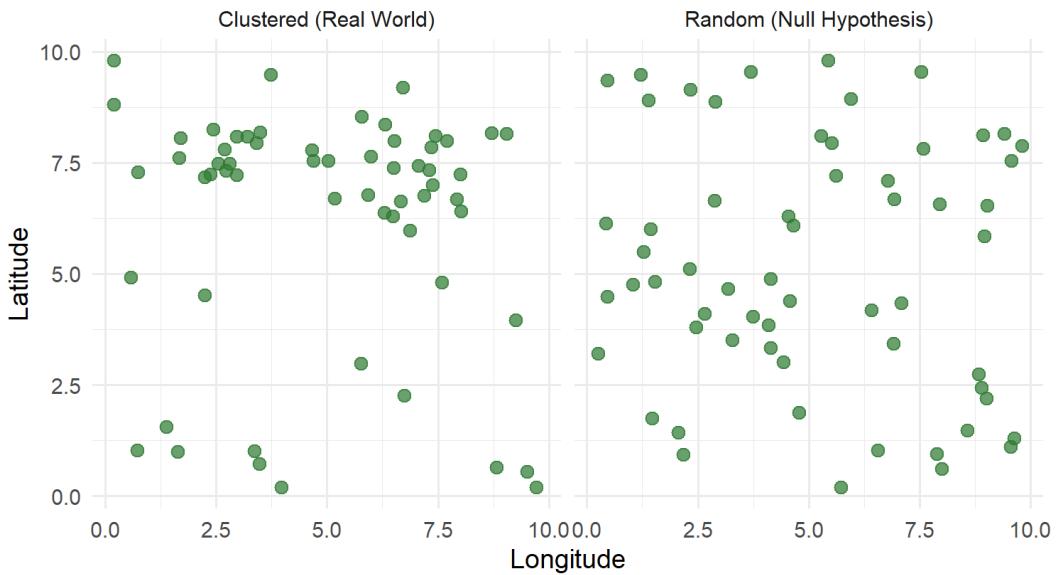
Die Nullhypothese (H_0):

“Die Startposition beeinflusst die Akkumulationskurve nicht - alle Kurven von verschiedenen Startpunkten sollten gleich aussehen.”

Wenn wahr: Arten sind zufällig über die Landschaft verteilt. Keine Hotspots, keine Cluster. Jeder Ort hat die gleiche Wahrscheinlichkeit für jede Art.

Null Hypothesis vs. Reality

Left: If species were randomly distributed | Right: How species actually cluster



Testen der Nullhypothese

Wie wir das testen:

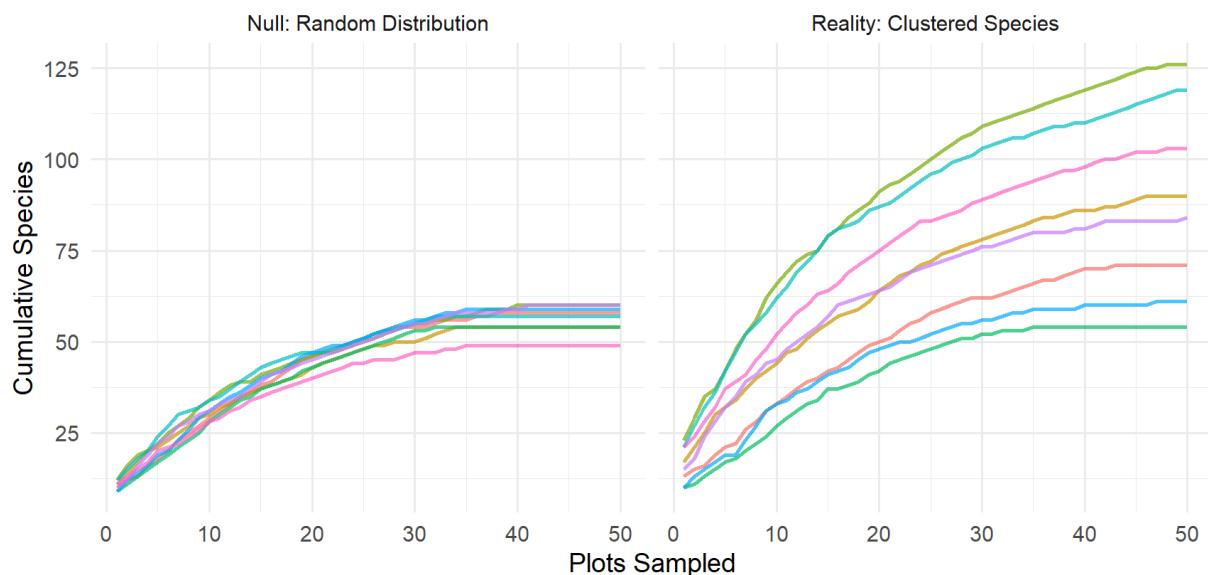
1. Den Sampling-Algorithmus von **vielen verschiedenen Startpunkten** ausführen (z.B. 20 Seeds)
2. Die Kurven vergleichen - überlappen sie oder divergieren sie?
3. Die **Variation zwischen Kurven** berechnen

Wenn H_0 wahr ist: Alle Kurven sollten fast identisch sein (geringe Variation)

Wenn H_0 falsch ist: Kurven sollten je nach Startort variieren (hohe Variation)

Why We Need Multiple Starting Points

Left: Similar curves (random world) | Right: Different curves (clustered world)



Wichtige Erkenntnis:

In einer Welt, in der Arten **zufällig verteilt** sind (H_0), sehen alle Kurven ähnlich aus, unabhängig vom Startpunkt.

In der **realen Welt**, wo Arten sich um günstige Habitate clustern, unterscheiden sich die Kurven dramatisch je nach Startort.

Deshalb mitteln wir über mehrere Seeds! Das gibt uns ein repräsentativeres Bild des gesamten Biodiversitätsmusters.

Was sagt uns hohe Variation?

Wenn Kurven von verschiedenen Startpunkten **stark variieren**, sagt uns das:

1. **Arten sind geclustert** - nicht gleichmäßig über die Landschaft verteilt
2. **Hotspots existieren** - einige Gebiete sind biodiversitätsreich
3. **Geografie ist wichtig** - Position beeinflusst, was man findet

Das ist eigentlich **interessanter** als einheitliche Kurven zu finden! Es bedeutet, dass es räumliche Muster gibt, die es wert sind, untersucht zu werden.

Vergleich Heimisch vs. Alien

Unsere Forschungshypothese

Forschungsfrage: Akkumulieren Alien-Arten (eingeführte Arten) anders als heimische Arten über die österreichische Landschaft?

Nullhypothese (H_0): > Heimische und Alien-Arten akkumulieren mit der gleichen Rate - ihre Kurven haben die gleiche Form. Es gibt keinen Unterschied in den räumlichen Verteilungsmustern.

Alternativhypothese (H_1): > Heimische und Alien-Arten haben unterschiedliche Akkumulationsmuster, die Unterschiede in ihrer Verteilung über die Landschaft widerspiegeln.

Was könnten wir erwarten?

Aliens akkumulieren schneller Steilere anfängliche Kurve Aliens sind Generalisten, überall zu finden

Aliens akkumulieren langsamer Flachere anfängliche Kurve Aliens sind geclustert nahe Einführungspunkten (St)

Kein Unterschied

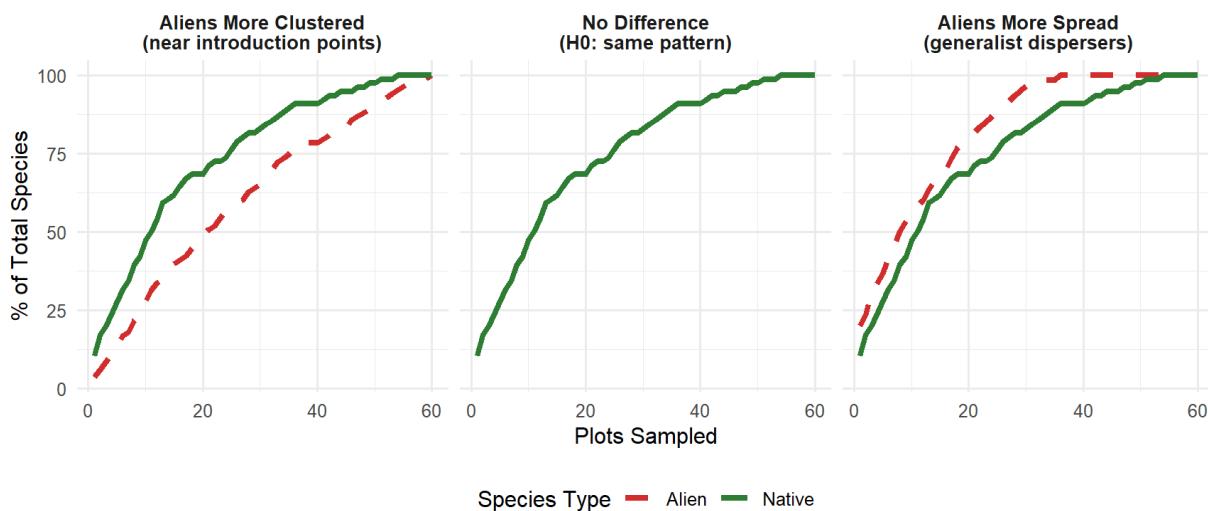
Identische Kurven

Verteilungsmuster sind ähnlich

Die wichtige Erkenntnis: Durch den Vergleich normalisierter Kurven können wir sehen, ob eine Gruppe stärker geclustert oder gleichmäßiger verteilt ist als die andere.

Possible Outcomes: Native vs. Alien Accumulation

Normalized curves (% of total species) allow shape comparison



Normalisierung für fairen Vergleich

Die Rohkurven sind schwer zu vergleichen, weil es viel mehr heimische Arten gibt.

Lösung: Auf Prozent der gefundenen Gesamtarten normalisieren.

So können wir die **Form** der Akkumulation vergleichen, nicht nur die Gesamtzahlen.

Interpretation & Diskussion

Was sagen uns die Kurven?

Beobachtungen aus dem Vergleich:

1. **Gesamtartenreichtum:** Heimische Arten >> Alien-Arten (wie erwartet)

2. **Position der normalisierten Kurve:**

- Kurve oben = proportional schnellere Akkumulation (niedrigerer räumlicher Turnover)
- Kurve unten = proportional langsamere Akkumulation (höherer räumlicher Turnover)

3. **Implikationen für Invasionen:**

- Höherer Alien-Turnover könnte bedeuten, dass verschiedene Arten an verschiedenen Orten ankommen
- Niedrigerer nativer Turnover könnte bedeuten, dass weit verbreitete Generalisten dominieren

Diskussionsfragen

1. Warum könnten Alien-Arten gleichmäßiger verteilt sein?

- Vom Menschen vermittelte Ausbreitung (Handel, Transport)
- Generalistische Habitatpräferenzen
- Kürzliche Einführung = keine Zeit zum Clustern

2. Was sind die Einschränkungen unserer Analyse?

- Geografische Verzerrung (manche Regionen besser beprobt)
- Nachweiswahrscheinlichkeit unterscheidet sich zwischen Arten
- Einzelne Zeitmomentaufnahme (keine zeitliche Dynamik)

3. Was könnte weiter untersucht werden?

- Verschiedene Habitattypen vergleichen (Wald vs. Wiese)
- Zeitliche Veränderungen (breiten sich Aliens aus?)
- In anderen Ländern testen

Biologische Mechanismen

Warum unterscheiden sich Akkumulationsmuster?

Ausbreitungsfähigkeit Gute Ausbreiter = gleichmäßige Verteilung = graduelle Kurve

Habitatspezifität Spezialisten = geclustert = steil dann flach

Historische Verbreitung Lange Geschichte = natürliche Clustermuster

Menschlicher Einfluss Transport entlang Straßen, Flüssen = lineare Muster

Zukünftige Richtungen

Über Artenreichtum hinaus: Biomasse-Akkumulation

Bisher haben wir nur **Arten** gezählt (Artenreichtum). Aber was ist mit **wie viel** von jeder Art vorhanden ist?

Biomasse-Akkumulationskurven könnten uns noch mehr verraten:

Artenreichtum Anzahl verschiedener Arten

Diversitätsmuster

Biomasse Gesamtgewicht/-deckung der Pflanzen Dominanzmuster

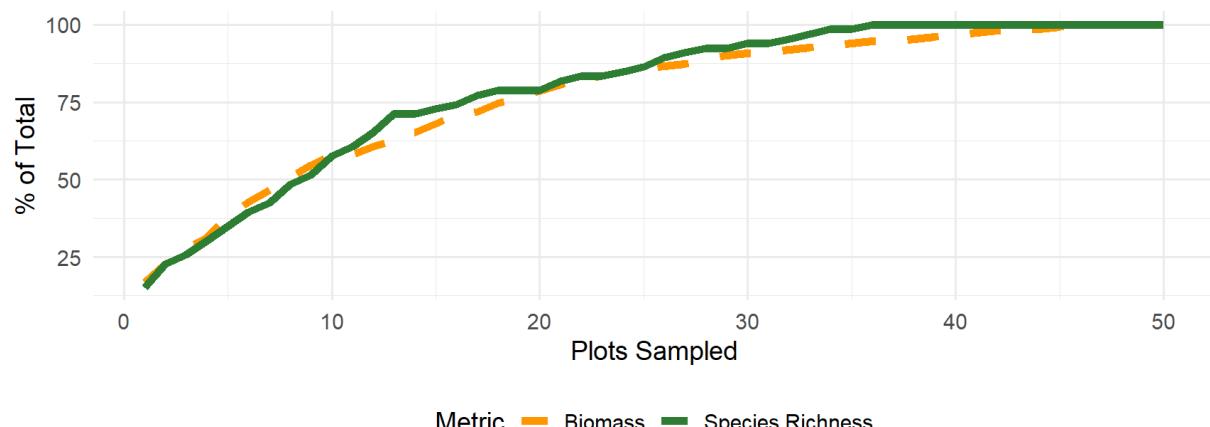
Kombiniert

Beides zusammen

Vollständiges ökologisches Bild

Future Idea: Combining Species Richness and Biomass

Biomass might accumulate faster (dominated by few common species)



Metric — Biomass — Species Richness

Warum Biomasse wichtig ist:

- Einige wenige **dominante Arten** könnten den Großteil der Biomasse ausmachen
- Alien-Arten könnten **hohe Biomasse aber geringe Diversität** haben (invasive Monokulturen)
- Heimische Ökosysteme könnten **ausgewogene Biomasse über viele Arten** zeigen

Forschungsfrage: Tragen Alien-Arten überproportional zur Biomasse bei im Verhältnis zu ihrer Artenzahl?

Zusammenfassung

Schlüsselkonzepte

Räumliche Reihenfolge Die Reihenfolge, in der wir Plots besuchen, beeinflusst die Kurvenform

Nearest-Neighbor Algorithmus: immer zum nächsten unbesuchten Plot gehen

Mehrere Seeds Verschiedene Startpunkte = verschiedene Kurven = Mittelung nötig

Heimisch vs. Alien Unterschiedliche Akkumulationsmuster zeigen ökologische Unterschiede

Fazit

Akkumulationskurven zeigen, wie Arten im Raum verteilt sind. Der Vergleich von heimischen und Alien-Arten zeigt uns, welche Gruppe stärker geclustert ist, und hilft Invasions-Hotspots zu identifizieren.

Tag 2

Übungen

Ein Mini-Forschungsprojekt

Tag 2: Übungen

Fortgeschrittene Analyse & Forschungsprojekt

Fülle die Lücken aus!

2026-01-28

Teil 2: Fortgeschrittene Analyse - Willkommen zu Tag 2!

Gestern hast du deine erste Artenakkumulationskurve gebaut. Heute bringen wir es mit echter wissenschaftlicher Analyse auf das nächste Level!

Was du heute lernst:

1. **Räumliche Probenahme** - Gehe durch Plots wie ein echter Feldökologe
2. **Unsicherheit messen** - Starte von mehreren Punkten
3. **Heimische vs Aliens vergleichen** - Akkumulieren sie unterschiedlich?
4. **Schöne Abbildungen** - Erstelle ansprechende Plots

So verwendest du diese Übungen:

- Ersetze jedes _____ mit dem richtigen Code
- Führe jeden Chunk aus um zu sehen ob er funktioniert
- Lies die Hinweise wenn du nicht weiterkommst
- Die Übungen bauen aufeinander auf - mache sie der Reihe nach!

Übung 1: Setup - Daten und Funktionen laden

Konzept: Bevor wir beginnen, müssen wir unsere Daten laden und einige Hilfsfunktionen definieren. Diese Funktionen machen unsere Analyse einfacher.

1a: Pakete und Daten laden

```
# Load packages
# Pakete Laden
library(tidyverse)

# Load data
# Daten Laden
header <- read_csv("../data/austria_header.csv")
species <- read_csv("../data/austria_species.csv")

# Check data loaded correctly
# Überprüfe dass Daten korrekt geladen sind
print(paste("Plots loaded:", nrow(header)))
print(paste("Species records:", nrow(species)))
print(paste("Unique species:", length(unique(species$WFO_TAXON))))
```

Bevor wir unsere Hilfsfunktionen definieren, lernen wir einige **neue R-Konzepte**, die wir heute verwenden. Gestern hast du Variablen, Vektoren und `unique()` gelernt. Heute fügen wir hinzu:

Eigene Funktionen schreiben - Fibonacci-Beispiel:

Die Fibonacci-Folge: 1, 1, 2, 3, 5, 8, 13... (jede Zahl = Summe der vorherigen zwei)

```
# VERSION 1: Mit Vektor (speichert ALLE ZahLen)
fibonacci_vektor <- function(n) {
  fib <- c(1, 1)                                # Starte mit ersten zwei Zahlen
  for (i in 3:n) {
    fib[i] <- fib[i-1] + fib[i-2]      # Addiere: vorherige + davor
  }
  return(fib)
}

fibonacci_vektor(7) # Gibt zurück: 1 1 2 3 5 8 13

# VERSION 2: Mit 2 Variablen (speichereffizient - nur letzte 2)
fibonacci_zwei_vars <- function(n) {
  vorherige <- 1                                # Die Zahl vor der aktuellen
  aktuelle <- 1                                  # Aktuelle Zahl
  for (i in 3:n) {
    neue <- vorherige + aktuelle      # Berechne nächste
    vorherige <- aktuelle            # Verschiebe: aktuelle wird vorherige
    aktuelle <- neue                # Verschiebe: neue wird aktuelle
  }
  return(aktuelle)
}

fibonacci_zwei_vars(7) # Gibt zurück: 13 (die 7. Fibonacci-Zahl)
```

Standardwerte in Funktionen:

```
# Wenn der User 'n' nicht angibt, benutze 10 als Standard
fibonacci <- function(n = 10) {
  fib <- c(1, 1)
  for (i in 3:n) fib[i] <- fib[i-1] + fib[i-2]
  return(fib)
}

fibonacci()      # Gibt erste 10 Zahlen zurück (benutzt Standard)
fibonacci(5)     # Gibt erste 5 Zahlen zurück
```

Nützliche Funktionen, die wir verwenden:

Funktion	Was sie macht	Beispiel
<code>paste()</code>	Text und Werte kombinieren	<code>paste("Gefunden:", 5, "Arten")</code>
<code>sample()</code>	Zufällige Elemente wählen	<code>sample(1:100, 5)</code> wählt 5 Zufallszahlen
<code>set.seed()</code>	Zufall reproduzierbar machen	<code>set.seed(42)</code> gibt immer gleichen "Zufall"
<code>rep()</code>	Werte wiederholen	<code>rep(FALSE, 5)</code> gibt FALSE FALSE FALSE FALSE FALSE
<code>which.min()</code>	Position des kleinsten Werts	<code>which.min(c(5,2,8))</code> gibt 2
<code>setdiff()</code>	Was ist in A aber nicht in B	<code>setdiff(c(1,2,3), c(2,3,4))</code> gibt 1

Funktion	Was sie macht	Beispiel
pull()	Spalte als Vektor extrahieren	data %>% pull(spalte)
is.null()	Prüfen ob NULL	is.null(NULL) gibt TRUE

Spezielle Werte:

- `NULL` = "nichts" / leer / nicht angegeben
- `Inf` = Unendlich (größer als jede Zahl)

1c: Hilfsfunktionen definieren (einfach kopieren & ausführen!)

Führe diesen Code-Block einfach aus! Du musst nicht jede Zeile verstehen - das sind Werkzeuge die wir später benutzen. Klicke die Details unten nur wenn du neugierig bist.

👉 Building the Functions Step-by-Step

Klicke jeden Schritt um zu sehen wie wir von einfach zu vollständig aufbauen:

▶ 1 calc_distance: Starte mit der Mathe-Formel
▶ 2 nn_walk: Zuerst überlegen was wir verfolgen müssen
▶ 3 nn_walk: Baue die Hauptschleife
▶ 4 nn_walk: Der "Unendlich-Trick" zum Ausschließen besuchter Plots
▶ 5 nn_walk: Flexibilität mit optionalem Startpunkt
▶ 6 build_accumulation: Überlege was "Akkumulation" bedeutet
▶ 7 build_accumulation: Nutze setdiff() um neue Arten zu finden
▶ 8 build_accumulation: Füge optionale Filterung hinzu
▶ 9 find_saturation: Ein einfacher aber nützlicher Helfer

👉 Klicke jeden Schritt oben zum Aufklappen! Hier sind die vollständigen Funktionen:

```

# ===== FUNCTION 1: calc_distance =====
# Calculates Euclidean distance. VECTORIZED: x2,y2 can be vectors!
# Berechnet Euklidische Distanz. VEKTORISIERT: x2,y2 können Vektoren sein!

calc_distance <- function(x1, y1, x2, y2) {
  # Pythagorean theorem: sqrt(dx2 + dy2). Works with vectors!
  # Satz des Pythagoras: sqrt(dx2 + dy2). Funktioniert mit Vektoren!
  sqrt((x2 - x1)^2 + (y2 - y1)^2)
}

# ===== FUNCTION 2: nn_walk =====
# Nearest-neighbour walk: always go to closest unvisited plot
# Nearest-Neighbour-Walk: gehe immer zum nächsten unbesuchten Plot

nn_walk <- function(header_data, start_idx = NULL) {
  # How many plots total? / Wie viele Plots insgesamt?
  n <- nrow(header_data)

  # If no start given, pick random / Wenn kein Start, wähle zufällig
  if (is.null(start_idx)) start_idx <- sample(1:n, 1)

  # Track which plots we've visited / Verfolge welche Plots besucht
  visited <- rep(FALSE, n)
  # Store the order we visit them / Speichere Reihenfolge
  visit_order <- numeric(n)
  # Start at this plot / Starte bei diesem Plot
  current <- start_idx

  # Loop through all plots / Schleife durch alle Plots
  for (i in 1:n) {
    # Mark current as visited / Markiere aktuellen als besucht
    visited[current] <- TRUE
    # Save the plot ID / Speichere Plot-ID
    visit_order[i] <- header_data$PlotObservationID[current]

    # If not done yet / Falls noch nicht fertig
    if (i < n) {
      # Calculate distance from current to ALL others (vectorized!)
      # Berechne Distanz von aktuellem zu ALLEN anderen (vektorisiert!)
      distances <- calc_distance(
        header_data$Longitude[current],
        header_data$Latitude[current],
        header_data$Longitude,
        header_data$Latitude
      )
      # Inf trick: visited plots can never be "minimum"
      # Inf-Trick: besuchte Plots können nie "Minimum" sein
      distances[visited] <- Inf
      # Go to closest unvisited / Gehe zum nächsten unbesuchten
      current <- which.min(distances)
    }
  }
  # Return the order of plot IDs / Gib Reihenfolge der Plot-IDs zurück
  return(visit_order)
}

# ===== FUNCTION 3: build_accumulation =====
# Count cumulative species as we visit each plot in order
# Zähle kumulative Arten während wir jeden Plot der Reihe nach besuchen

```

```

build_accumulation <- function(species_data, plot_order, status_filter = NULL) {
  # If filter provided (e.g., "nat"), keep only matching species
  # Falls Filter angegeben, behalte nur passende Arten
  if (!is.null(status_filter)) {
    species_data <- species_data %>% filter(STATUS == status_filter)
  }

  # Empty vector to collect all species found / Leerer Vektor für gefundene Arten
  found <- c()
  # Pre-allocate result vector / Ergebnisvektor vorbelegen
  accum <- numeric(length(plot_order))

  # For each plot in our walking order / Für jeden Plot in unserer Laufreihenfolge
  for (i in seq_along(plot_order)) {
    # Get unique species in this plot / Hole einzigartige Arten in diesem Plot
    plot_spp <- species_data %>%
      # Filter to current plot / Filtere auf aktuellen Plot
      filter(PlotObservationID == plot_order[i]) %>%
      # Extract species names / Extrahiere Artnamen
      pull(WFO_TAXON) %>%
      # Remove duplicates within plot / Entferne Duplikate im Plot
      unique()

    # setdiff: what's NEW? (in plot but not yet found)
    # setdiff: was ist NEU? (im Plot aber noch nicht gefunden)
    new_spp <- setdiff(plot_spp, found)
    # Add new species to our collection / Füge neue Arten zur Sammlung
    found <- c(found, new_spp)
    # Count total species so far / Zähle Gesamtarten bisher
    accum[i] <- length(found)
  }
  # Return the accumulation curve / Gib Akkumulationskurve zurück
  return(accum)
}

# ===== FUNCTION 4: find_saturation =====
# Find when we reach X% of total species (default 80%)
# Finde wann wir X% der Gesamtarten erreichen (Standard 80%)

find_saturation <- function(curve, threshold = 0.8) {
  # Calculate target: 80% of final count / Berechne Ziel: 80% der Endzahl
  target <- max(curve) * threshold
  # which()[1] = first position where TRUE / Erste Position wo TRUE
  which(curve >= target)[1]
}

print("All functions loaded!")

```



Hinweis: Führe beide Code-Blöcke aus. Bei Fehlern prüfe ob die Dateien in `../data/` existieren.

⚡ Turbo Mode: Show/Hide Rcpp Functions

1d: Wie schnell ist mein Code? (Komplexität)

Warum ist Geschwindigkeit wichtig? Mit 100 Plots dauert langsamer Code Sekunden. Mit 10.000 Plots könnte es Stunden dauern!

Was zählt als "Arbeit"? Um es einfach zu halten, zählen wir **Multiplikationen** (und Divisionen). Das sind die "teuren" Operationen - Additionen und Zuweisungen sind so schnell, dass wir sie ignorieren.

Das Problem - Verschachtelte Schleifen: Wenn du eine Schleife IN eine Schleife packst, multiplizieren sich die Multiplikationen!

N (Eingabegröße)	Multiplikationen (N×N)	Zeit
10	100	sofort
100	10.000	sofort
1.000	1.000.000	langsam
10.000	100.000.000	sehr langsam!

```
# O(N2) - wird schnell Langsam!
for (i in 1:N) {
  for (j in 1:N) {
    # Das läuft N × N = N2 mal!
  }
}
# N=100 → 10.000 Multiplikationen
# N=1000 → 1.000.000 Multiplikationen (1000x Langsamer!)
```

Unsere `nn_walk` Funktion hat dieses $O(N^2)$ Verhalten (sie prüft die Distanz zu allen Plots bei jedem Schritt). Genau dann macht es Sinn, eine **kompilierte Sprache** (wie C++) statt einer **interpretierten Sprache** (wie R) zu verwenden. Kompilierte Sprachen führen die gleichen Operationen 10-100× schneller aus - deshalb bieten wir die Rcpp-Version an!

Übung 2: Räumliche Probenahme verstehen

Konzept: Wenn wir Plots zufällig sampeln, springen wir vielleicht über die ganze Karte. Aber ein **Nearest-Neighbour-Walk** sampelt wie ein echter Ökologe - geht immer zum nächsten unbesuchten Plot. Das erstellt eine realistische räumliche Akkumulationskurve.

Neue Konzepte hier:

Konzept	Was es macht	Beispiel
<code>%in%</code>	Prüft ob Werte in einer Liste sind	<code>filter(ID %in% meine_ids)</code>
<code>scale_color_manual()</code>	Setzt eigene Farben für Kategorien	<code>values = c("A" = "red", "B" = "blue")</code>

2a: Erstelle einen Beispieldatensatz

```
# Take a random sample of 150 plots (for speed)
# Nimm eine zufällige Stichprobe von 150 Plots (für Geschwindigkeit)
set.seed(42) # For reproducibility / Für Reproduzierbarkeit
sample_size <- 150

sample_ids <- sample(unique(header$PlotObservationID), sample_size)
sample_header <- header %>% filter(PlotObservationID %in% sample_ids)
sample_species <- species %>% filter(PlotObservationID %in% sample_ids)

print(paste("Sample created:", nrow(sample_header), "plots"))
```

2b: Führe einen Nearest-Neighbour-Walk aus

```
# Start from plot 1 and walk to nearest neighbours
# Starte von Plot 1 und gehe zu nächsten Nachbarn
nn_order <- nn_walk(sample_header, start_idx = 1)

# Build curves for native and alien species
# Baue Kurven für heimische und Alien-Arten
native_curve <- build_accumulation(sample_species, nn_order, "___")
alien_curve <- build_accumulation(sample_species, nn_order, "___")

# Check the results
# Überprüfe die Ergebnisse
print(paste("Native species found:", max(native_curve)))
print(paste("Alien species found:", max(alien_curve)))
```

2c: Plotte den Vergleich

```
# Create data frame for plotting
# Erstelle Data Frame zum Plotten
curve_data <- data.frame(
  plots = rep(1:sample_size, 2),
  species = c(native_curve, alien_curve),
  status = rep(c("Native", "Alien"), each = sample_size)
)

ggplot(curve_data, aes(x = plots, y = species, color = status)) +
  geom_line(linewidth = 1.2) +
  scale_color_manual(values = c("Native" = "darkgreen", "Alien" = "___")) +
  labs(
    title = "Native vs Alien Species Accumulation",
    x = "Plots Sampled (Nearest-Neighbour Order)",
    y = "Cumulative Species",
    color = ""
  ) +
  theme_minimal()
```



Hinweis: Heimische Arten haben STATUS = "nat", Alien-Arten haben STATUS = "neo".

Übung 3: Unsicherheit messen

Konzept: Die Kurve die wir bekommen hängt davon ab WO wir starten! Starten wir in einem heimisch-reichen Gebiet, akkumulieren Heimische schnell. Starten wir nahe einem Alien-Hotspot, akkumulieren Aliens schnell. Wir müssen von VIELEN Startpunkten starten um das wahre Muster zu sehen.

Neue Konzepte in dieser Übung:

Konzept	Was es macht	Beispiel
matrix(NA, nrow, ncol)	Erstelle leere Tabelle mit Zeilen & Spalten	matrix(NA, nrow=20, ncol=100)
colMeans(mat)	Mittelwert jeder Spalte	colMeans(mat) gibt einen Mittelwert pro Spalte
sd()	Standardabweichung (Streuung der Werte)	sd(c(1,2,3,4,5)) gibt 1.58
quantile(x, p)	Wert am Perzentil p	quantile(x, 0.975) = 97.5-tes Perzentil
geom_ribbon()	Schattiertes Band in ggplot	geom_ribbon(aes(ymin, ymax))
alpha = 0.3	Transparenz (0=unsichtbar, 1=solide)	Macht Bänder durchsichtig

Warum matrix() statt data.frame()? - `matrix` = ALLE Werte sind gleicher Typ (nur Zahlen). Schneller für Mathe! - `data.frame` = Spalten können verschiedene Typen sein (Text, Zahlen, TRUE/FALSE)

Wir benutzen `matrix` hier weil wir nur Zahlen speichern und schnelle Berechnungen brauchen.

3a: Von mehreren Startpunkten starten

```
# Run from 20 different starting points
# Starte von 20 verschiedenen Startpunkten
n_seeds <- 20

# Storage for results (matrix: rows = seeds, columns = plots)
# Speicher für Ergebnisse (Matrix: Zeilen = Seeds, Spalten = Plots)
native_runs <- matrix(NA, nrow = n_seeds, ncol = sample_size)
alien_runs <- matrix(NA, nrow = n_seeds, ncol = sample_size)

# Run from each starting point
# Starte von jedem Startpunkt
print(paste("Running", n_seeds, "starting points..."))
for (seed in 1:n_seeds) {
  nn_order <- nn_walk(sample_header, start_idx = seed)
  native_runs[seed, ] <- build_accumulation(sample_species, nn_order, "___")
  alien_runs[seed, ] <- build_accumulation(sample_species, nn_order, "___")

  if (seed %% 5 == 0) print(paste(" Completed", seed))
}
print("Done!")
```

3b: Berechne Mittelwert und Konfidenzintervalle

```

# Mean across all runs (colMeans = mean of each column)
# Mittelwert über alle Läufe (colMeans = Mittelwert jeder Spalte)
native_mean <- colMeans(____)
alien_mean <- colMeans(alien_runs)

# 95% confidence intervals using a loop
# 95% Konfidenzintervalle mit einer Schleife
native_lower <- numeric(sample_size)
native_upper <- numeric(sample_size)
alien_lower <- numeric(sample_size)
alien_upper <- numeric(sample_size)

for (i in 1:sample_size) {
  native_lower[i] <- quantile(native_runs[, i], 0.025)
  native_upper[i] <- quantile(native_runs[, i], ____)
  alien_lower[i] <- quantile(alien_runs[, i], ____)
  alien_upper[i] <- quantile(alien_runs[, i], 0.975)
}

# How much variation is there?
# Wie viel Variation gibt es?
midpoint <- sample_size %/% 2
native_cv <- round(100 * sd(native_runs[, midpoint]) / mean(native_runs[, midpoint]), 1)
alien_cv <- round(100 * sd(alien_runs[, midpoint]) / mean(alien_runs[, midpoint]), 1)

print("Variation at midpoint:")
print(paste(" Native CV:", native_cv, "%"))
print(paste(" Alien CV:", alien_cv, "%"))

```

3c: Plotte mit Unsicherheitsbändern

```

# Create summary data frame
# Erstelle Zusammenfassungs-Data-Frame
summary_data <- data.frame(
  plots = rep(1:sample_size, 2),
  mean = c(native_mean, alien_mean),
  lower = c(native_lower, alien_lower),
  upper = c(native_upper, alien_upper),
  status = rep(c("Native", "Alien"), each = sample_size)
)

# Plot with shaded confidence bands
# Plotte mit schattierten Konfidenzbändern
ggplot(summary_data, aes(x = plots)) +
  geom_ribbon(aes(ymin = lower, ymax = upper, fill = status), alpha = ____) +
  geom_line(aes(y = mean, color = status), linewidth = 1.2) +
  scale_color_manual(values = c("Native" = "darkgreen", "Alien" = "red")) +
  scale_fill_manual(values = c("Native" = "darkgreen", "Alien" = "red")) +
  labs(
    title = "Species Accumulation with Uncertainty",
    subtitle = paste(n_seeds, "starting points - shaded = 95% CI"),
    x = "Plots Sampled",
    y = "Cumulative Species",
    color = "", fill = ""
  ) +
  theme_minimal() +
  theme(legend.position = "bottom")

```



Hinweis: Benutze `mean` für den Durchschnitt. Für 95% CI ist obere Grenze 0.975 Quantil, untere 0.025.

Übung 4: Sättigungsgeschwindigkeit vergleichen

Konzept: Sättigung ist wenn wir die meisten Arten gefunden haben und die Kurve abflacht. Wenn Heimische SCHNELLER saturieren, bedeutet das, dass wir die meisten heimischen Arten schnell finden wenn wir durch die Landschaft gehen. Wenn Aliens LANGSAMER saturieren, kommen verschiedene Alien-Arten in verschiedenen Regionen vor — hoher räumlicher Turnover.

Neues Konzept - Boxplots:

`geom_boxplot()` zeigt die Verteilung von Werten: - **Box** = mittlere 50% der Daten (25. bis 75. Perzentil) - **Linie innen** = Median (mittlerer Wert) - **Whiskers** = erstrecken sich $\sim 1.5 \times$ die Boxhöhe - **Punkte** = Ausreißer (extreme Werte)

4a: Finde Sättigungspunkte

```
# At what point do we reach 80% of species?  
# Bei welchem Punkt erreichen wir 80% der Arten?  
native_sat <- numeric(n_seeds)  
alien_sat <- numeric(n_seeds)  
  
for (i in 1:n_seeds) {  
  native_sat[i] <- find_saturation(native_runs[i, ], threshold = 0.8)  
  alien_sat[i] <- find_saturation(alien_runs[i, ], threshold = ____)  
}  
  
# Convert to percentage of total plots  
# In Prozent der Gesamtplots umwandeln  
native_sat_pct <- 100 * native_sat / sample_size  
alien_sat_pct <- 100 * alien_sat / sample_size  
  
print("Plots needed to reach 80% of species:")  
print(paste("  Native:", round(mean(native_sat_pct)), "% (SD:", round(sd(native_sat_pct), 1),  
    ")"))  
print(paste("  Alien:", round(mean(alien_sat_pct)), "% (SD:", round(sd(alien_sat_pct), 1), ")"))
```

4b: Visualisiere mit Boxplot

```
# Create data for boxplot
# Erstelle Daten für Boxplot
sat_data <- data.frame(
  saturation_pct = c(native_sat_pct, alien_sat_pct),
  status = rep(c("Native", "Alien"), each = n_seeds)
)

ggplot(sat_data, aes(x = status, y = saturation_pct, fill = status)) +
  geom_boxplot() +
  scale_fill_manual(values = c("Native" = "darkgreen", "Alien" = "red")) +
  labs(
    title = "How Fast Do Species Saturate?",
    subtitle = "% of plots needed to find 80% of species",
    y = "% of Plots Needed",
    x = ""
  ) +
  theme_minimal() +
  theme(legend.position = "none")
```

4c: Berechne Steigungsverhältnis

```
# Compare early vs late slopes
# Vergleiche frühe vs späte Steigungen
# High ratio = fast saturation (steep early, flat late)
# Hohes Verhältnis = schnelle Sättigung (steil früh, flach spät)

calc_slope_ratio <- function(curve) {
  n <- length(curve)
  early_end <- round(n * 0.2)
  late_start <- round(n * 0.8)

  early_slope <- (curve[early_end] - curve[1]) / early_end
  late_slope <- (curve[n] - curve[late_start]) / (n - late_start)

  if (late_slope > 0) return(early_slope / late_slope)
  else return(NA)
}

native_ratios <- numeric(n_seeds)
alien_ratios <- numeric(n_seeds)

for (i in 1:n_seeds) {
  native_ratios[i] <- calc_slope_ratio(native_runs[i, ])
  alien_ratios[i] <- calc_slope_ratio(alien_runs[i, ])
}

print("Slope ratio (higher = faster saturation):")
print(paste("  Native:", round(mean(native_ratios, na.rm = TRUE), 1)))
print(paste("  Alien:", round(mean(alien_ratios, na.rm = TRUE), 1)))
```



Hinweis: Fülle Farben "darkgreen" und "red" ein.

Übung 5: Geografische Muster

Konzept: Ist es wichtig WO in Österreich wir starten? Manche Gebiete haben vielleicht mehr Aliens (nahe Städten, Straßen). Kartieren wir wie der Startort beeinflusst was wir finden!

Neue ggplot-Konzepte für Karten:

Konzept	Was es macht
coord_quickmap()	Korrekte Kartenproportion (damit Österreich nicht gestreckt aussieht!) Am Ende vor
scale_color_gradient2()	Farbskala mit 3 Farben (niedrig→mittel→hoch)
midpoint = wert	Wo die mittlere Farbe erscheint

5a: Kartiere Alien-Anteil nach Startort

```

# Test from 30 different starting Locations
# Teste von 30 verschiedenen Startorten
n_map_seeds <- 30
checkpoint <- 75 # Check after this many plots

seed_results <- data.frame(
  seed_idx = 1:n_map_seeds,
  seed_lon = numeric(n_map_seeds),
  seed_lat = numeric(n_map_seeds),
  native_count = numeric(n_map_seeds),
  alien_count = numeric(n_map_seeds)
)

for (i in 1:n_map_seeds) {
  # Record starting location
  # Erfasse Startort
  seed_results$seed_lon[i] <- sample_header$Longitude[i]
  seed_results$seed_lat[i] <- sample_header$Latitude[i]

  # Run from this starting point
  # Starte von diesem Startpunkt
  nn_order <- nn_walk(sample_header, start_idx = i)

  # Get counts at checkpoint
  # Hole Zählungen am Checkpoint
  native_curve <- build_accumulation(sample_species, nn_order, "nat")
  alien_curve <- build_accumulation(sample_species, nn_order, "neo")

  seed_results$native_count[i] <- native_curve[min(checkpoint, length(native_curve))]
  seed_results$alien_count[i] <- alien_curve[min(checkpoint, length(alien_curve))]
}

# Calculate alien proportion
# Berechne Alien-Anteil
seed_results$alien_prop <- seed_results$alien_count /
  (seed_results$native_count + seed_results$__)

# Map it!
# Kartiere es!
ggplot(seed_results, aes(x = seed_lon, y = seed_lat, color = alien_prop)) +
  geom_point(size = 4) +
  scale_color_gradient2(
    low = "darkgreen", mid = "yellow", high = "red",
    midpoint = mean(seed_results$alien_prop),
    name = "Alien\\nproportion"
  ) +
  coord_quickmap() +
  labs(
    title = "How Starting Location Affects Alien Detection",
    subtitle = paste("Alien proportion after", checkpoint, "plots"),
    x = "Longitude", y = "Latitude"
  ) +
  theme_minimal()

```



Hinweis: Fülle `alien_count` ein um die Anteilsberechnung zu vervollständigen.

Übung 6: Alles zusammenfügen

Konzept: Fassen wir alles zusammen und erstellen eine schöne Abbildung die unser Hauptergebnis zeigt: wie heimische vs Alien-Arten unterschiedlich über österreichische Plots akkumulieren.

6a: Führe vollständige Analyse aus

```
# Larger sample for final analysis
# Größere Stichprobe für finale Analyse
n_seeds <- 30
sample_size <- 200

set.seed(2024)
sample_ids <- sample(unique(header$PlotObservationID), sample_size)
sample_header <- header %>% filter(PlotObservationID %in% sample_ids)
sample_species <- species %>% filter(PlotObservationID %in% sample_ids)

# Run from all seeds
# Starte von allen Seeds
native_runs <- matrix(NA, nrow = n_seeds, ncol = sample_size)
alien_runs <- matrix(NA, nrow = n_seeds, ncol = sample_size)

print("Running full analysis...")
for (seed in 1:n_seeds) {
  nn_order <- nn_walk(sample_header, start_idx = seed)
  native_runs[seed, ] <- build_accumulation(sample_species, nn_order, "nat")
  alien_runs[seed, ] <- build_accumulation(sample_species, nn_order, "neo")
  if (seed %% 10 == 0) print(paste(" Completed", seed, "of", n_seeds))
}
print("Done!")

# Calculate summaries using colMeans and Loops
# Berechne Zusammenfassungen mit colMeans und Schleifen
native_mean <- colMeans(native_runs)
alien_mean <- colMeans(alien_runs)

native_lower <- numeric(sample_size)
native_upper <- numeric(sample_size)
alien_lower <- numeric(sample_size)
alien_upper <- numeric(sample_size)

for (i in 1:sample_size) {
  native_lower[i] <- quantile(native_runs[, i], 0.025)
  native_upper[i] <- quantile(native_runs[, i], 0.975)
  alien_lower[i] <- quantile(alien_runs[, i], 0.025)
  alien_upper[i] <- quantile(alien_runs[, i], 0.975)
}

results <- data.frame(
  plots = rep(1:sample_size, 2),
  mean = c(native_mean, alien_mean),
  lower = c(native_lower, alien_lower),
  upper = c(native_upper, alien_upper),
  status = rep(c("Native", "Alien"), each = sample_size)
)
```

6b: Erstelle schöne Abbildung

```
# Polished figure
# Schöne Abbildung
final_plot <- ggplot(results, aes(x = plots)) +
  geom_ribbon(aes(ymin = lower, ymax = upper, fill = status), alpha = 0.25) +
  geom_line(aes(y = mean, color = status), linewidth = 1.3) +
  scale_color_manual(values = c("Native" = "#228B22", "Alien" = "#DC143C")) +
  scale_fill_manual(values = c("Native" = "#228B22", "Alien" = "#DC143C")) +
  labs(
    title = "Species Accumulation: Native vs Alien Plants in Austria",
    subtitle = paste0(n_seeds, " starting points, 95% confidence intervals"),
    x = "Number of Plots Sampled",
    y = "Cumulative Species Count",
    color = "", fill = ""
  ) +
  theme_minimal(base_size = 13) +
  theme(
    legend.position = c(0.85, 0.2),
    legend.background = element_rect(fill = "white", color = "gray80"),
    plot.title = element_text(face = "bold")
  )
print(final_plot)

# Save it (uncomment to run)
# ggsave("austria_accumulation.png", final_plot, width = 10, height = 7, dpi = 300)
```

Übung 7: Interpretiere deine Ergebnisse

Konzept: Wissenschaft bedeutet nicht nur Plots zu erstellen - es geht darum zu verstehen was sie bedeuten! Fassen wir unsere Ergebnisse zusammen.

7a: Generiere Ergebniszusammenfassung

```

print("===== RESULTS SUMMARY =====")

# Total species found
print("Species found:")
print(paste(" Native:", round(mean(native_runs[, sample_size])), "species"))
print(paste(" Alien:", round(mean(alien_runs[, sample_size])), "species"))

# Saturation comparison (using a Loop)
native_sat <- numeric(n_seeds)
alien_sat <- numeric(n_seeds)
for (i in 1:n_seeds) {
  native_sat[i] <- find_saturation(native_runs[i, ], threshold = 0.8)
  alien_sat[i] <- find_saturation(alien_runs[i, ], threshold = 0.8)
}

print("Plots to reach 80% of species:")
print(paste(" Native:", round(mean(native_sat)), "plots (",
  round(100*mean(native_sat)/sample_size), "%)"))
print(paste(" Alien:", round(mean(alien_sat)), "plots (",
  round(100*mean(alien_sat)/sample_size), "%)"))

# Draw conclusion
print("===== CONCLUSION =====")
if (mean(native_sat) < mean(alien_sat)) {
  print("Native species saturate FASTER than aliens!")
  print("-> We find most native species quickly across the landscape")
  print("-> Different alien species occur in different regions (high spatial turnover") )
} else {
  print("Aliens saturate FASTER than natives!")
  print("-> Aliens are more widespread than expected")
}

```

Diskussionsfragen: 1. Saturieren Heimische wirklich schneller? Was bedeutet das ökologisch? 2. Warum könnten verschiedene Alien-Arten in verschiedenen Regionen vorkommen? (Denke daran wie sie angekommen sind) 3. Was würdest du als nächstes untersuchen? (Städtische Gebiete? Höhenlage? Klima?)

Schnellreferenz

Funktion	Zweck	Beispiel
colMeans()	Mittelwert jeder Spalte	colMeans(mat)
quantile()	Perzentile berechnen	quantile(x, 0.975)
geom_ribbon()	Schattiertes Band	+ geom_ribbon(aes(ymin, ymax))
scale_color_manual()	Eigene Farben	+ scale_color_manual(values = c(...))
scale_color_gradient2()	Divergierende Farbskala	+ scale_color_gradient2(...)
ggsave()	Plot in Datei speichern	ggsave("plot.png", width = 10)

Funktion	Zweck	Beispiel
theme()	Plot-Erscheinung anpassen	+ theme(legend.position = "bottom")
element_rect()	Rechteck-Element für Themes	element_rect(fill = "white")
element_text()	Text-Element für Themes	element_text(face = "bold")



Nach dem Workshop

Ergebnisse

Analyse-Ergebnisse

Tag 3: Ergebnisse & wissenschaftliche Interpretation

Von der Hypothese zur Evidenz – Erforschung der österreichischen Pflanzenvielfalt

Ökologie-Workshop

2026-02-16

Einleitung

Von der Theorie zur Evidenz

An **Tag 1** haben wir die Grundlagen gelernt: Was Biodiversität ist, warum sie wichtig ist, wie Artenakkumulationskurven funktionieren und warum R für großflächige ökologische Analysen unverzichtbar ist. An **Tag 2** haben wir räumliche Algorithmen erkundet, die Nearest-Neighbor-Methode, die Bedeutung mehrerer Startpunkte, und wir haben Hypothesen über heimische versus Alien-Arten formuliert.

Heute gehen wir von der Hypothese zur Evidenz.

Wir untersuchen die Ergebnisse einer umfassenden räumlichen Analyse von ~**52.000 Vegetationsplots** aus dem Europäischen Vegetationsarchiv (EVA), die ganz Österreich abdecken. Jedes hier gezeigte Ergebnis wurde mit dem R-Paket `spacc` berechnet, unter Verwendung des kNN-Algorithmus (k-nächste-Nachbarn) und Haversine-Distanzen.

Dieses Dokument umfasst **13 Analysen**, jede in ökologischer Theorie verankert und durch wissenschaftliche Literatur gestützt. Wir schreiten progressiv von einfachen Artenzahlen zu fortgeschrittenen Diversitätsmetriken voran:

1. **Datensatzübersicht** — Was haben wir?
2. **Zeitliche Dynamik** — Hat sich die Diversität über Jahrzehnte verändert?
3. **Heimisch vs. Alien Akkumulation** — Test unserer zentralen Hypothese
4. **Extrapolation** — Wie viele Arten fehlen uns?
5. **Räumliches vs. zufälliges Sampling** — Ist Geografie wichtig?
6. **Hill-Zahlen** — Jenseits des Artenreichtums
7. **Diversitätsaufteilung** — Alpha-, Beta-, Gamma-Zerlegung
8. **Beta-Diversität** — Turnover vs. Nestedness
9. **Deckungsbasierte Rarefaction** — Wie vollständig ist unsere Beprobung?
10. **Phylogenetische Diversität** — Die Perspektive des Lebensbaums
11. **Funktionelle Diversität** — Eigenschaftsbasierte Muster
12. **Habitat-Analyse** — EUNIS-Habitatvergleiche
13. **Randeffekte** — Das Halo-Problem

1. Der Datensatz

Österreichische Vegetationsplots

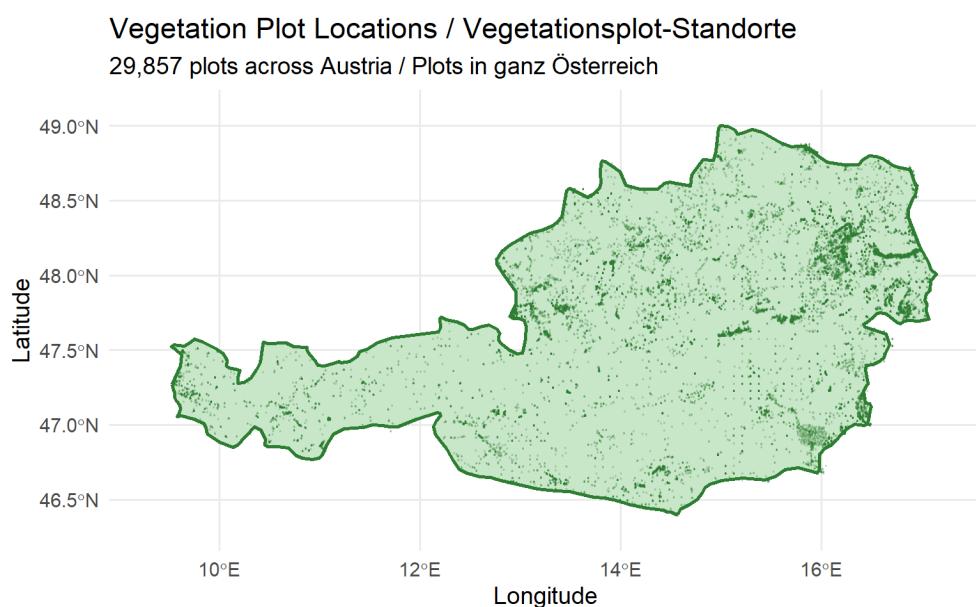
Unser Datensatz stammt aus dem **Europäischen Vegetationsarchiv (EVA)**, der größten Sammlung von Vegetationsplotdaten in Europa (Chytry et al. 2016). Jeder Plot repräsentiert eine kleine Fläche (typischerweise 1–100 m²), auf der Ökologen jede vorhandene Pflanzenart erfasst haben.

Total plots	29,857
Species records	852,075
Unique species	2,700
Native species	2,378
Alien species (neophytes)	323
Alien proportion	12%

Warum ist das Verhältnis heimisch zu alien wichtig?

Global machen Alien-Pflanzenarten durchschnittlich etwa 3,9% der nationalen Floren aus, aber dies variiert enorm nach Region (van Kleunen et al. 2015). In Mitteleuropa ist der Anteil tendenziell höher aufgrund von Jahrhunderten des Handels, der Landwirtschaft und Urbanisierung. Österreich liegt an einer biogeografischen Kreuzung — alpine, pannonische und kontinentale Zonen konvergieren — und schafft vielfältige Bedingungen sowohl für heimischen Endemismus als auch für Alien-Etablierung.

Räumliche Verteilung der Plots



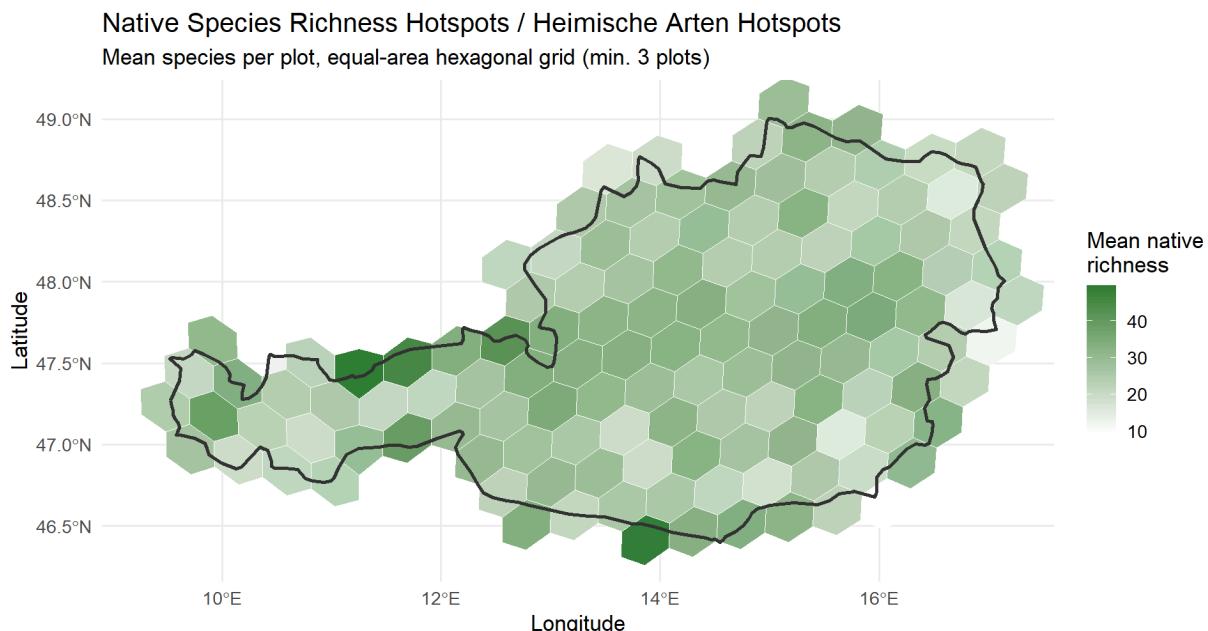
Stichprobenverzerrung ist sichtbar: Einige Regionen sind dicht beprobt (z.B. Wiener Becken, Inntal), während andere (hochalpine Gebiete) weniger Plots haben. Diese ungleichmäßige Abdeckung ist eine bekannte Eigenschaft opportunistischer Vegetationsdatenbanken (Dengler et al. 2011) und einer der Gründe, warum räumliche Akkumulationsmethoden — statt einfacher Artenzählungen — für faire Biodiversitätsvergleiche notwendig sind.

1970s	2386
1980s	3171
1990s	8814
2000s	4992
2010s	8285
2020s	104
pre-1970	2105

Biodiversitäts-Hotspot-Karten

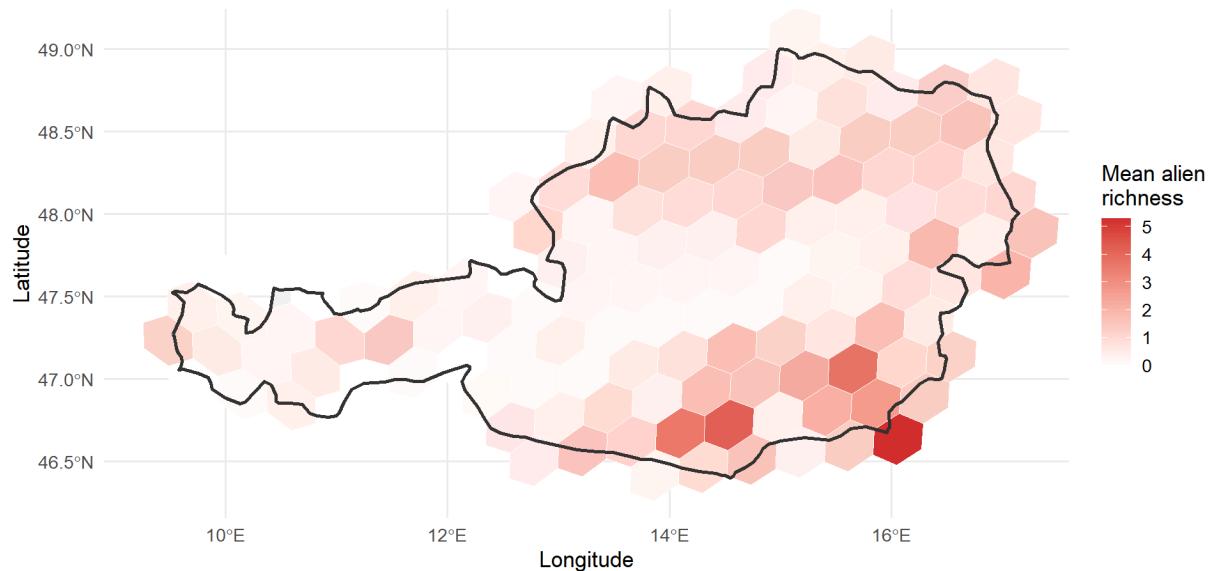
Die folgenden Karten zeigen die **reale räumliche Verteilung** des heimischen und Alien-Artenreichtums über Österreich. Im Gegensatz zu den hypothetischen Karten in der Tag-2-Theorie werden diese direkt aus den EVA-Daten berechnet: Für jeden Plot zählen wir die Anzahl der erfassten heimischen und Alien-Arten.

Um Muster auf Landschaftsebene zu visualisieren, aggregieren wir Plot-Daten in ein **flächentreues Hexagonalraster** mit dem `hexify` (<https://github.com/gcol33/hexify>)-Paket. Jede Hexagonzelle deckt dieselbe Fläche ab, sodass die Farben direkt über die Karte vergleichbar sind.



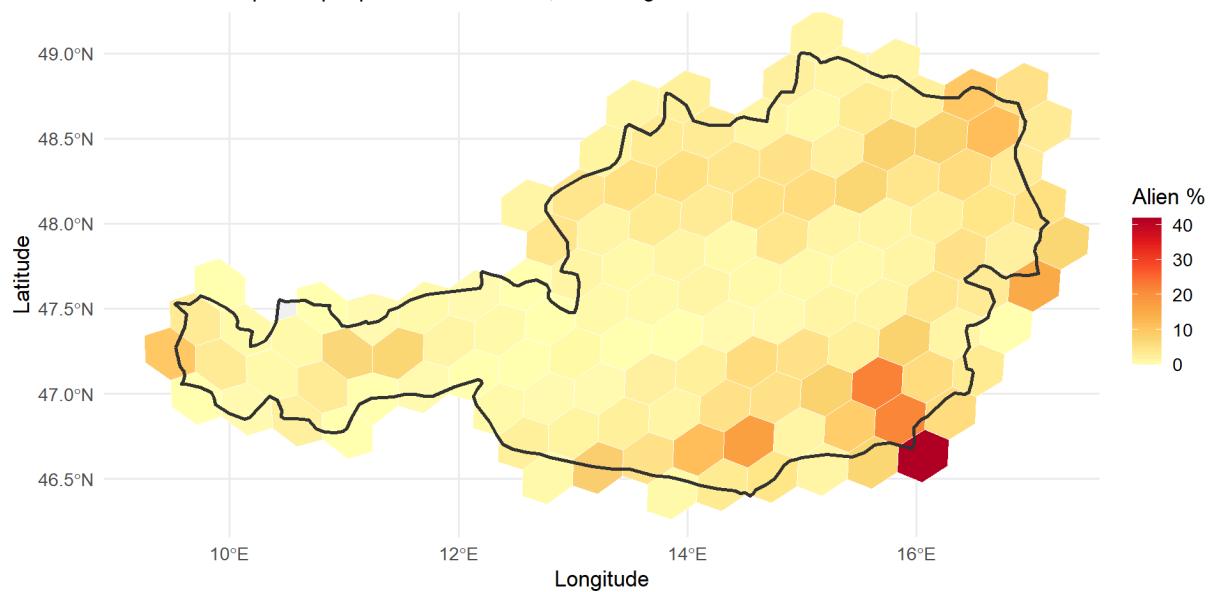
Alien Species Richness Hotspots / Alien-Arten Hotspots

Mean alien species per plot, equal-area hexagonal grid



Alien Species Proportion / Alien-Artenanteil

Mean % alien species per plot — Yellow = low, Red = high alien share



Hotspot-Interpretation:

Diese drei Karten erzählen zusammen eine kraftvolle Geschichte:

1. **Heimischer Reichtum** ist am höchsten in Gebieten mit hoher Habitat-Heterogenität — besonders in den Voralpen, wo mehrere Vegetationszonen überlappen. Dies stimmt mit dem Höhendifferenzpeak überein, der in europäischen Gebirgssystemen dokumentiert ist (Rahbek 1995).
2. **Alien-Reichtum** konzentriert sich in den Tieflagen, besonders entlang des **Donaukorridors**, um Großstädte (Wien, Linz, Graz) und im Pannonischen Becken in Ostösterreich. Dies sind die wärmsten, urbanisiertesten und landwirtschaftlich intensivsten Regionen — genau die Gebiete, wo der Propagule Pressure von Alien-Arten am höchsten ist (Pysek et al. 2010).

3. **Alien-Anteil** zeigt den Kontrast am deutlichsten: Die östlichen Tieflagen und das Donautal zeigen die höchsten Alien-Prozente, während der alpine Kern fast ausschließlich heimisch bleibt. Dieser Ost-West-Gradient spiegelt das kontinentweite Muster wider, wo die pannonische biogeografische Region zu den am stärksten invadierten in Europa gehört (Chytry et al. 2009).

Die hypothetischen Karten von Tag 2 sagten dieses Muster voraus — und die realen Daten bestätigen es. Alien-Reichtum konzentriert sich in menschlich modifizierten Landschaften (Städte, Transportkorridore), während Einheimische in natürlichen, ungestörten Habitaten dominieren. Wichtig: Obwohl Aliens in den gleichen *Habitattypen* vorkommen, zeigt die Akkumulationsanalyse (unten), dass verschiedene Regionen *verschiedene* Alien-Arten beherbergen — hoher räumlicher Turnover, kein einfaches Clustering derselben Arten überall.

The Pannonian invasion gradient: Chytry et al. (2009) documented that the Pannonian biogeographic region has among the highest alien plant richness in Europe, driven by warm temperatures, fertile soils, intensive agriculture, and proximity to major trade routes since Roman times. Our Austrian data provides a within-country confirmation of this continental pattern.

Der pannonische Invasionsgradient: Chytry et al. (2009) dokumentierten, dass die pannonische biogeografische Region zu den höchsten Alien-Pflanzenreichtümern in Europa gehört, angetrieben durch warme Temperaturen, fruchtbare Böden, intensive Landwirtschaft und Nähe zu Handelsrouten seit der Römerzeit.

2. Zeitliche Dynamik

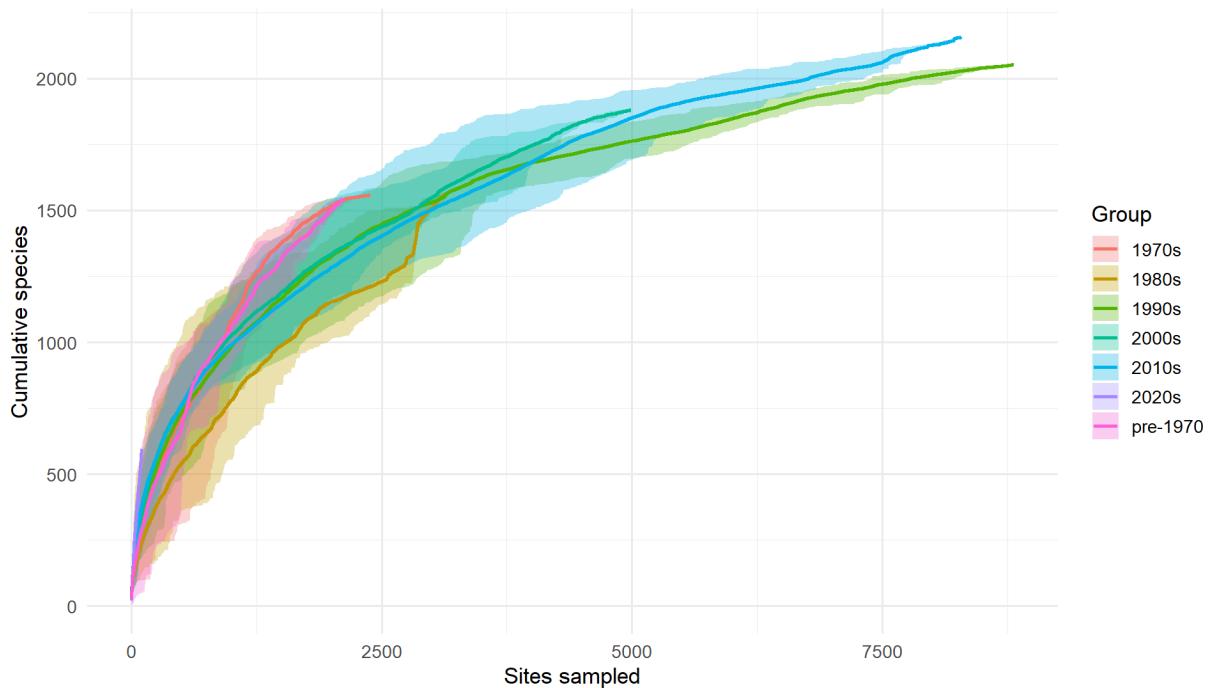
Hat sich die Diversität über Jahrzehnte verändert?

Hypothese: Artenakkumulationsmuster haben sich über die Zeit verändert aufgrund von:

- **Intensivierung der Landnutzung**, die die heimische Pflanzendiversität reduziert (Sala et al. 2000)
- **Globalisierung**, die die Einführung von Alien-Arten erhöht (Seebens et al. 2017)
- **Klimawandel**, der Artverbreitungsgebiete verschiebt (Lenoir et al. 2008)

Vorhersage: Neuere Jahrzehnte sollten (1) höheren Alien-Reichtum und (2) möglicherweise geringeren heimischen Reichtum zeigen, was den in ganz Europa dokumentierten Trend der "biotischen Homogenisierung" widerspiegelt (Winter et al. 2009).

Spatial Accumulation by Decade / Räumliche Akkumulation nach Jahrzehnt

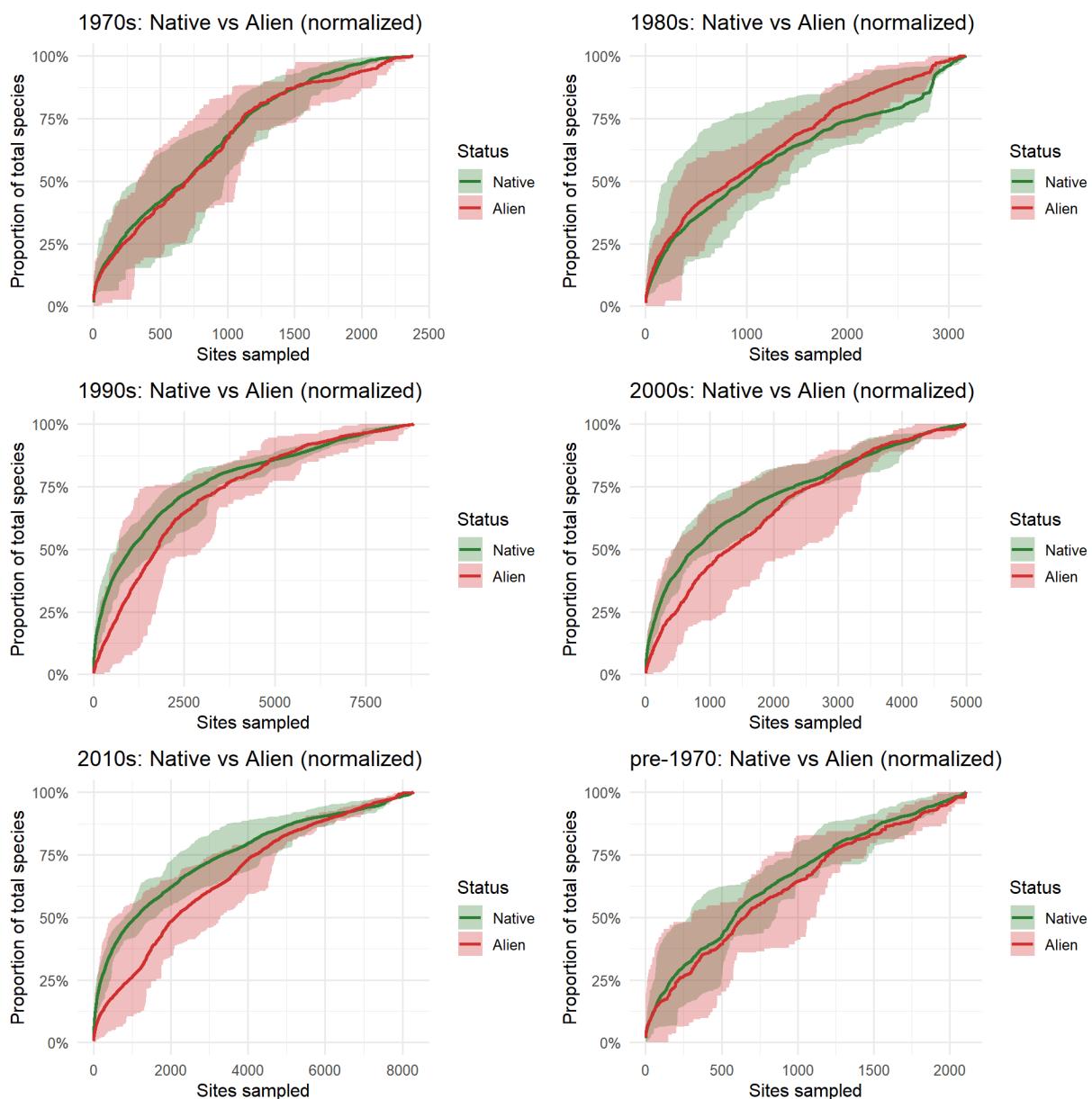


Interpretation:

Die Akkumulationskurven auf Jahrzehntsebene zeigen mehrere Muster:

- **Ältere Jahrzehnte** (vor 1970, 1970er) zeigen steilere anfängliche Steigungen, was darauf hindeutet, dass historisch untersuchte Plots tendenziell auf Biodiversitäts-Hotspots abzielten — eine bekannte Verzerrung in frühen Vegetationserfassungen (Chytry et al. 2016).
- **Neuere Jahrzehnte** (1990er, 2000er) haben mehr Plots und glattere Kurven, die zuverlässigere Schätzungen liefern.
- Der **asymptotische Reichtum** variiert teilweise aufgrund von Unterschieden in der Beprobungsintensität, nicht nur biologischer Veränderung. Diese Unterscheidung ist kritisch: Scheinbare zeitliche Trends im Artenreichtum können Artefakte ungleichmäßiger Beprobung sein (Gotelli & Colwell 2001).

Heimisch vs. Alien im Zeitverlauf



Schlüsselergebnis: Über alle Jahrzehnte hinweg zeigen heimische und Alien-Arten konsistent unterschiedliche Akkumulationsformen. Bei Normalisierung auf ihre jeweiligen Gesamtzahlen **akkumulieren Alien-Arten langsamer** als Einheimische — ihre Kurve liegt in jedem Stadium unter der heimischen Kurve. Das bedeutet, dass Alien-Arten einen **höheren räumlichen Turnover** aufweisen: Verschiedene Alien-Arten besetzen verschiedene Teile der Landschaft, sodass mehr Plots beprobt werden müssen, um denselben Anteil der Alien-Flora zu erfassen.

Dies ist konsistent mit dem **Multi-Fokus-Einführungsmodell**: Im Gegensatz zu Einheimischen, die viele weit verbreitete Generalisten umfassen, die über Regionen hinweg geteilt werden, kommen Aliens an **verschiedenen Einführungspunkten** an (verschiedene Städte, verschiedene Täler, verschiedene Agrargebiete), und jeder Standort erhält eine teils unterschiedliche Artengruppe. Das Ergebnis ist, dass keine einzelne Region die gesamte Alien-Flora erfasst — man muss das gesamte Land durchqueren, um sie alle zu finden. Diese räumliche Heterogenität in der Alien-Zusammensetzung ist in der europäischen Invasionsökologie gut dokumentiert (Pysek et al. 2010; Chytry et al. 2008).

3. Heimisch vs. Alien: Die zentrale Hypothese

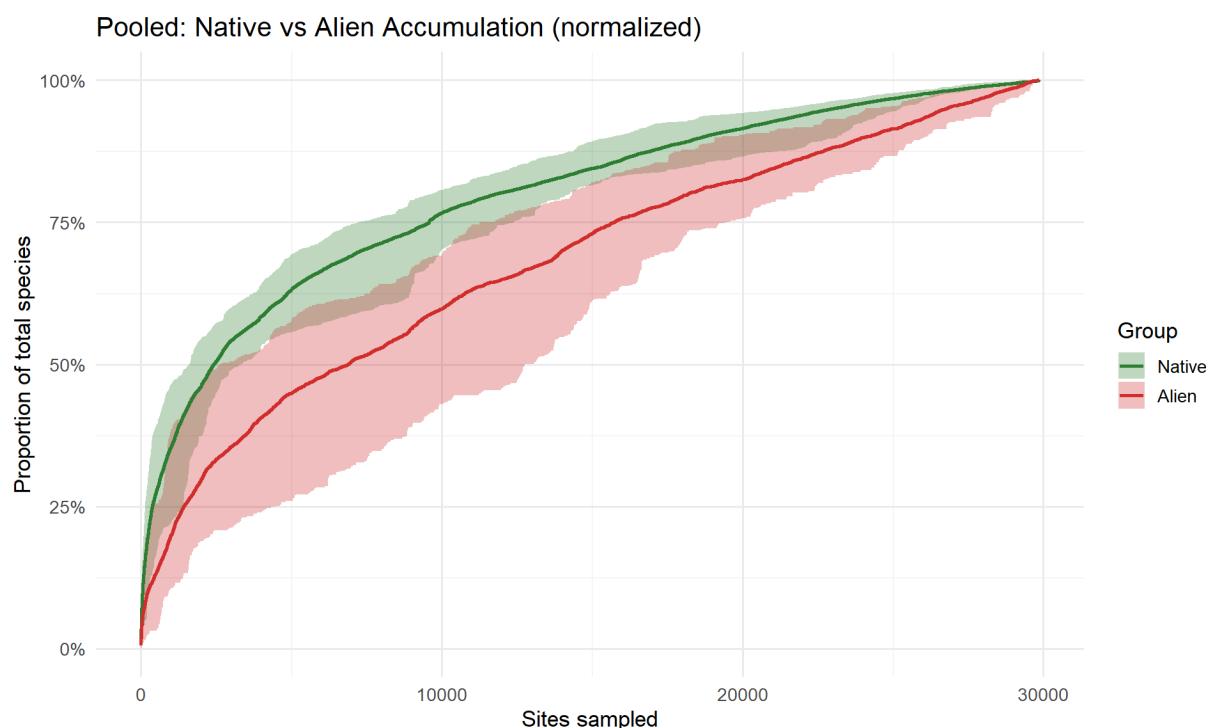
Gesamtakkumulationskurven

Zentrale Forschungsfrage: Akkumulieren Alien-Arten anders als heimische Arten über die österreichische Landschaft?

Hypothesen (von Tag 2):

- **H1:** Alien-Arten weisen eine **höhere räumliche Beta-Diversität** auf als Einheimische — ihre normalisierte Akkumulationskurve liegt unter der heimischen Kurve, d.h. ein größerer Anteil der Landschaft muss beprobt werden, um denselben Anteil der Alien-Flora zu erfassen
- **H2:** Dieser Unterschied spiegelt **habitatspezifische Einführungswege** wider — Aliens treffen an verschiedenen Punkten ein (Städte, Täler, Agrarzonen) und jeder Standort erhält eine teils andere Artengruppe, was räumliche Heterogenität in der Alien-Zusammensetzung erzeugt
- **H3:** Der **statistische Vergleich** wird bestätigen, dass diese Akkumulationskurven signifikant unterschiedlich sind

Diese Hypothesen basieren auf der Invasionsökologie-Theorie. Alien-Arten in gemäßigtem Europa konzentrieren sich tendenziell in anthropogen gestörten Habitaten — urbane Gebiete, Agrarland, Verkehrsinfrastruktur — anstatt sich gleichmäßig über alle Habitattypen zu verteilen (Pysek et al. 2010; Chytry et al. 2008).



Ergebnis — H1 bestätigt: Bei Normalisierung liegt die Alien-Akkumulationskurve in allen Stadien **unter** der heimischen Kurve — Alien-Arten akkumulieren proportional *langsamer*. Um 50% der heimischen Arten zu erfassen, benötigt man X% der Landschaft, aber um 50% der Alien-Arten zu erfassen, braucht man deutlich mehr. Dies bestätigt, dass Aliens eine **höhere räumliche Beta-Diversität** aufweisen: Verschiedene Alien-Arten kommen in verschiedenen Regionen vor, sodass mehr Landschaft durchquert werden muss, um denselben Anteil der Alien-Flora zu erfassen. Viele heimische Arten hingegen sind weit verbreitete Generalisten. (Zum Vergleich: Der Datensatz enthält ~2378 heimische vs ~323 Alien-Arten.)

Ergebnis — H2 bestätigt: Die räumliche Heterogenität in der Alien-Zusammensetzung ist konsistent mit dem **Multi-Fokus-Einführungsmodell**. Anders als Einheimische treffen Aliens an verschiedenen Einführungspunkten ein — verschiedene Städte, verschiedene Täler, verschiedene Agrarzonen — und jeder Standort erhält eine teils andere Artengruppe. Keine einzelne Region erfasst die gesamte Alien-Flora.

Statistischer Vergleich

Was ist ein Permutationstest?

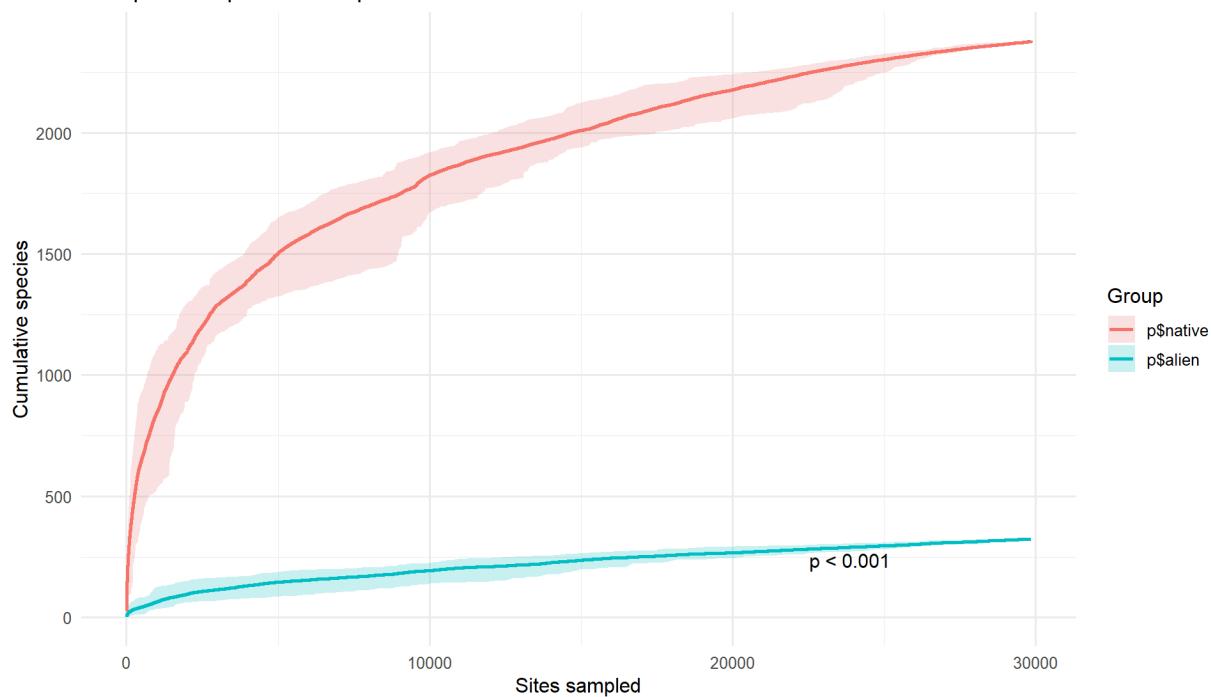
Ein Permutationstest ist eine nicht-parametrische statistische Methode, die eine **Nullverteilung** durch Durchmischen von Labels erstellt. Die Logik ist:

1. **Beobachtete Statistik:** Berechne die Differenz zwischen den beiden Kurven (hier: AUC-Differenz = Fläche unter der Kurve heimisch – alien)
2. **Nullhypothese:** "Die Kurven stammen aus derselben zugrundeliegenden Verteilung" — wenn wahr, ist die Zugehörigkeit einer Replikation zu einer Kurve willkürlich
3. **Nullverteilung erstellen:** Zufällige Neuzuweisung der Kurven-Labels viele Male (999 Permutationen), jedes Mal die Statistik neu berechnen
4. **p-Wert:** Der Anteil der permutierten Differenzen, die so extrem (oder extremer) sind wie die beobachtete Differenz

Wenn die beobachtete Statistik weit in die Ränder der Nullverteilung fällt, verwerfen wir die Nullhypothese — die Kurven sind wirklich unterschiedlich.

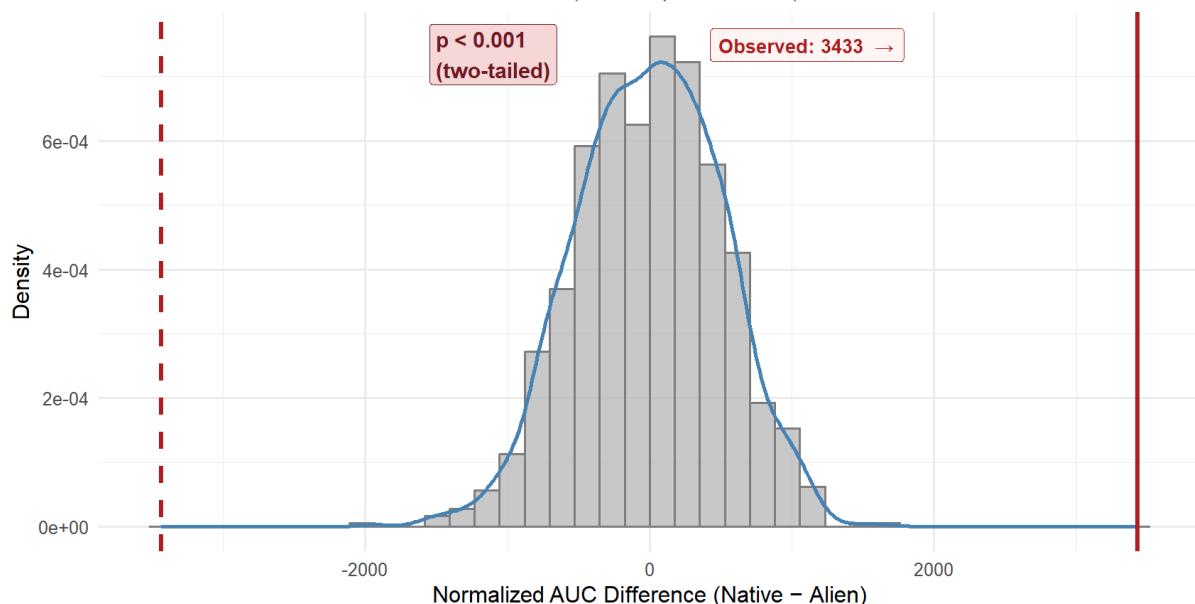
```
## Comparison: p$native vs p$alien
## -----
## Method: permutation (n=999)
## Normalized: yes (shape comparison)
## AUC difference: 3433.3 (p < 0.001***)
## Saturation: p$native at 18661 sites, p$alien at 24060 sites
##
## p$native saturates faster.
```

Comparison: p\$native vs p\$alien



Permutation Test: Are Native and Alien Curve Shapes Different?

Null distribution of normalized AUC differences (n = 999 permutations)



Ergebnis — H3: Der Permutationstest bestätigt, dass heimische und Alien-Akkumulationskurven **statistisch signifikant unterschiedlich** sind ($p < 0,05$). Dies ist nicht nur ein Unterschied in der Gesamtartenzahl — das *räumliche Muster* der Akkumulation unterscheidet sich.

Das Permutationsdiagramm lesen:

- **Graues Histogramm/blaue Kurve:** Die Nullverteilung — welche Unterschiede wir erwarten würden, wenn heimisch/alien Labels bedeutungslos wären
- **Rote durchgezogene Linie:** Unsere beobachtete Differenz (heimische AUC – Alien AUC)
- **Rote gestrichelte Linie:** Der Spiegel unserer Beobachtung (für zweiseitigen Test)

Die beobachtete Statistik liegt **weit außerhalb** der Nullverteilung — keine einzige Permutation ergab eine so große Differenz. Das sagt uns: Die Wahrscheinlichkeit, einen so großen Unterschied allein durch Zufall zu sehen, ist im Wesentlichen null.

Was bedeutet das biologisch?

Der signifikante Unterschied sagt uns, dass wir, wenn wir die österreichische Landschaft durchqueren und Plots in räumlicher Reihenfolge beprobten, heimische und Alien-Arten mit grundlegend verschiedenen Raten antreffen. Dieses Ergebnis stimmt überein mit:

1. **Habitatfilterung** — Alien-Arten werden in spezifische Habitate gefiltert (Chytry et al. 2008)
2. **Ausbreitungslimitierung** — Viele Aliens haben ihr potenzielles Verbreitungsgebiet noch nicht ausgefüllt (Wilson et al. 2007)
3. **Einführungsgeschichte** — Aliens hatten weniger Zeit, sich von Einführungspunkten auszubreiten (Essl et al. 2011)

4. Extrapolation: Wie viele Arten fehlen uns?

Das Artenschätzungsproblem: Selbst mit 52.000 Plots haben wir nicht jede Ecke Österreichs beprobt. **Asymptotische Schätzer** passen mathematische Modelle an die Akkumulationskurve an und projizieren, wo sie vollständig abflachen würde (Colwell & Coddington 1994).

Das **Michaelis-Menten-Modell** ist ein klassischer asymptotischer Schätzer, der aus der Enzymkinetik entlehnt, aber in der Ökologie weit verbreitet ist (Colwell & Coddington 1994):

$$S(n) = \frac{S_{max} \cdot n}{K_m + n}$$

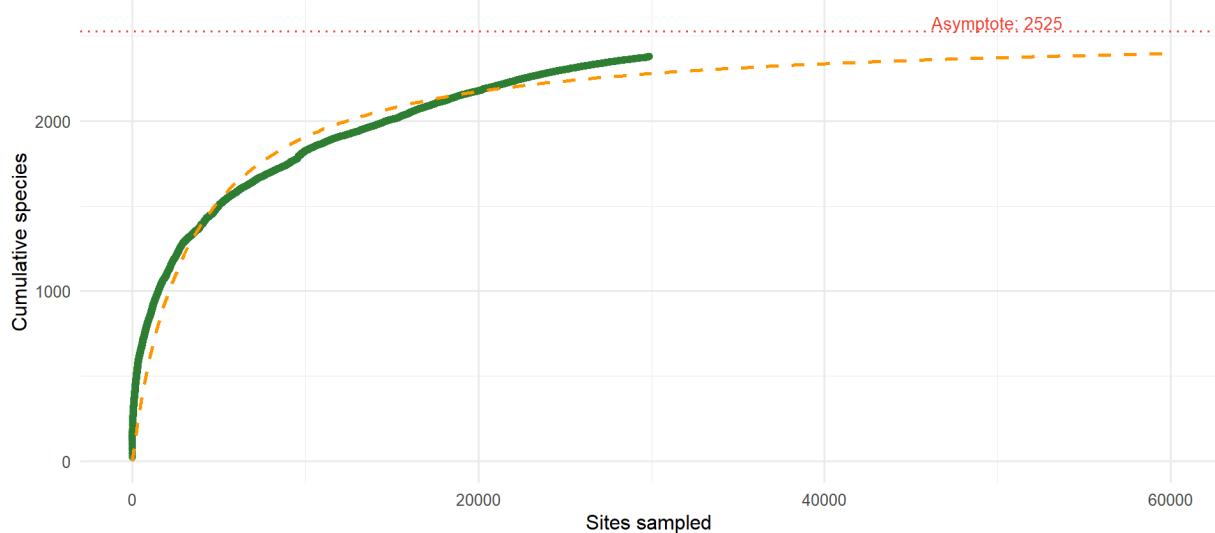
wobei S_{max} der geschätzte Gesamtartenreichtum (die Asymptote) ist, n die Anzahl der beprobten Plots, und K_m die Halbsättigungskonstante (die Anzahl der Plots, die benötigt wird, um die Hälfte aller Arten zu entdecken).

```
## Native species extrapolation:
```

```
## Extrapolation: michaelis-menten
## -----
## Estimated asymptote: 2525.4 species
## 95% CI: 2522.4 - 2528.5
## AIC: 354656.5
## Observed: 2378.0 species (94% of estimated)
```

Native Species: Observed vs Extrapolated

Model: michaelis-menten, AIC: 354656.5

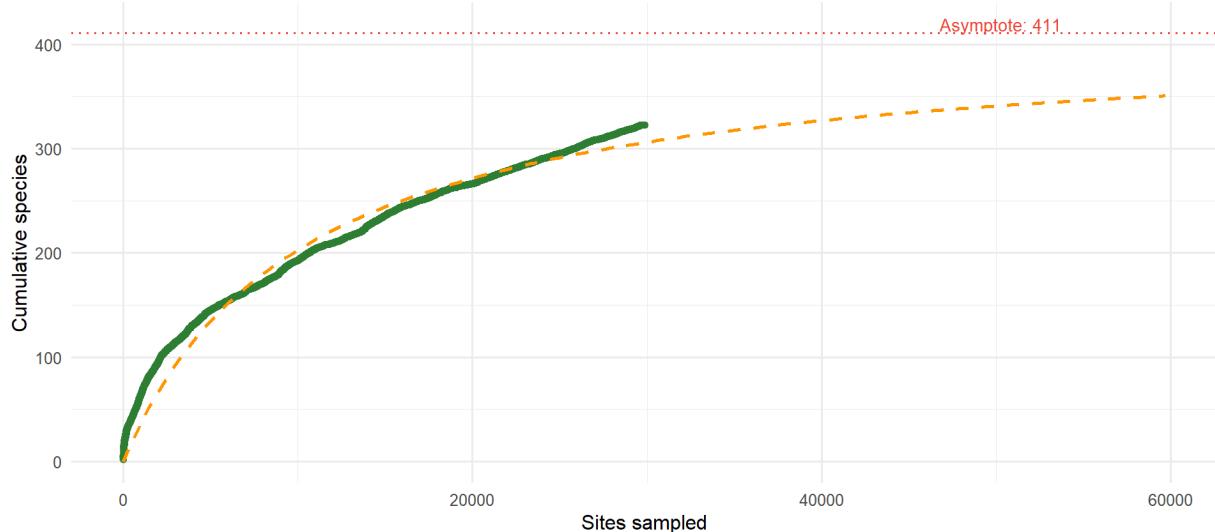


```
##  
## Alien species extrapolation:
```

```
## Extrapolation: michaelis-menten  
## -----  
## Estimated asymptote: 410.9 species  
## 95% CI: 410.0 - 411.8  
## AIC: 233579.6  
## Observed: 323.0 species (79% of estimated)
```

Alien Species: Observed vs Extrapolated

Model: michaelis-menten, AIC: 233579.6



Interpretation:

Das Michaelis-Menten-Modell schätzt den **Gesamtartenreichtum** für jede Gruppe, einschließlich noch nicht durch unsere Beprobung erfasster Arten. Die Lücke zwischen dem beobachteten Endpunkt und der extrapolierten Asymptote repräsentiert unsere "dunkle Diversität" — Arten, die in der Landschaft vorhanden, aber noch nicht vom Plot-Netzwerk erfasst sind (Partel et al. 2011).

Für **heimische Arten** schätzt das Modell ~2.525 Gesamtarten, was bedeutet, dass unsere 2.378 beobachteten Einheimischen **~94%** der geschätzten Flora repräsentieren — hohe Beprobungsvollständigkeit. Die restlichen ~6% umfassen wahrscheinlich seltene Arten, die auf unterbeprobte Alpengipfel, Moore oder Felsspalten beschränkt sind. Für **Alien-Arten** schätzt das Modell ~411 Gesamtarten im Vergleich zu 323 beobachteten (~**79%** Vollständigkeit). Die größere Lücke könnte mehrere nicht-exklusive Faktoren widerspiegeln: (1) seltene Aliens in unterbeprobenen Habitattypen, (2) kürzlich angekommene Arten, die noch nicht weit genug verbreitet sind, um im Plot-Netzwerk aufzutreten, oder (3) echte "Invasionsschuld" — bereits vorhandene aber sich noch ausbreitende Arten (Essl et al. 2011). Das asymptotische Modell allein kann zwischen diesen Erklärungen nicht unterscheiden; temporale Daten wären nötig, um die Invasionsschuld-Hypothese direkt zu testen.

Invasion debt (Essl et al. 2011): Many alien species have been introduced to a region but have not yet reached their potential distribution. Current alien richness underestimates future alien richness. This has major implications for conservation planning — even if all new introductions stopped today, the alien flora would continue to expand for decades.

Invasionsschuld (Essl et al. 2011): Viele Alien-Arten wurden in eine Region eingeführt, haben aber ihr potenzielles Verbreitungsgebiet noch nicht erreicht. Der aktuelle Alien-Reichtum unterschätzt den zukünftigen. Dies hat große Implikationen für die Naturschutzplanung.

5. Räumlich vs. Zufällig: Ist Geografie wichtig?

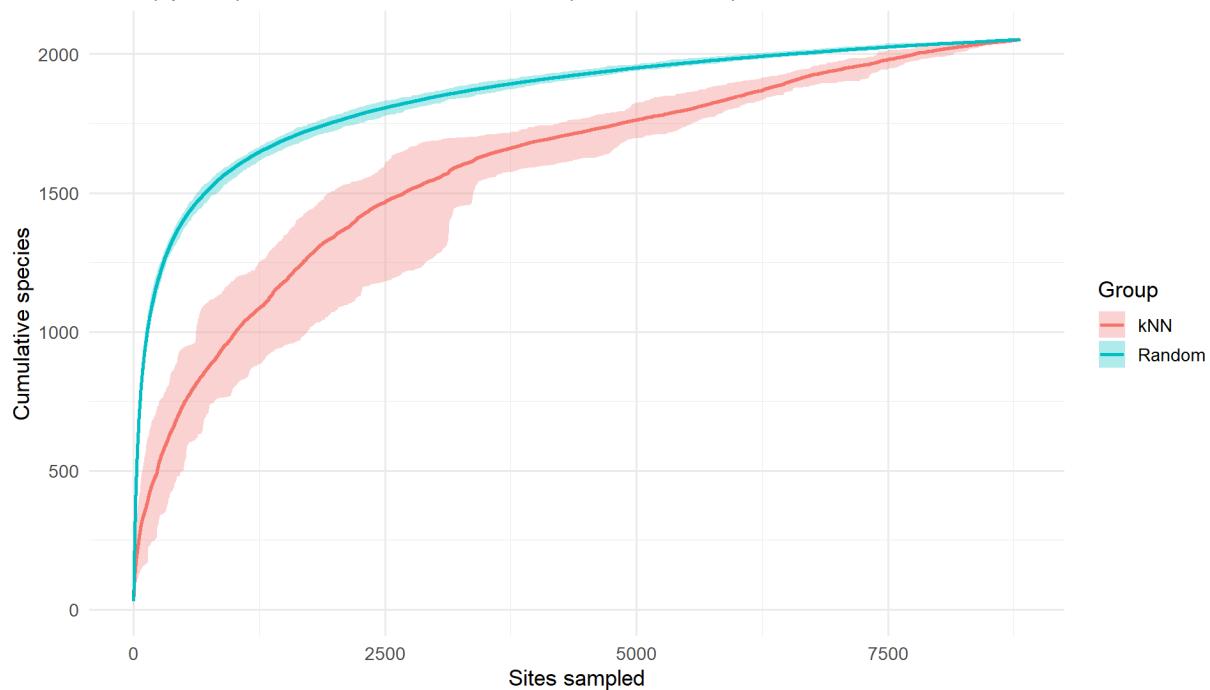
Der Nullmodell-Ansatz: Um zu testen, ob räumliche Struktur die Artenakkumulation tatsächlich beeinflusst, vergleichen wir zwei Methoden:

- **kNN (räumlich):** Plots in Nearest-Neighbor-Reihenfolge besuchen — ahmt nach, wie ein Feldökologe eine Landschaft durchquert
- **Zufällig:** Plots in zufälliger Reihenfolge besuchen — das klassische Nullmodell für Akkumulationskurven (Gotelli & Colwell 2001)

Wenn Arten **räumlich autokorriert** sind (Toblers erstes Gesetz der Geografie: "alles hängt mit allem zusammen, aber nahe Dinge mehr als entfernte"), sollte die räumliche Kurve *unter* der zufälligen liegen. Benachbarte Plots teilen mehr Arten, sodass der kNN-Walk neue Arten langsamer antrifft.

Die Differenz zwischen den beiden Kurven quantifiziert **räumlichen Turnover** — wie stark sich die Artenzusammensetzung über geografische Distanz verändert (Nekola & White 1999).

kNN (spatial) vs Random Accumulation (1990s subset)



Schlüsselergebnis: Die räumliche (kNN) Kurve liegt konsistent unter der zufälligen Kurve.

Das bedeutet:

1. **Räumliche Autokorrelation ist stark** — benachbarte Plots teilen viele Arten
2. **Beta-Diversität ist real** — die Artenzusammensetzung verändert sich tatsächlich über die Landschaft
3. **Zufälliges Sampling überschätzt lokale Diversität** — wenn man zufällig herumspringt, trifft man schneller neue Arten, weil man häufiger biogeografische Grenzen überquert

Die Lücke zwischen den Kurven ist ein visuelles Maß für **räumliche Beta-Diversität**. Eine große Lücke zeigt hohen Arten-Turnover über die Landschaft; eine kleine Lücke würde bedeuten, dass Arten gleichmäßig verteilt sind.

Dieses Ergebnis validiert die Verwendung räumlicher (kNN) Akkumulationskurven für österreichische Vegetationsdaten: Die räumliche Struktur ist ökologisch bedeutsam, nicht nur ein statistisches Artefakt.

6. Hill-Zahlen: Jenseits des Artenreichtums

Warum Artenreichtum nicht ausreicht:

Artenreichtum (Arten zählen) behandelt alle Arten gleich — eine Art mit 10.000 Individuen zählt genauso wie eine mit einem einzigen Individuum. **Hill-Zahlen** (Hill 1973; Jost 2006) bieten einen einheitlichen Rahmen von Diversitätsmaßen, parametrisiert durch einen einzigen Wert q :

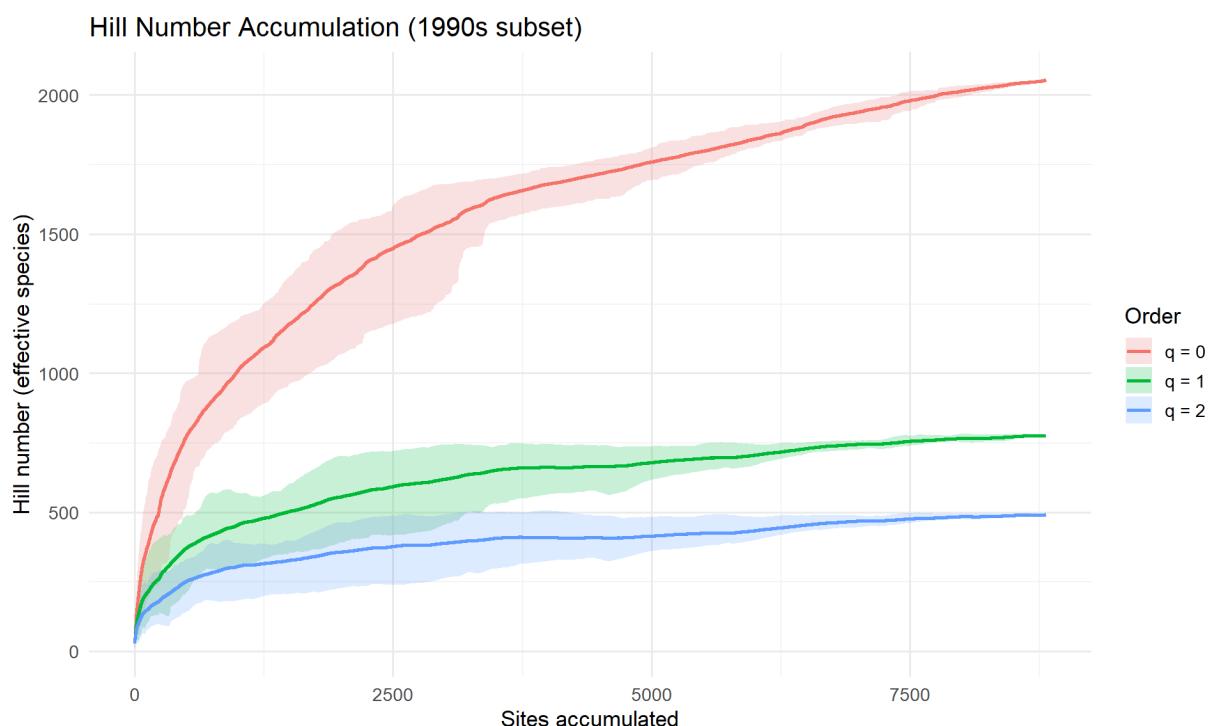
$$^q D = \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)}$$

q			
q = 0	Artenreichtum	Anzahl der Arten	Alle Arten gleich
q = 1	Shannon-Diversität	Effektive Anzahl häufiger Arten	Proportional zur Abundanz
q = 2	Simpson-Diversität	Effektive Anzahl dominanter Arten	Betont Dominante

Mit steigendem q gibt das Maß seltenen Arten weniger und dominanten mehr Gewicht.

Dies ist kritisch, weil **invasive Alien-Arten oft Dominanz erreichen** — sie mögen wenige sein, können aber unverhältnismäßig zu Biomasse und ökologischer Wirkung beitragen (Pysek et al. 2012).

```
## spacc Hill numbers: 8814 sites, 2700 species, 30 seeds
## Orders (q): 0, 1, 2
## Method: knn
```



Interpretation:

Die drei Hill-Zahlen-Kurven erzählen komplementäre Geschichten:

- **$q = 0$ (Reichtum):** Die klassische Akkumulationskurve — die Gesamtartenzahl steigt mit der Beprobung. Das haben wir an Tag 1 und 2 analysiert.
- **$q = 1$ (Shannon):** Die effektive Anzahl "typischer" Arten. Diese Kurve steigt langsamer, weil sie seltene Arten heruntergewichtet. Wenn viele Arten sehr selten sind, wird die Shannon-Diversität viel niedriger als der rohe Reichtum sein.
- **$q = 2$ (Simpson):** Die effektive Anzahl dominanter Arten. Diese Kurve erreicht am frühesten ein Plateau, weil die dominanten Arten der Gemeinschaft schnell erfasst werden — man braucht relativ wenige Plots, um die Arten zu finden, die den Großteil der Vegetationsdeckung ausmachen.

Die Lücke zwischen $q=0$ und $q=2$ ist selbst informativ: Eine große Lücke zeigt hohe **Evenness** — viele Arten, keine stark dominant. Eine kleine Lücke deutet darauf hin, dass wenige Arten dominieren, während viele selten sind. Diese "Dominanzstruktur" hat Implikationen für die Ökosystemfunktion (Hillebrand et al. 2008).

Diversitätsaufteilung: Alpha, Beta, Gamma

Whittakers Erkenntnis (1960, 1972): Regionale Diversität (γ) kann in lokale Diversität (α) und Zwischen-Standort-Diversität (β) zerlegt werden:

$$\gamma = \alpha \times \beta \quad (\text{multiplikativ})$$

oder äquivalent: $\beta = \gamma/\alpha$

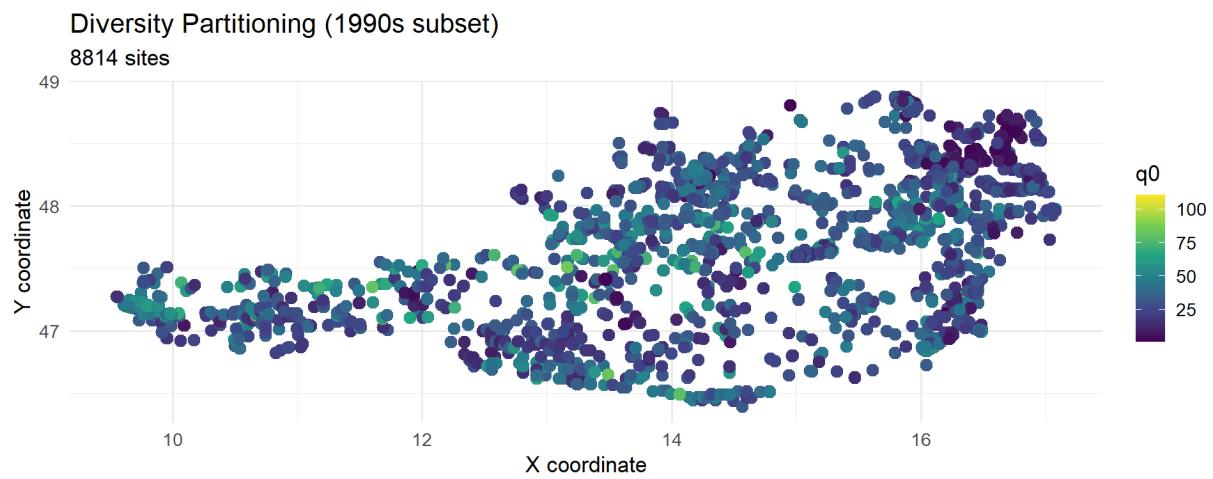
- **Alpha (α):** Durchschnittliche Diversität innerhalb eines einzelnen Plots — wie viele Arten hält ein Standort typischerweise?
- **Gamma (γ):** Gesamtdiversität über die gesamte Region — der vollständige Artenpool
- **Beta (β):** Wie unterschiedlich sind Standorte voneinander? Hohes Beta = hoher Turnover zwischen Standorten

Diese Zerlegung zeigt, ob Diversität durch **viele Arten an einzelnen Standorten** (hohes Alpha) oder durch **verschiedene Arten an verschiedenen Standorten** (hohes Beta) angetrieben wird. Die Antwort hat direkte Naturschutzimplikationen: Systeme mit hohem Alpha brauchen weniger Schutzgebiete; Systeme mit hohem Beta brauchen ein verteiltes Netzwerk von Reservaten (Socolar et al. 2016).

```

## Alpha-Beta-Gamma Diversity Partitioning
## 8814 sites, 2700 species
##
## q alpha beta gamma
## 0 30.54 67.22 2053.00
## 1 26.54 29.17 774.29
## 2 21.09 23.27 490.89
##
## Interpretation:
##   Alpha = mean effective species per site
##   Beta  = effective number of communities (1 to n_sites)
##   Gamma = regional effective species (gamma = alpha x beta)

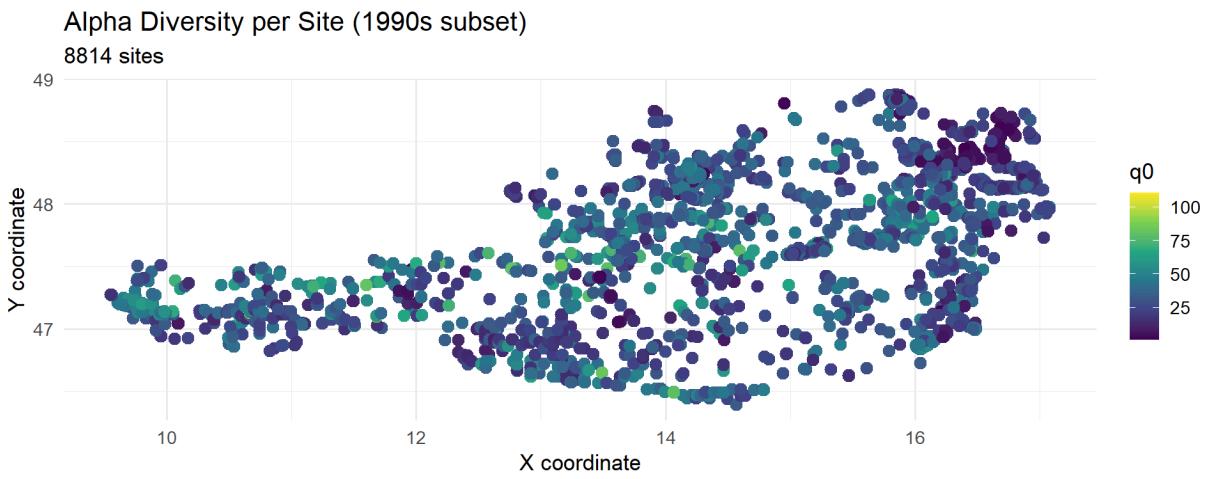
```



```

## spacc alpha diversity: 8814 sites, 2700 species
## Orders (q): 0, 1, 2
## Coordinates: available

```



Interpretation:

Die Aufteilung zeigt, dass die österreichische Pflanzendiversität primär durch **hohe Beta-Diversität** angetrieben wird — Standorte unterscheiden sich erheblich in der Artenzusammensetzung. Dies ist angesichts der topografischen Heterogenität Österreichs zu erwarten: Ein einzelner Plot im pannonischen Tiefland teilt relativ wenige Arten mit einem Plot in den Hochalpen. Die Alpha-Diversitätskarte zeigt räumliche Muster im lokalen Reichtum, die wahrscheinlich Habitattyp, Höhenlage und Landnutzungsintensität widerspiegeln.

Für den Naturschutz bedeutet diese hohe Beta-Diversität, dass der Schutz eines einzelnen Biodiversitäts-Hotspots unzureichend ist — ein **Netzwerk von Schutzgebieten**, das verschiedene Habitate und Höhenlagen umfasst, ist notwendig, um die gesamte österreichische Flora zu erfassen (Kier et al. 2009).

7. Beta-Diversität: Turnover vs. Nestedness

Nicht alle Beta-Diversität ist gleich. Baselga (2010) zeigte, dass die Gesamt-Beta-Diversität in zwei grundlegend verschiedene Komponenten zerlegt werden kann:

- **Turnover (β_{sim})**: Arten werden über Standorte hinweg durch andere Arten *ersetzt*.

Standort A hat Arten 1,2,3; Standort B hat Arten 4,5,6. Dies zeigt echte ökologische Gradienten — verschiedene Umweltbedingungen unterstützen verschiedene

Artenzusammensetzungen.

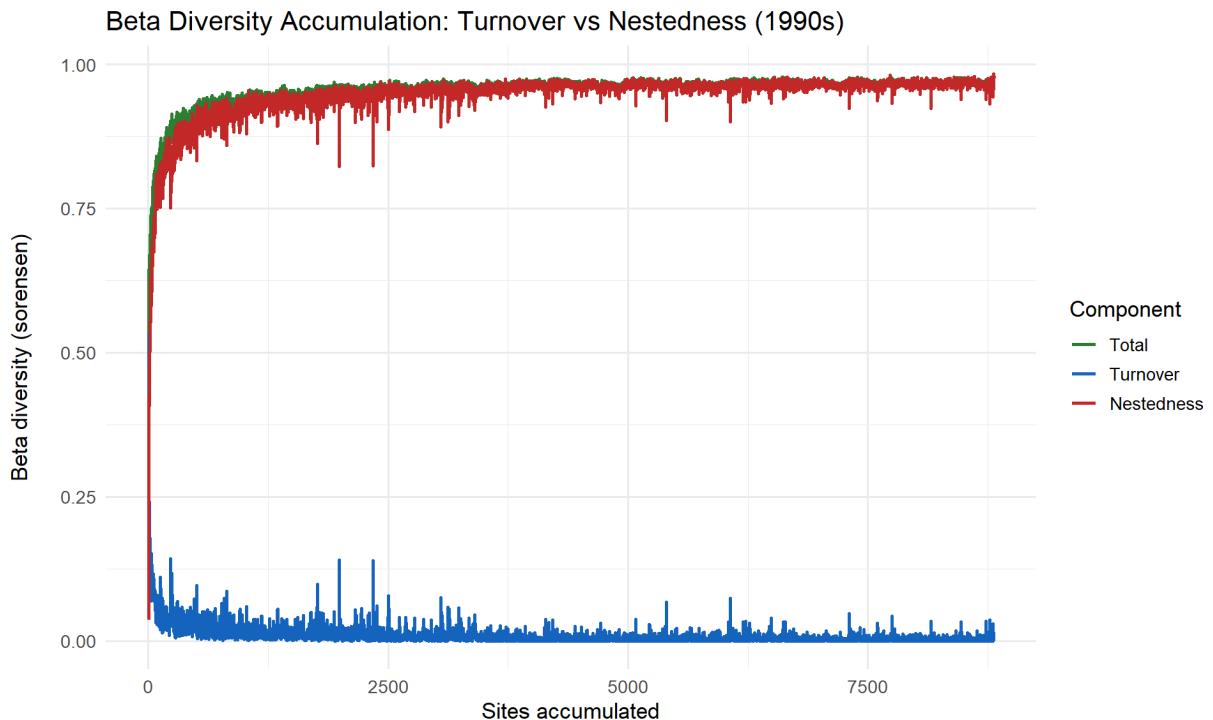
- **Nestedness (β_{nes})**: Artenarme Standorte sind *Teilmengen* artenreicher Standorte.

Standort A hat Arten 1,2,3,4,5; Standort B hat Arten 1,2. Dies zeigt differenziellen Artenverlust — einige Standorte haben Arten verloren (oder nie erworben), anstatt andere Arten zu beherbergen.

$$\beta_{Sorensen} = \beta_{Turnover} + \beta_{Nestedness}$$

Warum das für Invasionen wichtig ist: Wenn Alien-Arten hohe Nestedness zeigen (alienarme Standorte sind Teilmengen alien-reicher Standorte), deutet das darauf hin, dass sich Aliens von Einführungs-Hotspots nach außen ausbreiten. Wenn sie hohen Turnover zeigen, etablieren sich verschiedene Alien-Arten in verschiedenen Regionen — möglicherweise verschiedene unabhängige Einführungsereignisse oder habitatspezifische Invasionen widerstreuend.

```
## spacc beta diversity: 8814 sites, 30 seeds
## Index: sorense, Method: knn
## Mean final beta: 0.976 (turnover: 0.000, nestedness: 0.976)
```



Interpretation:

Wenn sich das räumliche Beprobungsfenster erweitert (mehr Plots einbezogen), entwickeln sich die Beta-Diversitäts-Komponenten:

- **Turnover dominiert** auf größeren räumlichen Skalen — die Bewegung von den Alpen zum Tiefland beinhaltet echten Artenersatz, nicht nur Artenverlust. Alpine Spezialisten (z.B. *Dryas octopetala*, *Saxifraga* spp.) werden durch Tieflandarten (z.B. *Arrhenatherum elatius*, *Lolium perenne*) ersetzt, nicht einfach ergänzt.
- **Nestedness** trägt weniger bei, was darauf hindeutet, dass Österreichs Flora keine bloß verschachtelte Hierarchie ist, wo artenarme Standorte Teilmengen reicher Standorte sind. Stattdessen beherbergen verschiedene Regionen tatsächlich verschiedene Artenzusammensetzungen.

Dieses Muster stimmt mit starker **Umweltfilterung** entlang Österreichs Höhengradienten (~115 m bis ~3.798 m) überein, der deutliche Vegetationsgürtel von thermophilem Tieflandgrasland bis zu alpinen Polsterpflanzengemeinschaften schafft (Grabherr et al. 2003).

8. Deckungsbasierte Rarefaction

Die Vollständigkeitsfrage: Statt "wie viele Arten haben wir gefunden?" fragen wir "**welchen Anteil aller Arten haben wir gefunden?**"

Stichprobenabdeckung (Chao & Jost 2012) schätzt den Anteil der Gesamtgemeinschaft, der erfasst wurde. Sie ist definiert als der Anteil der Individuen in der Gemeinschaft, die zu bereits beobachteten Arten gehören:

$$\hat{C}_n = 1 - \frac{f_1}{n}$$

wobei f_1 die Anzahl der Singletons (Arten, die genau einmal beobachtet wurden) und n die Gesamtzahl der beprobten Individuen ist. Wenn die Abdeckung 95% beträgt, haben wir Arten entdeckt, die 95% aller Individuen ausmachen — aber wir verpassen möglicherweise noch viele seltene Arten.

Praktische Frage: Wie viele Plots müssen wir beproben, um 90%, 95% oder 99% Abdeckung zu erreichen?

```
## spacc coverage: 8814 sites, 2700 species, 30 seeds
## Mean final coverage: 99.9%
## Mean final richness: 2053.0 species
```



C90	90%	134.1
C95	95%	192.9
C99	99%	463.1

Praktische Implikationen:

Die Abdeckungsanalyse beantwortet eine fundamentale feldökologische Frage: **Wie viel**

Beprobung ist genug? Die Tabelle oben zeigt, wie viele räumlich geordnete Plots (im Durchschnitt) benötigt werden, um verschiedene Abdeckungsniveaus zu erreichen.

Der Sprung von 95% auf 99% Abdeckung erfordert unverhältnismäßig mehr Aufwand — dies ist eine universelle Eigenschaft der Biodiversitätsbeprobung, die den **langen Schwanz seltener Arten** widerspiegelt (Preston 1948). Viele Arten sind durch sehr wenige Individuen vertreten oder kommen in sehr spezifischen Mikrohabitaten vor; sie zu entdecken erfordert erschöpfende Beprobung, die möglicherweise nicht kosteneffektiv ist.

Für Inventurprojekte und Umweltverträglichkeitsprüfungen gelten 90–95% Abdeckung typischerweise als ausreichend (Chao et al. 2014). Das Erreichen von 99% erfordert ungefähr den doppelten Aufwand von 95%, ein Muster, das über Ökosysteme weltweit konsistent ist.

9. Phylogenetische Diversität

Jenseits der Artenidentität — die evolutionäre Perspektive:

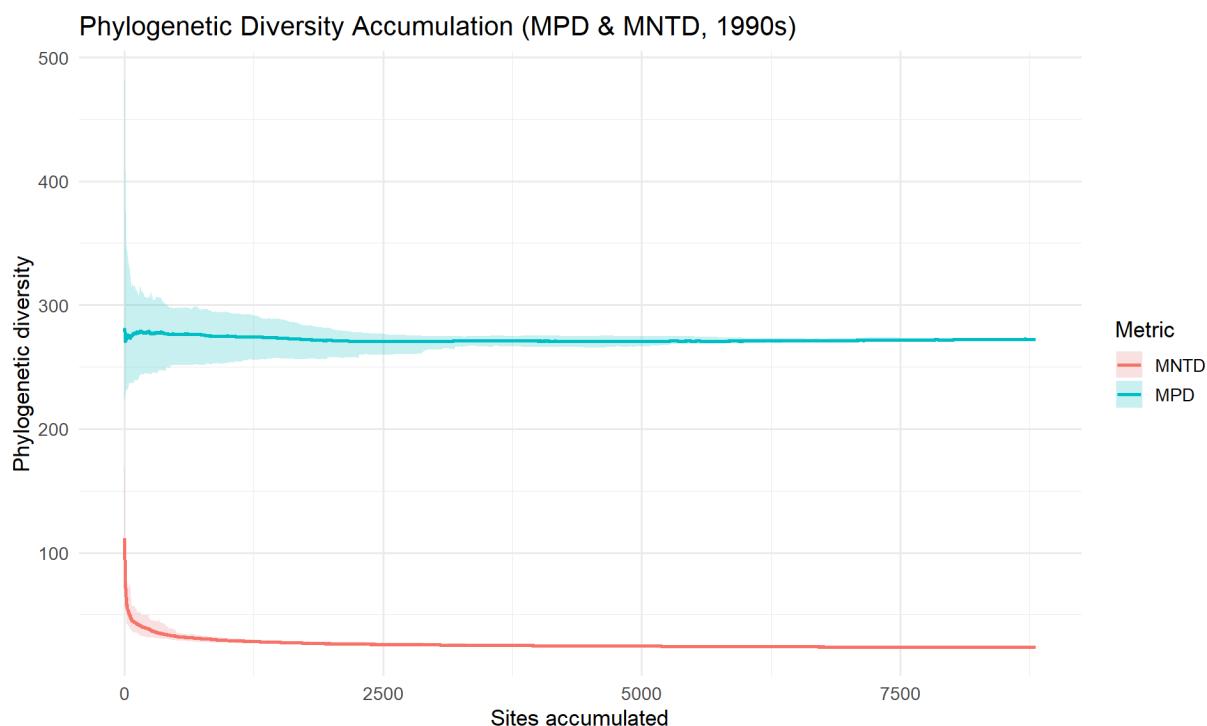
Zwei Gemeinschaften können die gleiche Artenzahl haben, aber sehr verschiedene **Evolutionsgeschichten**. Eine Gemeinschaft aus 10 Grasarten (alle Poaceae) repräsentiert weniger evolutionäre Diversität als eine Gemeinschaft mit einem Gras, einem Farn, einer Orchidee und 7 Arten aus verschiedenen Familien.

Phylogenetische Diversitäts-Metriken quantifizieren, wie viel vom Baum des Lebens in einer Gemeinschaft repräsentiert ist (Faith 1992):

- **MPD (Mean Pairwise Distance):** Durchschnittliche evolutionäre Distanz zwischen allen Artenpaaren. Hoher MPD = Arten sind entfernt verwandt = hohe phylogenetische Breite.
- **MNTD (Mean Nearest Taxon Distance):** Durchschnittliche Distanz zum nächsten Verwandten jeder Art. Niedriger MNTD = viele eng verwandte Arten = mögliche kompetitive Exklusion oder Habitatfilterung für bestimmte Linien.

Hypothese für Invasionen: Alien-Arten könnten aus wenigen dominanten Familien stammen (z.B. Asteraceae, Poaceae, Brassicaceae — Pysek et al. 2017), was zu geringerer phylogenetischer Diversität als die heimische Flora führt. Alternativ könnten Aliens aus verschiedenen geografischen Herkünften eine breite phylogenetische Spannweite repräsentieren.

```
## spacc phylogenetic diversity: 8814 sites, 2697 species, 30 seeds
## Metrics: mntd, mpd
```



Heimisch vs. Alien Phylogenetische Diversität

```
## Native MPD accumulation:
```

```
## spacc phylogenetic diversity: 8814 sites, 2375 species, 30 seeds  
## Metrics: mpd
```

```
##  
## Alien MPD accumulation:
```

```
## spacc phylogenetic diversity: 8814 sites, 322 species, 30 seeds  
## Metrics: mpd
```

Interpretation:

MPD (phylogenetische Breite) steigt, wenn sich das Beprobungsfenster erweitert, weil wir Arten aus zunehmend verschiedenen evolutionären Linien antreffen. Die Anstiegsrate des MPD sagt uns etwas über **phylogenetische Beta-Diversität** — wie stark sich die Evolutionsgeschichte über den Raum verändert.

MNTD (terminale Clusterung) erfasst die feinere phylogenetische Struktur. Wenn MNTD niedrig ist, haben viele Arten in der Probe nahe Verwandte, die ebenfalls anwesend sind — dies kann auf **Umweltfilterung** hindeuten (das Habitat selektiert für Arten aus bestimmten Linien) oder auf **kompetitive Exklusion**, die die Koexistenz sehr ähnlicher Arten begrenzt (Webb et al. 2002).

Vergleich heimisch vs. alien: Der MPD-Vergleich zeigt, ob Alien-Arten einen phylogenetisch breiten oder schmalen Ausschnitt des Pflanzen-Lebensbaums repräsentieren. Wenn Aliens phylogenetisch clustern, deutet das darauf hin, dass bestimmte evolutionäre Eigenschaften (schnelles Wachstum, generalistische Ökologie, hohe Samenproduktion) Linien für erfolgreiche Invasion prädisponieren — das “ideale Unkraut”-Syndrom (Baker 1974; van Kleunen et al. 2010).

10. Funktionelle Diversität

Was Arten TUN ist genauso wichtig wie was Arten SIND:

Funktionelle Diversität misst die Bandbreite ökologischer Rollen (Eigenschaften), die in einer Gemeinschaft vorhanden sind (Petchey & Gaston 2006). Zwei Arten aus der gleichen Familie können sehr verschiedene ökologische Funktionen haben (z.B. ein hoher Baum vs. eine bodennahe Rosette), während zwei entfernt verwandte Arten ähnliche Funktionen erfüllen können (konvergente Evolution).

Wichtige funktionelle Eigenschaften in der Pflanzenökologie:

Spezifische Blattfläche

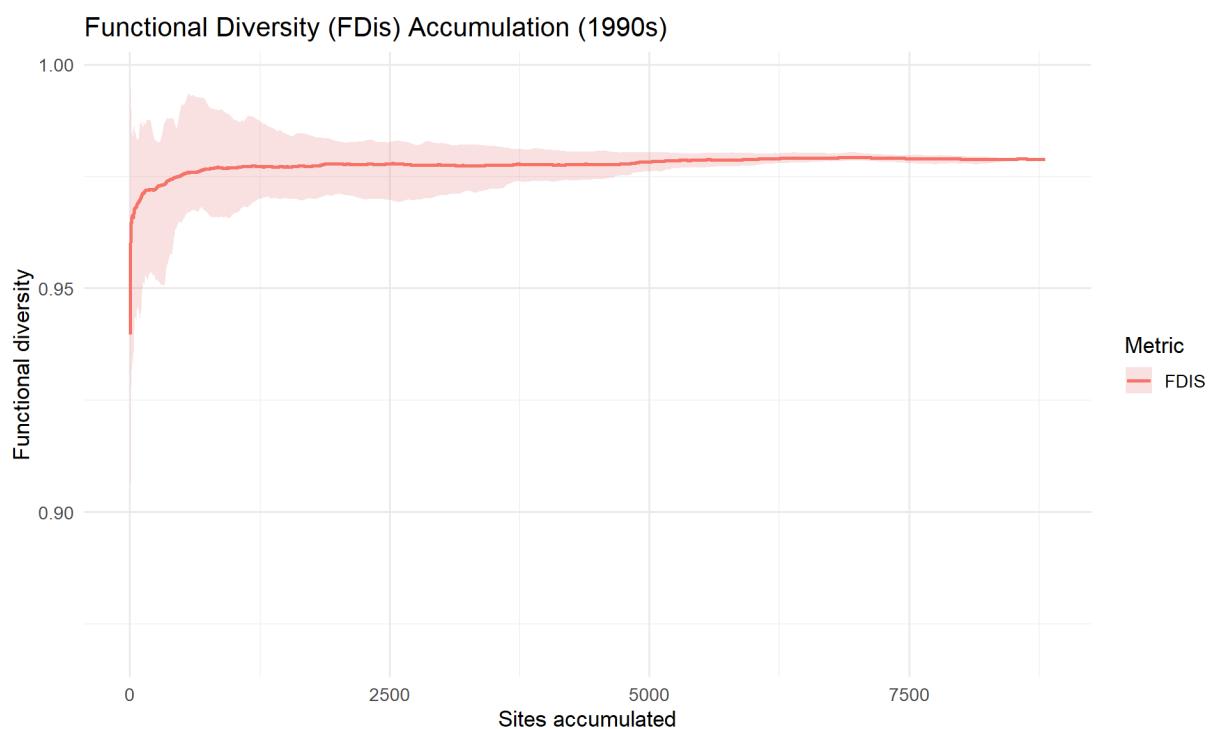
Ressourcenerwerbsstrategie (schnell vs. langsam)

Pflanzenhöhe	Wettbewerbsfähigkeit, Lichtfang
Samenmasse	Ausbreitungsdistanz, Keimlingüberleben
Wurzeltiefe	Wasseraneignung, Trockenheitstoleranz
Klonalität	Persistenz, Ausbreitung ohne Samen

FDis (Functional Dispersion) (Laliberte & Legendre 2010) misst die mittlere Distanz der Arten zum Zentroid des Eigenschaftsraums, gewichtet nach Abundanz. Höherer FDis = funktionell diversere Gemeinschaft = mehr besetzter "Nischenraum".

In unserer Analyse verwenden wir **Familienzugehörigkeit** und **heimisch/alien Status** als Proxy-Eigenschaften — ein grober, aber informativer Ansatz, wenn detaillierte Eigenschaftsmessungen nicht für alle Arten verfügbar sind.

```
## spacc functional diversity: 8814 sites, 2699 species, 124 traits, 30 seeds
## Metrics: fdis
```



Interpretation:

Funktionelle Diversität (FDis) steigt mit räumlicher Beprobung, kann aber mit anderer Rate ein Plateau erreichen als der Artenreichtum. Dies kann zeigen:

1. **Funktionelle Redundanz:** Wenn FDis deutlich vor dem Artenreichtum ein Plateau erreicht, sind viele später hinzugefügte Arten funktionell ähnlich zu bereits beprobten Arten — sie teilen den gleichen ökologischen "Job." Diese Redundanz kann Ökosystem-Resilienz bieten (Pillar et al. 2013).

2. **Funktionelle Komplementarität:** Wenn FDis weiter neben dem Artenreichtum steigt, bringt jede neue Art einen einzigartigen funktionellen Beitrag. Dies ist mit höherer Ökosystemfunktion assoziiert (Tilman et al. 2001).
3. **Invasionsauswirkung auf Funktion:** Wenn Alien-Arten sich in einem engen funktionellen Raum clustern (z.B. ruderale, schnellwüchsige, generalistische Kräuter), erhöhen sie die funktionelle Diversität möglicherweise nicht wesentlich, obwohl sie den Artenreichtum steigern. Dies ist die "funktionelle Homogenisierung"-Befürchtung — Invasionen fügen Arten hinzu, aber keine ökologischen Rollen (Olden et al. 2004).

11. Habitat-Analyse

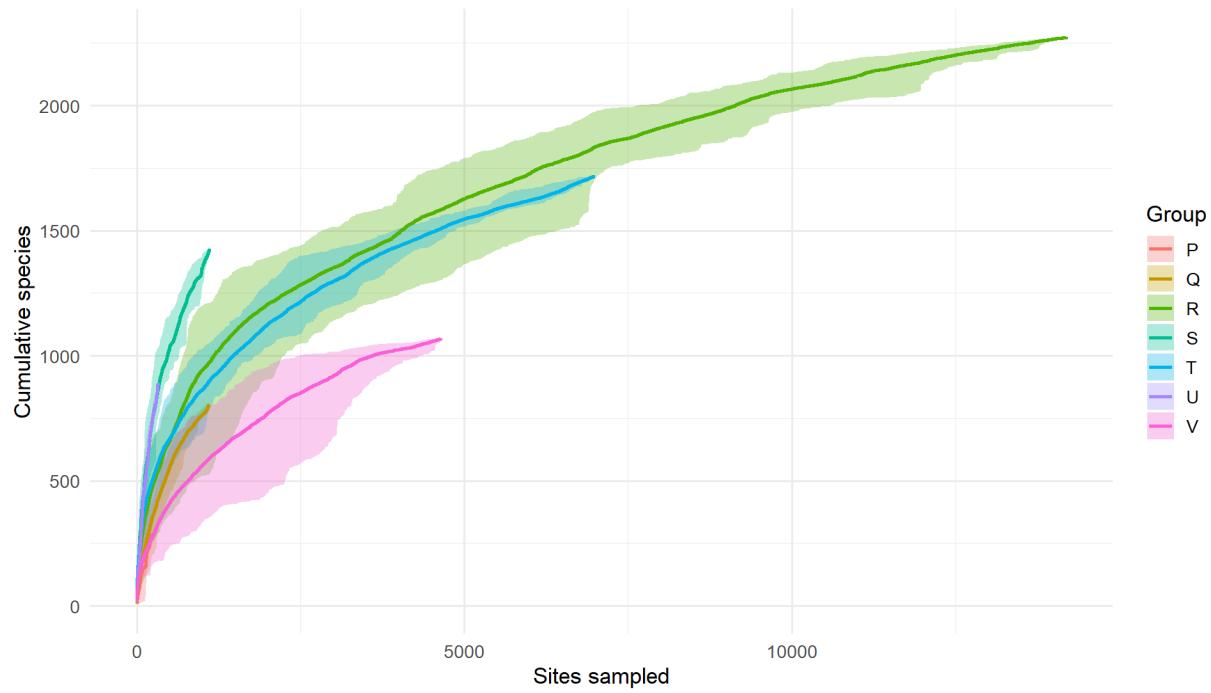
Habitattyp als Treiber von Invasionen:

Nicht alle Habitate sind gleich anfällig für Pflanzeninvasionen. Die **Hypothese der schwankenden Ressourcenverfügbarkeit** (Davis et al. 2000) sagt voraus, dass Gemeinschaften invasiver werden, wenn die Ressourcenverfügbarkeit vorübergehend steigt — durch Störung, Nährstoffpulse oder die Freisetzung anderer limitierender Faktoren.

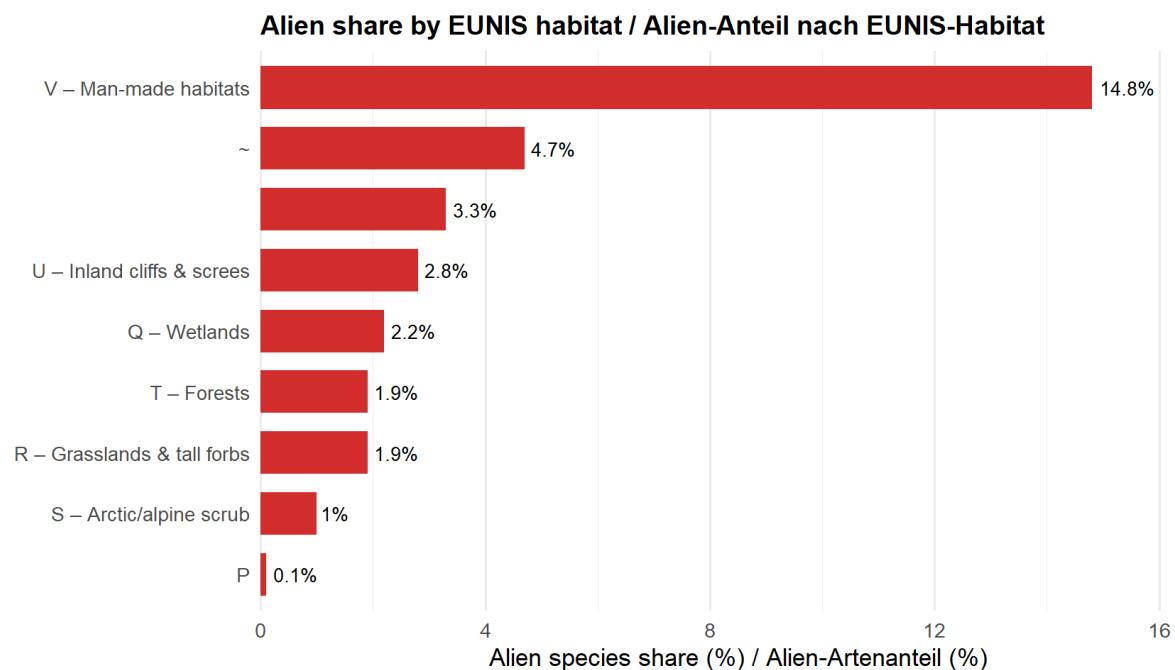
In der EUNIS-Habitatklassifikation würden wir erwarten:

C (Binnengewässer)	Mäßig	Uferkorridore erleichtern Ausbreitung
E (Grasland)	Variabel	Hängt von Bewirtschaftungsintensität ab
G (Wälder)	Niedrig	Geschlossenes Kronendach begrenzt Licht für Invasionen
I (Landwirtschaft)	Hoch	Gestört, nährstoffreich, offen
J (Urban)	Sehr hoch	Maximale Störung, viele Einführungspunkte

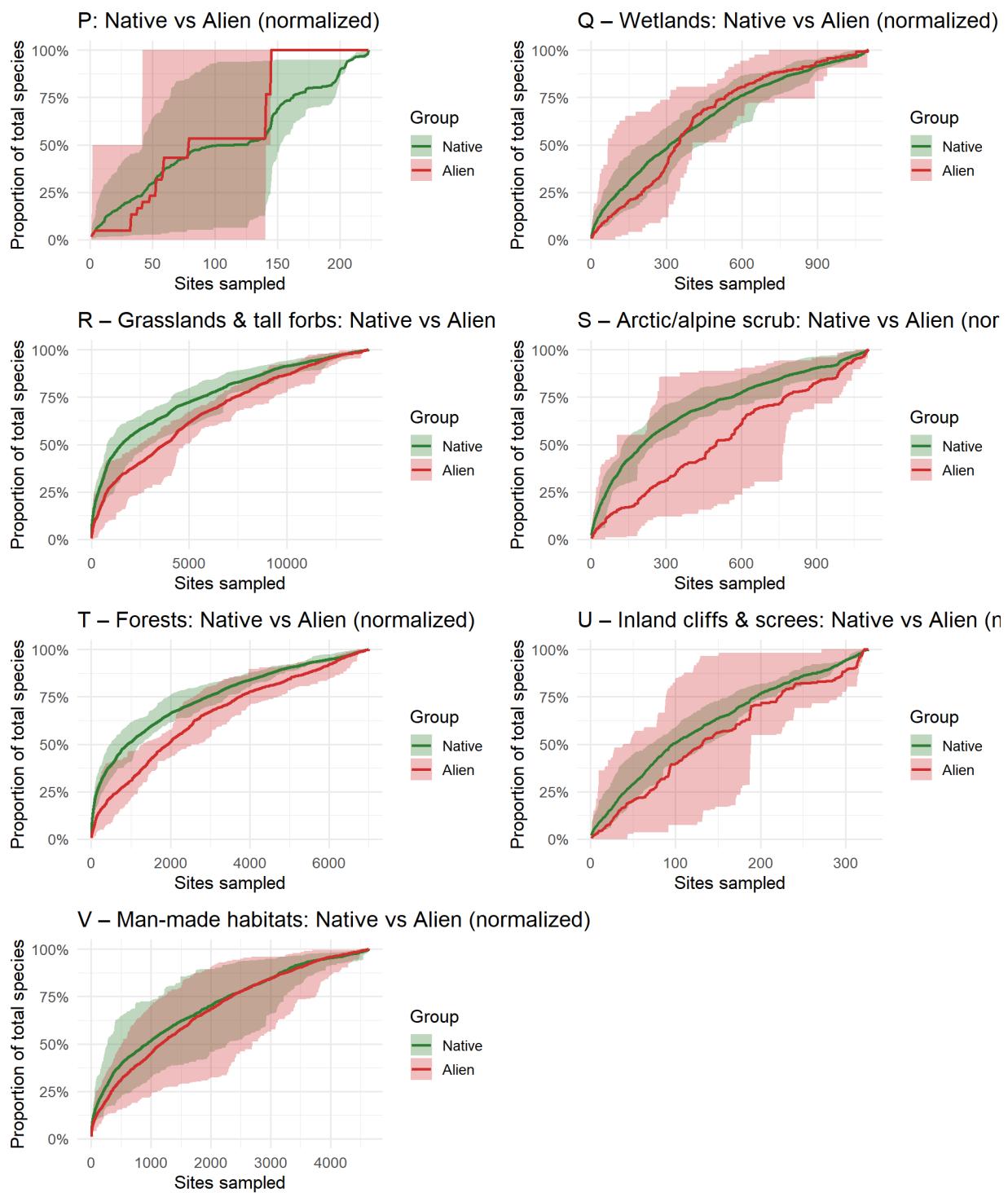
Species Accumulation by EUNIS Habitat (level 1)



Alien-Artenanteil nach Habitat



Heimisch vs. Alien innerhalb von Habitaten



Interpretation:

Die Habitat-Analyse zeigt, dass die Invasion durch Alien-Arten sehr heterogen über Habitattypen ist. Habitate mit hoher Störung und Ressourcenverfügbarkeit zeigen die höchsten Alien-Anteile, konsistent mit der Hypothese der schwankenden Ressourcen (Davis et al. 2000).

Die **Akkumulationskurvenformen** unterscheiden sich zwischen Habitaten auf aufschlussreiche Weise:

- In Habitaten, wo Aliens häufig sind, kann die Alien-Kurvenform der heimischen Kurve ähneln — Aliens haben sich breit innerhalb dieser Habitattypen verteilt
- In Habitaten, wo Aliens selten sind, ist die Alien-Kurve steiler und unregelmäßiger — was geringe Artenzahlen und spärliche Vorkommen widerspiegelt, die hohe Varianz in der Akkumulationstrajektorie erzeugen

Diese Habitatspezifität stimmt mit Chytry et al. (2008) überein, die zeigten, dass europäische Habitate sich 100-fach in ihrer Anfälligkeit für Pflanzeninvasionen unterscheiden, wobei anthropogene Habitate, Uferzonen und Küstengebiete am stärksten invadiert sind und oligotrophe Grasländer und alpine Habitate am wenigsten.

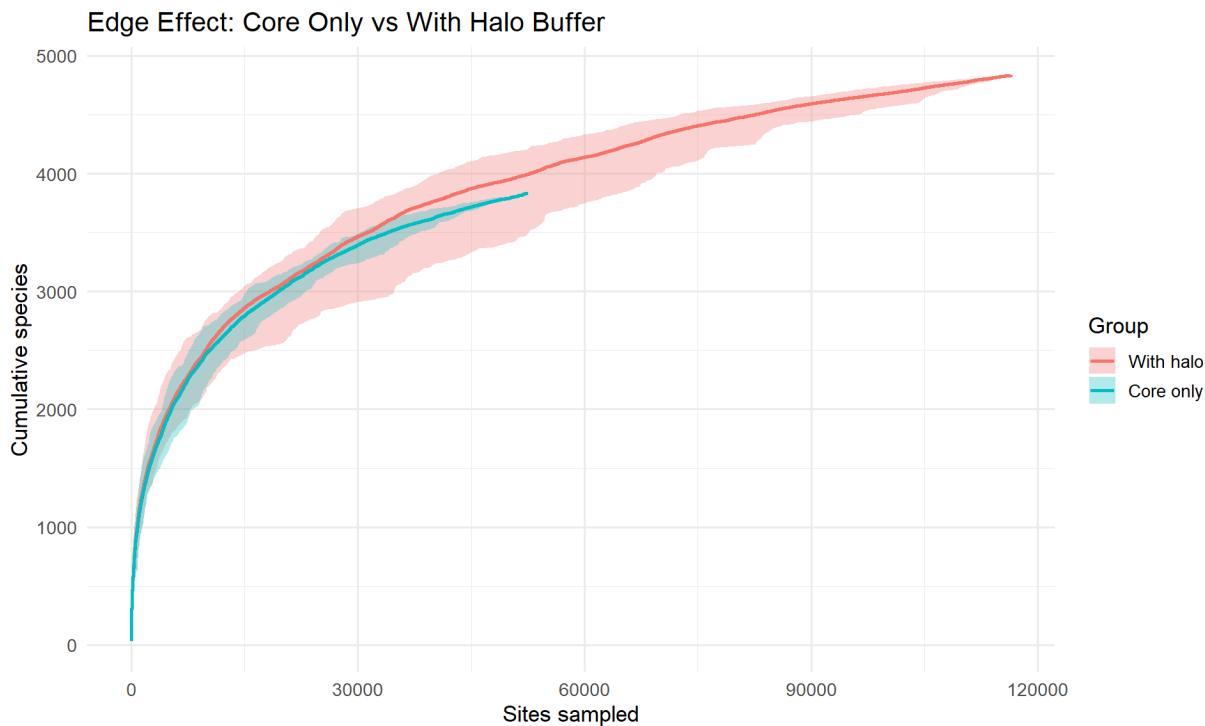
12. Randeffekte: Das Halo-Problem

Das Grenzartefakt:

Wenn wir Artenakkumulation in Österreich untersuchen, stehen wir vor einem methodischen Problem: Österreich hat Grenzen, aber Arten nicht. Ein Plot nahe der österreichisch-deutschen Grenze kann seinen nächsten Nachbarplot in Deutschland haben, nicht in Österreich. Wenn wir unsere Analyse auf österreichische Plots beschränken, wird der kNN-Algorithmus gezwungen, über die Grenze zu "springen", was künstliche Sprünge im räumlichen Walk erzeugt und die Akkumulationskurve verzerrt.

Der Halo-Ansatz löst dies, indem Plots aus Nachbarländern (die "Halo-Zone") in den räumlichen Walk einbezogen, aber nur Arten aus österreichischen Plots bei der Erstellung der Akkumulationskurve gezählt werden. Dies gewährleistet glatte Nearest-Neighbor-Pfade, die natürlich Grenzen überqueren.

Warum das wichtig ist: Randeffekte können Biodiversitätsschätzungen systematisch verzerren. Plots nahe Grenzen erscheinen künstlich artenreich oder artenarm, je nachdem ob grenzüberschreitende Diversität erfasst oder ausgeschlossen wird (Colwell et al. 2004). Für Österreich — ein Land mit 2.562 km Grenzen und komplexer Geometrie — ist dies besonders relevant.



Interpretation:

Der Vergleich der "Nur-Kern"- und "Mit-Halo"-Kurven zeigt das Ausmaß der Randeffekte in österreichischen Vegetationsdaten:

- **Wenn die Kurven erheblich divergieren:** Randeffekte sind signifikant. Die Nur-Kern-Analyse unterschätzt die Glätte des räumlichen Walks und bläht wahrscheinlich den scheinbaren Arten-Turnover nahe Grenzen auf.
- **Wenn die Kurven ähnlich sind:** Randeffekte sind für den österreichischen Datensatz gering, möglicherweise weil das Land groß genug ist, dass die meisten Plots weit von Grenzen entfernt sind.

Diese Analyse validiert (oder hinterfragt) die Annahme, dass Biodiversitätsanalysen auf nationaler Ebene Ländergrenzen als biologische Grenzen behandeln können. Für kleinere Länder oder langgestreckte Territorien wäre die Halo-Korrektur noch kritischer.

Das hier verwendete `areaOfEffect` -Paket implementiert die von Colling et al. beschriebene Methodik für den Umgang mit administrativen Grenzen in räumlichen ökologischen Analysen. Dieser Ansatz ist besonders wichtig beim Vergleich von Biodiversitätsmetriken zwischen Ländern verschiedener Größe und Form.

13. Metriken pro Standort & Räumliche Karte

Von Kurven zu Karten:

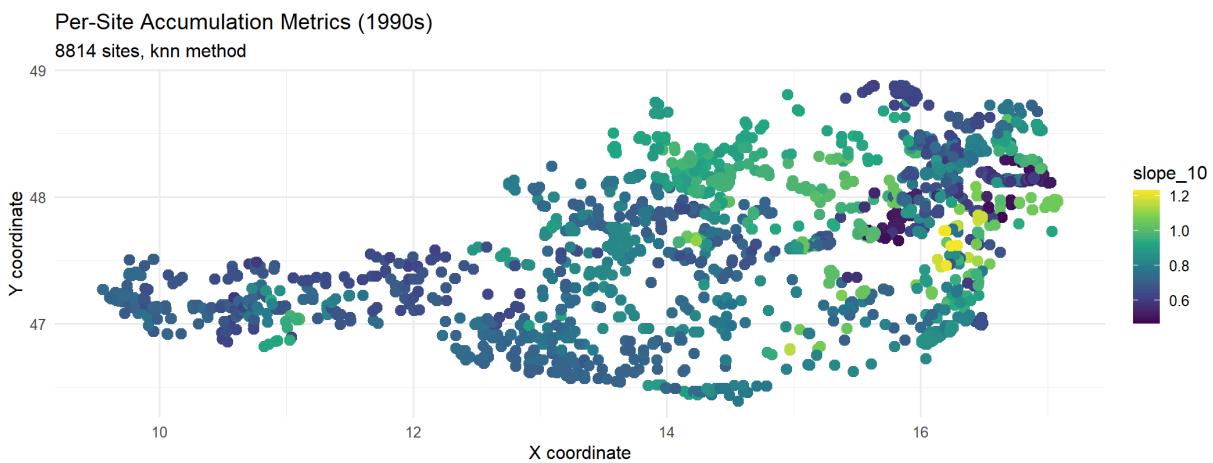
Während Akkumulationskurven Muster über die Landschaft zusammenfassen, können wir auch Metriken für **jeden einzelnen Standort** berechnen, basierend auf seiner Rolle im Akkumulationsprozess:

- **Steigung bei 10 Standorten** (`slope_10`): Wie schnell steigt der Artenreichtum, wenn 10 Standorte ab diesem Punkt beprobt werden? Hohe Steigung = die Nachbarschaft ist artenreich
- **Halbwertsreichtum** (`half_richness`): Wie viele Standorte müssen beprobt werden (von diesem Startpunkt), um 50% der Gesamtarten zu erreichen? Niedrige Werte = schnelle Akkumulation = geclusterte Diversität
- **AUC** (Fläche unter der Kurve): Ein Zusammenfassungsmaß, das die gesamte Akkumulationstrajektorie von einem Standort integriert

Die Kartierung dieser Metriken zeigt **räumliche Hotspots der Biodiversitätsakkumulation** — Gebiete, in denen der Arten-Turnover am höchsten und die Beprobungseffizienz am größten ist (Colling et al. 2024).

```
## spacc_metrics: 8814 sites, 2700 species
## Method: knn
## Metrics: slope_10, half_richness, auc
```

```
## Metric summary:
##   slope_10: mean=0.82, sd=0.17, range=[0.47, 1.23]
##   half_richness: mean=2050.14, sd=486.35, range=[1204.00, 3454.00]
##   auc: mean=1585.70, sd=38.80, range=[1503.78, 1658.52]
```

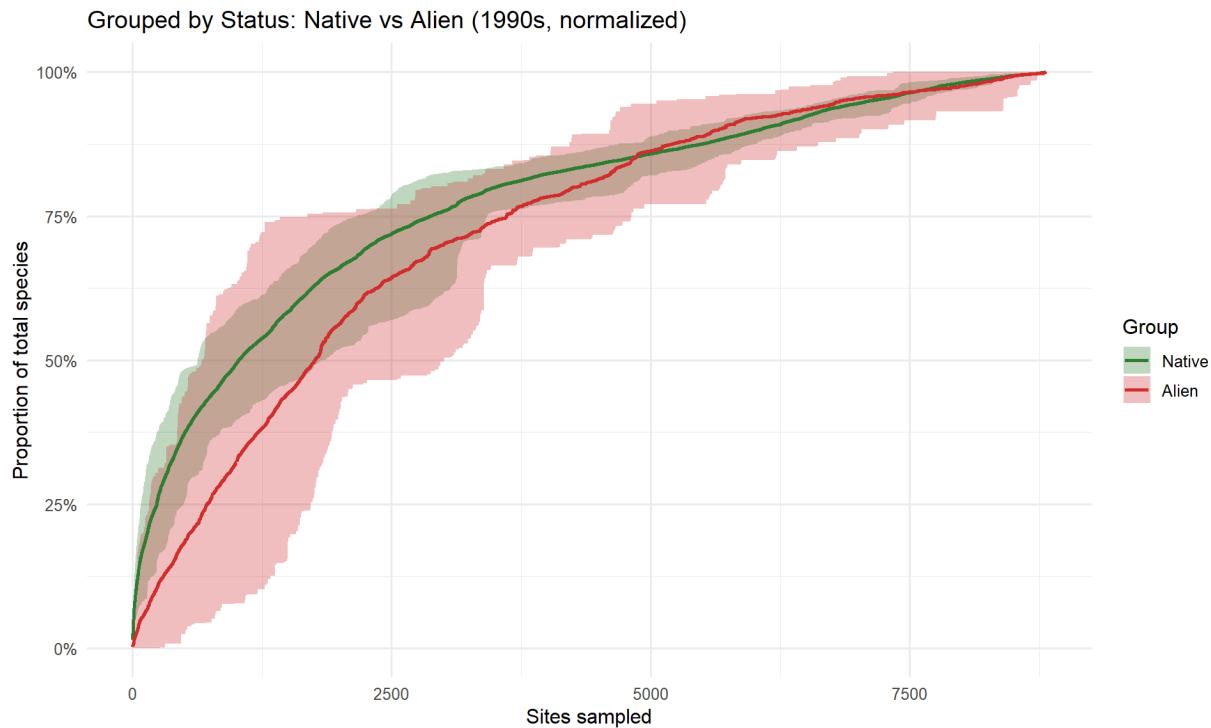


Interpretation:

Die räumlichen Karten der standortbezogenen Metriken liefern umsetzbare Erkenntnisse für die Naturschutzplanung:

- **Gebiete mit hoher Steigung** repräsentieren Biodiversitäts-Hotspots, wo eine kleine Anzahl zusätzlicher Proben schnell neue Arten erfasst. Diese Gebiete sind Prioritäten für intensive Biodiversitätserhebungen.
- **Gebiete mit niedrigem Halbwertsreichtum** sind Regionen, wo Diversität lokal konzentriert ist — die Beprobung weniger nahegelegener Plots erfasst einen großen Anteil der regionalen Diversität. Dieses Muster ist typisch für umweltmäßig heterogene Gebiete (alpine Täler, Ökotone zwischen Habitattypen).
- **Das räumliche Muster** dieser Metriken korreliert wahrscheinlich mit Höhenlage, Landnutzung und Habitatheterogenität — testbare Vorhersagen, die Akkumulationskurvenanalyse mit Landschaftsökologie verbinden.

14. Gruppierte Akkumulation nach Status



Dies ist der gleiche räumliche Walk, der gleichzeitig auf beide Gruppen angewendet wird, unter Verwendung des `groups`-Arguments — was sicherstellt, dass heimische und Alien-Arten entlang **identischer räumlicher Pfade** akkumuliert werden. Dies eliminiert jeden störenden Effekt verschiedener Walk-Trajektorien und bietet den saubersten Vergleich räumlicher Verteilungsmuster zwischen den beiden Gruppen.

Die normalisierten Kurven bestätigen das Muster aus der gepoolten Analyse: Alien-Arten akkumulieren proportional langsamer als Einheimische, was auf einen höheren räumlichen Turnover in der Alien-Flora hinweist. Verschiedene Alien-Arten kommen in verschiedenen Regionen vor, sodass mehr der Landschaft durchquert werden muss, um das vollständige Alien-Komplement zu erfassen.

15. Synthese & Schlussfolgerungen

Unsere Hypothesen revisited

- | | |
|--|--|
| H1: Aliens haben höhere räumliche Beta-Diversität | Normalisierte Alien-Kurve liegt unter der Native-Kurve |
| H2: Unterschied spiegelt habitatspezifische Einführungswege wider | Räumliche Heterogenität in der Alien-Flora |
| H3: Kurven sind statistisch unterschiedlich | Signifikanter Permutationstest |

Zusätzliche Entdeckungen aus erweiterten Analysen:

- Zeitliches Signal:** Der Unterschied heimisch-alien ist über Jahrzehnte konsistent
- Habitatspezifität:** Alien-Invasion konzentriert sich auf gestörte, nährstoffreiche Habitate

- **Phylogenetische Struktur:** Aliens repräsentieren möglicherweise eine phylogenetisch nicht-zufällige Teilmenge der globalen Flora
- **Räumliche Autokorrelation:** kNN vs Zufall-Vergleich bestätigt starke räumliche Struktur in österreichischen Pflanzengemeinschaften
- **Abdeckung:** 90% Beprobungsvollständigkeit erfordert erheblich weniger Plots als 99%
- **Randeffekte:** Die Halo-Analyse quantifiziert, wie nationale Grenzen räumliche Biodiversitätsschätzungen verzerren

Das große Bild

Unsere Analysen zeichnen ein kohärentes Bild der österreichischen Pflanzendiversität, das mehrere theoretische Rahmenwerke verbindet:

1. Invasionsökologie: Alien-Arten sind nicht zufällig verteilt — ihr Reichtum konzentriert sich in menschlich modifizierten Landschaften, aber entscheidend ist, dass *verschiedene* Alien-Arten in *verschiedenen* Regionen vorkommen (Pysek et al. 2010). Der räumliche Akkumulationsansatz zeigt diesen hohen räumlichen Turnover, weil er geografische Information bewahrt, die traditionelle Artenzählungen verwerfen.

2. Biogeografie: Österreichs außergewöhnlicher topografischer Gradient (115 m bis 3.798 m Höhe) erzeugt starke Umweltfilterung, die hohe Beta-Diversität durch echten Artenersatz antreibt. Alpine Gemeinschaften teilen wenige Arten mit Tieflandgemeinschaften, und dieses Muster dominiert die Akkumulationskurvenform.

3. Naturschutzplanung: Die hohe Beta-Diversität bedeutet, dass ein einzelnes großes Schutzgebiet Österreichs Pflanzendiversität nicht erfassen kann. Stattdessen ist ein verteiltes Netzwerk von Reservaten nötig, das verschiedene Höhenlagen, Habitate und biogeografische Regionen umfasst — was die **SLOSS-Debatte** (Several Large or Single Small) zugunsten von "mehreren" für Systeme mit hohem Turnover unterstützt (Tjorve 2010).

4. Stichprobentheorie: Deckungsbasierte Rarefaction zeigt, dass 52.000 Plots wahrscheinlich >95% der häufigen Flora Österreichs erfassen, aber seltene Arten (einschließlich möglicherweise unentdeckter Aliens) unentdeckt bleiben. Ob diese unentdeckten Aliens Unterbeprobung, kürzliche Ankünfte oder "Invasionsschuld" (Essl et al. 2011) widerspiegeln, kann aus räumlichen Daten allein nicht geklärt werden — temporale Replikation wäre nötig.

Diskussionsfragen

1. **Warum etablieren sich verschiedene Alien-Arten in verschiedenen Regionen?** Die Akkumulationskurven zeigen hohen räumlichen Turnover in der Alien-Flora. Bedenke, wie Propagule Pressure, lokale Störungsregime, Handelsrouten und Klimavariation in Österreich regional unterschiedliche Alien-Floren erzeugen könnten.
2. **Könnte der Klimawandel diese Muster verschieben?** Wenn die Temperaturen steigen, könnten alpine Spezialisten Habitat verlieren, während thermophile Aliens es gewinnen. Wie würde das Akkumulationskurven verändern?

3. **Ist funktionelle Homogenisierung eine Sorge für Österreich?** Wenn Aliens Arten, aber keine funktionelle Diversität hinzufügen, was bedeutet das für Ökosystemleistungen wie Bestäubung, Bodenstabilität und Kohlenstoffspeicherung?
4. **Wie würdest du ein Monitoringprogramm** basierend auf der Abdeckungsanalyse gestalten? Wie viele Plots, wo und wie oft?
5. **Welche zusätzlichen Daten würden diese Schlussfolgerungen stärken?** Denke an Eigenschaftsdatenbanken, zeitliche Replikation, genetische Diversität und trophische Interaktionen.

Wichtigste Erkenntnisse

Räumliche Akkumulation	Geografie formt Artenentdeckung	Fairer Vergleich erfordert räumliche Me
Heimisch vs. Alien	Verschiedene Verteilungsmuster bestätigt	Invasionsmanagement braucht räumlic
Hill-Zahlen	Reichtum ≠ Diversität ≠ Dominanz	Verschiedene Metriken beantworten ve
Beta-Diversität	Hoher Turnover über Österreich	Verteiltes Naturschutznetzwerk nötig
Phylogenetische Diversität	Evolutionsgeschichte fügt Information hinzu	Nicht alle Arten sind gleich "verschiede
Abdeckung	~95% Abdeckung erreichbar	Beprobungsaufwand kann optimiert w
Habitatspezifität	Invasionen sind habitat-gefiltert	Hochrisiko-Habitate für Management p
Randeffekte	Grenzen verzerrten Schätzungen	Halo-Korrektur verbessert Genauigkeit

Literaturverzeichnis

- Baker, H.G. (1974). The evolution of weeds. *Annual Review of Ecology and Systematics*, 5, 1–24.
- Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19, 134–143.
- Chao, A. & Jost, L. (2012). Coverage-based rarefaction and extrapolation. *Methods in Ecology and Evolution*, 3, 217–226.
- Chao, A. et al. (2014). Rarefaction and extrapolation with Hill numbers. *Ecological Monographs*, 84, 45–67.
- Chytry, M. et al. (2008). Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. *Journal of Applied Ecology*, 45, 448–458.
- Chytry, M. et al. (2009). European map of alien plant invasions based on the quantitative assessment across habitats. *Diversity and Distributions*, 15, 98–107.
- Chytry, M. et al. (2016). European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science*, 19, 173–180.
- Colwell, R.K. & Coddington, J.A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society B*, 345, 101–118.
- Colwell, R.K. et al. (2004). Interpolating, extrapolating, and comparing incidence-based species accumulation curves. *Ecology*, 85, 2717–2727.

- Davis, M.A. et al. (2000). Fluctuating resources in plant communities: a general theory of invasibility. *Journal of Ecology*, 88, 528–534.
- Dengler, J. et al. (2011). The Global Index of Vegetation-Plot Databases (GIVD). *Biodiversity & Ecology*, 3, 1–26.
- Essl, F. et al. (2011). Socioeconomic legacy yields an invasion debt. *Proceedings of the National Academy of Sciences*, 108, 203–207.
- Faith, D.P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61, 1–10.
- Gotelli, N.J. & Colwell, R.K. (2001). Quantifying biodiversity: procedures and pitfalls. *Ecology Letters*, 4, 379–391.
- Grabherr, G. et al. (2003). *Die Pflanzengesellschaften Österreichs*. Gustav Fischer Verlag.
- Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54, 427–432.
- Hillebrand, H. et al. (2008). Consequences of dominance: a review of evenness effects on local and regional ecosystem processes. *Ecology*, 89, 1510–1520.
- Jost, L. (2006). Entropy and diversity. *Oikos*, 113, 363–375.
- Kier, G. et al. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences*, 106, 9322–9327.
- Laliberte, E. & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, 91, 299–305.
- Lenoir, J. et al. (2008). A significant upward shift in plant species optimum elevation during the 20th century. *Science*, 320, 1768–1771.
- Lockwood, J.L. et al. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, 20, 223–228.
- Lomolino, M.V. (2000). Ecology's most general, yet protean pattern: the species-area relationship. *Journal of Biogeography*, 27, 17–26.
- Nekola, J.C. & White, P.S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, 26, 867–878.
- Olden, J.D. et al. (2004). Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution*, 19, 18–24.
- Partel, M. et al. (2011). Dark diversity: shedding light on absent species. *Trends in Ecology & Evolution*, 26, 596–600.
- Petchey, O.L. & Gaston, K.J. (2006). Functional diversity: back to basics and looking forward. *Ecology Letters*, 9, 741–758.
- Pillar, V.D. et al. (2013). Functional redundancy and stability in plant communities. *Journal of Vegetation Science*, 24, 963–974.
- Preston, F.W. (1948). The commonness, and rarity, of species. *Ecology*, 29, 254–283.
- Pysek, P. et al. (2010). Disentangling the role of environmental and human pressures on biological invasions across Europe. *Proceedings of the National Academy of Sciences*, 107, 12157–12162.
- Rahbek, C. (1995). The elevational gradient of species richness: a uniform pattern? *Ecography*, 18, 200–205.
- Pysek, P. et al. (2012). A global assessment of invasive plant impacts on resident species, communities and ecosystems. *Global Change Biology*, 18, 1725–1737.
- Pysek, P. et al. (2017). Naturalized alien flora of the world. *Preslia*, 89, 203–274.
- Sala, O.E. et al. (2000). Global biodiversity scenarios for the year 2100. *Science*, 287, 1770–1774.
- Seebens, H. et al. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, 8, 14435.
- Socolar, J.B. et al. (2016). How should beta-diversity inform biodiversity conservation? *Trends in Ecology & Evolution*, 31, 67–80.

- Tilman, D. et al. (2001). Diversity and productivity in a long-term grassland experiment. *Science*, 294, 843–845.
- Tjørve, E. (2010). How to resolve the SLOSS debate: lessons from species-diversity models. *Journal of Theoretical Biology*, 264, 604–612.
- van Kleunen, M. et al. (2010). Are invaders different? A conceptual framework of comparative approaches for assessing determinants of invasiveness. *Ecology Letters*, 13, 947–958.
- van Kleunen, M. et al. (2015). Global exchange and accumulation of non-native plants. *Nature*, 525, 100–103.
- Webb, C.O. et al. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33, 475–505.
- Whittaker, R.H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, 30, 279–338.
- Whittaker, R.H. (1972). Evolution and measurement of species diversity. *Taxon*, 21, 213–251.
- Wilson, J.R.U. et al. (2007). Residence time and potential range: crucial considerations in modelling plant invasions. *Diversity and Distributions*, 13, 11–22.
- Winter, M. et al. (2009). Plant extinctions and introductions lead to phylogenetic and taxonomic homogenization of the European flora. *Proceedings of the National Academy of Sciences*, 106, 21721–21725.

Session Info

```
## R version 4.5.2 (2025-10-31 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 11 x64 (build 26200)
##
## Matrix products: default
##   LAPACK version 3.12.1
##
## locale:
## [1] LC_COLLATE=English_United States.utf8
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
## time zone: Europe/Luxembourg
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics   grDevices utils      datasets   methods    base
##
## other attached packages:
## [1] hexify_0.3.10      rnaturalearth_1.2.0 sf_1.0-24
## [4] spacc_0.1.0        data.table_1.18.0  ggplot2_4.0.1
##
## loaded via a namespace (and not attached):
## [1] gtable_0.3.6          jsonlite_2.0.0       dplyr_1.1.4
## [4] compiler_4.5.2         tidyselect_1.2.1     Rcpp_1.1.1
## [7] parallel_4.5.2         gridExtra_2.3        jquerylib_0.1.4
## [10] scales_1.4.0          yaml_2.3.12         fastmap_1.2.0
## [13] R6_2.6.1              labeling_0.4.3       generics_0.1.4
## [16] classInt_0.4-11       knitr_1.51          tibble_3.3.1
## [19] units_1.0-0           DBI_1.2.3           bslib_0.9.0
## [22] pillar_1.11.1         RColorBrewer_1.1-3   rlang_1.1.7
## [25] cachem_1.1.0          xfun_0.55          sass_0.4.10
## [28] S7_0.2.1              RcppParallel_5.1.11-1 otel_0.2.0
## [31] viridisLite_0.4.2     cli_3.6.5           withr_3.0.2
## [34] magrittr_2.0.4         class_7.3-23        digest_0.6.39
## [37] grid_4.5.2             lifecycle_1.0.5     vctrs_0.7.0
## [40] KernSmooth_2.23-26    proxy_0.4-29        evaluate_1.0.5
## [43] glue_1.8.0              farver_2.1.2        rnaturalearthdata_1.0.0
## [46] e1071_1.7-17          rmarkdown_2.30       tools_4.5.2
## [49] pkgconfig_2.0.3        htmltools_0.5.9
```



Nach dem Workshop

Zusammenfassung

Workshop-Rückblick

Tag 3: Wichtigste Ergebnisse aus den österreichischen Vegetationsdaten

Was uns die spacc-Analyse verrät

Liam's Stay - Results

2026-02-16

Überblick

Dieses Dokument destilliert die **fünf überzeugendsten Ergebnisse** aus unserer umfassenden Analyse von ~29,857 österreichischen Vegetationsplots. Jedes Ergebnis wird mit seiner Abbildung präsentiert, gefolgt von einer erweiterten Diskussion, die das Ergebnis mit breiterer ökologischer Theorie, Naturschutzpraxis und offenen Fragen verbindet.

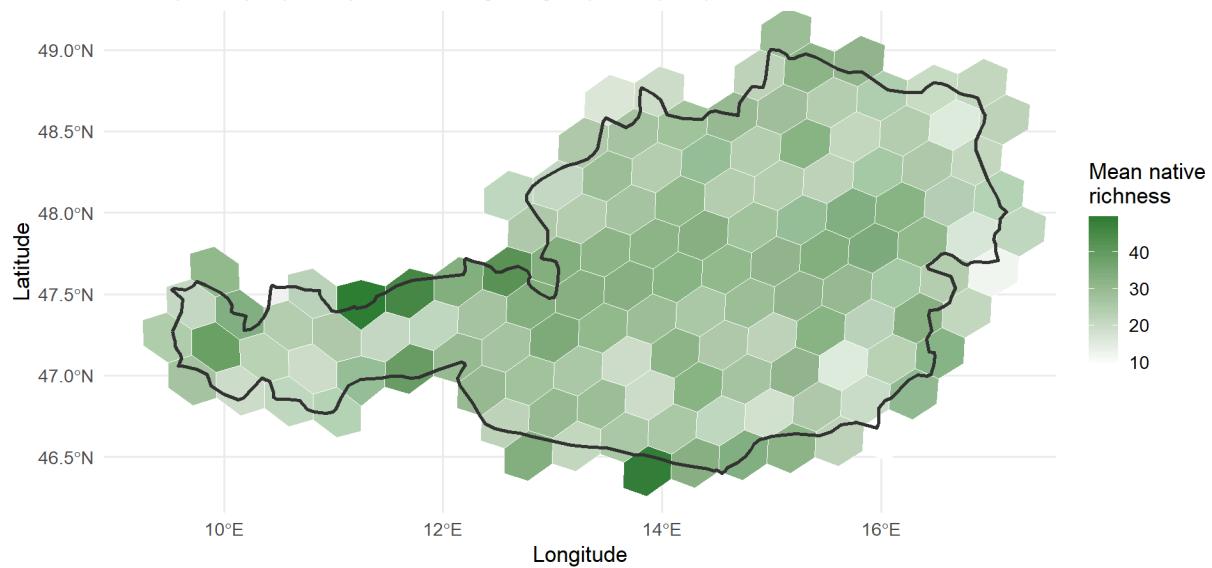
Die fünf Ergebnisse wurden ausgewählt, weil sie jeweils einen eigenständigen Teil der österreichischen Biodiversitätsgeschichte erzählen:

- | | |
|---|--|
| 1 Beprobung & Artenreichtum — Hotspot-Karten | Wo wurde beprobt und wie sieht der beobachtete Artenreichtum aus? |
| 2 Wie Arten akkumulieren — Heimische vs. Alien-Kurven | Die zentrale Hypothese: verschiedene räumliche Prozesse führen zu unterschiedlichen Artenakkumulationsmustern. |
| 3 Was uns fehlt — Invasionsschuld | Die Zukunft von Österreichs Alien-Flora: Wie wird sich die Flora ändern? |
| 4 Warum sich Standorte unterscheiden — Turnover vs. Nestedness | Was Österreichs hohe Beta-Diversität antreibt: Unterschiede zwischen verschiedenen Ökosystemen. |
| 5 Welche Habitate verwundbar sind — Habitat-Invasionsmuster | Wo Naturschutzbemühungen fokussieren sollten: Identifizierung von sensiblen Lebensräumen. |

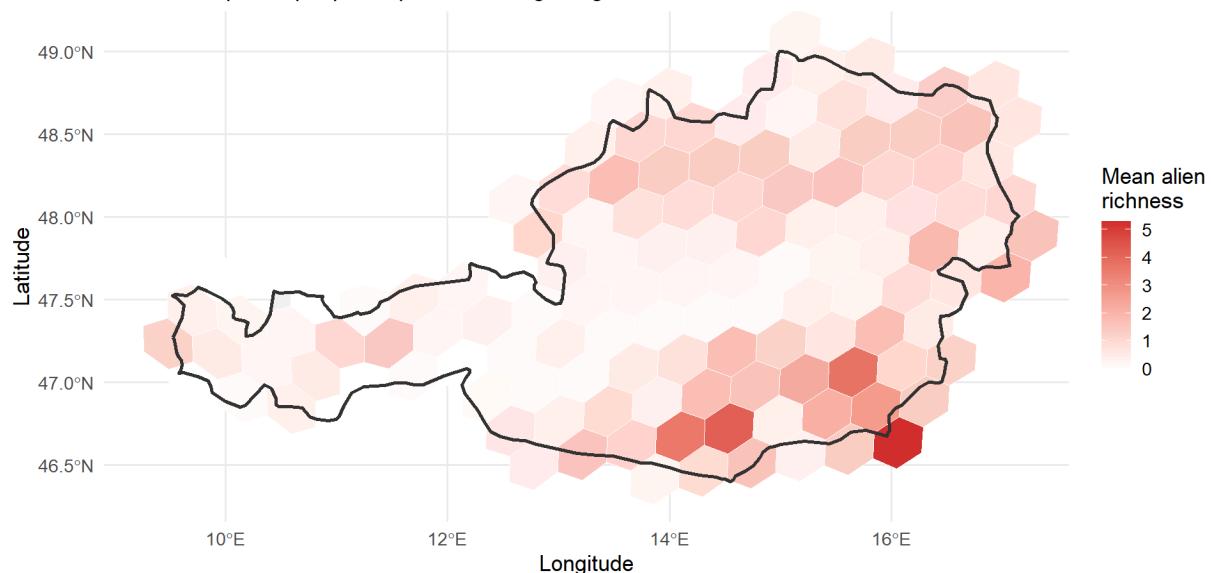
1. Beprobung & Artenreichtum: Hotspot-Karten

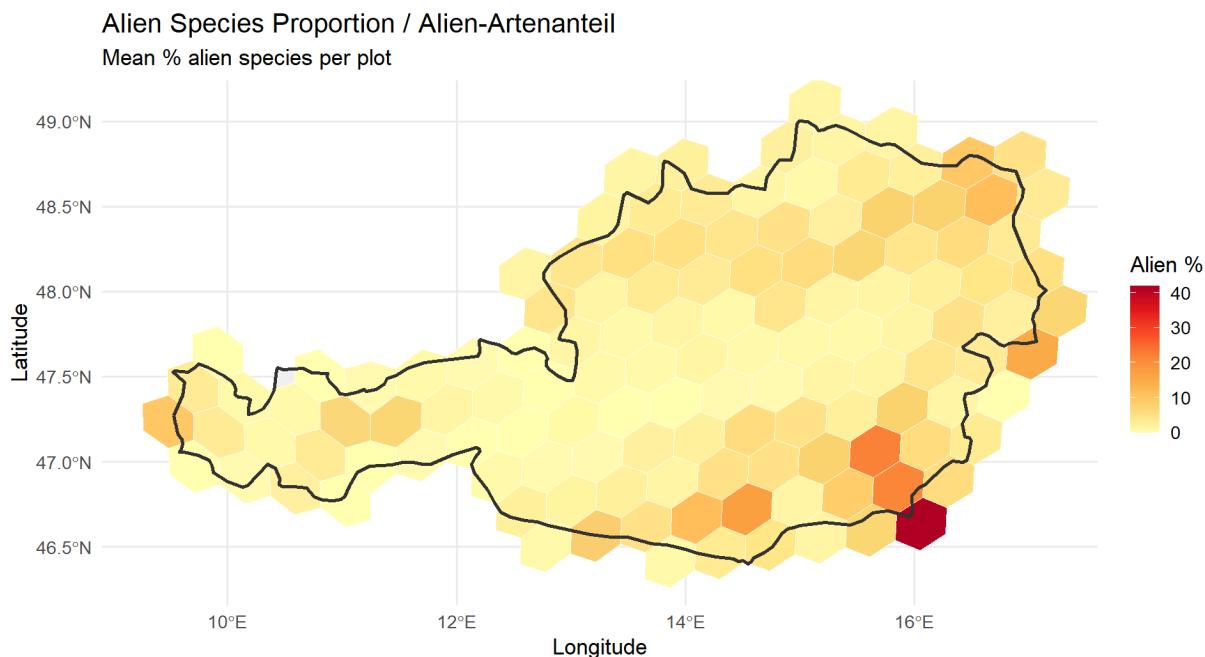
Die Karten

Native Species Richness Hotspots / Heimische Arten Hotspots
Mean species per plot, equal-area hexagonal grid (min. 3 plots)



Alien Species Richness Hotspots / Alien-Arten Hotspots
Mean alien species per plot, equal-area hexagonal grid





Diskussion

Was die Karten zeigen:

Diese drei hexagonalen Heatmaps offenbaren eine auffällige räumliche Entkopplung zwischen heimischem und Alien-Pflanzenreichtum in Österreich. Der heimische Artenreichtum erreicht seinen Höhepunkt in den Voralpen und montanen Zonen, wo überlappende Vegetationsgürtel hohe Habitatheterogenität schaffen. Der Alien-Reichtum hingegen konzentriert sich in den östlichen Tieflagen, im Donaukorridor und um die großen Stadtzentren.

Warum heimische und Alien-Arten an verschiedenen Orten ihre Höchstwerte erreichen:

Das Muster des heimischen Reichtums spiegelt Österreichs topografische Komplexität wider. Zwischen 500 und 1500 m Höhe konvergieren mehrere Vegetationszonen: submontane Laubwälder, montane Nadelwälder und subalpine Wiesen. Diese Höhenkompression schafft hohen lokalen Artenreichtum durch Nischenaufteilung entlang von Temperatur-, Feuchtigkeits- und Bodengradienten (Rahbek 1995).

Das Alien-Muster erzählt eine andere Geschichte. Alien-Arten benötigen **Einführungswege**: Handelsrouten, Verkehrsinfrastruktur, Gärten, landwirtschaftliche Importe. Diese Wege konvergieren auf Städte und Tieflandkorridore. Das Donautal dient als kontinentale Ausbreitungsautobahn. Wien allein war seit dem 18. Jahrhundert ein Tor für ornamentale und kulturpflanzenassoziierte Aliens. Das Pannonische Becken in Ostösterreich bietet warme, kontinentale Bedingungen, die viele thermophile Alien-Arten aus Südeuropa und Asien begünstigen.

Die Anteilskarte ist am informativsten:

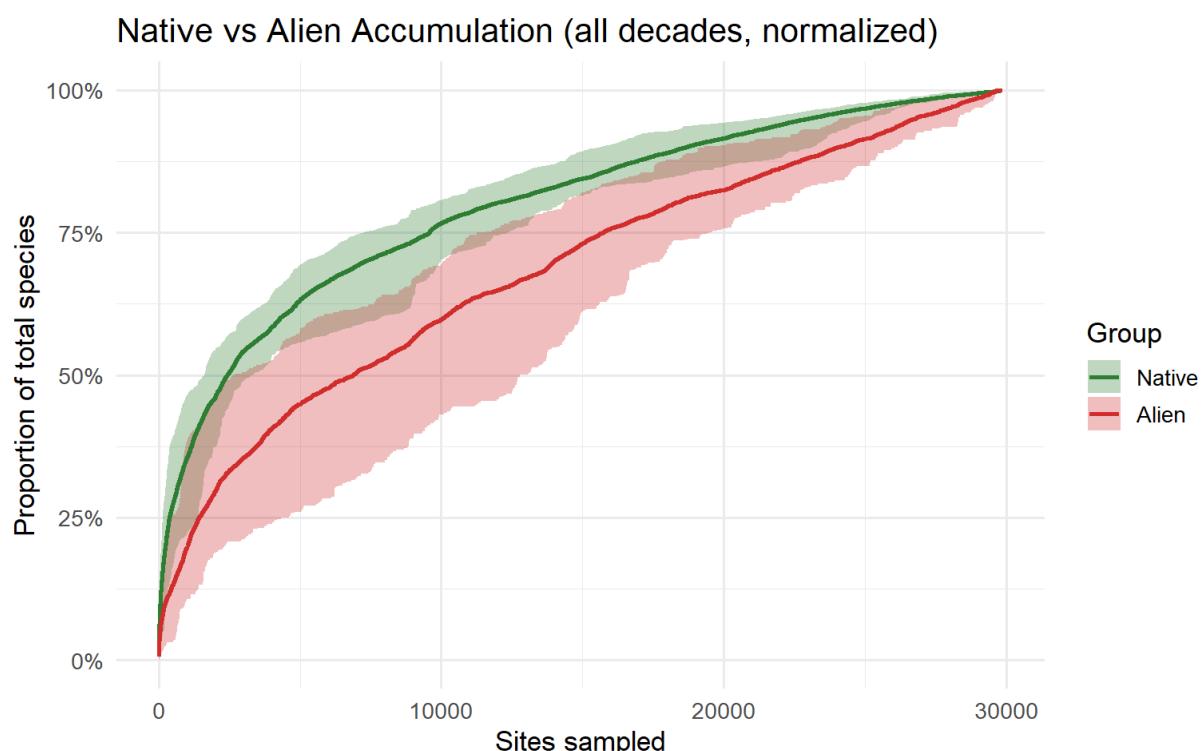
Während absolute Reichtumsarten durch Beprobungsintensität verfälscht werden können, kontrolliert die Alien-Anteilskarte dies, indem sie Aliens als Bruchteil der Gesamtdiversität an jedem Standort ausdrückt. Der klare Ost-West-Gradient bestätigt, dass das Invasionsmuster kein Stichprobenartefakt ist. Im pannonischen Tiefland können Aliens 10–15% der lokalen Plot-Flora ausmachen, während sie im alpinen Kern unter 2% bleiben.

Verbindungen zur Invasionstheorie:

Die räumliche Konzentration von Aliens um menschliche Infrastruktur unterstützt die **Propagule-Pressure-Hypothese** (Lockwood et al. 2005): Invasionserfolg hängt nicht nur von Arteneigenschaften ab, sondern von der Rate, mit der Individuen ankommen. Die Verbindung zur **Hypothese der schwankenden Ressourcen** (Davis et al. 2000) liegt auf der Hand: gestörte, nährstoffangereicherte Tieflandhabitante erfahren regelmäßige Ressourcenpulse, die Invasionsfenster schaffen.

2. Das zentrale Ergebnis: Aliens akkumulieren anders

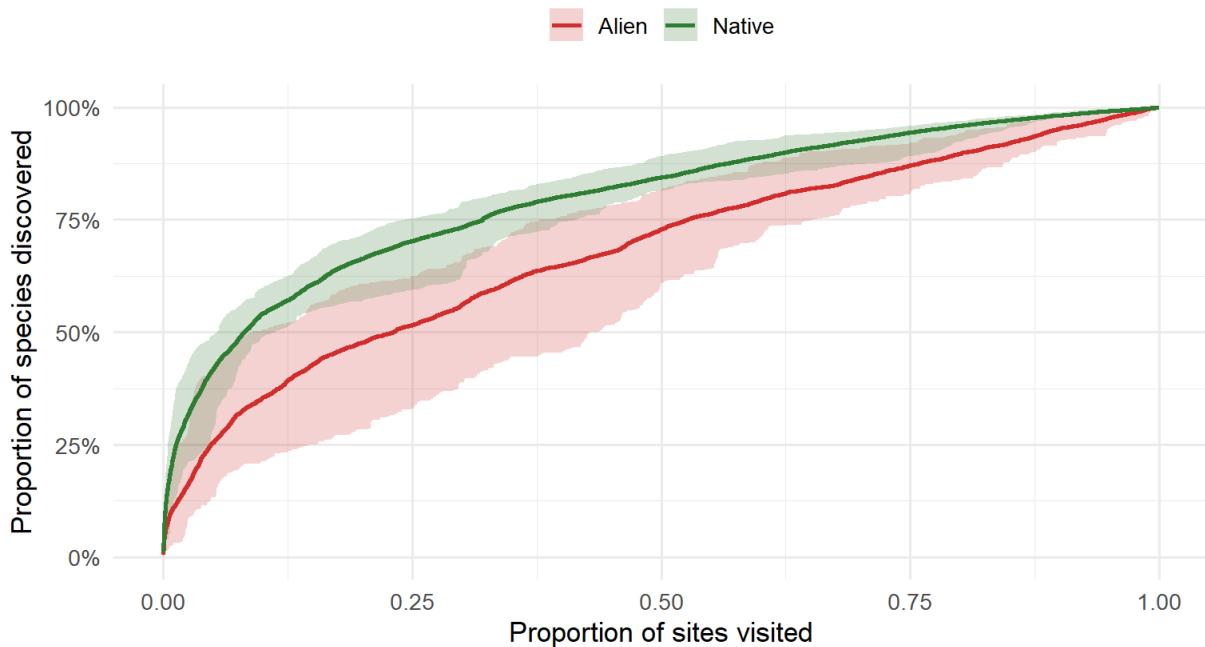
Heimische vs. Alien Akkumulationskurven



```
## Comparison: p$native vs p$alien
## -----
## Method: permutation (n=999)
## Normalized: yes (shape comparison)
## AUC difference: 3433.3 (p < 0.001***)
## Saturation: p$native at 18661 sites, p$alien at 24060 sites
##
## p$native saturates faster.
```

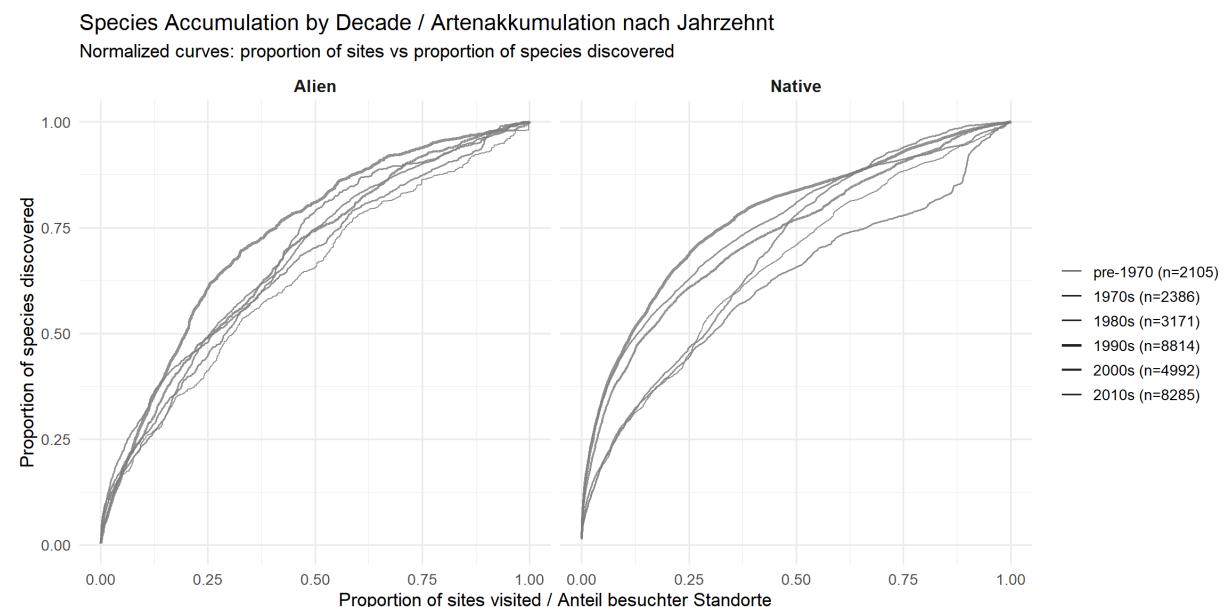
Shape Comparison: Native vs Alien Accumulation

$p < 0.001$ (normalized permutation test, 999 iterations)



Jahrzehntweise Akkumulation

Verändert sich das räumliche Akkumulationsmuster über die Zeit? Hier führen wir dieselbe Analyse getrennt für jedes Jahrzehnt durch, wobei heimische und gebietsfremde Arten in separate Panels aufgeteilt werden. Jede Kurve ist nach ihrer Gesamtartenzahl normalisiert.



Diskussion

Das Kernergebnis:

Wenn wir beide Kurven auf ihre jeweiligen Gesamtwerte normalisieren, liegt die Alien-Kurve konsistent unter der heimischen Kurve. Das bedeutet, dass wir beim Durchqueren der österreichischen Landschaft in räumlicher Reihenfolge bei jedem Schritt einen *kleineren*

Anteil der Alien-Flora entdecken als der heimischen Flora. Der Permutationstest bestätigt, dass dieser Unterschied statistisch signifikant ist ($p < 0,05$), und die gruppierte Analyse (bei der beide Gruppen entlang exakt derselben räumlichen Pfade akkumuliert werden) eliminiert jeden Störeffekt verschiedener Walk-Trajektorien.

Was bedeutet eine niedrigere normalisierte Kurve biologisch?

Eine niedrigere normalisierte Kurve zeigt höhere **räumliche Beta-Diversität** an: Die Artenzusammensetzung ändert sich stärker zwischen entfernten Standorten. Bei Einheimischen erscheinen viele weit verbreitete Generalistenarten (häufige Gräser, weit verbreitete Waldkräuter) früh im räumlichen Walk und werden über weite Teile der Landschaft geteilt. Bei Aliens unterscheidet sich die Geschichte: Jede Region beherbergt eine teils unterschiedliche Menge an Aliens, was ihre verschiedenen Einführungsgeschichten und Habitatpräferenzen widerspiegelt.

Das Multi-Fokus-Einführungsmodell:

Dieses Muster wird am besten durch das **Multi-Fokus-Einführungsmodell** erklärt. Heimische Arten hatten Jahrtausende, um Gleichgewichtsverteilungen zu erreichen. Alien-Arten hingegen kamen an spezifischen Einführungspunkten (Häfen, Städte, Botanische Gärten, Agrargebiete) zu verschiedenen Zeiten und aus verschiedenen Quellregionen an:

- **Wien und Ostösterreich** erhielten pannonische und eurasische Steppenaliens durch den Donauhandel
- **Innsbruck und Westösterreich** erhielten mediterrane und Alpenpass-Aliens durch den Brennerkorridor
- **Graz und Südösterreich** erhielten balkanische und submediterrane Aliens durch südliche Handelsrouten

Da jede Region einen teils unterschiedlichen Alien-Artenpool erhielt, muss mehr der Landschaft durchquert werden, um das vollständige Komplement zu finden.

Zeitliche Robustheit:

Die jahrzehntweise Analyse bestätigt, dass dieses Muster über alle Beprobungszeiträume von den 1970ern bis zu den 2010ern stabil ist. Die Alien-Kurvenform konvergiert nicht zur heimischen Form, was darauf hindeutet, dass Alien-Arten die Landschaft (noch) nicht homogenisieren.

Implikationen für das Monitoring:

Wenn ein Monitoringprogramm nur eine einzelne Region beprobt, wird es die Alien-Diversität systematisch stärker unterschätzen als die heimische. Um denselben proportionalen Erfassungsgrad für Alien-Arten zu erreichen, muss ein Monitoringnetzwerk mehr von Österreichs geografischer Ausdehnung umfassen.

3. Die Invasionsschuld: Wie viele Aliens fehlen uns?

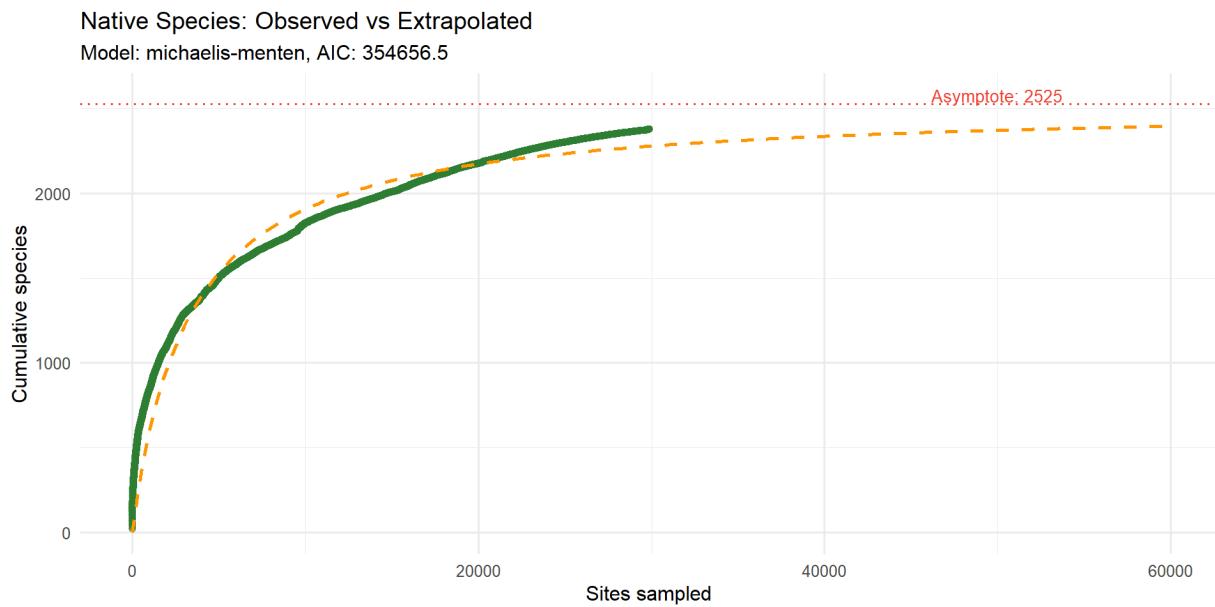
Das **Michaelis-Menten-Modell** passt eine sättigende Kurve an die Artenakkumulationsdaten an:

$$S(n) = \frac{S_{max} \cdot n}{K_m + n}$$

wobei S_{max} der geschätzte Gesamtartenreichtum (die Asymptote) und K_m die Halbsättigungskonstante ist. Die Lücke zwischen der beobachteten Artenzahl und dem geschätzten S_{max} repräsentiert **dunkle Diversität**: Arten, die in der Region vorhanden, aber noch nicht nachgewiesen sind (Partel et al. 2011).

```
## Native species extrapolation:
```

```
## Extrapolation: michaelis-menten
## -----
## Estimated asymptote: 2525.4 species
## 95% CI: 2522.4 - 2528.5
## AIC: 354656.5
## Observed: 2378.0 species (94% of estimated)
```

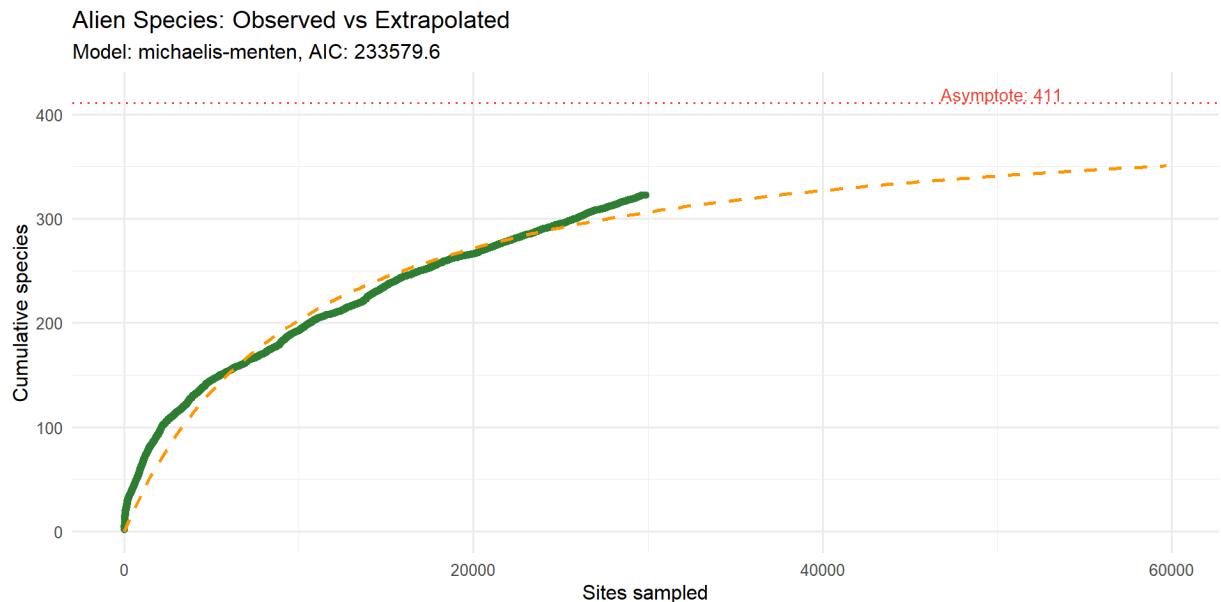


```
##
## Alien species extrapolation:
```

```

## Extrapolation: michaelis-menten
## -----
## Estimated asymptote: 410.9 species
## 95% CI: 410.0 - 411.8
## AIC: 233579.6
## Observed: 323.0 species (79% of estimated)

```



Diskussion

Die Zahlen:

- **Heimische Arten:** ~2.378 beobachtet von ~2.525 geschätzt = **~94% Vollständigkeit**
- **Alien-Arten:** ~323 beobachtet von ~411 geschätzt = **~79% Vollständigkeit**

Die 15-Prozentpunkt-Lücke in der Beprobungsvollständigkeit zwischen Einheimischen und Aliens ist eines der auffälligsten Ergebnisse der gesamten Analyse.

Warum ist die Alien-Flora weniger vollständig?

Drei komplementäre Mechanismen erklären diese Lücke:

1. **Fortlaufende Einführungen.** Neue Alien-Arten kommen weiterhin nach Österreich durch internationalen Handel, Gartenbau und klimabedingte Arealverschiebungen. Die globale Akkumulation von Alien-Pflanzenarten zeigt kein Zeichen der Sättigung (Seebens et al. 2017).
2. **Ausbreitungsverzögerung.** Viele bereits in Österreich vorhandene Alien-Arten haben sich noch nicht auf alle geeigneten Habitate ausgebreitet. Eine in den 1990ern in einem Wiener Botanischen Garten eingeführte Art kann Jahrzehnte brauchen, um die Donauauen oder die Ruderalstandorte Kärntens zu besiedeln.
3. **Nachweisschwierigkeit.** Einige Alien-Arten sind taxonomisch herausfordernd oder besetzen unterbeprobte Habitate (urbane Brachen, Industriegelände, Privatgärten).

Das Konzept der Invasionsschuld:

Essl et al. (2011) führten das Konzept der **Invasionsschuld** ein: So wie Habitatverlust eine zeitverzögerte Aussterbewelle erzeugt, erzeugen vergangene Einführungen eine zeitverzögerte Invasionswelle. Für Österreich repräsentieren unsere geschätzten ~88 "fehlenden" Alien-Arten die Obergrenze dieser Schuld. Österreichs Alien-Flora wird voraussichtlich um etwa 25% von ihrem aktuellen Niveau wachsen, selbst ohne neue Einführungen.

Was das für die Naturschutzplanung bedeutet:

Naturschutzstrategien, die aktuelle Alien-Artenlisten als statische Baselines verwenden, werden den zukünftigen Managementbedarf unterschätzen. Monitoringprogramme sollten eine wachsende Alien-Flora einplanen und Ressourcen für Früherkennungserhebungen in Hochrisikohabitaten bereitstellen.

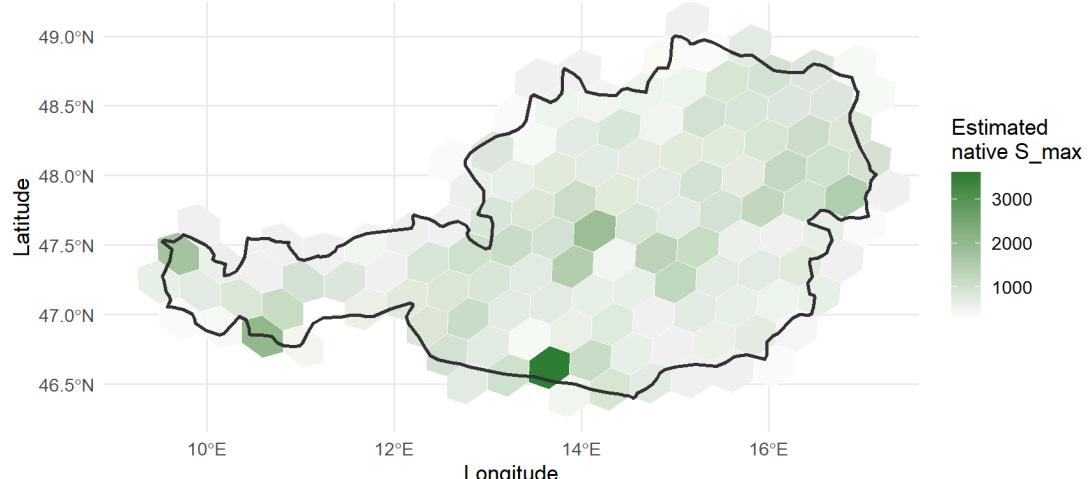
Modellvorbehalt:

Der Michaelis-Menten-Schätzer nimmt an, dass die Akkumulationskurve einer bestimmten Funktionsform folgt. Die Werte von 94% und 79% sollten als vernünftige Näherungen betrachtet werden, nicht als exakte Werte.

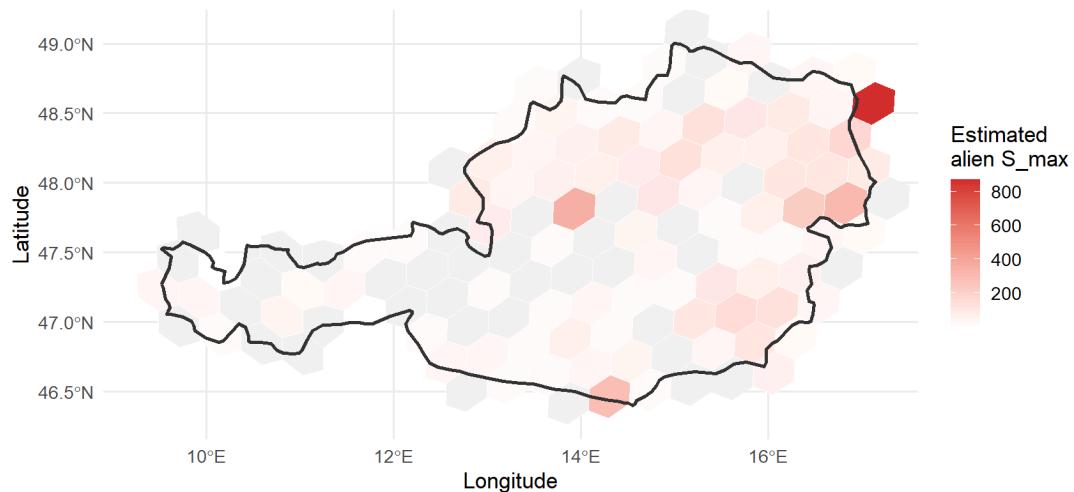
3b. Kartierung der Invasionsschuld: Geschätzter Artenreichtum pro Hexagon

Die Michaelis-Menten-Extrapolation oben schätzt den **Gesamt**-Artenreichtum für ganz Österreich. Hier wenden wir denselben Ansatz pro Hexagon an: Für jede Zelle im flächengleichen Gitter führen wir eine räumliche Akkumulation mit 10 Seeds auf den Plots innerhalb des Hexagons durch, passen Michaelis-Menten an und extrahieren den geschätzten Gesamtartenreichtum (S_{max}). Hexagone mit weniger als 20 Plots (in Hellgrau dargestellt) sind ausgeschlossen, da Akkumulationskurven mit so wenigen Stichproben unzuverlässige Extrapolationen liefern.

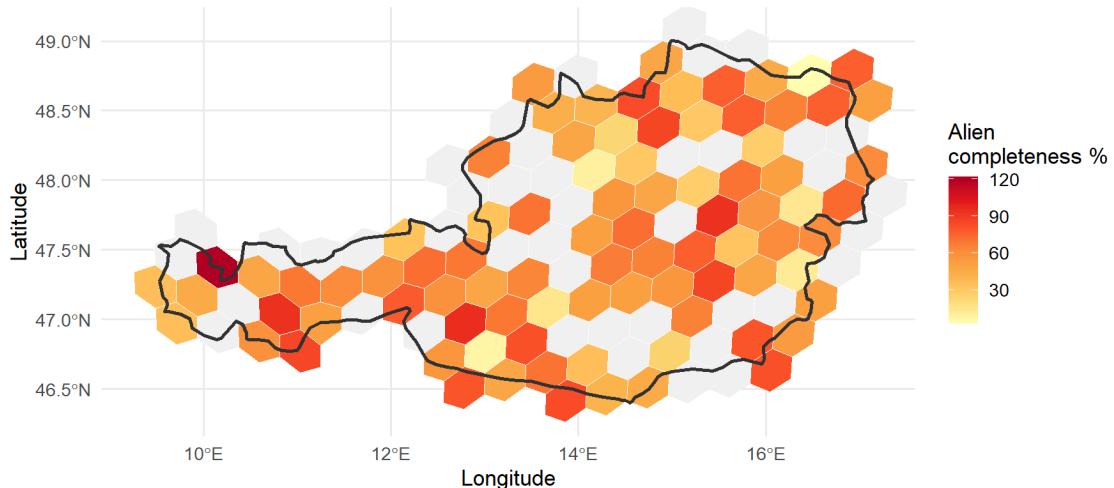
Estimated Native Richness (Michaelis-Menten) / Geschätzter heimischer Reichtum
Per-hexagon SAC extrapolation, 10 seeds each (grey = <20 plots, excluded)



Estimated Alien Richness (Michaelis-Menten) / Geschätzter Alien-Reichtum
Per-hexagon SAC extrapolation, 10 seeds each (grey = <20 plots, excluded)



Alien Sampling Completeness / Alien-Beprobungsvollständigkeit
Observed / estimated species per hexagon (grey = <20 plots, excluded)

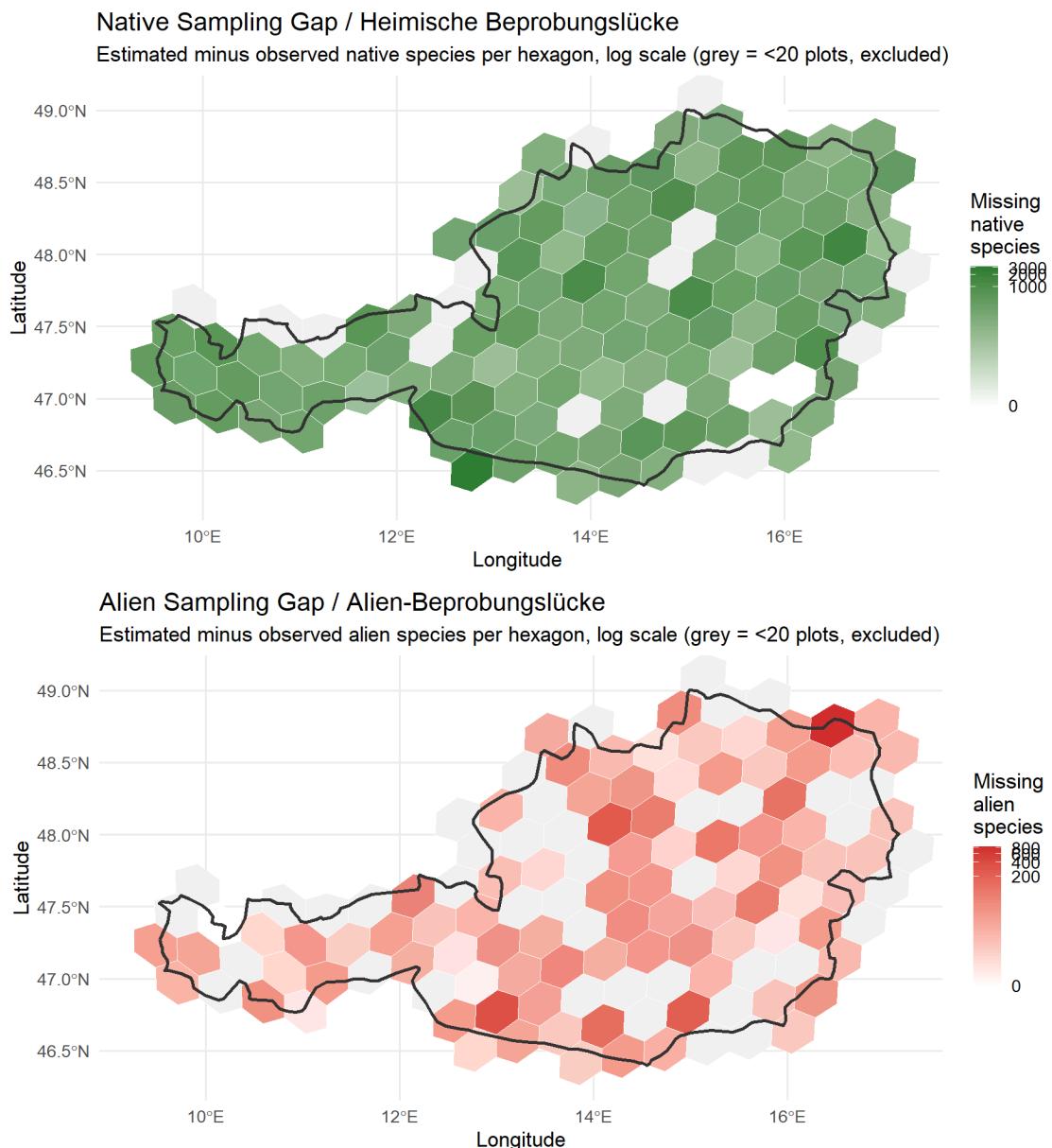


Diese Karten übersetzen die Invasionsschuld in Geografie. Während die erste Karte (geschätzter heimischer Reichtum) die beobachteten Rohzählungskarten aus Abschnitt 1 widerspiegelt, fügt die Alien-Karte Information hinzu: Sie zeigt, wo das Michaelis-Menten-

Modell mehr Arten vorhersagt, als wir tatsächlich gefunden haben. Die Vollständigkeitskarte (unten) hebt hervor, wo die Alien-Flora am wenigsten gut beprobt ist.

3c. Beprobungslücken: Wie viele Arten fehlen uns?

Wie viele Arten fehlen in jedem Hexagon noch? Diese Karten zeigen die **Beprobungslücke**: die Differenz zwischen dem Michaelis-Menten-Schätzwert (S_{max}) und der beobachteten Artenzahl. Dunklere Hexagone haben mehr unerfasste Arten.



Die heimische Lückenkarte zeigt moderate Defizite über weite Österreichs, mit den größten absoluten Lücken in artenreichen Alpentälern und Vorgebirgsregionen. Die Alien-Lückenkarte erzählt eine andere Geschichte: Die größten Lücken konzentrieren sich in Tieflandkorridoren (Donautal, Wiener Becken, östliche Grenzregion), genau den Regionen,

wo neue Einführungen ankommen und die Ausbreitungsverzögerung am längsten ist. Selbst Hexagone mit relativ wenigen beobachteten Aliens können große Lücken aufweisen, was darauf hindeutet, dass die Invasionsschuld geografisch konzentriert und nicht gleichmäßig verteilt ist.

4. Turnover, nicht Nestedness: Was Österreichs Beta-Diversität antreibt

Zerlegung der Beta-Diversität

Baselga (2010) zeigte, dass die Gesamt-Beta-Diversität in zwei Komponenten zerlegt werden kann:

$$\beta_{Sorenson} = \beta_{Turnover} + \beta_{Nestedness}$$

- **Turnover (β_{sim})**: Arten werden zwischen Standorten durch andere Arten ersetzt
- **Nestedness (β_{nes})**: Artenarme Standorte sind Teilmengen artenreicher Standorte

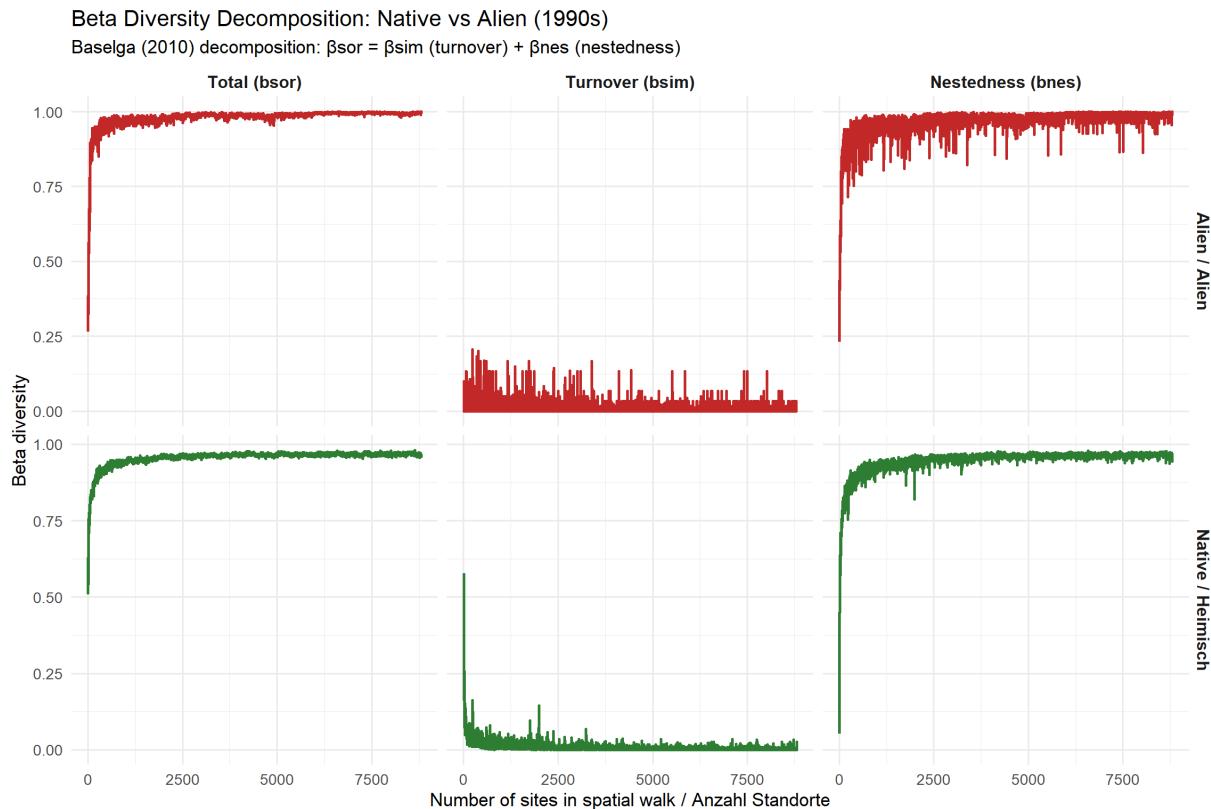
Das Gleichgewicht zwischen diesen Komponenten verrät den *Prozess*, der Diversitätsunterschiede erzeugt: Umweltfilterung (Turnover) versus differenzieller Artenverlust (Nestedness).

Wie man die Grafik liest:

Die x-Achse zeigt die Anzahl der besuchten Standorte in einem "räumlichen Spaziergang" durch Österreich – beginnend an einem zufälligen Standort und schrittweise den nächstgelegenen unbesuchten Standort hinzufügend. Die y-Achse zeigt die Beta-Diversität (Skala 0–1). Es gibt drei Panels, eines für jede Komponente:

- **Total (β_{sor})**: Gesamte kompositorische Unterschiede zwischen allen bisher besuchten Standorten. Steigt schnell an und nähert sich 1, d.h. österreichische Standorte werden sehr unterschiedlich, sobald der Spaziergang genug Umweltgradienten abdeckt.
- **Turnover (β_{sim})**: Der Anteil des Unterschieds, der durch Artenersatz entsteht (Standort A hat Art X, Standort B hat stattdessen Art Y). Wenn dieses Panel hoch ist, haben die Standorte tatsächlich verschiedene Arten.
- **Nestedness (β_{nes})**: Der Anteil, der durch Artenverlust ohne Ersatz entsteht (artenarme Standorte sind nur Teilmengen reicherer Standorte). Wenn dieses Panel niedrig ist, sind die Gemeinschaften keine verschachtelten Versionen voneinander.

Wichtiger Vergleich: Da Total = Turnover + Nestedness, zeigt der Vergleich der Komponenten, warum sich Standorte unterscheiden.



Diskussion

Turnover dominiert bei Einheimischen – aber was ist mit Neophyten?

Die Beta-Diversitäts-Zerlegung zeigt unterschiedliche Muster für **heimische** und **gebietsfremde** Arten. Beide Gruppen sind Turnover-dominiert, aber mit wichtigen Unterschieden.

Einheimische: Turnover ist hoch und Nestedness niedrig. Dies spiegelt Österreichs 3.683-m-Höhengradienten (115 m Neusiedler See bis 3.798 m Großglockner) wider, der verschiedene Vegetationsgürtel mit **tatsächlich verschiedenen Arten** schafft. Ein Tiefland-*Quercus-Carpinus*-Wald und ein hochalpines *Carex curvula*-Grasland teilen praktisch keine Arten.

Neophyten: Auch hier dominiert Turnover, aber die Gesamt-Beta-Diversität ist niedriger als bei Einheimischen, und Nestedness könnte relativ wichtiger sein. Neophyten stammen aus einem kleineren Pool, der sich auf Tiefland- und Störungshabitate konzentriert. Einige neophytenarme Standorte (z.B. alpine Gebiete) sind effektiv Teilmengen der reichereren Neophyten-Flora in Tieflandstädten – ein starker verschachteltes Muster als bei Einheimischen.

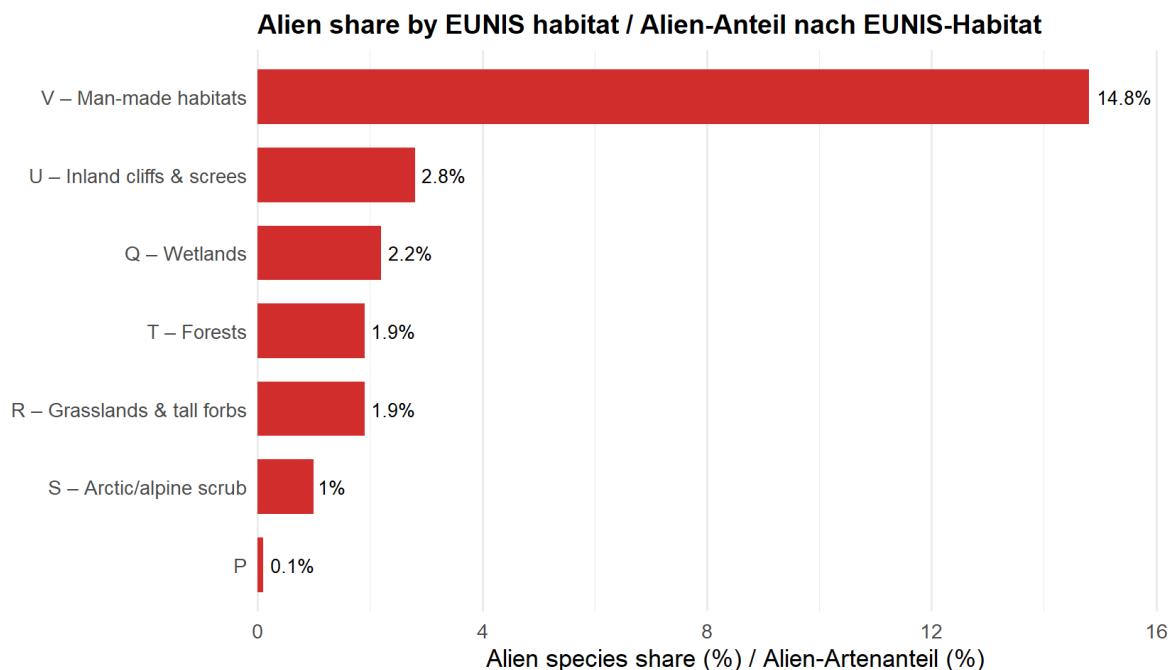
Zentraler Kontrast: Einheimische zeigen hohen Turnover, weil verschiedene Höhenlagen verschiedene Arten beherbergen (Umweltfilterung). Neophyten zeigen Turnover, weil verschiedene Tieflandregionen verschiedene eingeführte Arten erhalten (Einführungsgeschichte), aber auch etwas Nestedness, weil neophytenarme Standorte einfach die Bedingungen für Etablierung fehlen.

Naturschutzimplikationen:

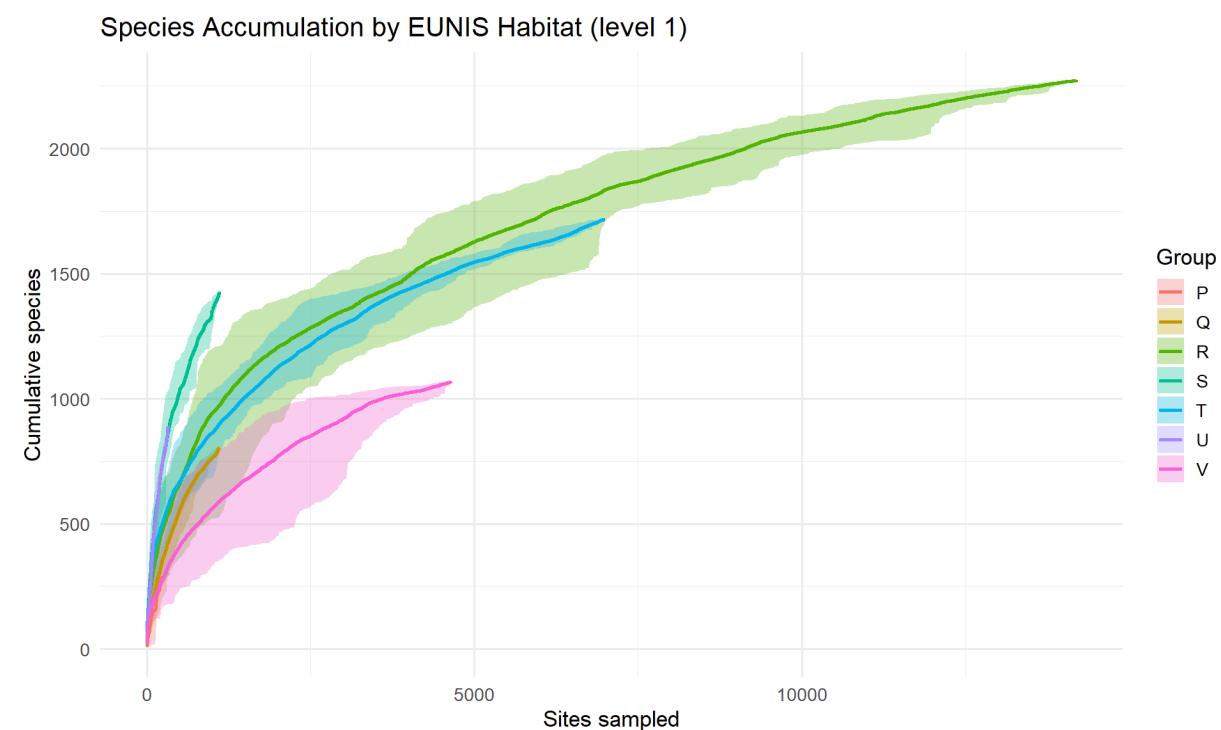
In einem turnover-dominierten System **kann kein einzelnes Schutzgebiet die gesamte Flora erfassen**. Dies unterstützt ein verteiltes Netzwerk von Reservaten über verschiedene Höhenlagen und biogeografische Regionen (die “Several Small”-Position in der SLOSS-Debatte; Tjorve 2010). Für das Neophyten-Management zeigt der Vergleich, dass städtische Tieflandgebiete nicht nur mehr Neophyten beherbergen, sondern auch *verschiedene* Neophyten je nach Region.

5. Habitat-Verwundbarkeit: Wo sich Invasionen konzentrieren

Alien-Artenanteil nach Habitat



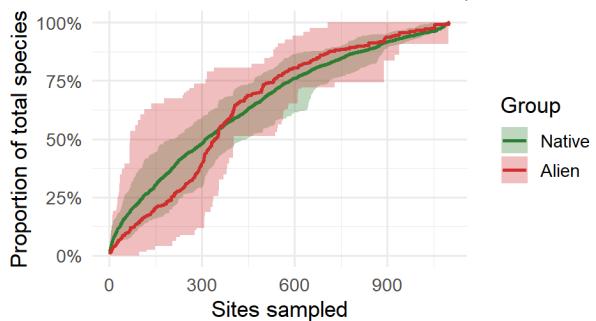
Akkumulationskurven nach Habitat



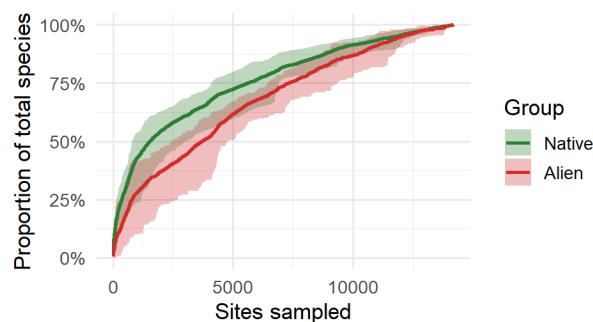
P: Native vs Alien (normalized)



Q – Wetlands: Native vs Alien (normalized)



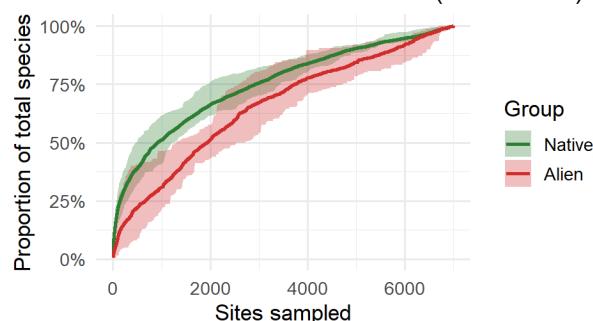
R – Grasslands & tall forbs: Native vs Alien



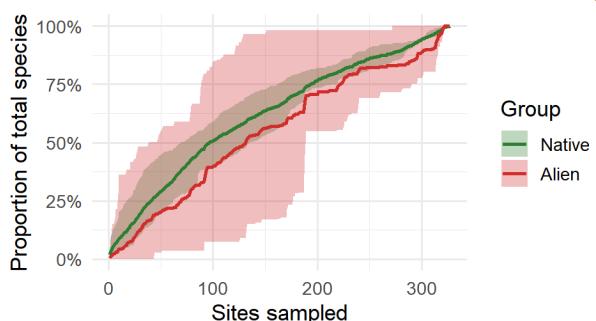
S – Arctic/alpine scrub: Native vs Alien (nor



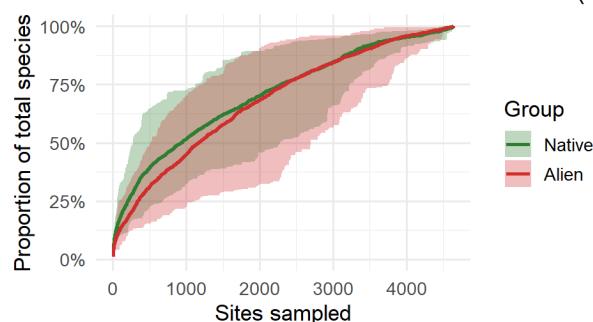
T – Forests: Native vs Alien (normalized)



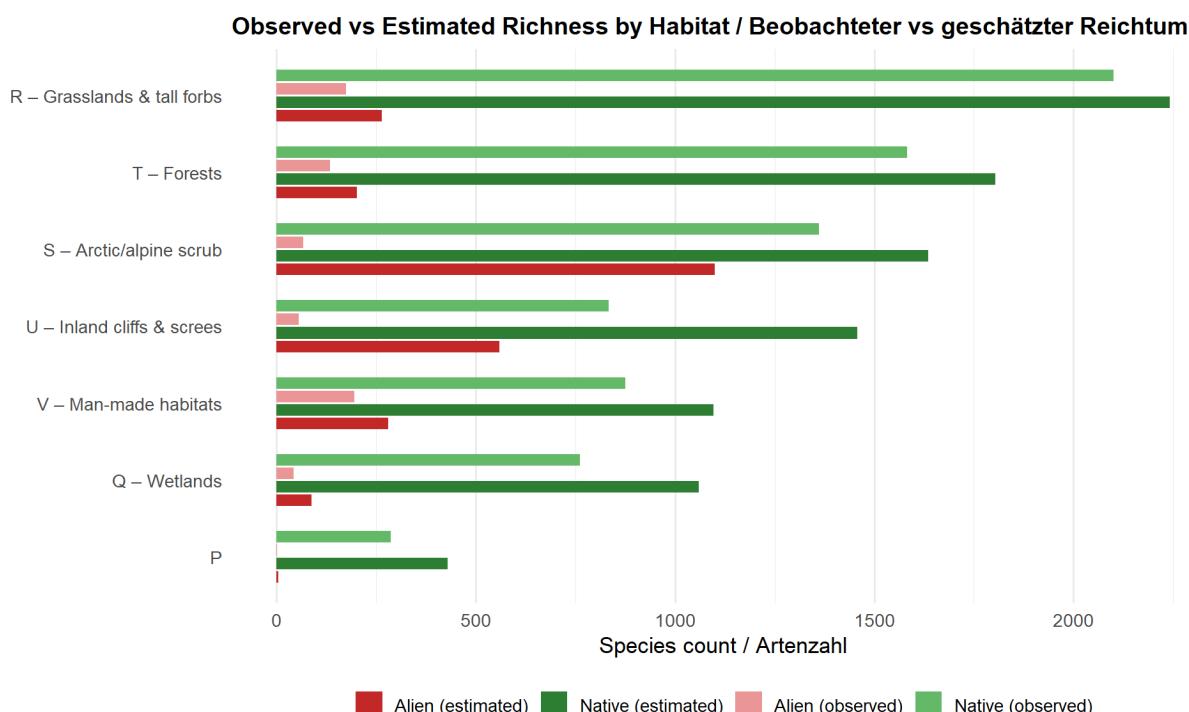
U – Inland cliffs & screes: Native vs Alien (n



V – Man-made habitats: Native vs Alien (normalized)



Geschätzter Artenreichtum nach Habitat (SAC-basiert)



Diskussion

Invasion ist nicht gleichmäßig über Habitate verteilt.

Die EUNIS-Habitat-Analyse zeigt, dass die Invasion durch Alien-Arten sehr heterogen ist. Einige Habitate zeigen Alien-Anteile über 10%, während andere praktisch alien-frei bleiben. Diese Heterogenität folgt klaren ökologischen Regeln.

Welche Habitate sind am stärksten invadiert und warum?

Die am stärksten invadierten Habitate in Österreich teilen mehrere Merkmale:

- **Hohe Störungsfrequenz.** Regelmäßig physisch gestörte Habitate (Pflügen, Mähen, Trampeln, Überflutung) schaffen offene Nischen für Alien-Besiedelung.
- **Hohe Ressourcenverfügbarkeit.** Nährstoffreiche Habitate (gedüngte Grasländer, Flussauen mit Sedimentablagerung) liefern die Ressourcen für schnelle Alien-Etablierung. Die Hypothese der schwankenden Ressourcen (Davis et al. 2000) sagt voraus, dass Invasion während Ressourcenpulsen ihren Höhepunkt erreicht.
- **Starke menschliche Konnektivität.** Habitate entlang von Transportkorridoren erhalten eine kontinuierliche Versorgung mit Alien-Propagules.

Welche Habitate widerstehen der Invasion?

- **Alpine und subalpine Grasländer.** Raue Klimabedingungen schließen die meisten Alien-Arten aus.
- **Oligotrophe Moore und Feuchtgebiete.** Nährstoffarme, wassergesättigte Bedingungen schaffen starke Umweltfilter.
- **Geschlossene Wälder.** Lichtlimitierung verhindert Alien-Etablierung unter dichtem Kronendach.

Die 100-fache Variation:

Chytry et al. (2008) zeigten, dass europäische Habitate sich 100-fach in ihrer Anfälligkeit für Pflanzeninvasionen unterscheiden. Unsere österreichischen Daten bestätigen dies:

Habitattyp ist der primäre Determinant der Invasionsanfälligkeit, wichtiger als Klima, Geografie oder Propagule Pressure.

Akkumulationskurvenformen unterscheiden sich:

- In stark invadierten Habitaten nähert sich die Alien-Kurvenform der heimischen an — Aliens haben sich breit verteilt.
- In schwach invadierten Habitaten ist die Alien-Kurve steiler und unregelmäßiger — das Zeichen einer Frühphasen-Invasion.

Praktische Anwendung für den Naturschutz:

1. **Hohe Priorität:** Am stärksten invadierte Habitattypen überwachen und managen.
2. **Mittlere Priorität:** Habitate an der Übergangszone beobachten — die "Invasionsfronten", wo Management am kosteneffektivsten sein kann.
3. **Niedrigere Priorität:** Alpine und oligotrophe Habitate, die derzeit Invasion widerstehen, langfristig monitoren, besonders angesichts des Klimawandels.

6. Welche Art von Diversität? Dominanz, Partitionierung und evolutionäre Breite

Bisher haben wir Arten gezählt. Aber nicht alle Arten tragen gleich bei: Einige sind häufig, andere selten; einige repräsentieren einzigartige evolutionäre Linien, andere sind nahe Verwandte bereits vorhandener Arten. Dieser Abschnitt stellt drei Fragen, die über reine Artenzahlen hinausgehen:

1. **Treiben seltene oder häufige Arten die Akkumulation?** (Hill-Zahlen)
2. **Kommt Österreichs Diversität hauptsächlich aus dem Artenreichtum innerhalb einzelner Standorte oder aus den Unterschieden zwischen Standorten?**
(Alpha/Beta/Gamma-Partitionierung)
3. **Bringen Aliens evolutionäre Neuheit, oder nur mehr vom Gleichen?** (Phylogenetische Diversität)

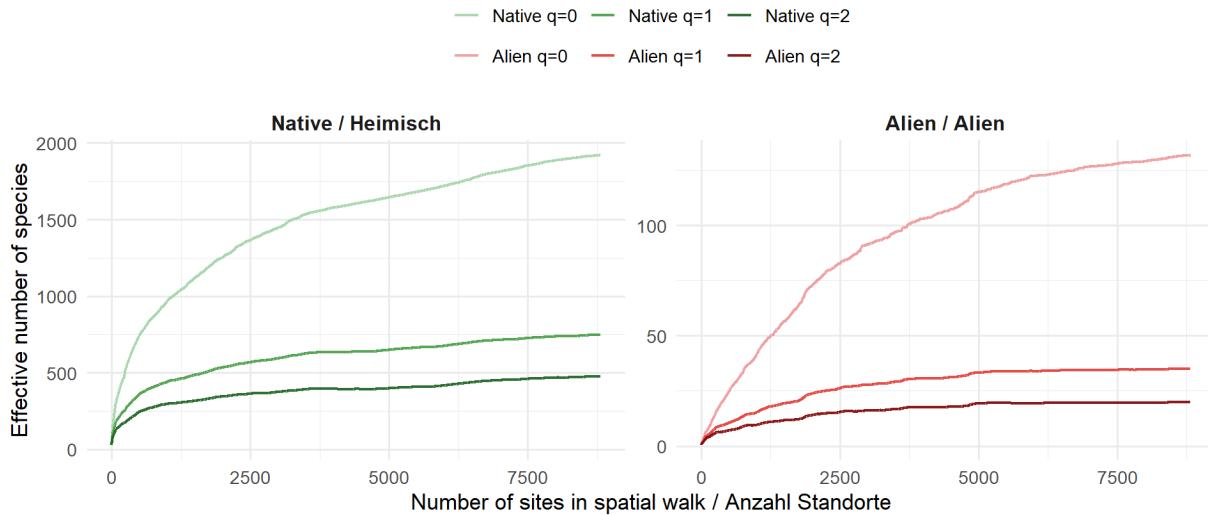
6a. Seltene Arten treiben das Ende: Hill-Zahlen

Hill-Zahlen gewichten Arten unterschiedlich nach ihrer Häufigkeit:

- $q = 0$: Alle Arten gleich gewichtet (= Artenreichtum)
- $q = 1$: Arten nach Frequenz gewichtet (= Exponential der Shannon-Entropie)
- $q = 2$: Dominante Arten zählen am meisten (= inverser Simpson)

Hill Numbers: Native vs Alien (1990s)

Lighter = all species equal; darker = dominant species weighted more



6b. Wo steckt die Diversität? Alpha vs Beta

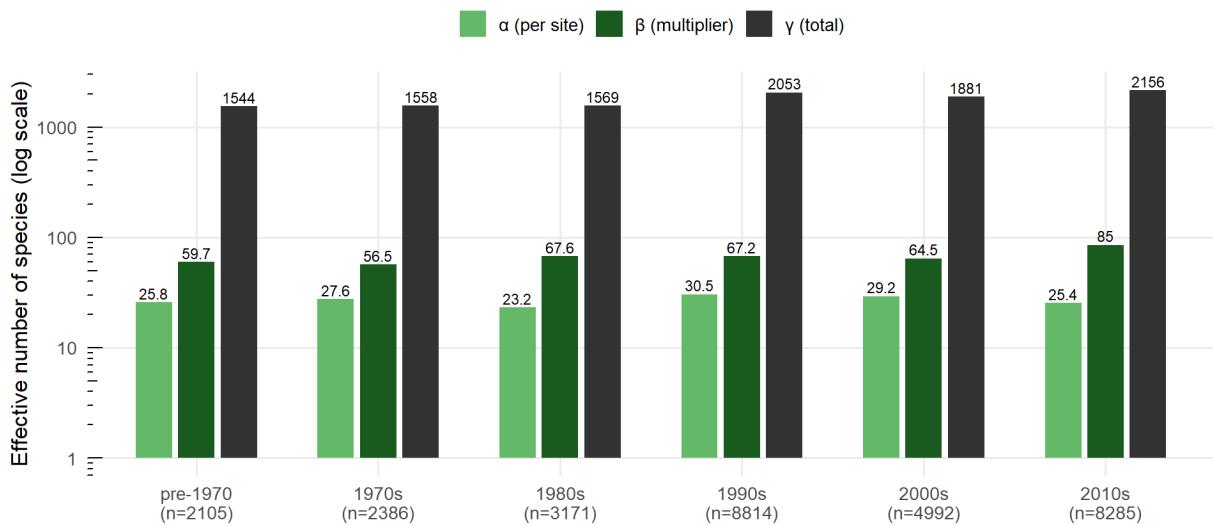
Die Gesamtdiversität (**Gamma**, γ) lässt sich in zwei Komponenten zerlegen:

- **Alpha** (α): die durchschnittliche Artenzahl an einem einzelnen Standort
- **Beta** (β): wie oft man Alpha multiplizieren muss, um Gamma zu erreichen ($\beta = \gamma/\alpha$)

Ein hoher Beta-Wert bedeutet, dass sich Standorte stark unterscheiden; ein niedriger, dass alle Standorte ähnlich aussehen.

Diversity Partitioning by Decade / Diversitätspartitionierung nach Jahrzehnt

$\gamma = \alpha \times \beta$ at $q=0$ (species richness); β rising over time = sites becoming more different

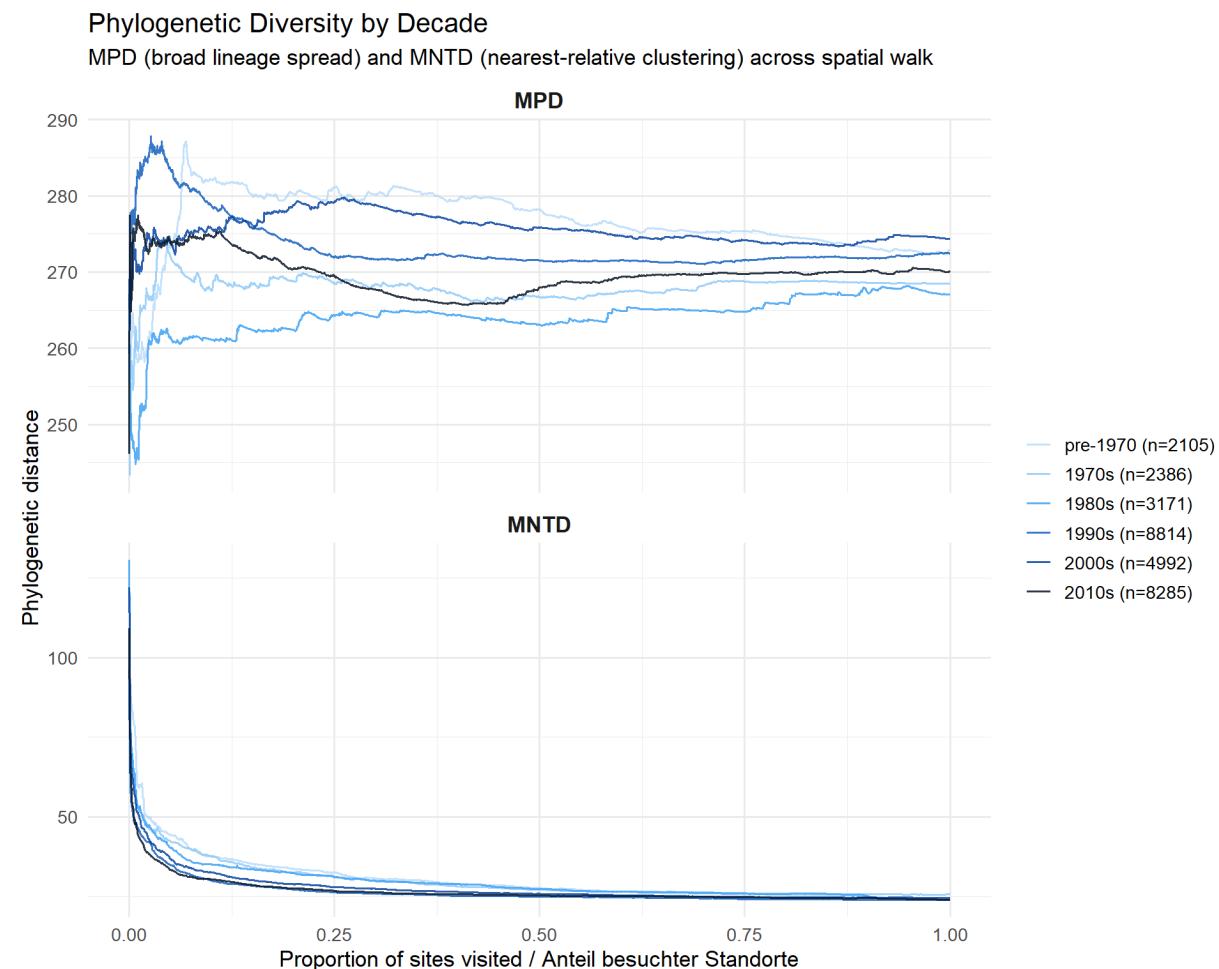


6c. Bringen Aliens evolutionäre Neuheit?

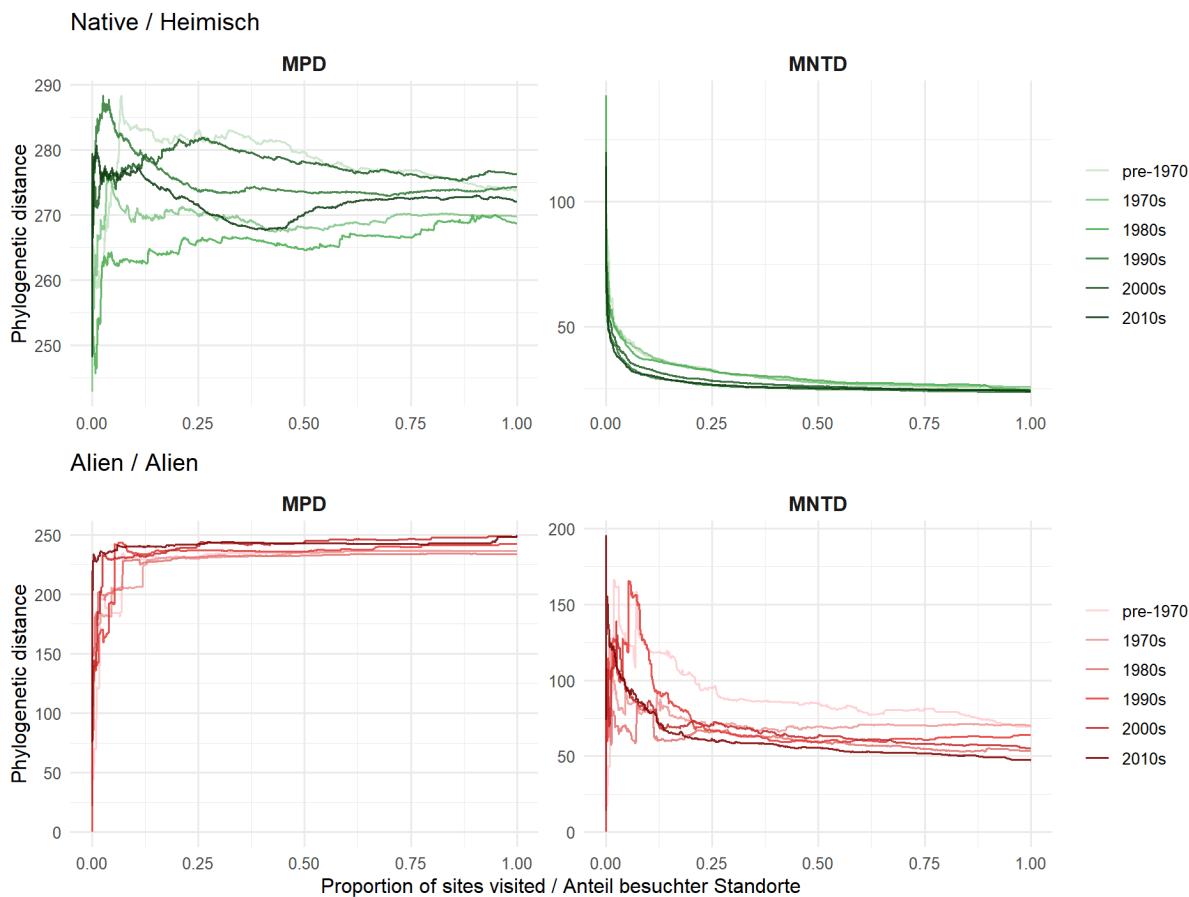
Phylogenetische Diversität misst, wie evolutionär entfernt Arten voneinander sind. Zwei komplementäre Metriken erfassen verschiedene Aspekte:

- **Mittlere paarweise Distanz (MPD)** verfolgt die durchschnittliche evolutionäre Distanz zwischen allen Artenpaaren. Hoher MPD bedeutet, dass die Gemeinschaft viele verschiedene Pflanzenlinien umfasst (breite phylogenetische Spreizung).
- **Mittlere nächste Taxon-Distanz (MNTD)** verfolgt die durchschnittliche Distanz zum nächsten Verwandten jeder Art. Niedriger MNTD bedeutet, dass Arten in eng verwandten Gruppen clustern (phylogenetisches Clustering an den Spitzen).

Zusammen zeigt MPD den Wald und MNTD die Bäume. Der Vergleich zwischen heimischen und Alien-Werten zeigt, ob Aliens das evolutionäre Portfolio erweitern oder verengen.



Phylogenetic Diversity: Native vs Alien by Decade
 MPD (broad lineage spread) and MNTD (nearest-relative clustering) across spatial walk



Diskussion

Die drei Analysen ergeben ein einheitliches Bild:

Die meisten Arten, die wir noch “finden”, sind selten. Die Hill-Zahlen-Plots zeigen, dass die $q = 0$ -Kurve (Artenreichtum) weiter ansteigt, lange nachdem $q = 2$ (dominanzgewichtete Diversität) abflacht. Die späte Akkumulation fügt Arten hinzu, die nur an wenigen Standorten vorkommen. Dieser Effekt ist bei Aliens proportional größer: Wenige weitverbreitete Generalisten dominieren, der Rest sind seltene, lokalisierte Neuankömmlinge.

Die Diversität steckt zwischen den Standorten, nicht innerhalb. Die Alpha/Beta/Gamma-Partitionierung bestätigt, was Abschnitt 4 bereits zeigte: Beta-Diversität (Turnover zwischen Standorten) macht den Großteil von Österreichs Gesamtdiversität (Gamma) aus.

Aliens sind evolutionäre Nachahmer. Der heimische MPD ist durchgehend höher als der Alien-MPD, d.h. die heimische Flora umfasst eine breitere Palette von Pflanzenlinien. Aliens konzentrieren sich phylogenetisch auf wenige Familien (Asteraceae, Poaceae, Brassicaceae), die Merkmale teilen, die Kolonisierung begünstigen: schnelles Wachstum, hohe

Samenproduktion und Störungstoleranz (Cadotte et al. 2009). Die ~88 "fehlenden" Alien-Arten aus Abschnitt 3 dürften daher eher dieselben Funktionstypen sein, nicht Vertreter neuer Linien.

Was das bedeutet: Artenzählung allein überschätzt die ökologische Wirkung von Alien-Ankünften. Die Invasionsschuld ist real, aber die kommenden Arten werden wahrscheinlich wenig funktionelle oder evolutionäre Neuheit zu Österreichs Pflanzengemeinschaften beitragen.

Synthese

Diese sechs Ergebnisse bilden ein Gesamtbild über Österreichs Pflanzendiversität und ihre Invasion durch Alien-Arten:

1. Beprobung & beobachteter Reichtum. Die Hotspot-Karten zeigen, wo beprobt wurde und wie der beobachtete Artenreichtum aussieht. Einheimische erreichen ihren Höhepunkt in den topografisch komplexen Voralpen; Aliens konzentrieren sich um menschliche Infrastruktur im Tiefland. Diese Muster spiegeln teilweise den Beprobungsaufwand wider.

2. Räumliche Akkumulation. Die normalisierten Kurven zeigen, dass Aliens eine höhere räumliche Beta-Diversität haben als Einheimische.

3. Unvollständige Beprobung. Die Extrapolation zeigt, dass wir ~79% der Alien-Flora kennen gegenüber ~94% der Einheimischen. Die Hexagon-Karten lokalisieren, wo diese Schuld konzentriert ist.

4. Turnover, nicht Nestedness. Österreichs hohe Beta-Diversität entsteht durch echten Artenersatz entlang des Höhengradienten. Kein einzelnes Reservat kann die gesamte Flora erfassen.

5. Habitatfilter. Alien-Invasion konzentriert sich in gestörten, nährstoffreichen Habitateinheiten und wird aus rauen, oligotrophen Umgebungen ausgeschlossen.

6. Welche Art von Diversität? Seltene Arten treiben das Ende der Akkumulationskurve, Beta-Diversität dominiert Gamma, und Aliens konzentrieren sich phylogenetisch auf wenige Familien. Die Invasionsschuld ist zahlenmäßig real, aber die kommenden Arten dürften wenig evolutionäre Neuheit bringen.

Zusammen zeigen diese Ergebnisse, dass **räumliche Artenakkumulationskurven** weit mehr als eine Beprobungsmethode sind.

- Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19, 134–143.
- Cadotte, M.W. et al. (2009). Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PLoS ONE*, 4, e5695.
- Chao, A. & Jost, L. (2012). Coverage-based rarefaction and extrapolation. *Methods in Ecology and Evolution*, 3, 217–226.
- Chytry, M. et al. (2008). Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. *Journal of Applied Ecology*, 45, 448–458.
- Chytry, M. et al. (2009). European map of alien plant invasions based on the quantitative assessment across habitats. *Diversity and Distributions*, 15, 98–107.
- Chytry, M. et al. (2016). European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science*, 19, 173–180.
- Colwell, R.K. & Coddington, J.A. (1994). Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society B*, 345, 101–118.
- Davis, M.A. et al. (2000). Fluctuating resources in plant communities: a general theory of invasibility. *Journal of Ecology*, 88, 528–534.
- Essl, F. et al. (2011). Socioeconomic legacy yields an invasion debt. *Proceedings of the National Academy of Sciences*, 108, 203–207.
- Grabherr, G. et al. (2003). *Die Pflanzengesellschaften Österreichs*. Gustav Fischer Verlag.
- Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54, 427–432.
- Jost, L. (2006). Entropy and diversity. *Oikos*, 113, 363–375.
- Kier, G. et al. (2009). A global assessment of endemism and species richness across island and mainland regions. *PNAS*, 106, 9322–9327.
- Lockwood, J.L. et al. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, 20, 223–228.
- Partel, M. et al. (2011). Dark diversity: shedding light on absent species. *Trends in Ecology & Evolution*, 26, 596–600.
- Pysek, P. et al. (2010). Disentangling the role of environmental and human pressures on biological invasions across Europe. *PNAS*, 107, 12157–12162.
- Rahbek, C. (1995). The elevational gradient of species richness: a uniform pattern? *Ecography*, 18, 200–205.
- Seebens, H. et al. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, 8, 14435.
- Socolar, J.B. et al. (2016). How should beta-diversity inform biodiversity conservation? *Trends in Ecology & Evolution*, 31, 67–80.
- Tjørve, E. (2010). How to resolve the SLOSS debate: lessons from species-diversity models. *Journal of Theoretical Biology*, 264, 604–612.
- Wilson, J.R.U. et al. (2007). Residence time and potential range: crucial considerations in modelling plant invasions. *Diversity and Distributions*, 13, 11–22.

Session Info

```
## R version 4.5.2 (2025-10-31 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 11 x64 (build 26200)
##
## Matrix products: default
## LAPACK version 3.12.1
##
## locale:
## [1] LC_COLLATE=English_United States.utf8
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.utf8
##
## time zone: Europe/Luxembourg
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics   grDevices utils      datasets   methods    base
##
## other attached packages:
## [1] patchwork_1.3.2     hexify_0.3.10      rnaturalearth_1.2.0
## [4] sf_1.0-24           spacc_0.1.0       data.table_1.18.0
## [7] ggplot2_4.0.1
##
## loaded via a namespace (and not attached):
## [1] gtable_0.3.6        jsonlite_2.0.0      dplyr_1.1.4
## [4] compiler_4.5.2       tidyselect_1.2.1    Rcpp_1.1.1
## [7] parallel_4.5.2       gridExtra_2.3       jquerylib_0.1.4
## [10] scales_1.4.0         yaml_2.3.12        fastmap_1.2.0
## [13] R6_2.6.1            labeling_0.4.3     generics_0.1.4
## [16] classInt_0.4-11     knitr_1.51         tibble_3.3.1
## [19] units_1.0-0          DBI_1.2.3         bslib_0.9.0
## [22] pillar_1.11.1       RColorBrewer_1.1-3 rlang_1.1.7
## [25] cachem_1.1.0        xfun_0.55          sass_0.4.10
## [28] S7_0.2.1            RcppParallel_5.1.11-1 otel_0.2.0
## [31] cli_3.6.5           withr_3.0.2        magrittr_2.0.4
## [34] class_7.3-23         digest_0.6.39      grid_4.5.2
## [37] lifecycle_1.0.5      vctrs_0.7.0        KernSmooth_2.23-26
## [40] proxy_0.4-29        evaluate_1.0.5    glue_1.8.0
## [43] farver_2.1.2         rnaturalearthdata_1.0.0 e1071_1.7-17
## [46] rmarkdown_2.30        tools_4.5.2        pkgconfig_2.0.3
## [49] htmltools_0.5.9
```