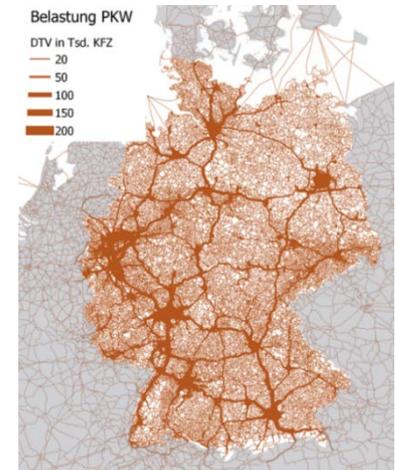
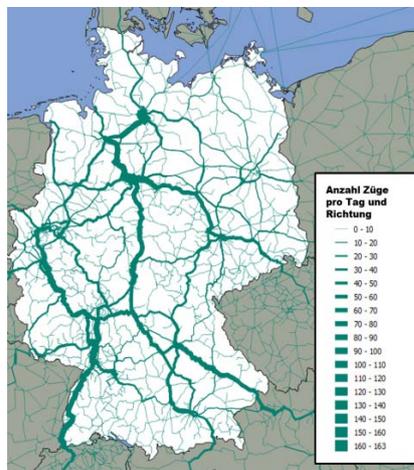
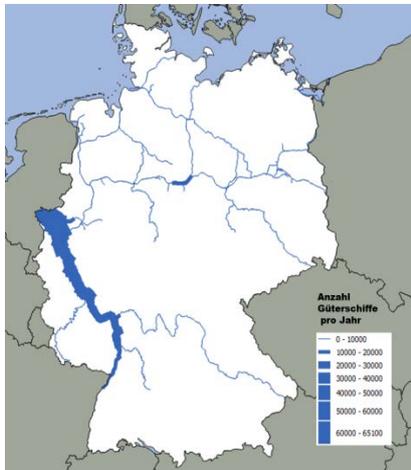
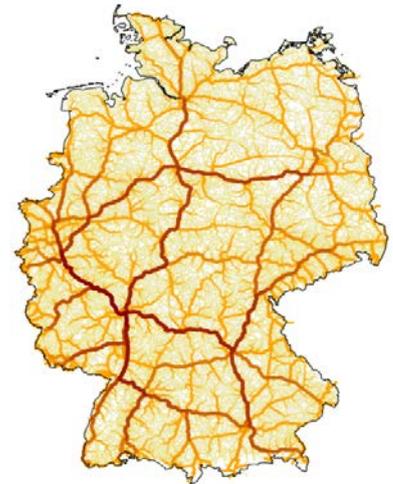
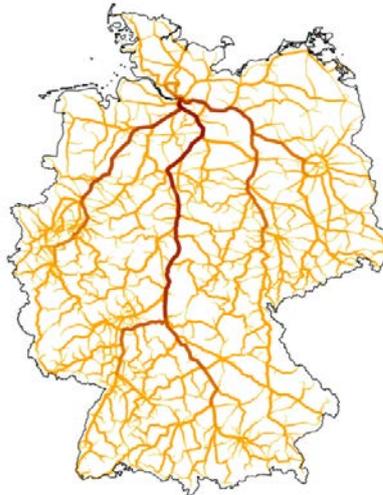
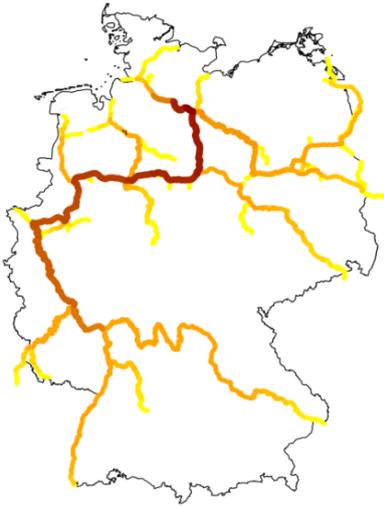




Berichte
des Deutschen Zentrums
für Schienenverkehrsforschung

Bericht 5 (2020)

Modellierung der Einfuhr und Verbreitung von invasiven Arten durch Verkehrsträger



Berichte des Deutschen Zentrums
für Schienenverkehrsforschung, Nr. 5 (2020)
Projektnummer 2017-U-10-1210

Modellierung der Einfuhr und Verbreitung von invasiven Arten durch Verkehrsträger

von

Hanno Seebens, Maurizio Bagnara
Senckenberg Biodiversität und Klima Forschungszentrum, Frankfurt

Robert Burg, Stefan Schrempp
SSP Consult Beratende Ingenieure GmbH, Waldkirch

Im Auftrag des Deutschen Zentrums für Schienenverkehrsforschung beim Eisenbahn-Bundesamt

Impressum

HERAUSGEBER

Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt

August-Bebel-Straße 10
01219 Dresden

www.dzsf.bund.de

DURCHFÜHRUNG DER STUDIE

Senckenberg Biodiversität und Klima Forschungszentrum
Senckenberganlage 25
60325 Frankfurt

ABSCHLUSS DER STUDIE

April 2019

REDAKTION

Eisenbahn-Bundesamt
Dr. Marion Leiblein-Wild, Referat Umwelt

TITELBILD

Projizierte Wahrscheinlichkeit der Ausbreitung und Etablierung von *Senecio inaequidens* basierend auf den Projektergebnissen.

PUBLIKATION ALS PDF

<https://www.dzsf.bund.de/Forschungsergebnisse/Forschungsberichte>

ISSN 2629-7973

[doi: 10.48755/dzsf.210013.01](https://doi.org/10.48755/dzsf.210013.01)

Dresden, November 2020

Inhaltsverzeichnis

Kurzbeschreibung	7
Abstract	9
Glossar	10
1 Einleitung	12
2 Methodik	14
2.1 Ermittlung der verfügbaren Datengrundlagen und Datenbeschaffung (AP 1)	14
2.2 Datenaufbereitung (AP 2)	17
2.2.1 Verkehrsmodellierung	18
2.2.2 Habitatcharakteristiken	25
2.2.3 Verbreitungsdaten	27
2.3 Modellentwicklung (AP 3)	30
2.3.1 Überblick	30
2.3.2 Modellstruktur	31
2.3.3 Kalibrierung und Validierung des Modells	40
2.4 Modellannahmen	41
3 Ergebnisse und Diskussion	44
3.1 Senecio inaequidens	44
3.2 Corbicula fluminea	47
3.3 Dikerogammarus villosus	49
3.4 Spezifische Fragestellungen	51
3.4.1 Einfallstore der Invasion	51
3.4.2 Vektoren	52
3.4.3 Mögliche Umschlagpunkte	52
3.4.4 Ausbreitungsrouten	52
3.4.5 Geschwindigkeiten der Ausbreitungen	53
3.4.6 „Loads“ und Schwerpunkte der Invasion	54
3.5 Analyse der Modellergebnisse	55
4 Weitere Schritte	62
4.1 Wissens- und Datenlücken	62
4.2 Ausblick	63
5 Abbildungsverzeichnis	66

6	Tabellenverzeichnis.....	68
7	Quellenverzeichnis	69
8	Anhänge	71
8.1	Detailergebnisse der Kalibrierung	71
8.1.1	Senecio inaequidens.....	72
8.1.2	Corbicula fluminea	74
8.1.3	Dikerogammarus villosus.....	75
8.2	Handbuch CASPIAN	76

Kurzbeschreibung

Das Ziel dieser Studie ist es, die Ausbreitung und Etablierung von gebietsfremden Arten (Neobiota) durch Verkehrsträger in Deutschland vorherzusagen und zu quantifizieren. Hierzu wurde ein Computermodell (CASPIAN) entwickelt, welches die Ausbreitung und Etablierung von Neobiota über die Verkehrsträger Straße, Schiene und Binnenwasserstraße deutschlandweit simuliert. Das Modell berücksichtigt eine Vielzahl von Ausbreitungswegen, die mit Verkehrsträgern und Verkehr assoziiert sind. Dies beinhaltet i) die Nutzung von Infrastrukturelementen als Wanderkorridore, ii) die Ausbreitung über die Anheftung an Transportmittel wie Personenkraftwagen, Lastkraftwagen, Zügen und Schiffen, iii) die Ausbreitung durch Fahrtwind und iv) die Ausbreitung in Containern und Paletten. Das Computermodell ist modular aufgebaut, so dass einzelne Ausbreitungswege durch den Nutzer ein- bzw. ausgeschaltet werden können.

Die Grundlage der Computersimulationen bilden umfangreiche Datensätze zu Verkehrsflüssen, die aus routenfähigen Netzmodellen zum Verkehrsträger Straße entnommen wurden und aus Originaldaten zu Strömen des Schienenverkehrs, der Binnenschifffahrt und des Güterverkehrs. Dies ermöglicht eine detaillierte Darstellung des deutschlandweiten Verkehrsaufkommens für Personen- und Güterverkehr auf verschiedenen hierarchischen Ebenen. Basierend auf den Verkehrs- und Güterdaten wird die Wahrscheinlichkeit der Ausbreitung einer Art verkehrsträgerübergreifend für jeden Ausbreitungsweg separat berechnet. So wurde z.B. die Ausbreitung durch Fahrtwind über den kombinierten Schienen- und Straßenverkehr ermittelt. Güterströme werden über alle drei Verkehrsträger berücksichtigt. Die Gesamtwahrscheinlichkeit einer Ausbreitung ergibt sich aus den Ergebnissen aller berücksichtigten Wege.

Die Möglichkeit der Etablierung einer neuen Art an einem neuen Standort wird mit Hilfe von Datensätzen zur Charakterisierung der lokalen Gegebenheiten und Arteigenschaften abgeschätzt. Hierzu werden hochaufgelöste Landbedeckungsdaten zur Charakterisierung der terrestrischen Habitate und Informationen der deutschlandweiten Netzwerke der Pegelmessstationen für die aquatischen Habitate verwendet. Das Modell wurde mit drei Datensätzen zur beobachteten Ausbreitung von Neobiota in Deutschland kalibriert und validiert. Das kalibrierte Modell erlaubt die Berechnung der Ausbreitungswahrscheinlichkeit von jedem Standort der Verkehrsinfrastruktur in Deutschland über die entsprechenden Verkehrsträger. Hierbei ermöglicht das Modell sowohl die Simulation der Ausbreitung einzelner Arten als auch die Quantifizierung und Darstellung der generellen Haupttrouten der Ausbreitung und Schwerpunkte der Vorkommen in Deutschland.

Die Modellergebnisse bilden die Basis für eine weiterführende statistische Analyse, bei der die häufigsten Ausbreitungsrouten innerhalb Deutschlands über verschiedene Transportnetzwerke und von verschiedenen Startpunkten ausgehend ermittelt werden. Dies ermöglicht eine Priorisierung wichtiger Ausbreitungsrouten für Neobiota innerhalb Deutschlands.

Die Ergebnisse der Analyse werden in Bild und Schrift dargestellt. Darüber hinaus werden Lücken der Datenverfügbarkeit aufgezeigt, die eine Implementierung von Ausbreitungswegen im Rahmen dieses Forschungsprojektes erschweren bzw. nicht möglich machten. Hier besteht besonderer Handlungsbedarf bei der Erfassung von Vorkommen der Arten in Deutschland. Diese Daten liegen vor allem für den terrestrischen Bereich nur sehr unvollständig vor. Darüber hinaus werden Möglichkeiten aufgezeigt, wie das Modell erweitert und sinnvoll weiterentwickelt werden kann.

Das Computermodell liegt als Softwarepaket in der frei verfügbaren Softwareumgebung R zusammen mit einem Handbuch zur Anwendung vor und wird nach Abschluss des Projekts öffentlich verfügbar sein.

Abstract

The aim of this study was to quantify and predict the spread and the establishment of non-native species (neobiota) along German infrastructure and by traffic flows. For this purpose, a computer model (CASPIAN) was developed to simulate both the spread and the establishment of neobiota Germany-wide by traffic (roads, railways and inland canals). The model considers a multitude of pathways of spread related to traffic: i) natural migration along infrastructure elements, ii) attachment to vehicles such as cars, trucks, trains and ships, iii) dispersal through air stream and iv) dispersal in containers and wooden pallets. The computer model exhibits a modular structure, which allows the user to switch individual pathways of dispersal on and off, respectively.

The basis for the simulations represents comprehensive data sets of traffic flows obtained from traffic models, which are capable of simulating routes of vehicle movements on roads, and original data for railways and inland shipping, and the respective intensities of traffic and commodity flows. These data sets allowed a detailed representation of nation-wide flows of passenger and freight transport along different hierarchical levels of organisation. This information was utilised to calculate probabilities of spread across traffic networks for each dispersal pathway individually. For example, the spread via air stream was calculated for the combined network of roads and railways. Commodity flows are considered across all three types of traffic networks. The total probability of spread is obtained from the results of all selected dispersal pathway.

The potential establishment of a new species at a new site is calculated based on data sets of local environmental conditions and species preferences. High-resolution remote sensing products of land cover types have been used to characterise terrestrial habitats, while information from individual monitoring sites of the nation-wide network of permanent river measuring stations has been used for the aquatic habitats. The model was calibrated and validated with three different data sets of reported spreading dynamics of neobiota in Germany.

The calibrated model enables the calculation of probabilities of spread starting from any point of an infrastructure in Germany along the selected traffic network. The model is capable of simulating the spread of individual species and the quantification of major routes and loads of spread as well as hot spots of occurrence in Germany. Simulation results were investigated in an analysis of shortest paths to determine the main routes of spread within Germany for both across transport network and for individual networks from various starting points.

Study results are presented as figures and text. Major knowledge gaps, which hindered a more detailed implementation and calibration of additional pathways, are highlighted and discussed. The most severe limitation for model developed, which was identified in this project, was the lack of comprehensive high-quality data of species occurrence particularly in the terrestrial realm. Potential ways forwards to improve and extend the model are presented and discussed.

The model is implemented as an open source software package in the statistical language R, which allows other users to apply, test and extend the model. Together with a manual, the model will publicly accessible through an online repository. Potential extensions and next steps are discussed at the end of this report.

Glossar

- Neobiota:** Neobiota, oder auch gebietsfremde Arten, stellen Organismen dar, die durch den Menschen in Gebiete außerhalb ihres ursprünglichen Verbreitungsgebiets gelangten (Blackburn et al., 2011). Dies kann bewusst oder unbewusst erfolgen. Der Begriff Neobiota umfasst alle Arten, die nach 1500 eingeführt wurden.
- Invasive Arten:** Invasive Arten stellen Neobiota dar, deren Einführung zu negativen Auswirkungen führten (Simberloff et al., 2013). Die Auswirkungen umfassen ein weites Spektrum an möglichen Folgen wie z.B. die Verdrängung einheimischer Arten oder Beeinträchtigung der Wirtschaft, der Gesellschaft oder der Gesundheit.
- Ausbreitungsweg:** Ein Ausbreitungsweg ist der Mechanismus, über den eine Art von einem Ort zum nächsten gelangt. Dies kann passiv geschehen, wie z.B. über den Transport mit einem Fahrzeug oder durch Fahrtwind, oder aktiv, wie z.B. die natürliche Ausbreitung entlang eines Infrastrukturelements (z.B. Straßenabschnitt).
- Link:** Ein Link oder auch Kanten ist ein Bestandteil eines Netzwerks und bezeichnet die Verbindung zwischen zwei Knoten (s. Knoten). In diesem Forschungsvorhaben stellen Links Straßen-, Schienen- oder Bundeswasserstraßensegmente dar. Ein Link weist immer einheitliche Charakteristiken auf. So ändert sich die z.B. die Verkehrsdichte oder die Geschwindigkeitsbegrenzung entlang eines Links nicht.
- Knoten:** Ein Knoten ist ein Bestandteil eines Netzwerks und bezeichnet Berührungspunkte mehrerer Links (s. Link). In den Verkehrsnetzwerken stellt ein Knoten z.B. eine Kreuzung dar. Ein Knoten kann aber auch zwischen zwei einzelnen Links liegen, wenn sich Charakteristiken der Links, wie z.B. Geschwindigkeitsbeschränkungen, ändern.
- Netzwerk:** Ein Netzwerk besteht aus mehreren Knoten, die über Links miteinander verbunden sind. In diesem Forschungsvorhaben sind dies die Verkehrsnetzwerke.
- Ströme:** Ströme bezeichnen Informationen zur gerichteten Bewegung von Einheiten (z.B. Güter, Personen, Fahrzeuge) unter Angaben von Quelle, Ziel und Anzahl der Einheiten.
- „Dispersal kernel“:** Ein dispersal kernel beschreibt die Wahrscheinlichkeit der Ausbreitung eines Individuums einer Art über eine gewisse Distanz. In der Regel werden kontinuierliche Funktionen verwendet, die die Verteilung der Wahrscheinlichkeit als Funktion der geographischen Distanz beschreiben.
- Kürzester Pfad („shortest path“):** Ein kürzester Pfad bezeichnet in einem Netzwerk den kürzesten Weg zwischen zwei Knoten. Häufig wird dies über die geographische Distanz aller Links berechnet und kann dann auch als kürzeste Route bezeichnet werden. Es können aber auch andere Variablen zur Berechnung der kürzesten Pfade verwendet werden wie z.B. die Verkehrsintensitäten oder Geschwindigkeitsbegrenzung, was häufig als schnellste Route bezeichnet wird. In diesem Forschungsvorhaben werden die kürzesten Pfade über die Wahrscheinlichkeit der Ausbreitung berechnet. Der „kürzeste Pfad“ stellt damit den Weg zwischen zwei Knoten dar, der insgesamt über alle Links die höchste Wahrscheinlichkeit der Ausbreitung von allen möglichen Wegen hat.

- Modellgüte:** Die Modellgüte beschreibt ein Maß, mit dem die Genauigkeit eines Modells beschrieben wird, mit der es Messdaten vorhersagen kann. Zum Beispiel ist die Modellgüte hoch, wenn ein Modell sehr gut die beobachteten Vorkommen einer Art reproduzieren kann. Umgekehrt ist die Modellgüte niedrig, wenn das Modell das Vorkommen der Art nicht vorhersagen kann.
- Konvergenz:** Allgemein beschreibt Konvergenz das Zusammenlaufen von Linien. In diesem Forschungsvorhaben wird damit die Annäherung der Modellvorhersagen an die beobachteten Werte der jeweiligen Datensätze beschrieben. Die Modellvorhersagen ändern sich bei Verwendung unterschiedlicher Modellparameter. Beim Anpassen der Modellvorhersagen an die beobachteten Vorkommen der Arten werden die Modellparameter so lange variiert, bis sich keine Verbesserung der Modellgüte mehr einstellt. Wenn sich die Modellgüte über viele neu ausgewählte Parameterkombinationen nicht mehr verbessert, spricht man von Konvergenz.
- Kalibrierung:** Die Kalibrierung beschreibt die Auswahl der Modellparameter anhand eines Vergleichs der Modellvorhersagen mit Messdaten. Bei der Kalibrierung wird für unterschiedliche Parametersätze die Modellgüte ermittelt und der Parametersatz ausgewählt, der die beste Modellgüte ergibt. Dies wird in einem iterativen Verfahren so lange wiederholt, bis die Modellgüte konvergiert (s. Konvergenz), was bedeutet, dass sich die Modellgüte nicht weiter verbessern lässt.
- Validierung:** Die Validierung bezeichnet den Test eines kalibrierten Modells. Die Validierung findet daher nach der Ermittlung der Werte der Modellparameter (s. Kalibrierung) statt. Zur Validierung eines Modells wird die Modellgüte idealerweise mit Messdaten berechnet, die nicht für die Kalibrierung verwendet wurden. Der Vergleich der Modellgüte aus der Validierung mit der Modellgüte aus der Kalibrierung erlaubt Rückschlüsse darüber, wie gut die Modellvorhersagen mit neuen Daten ist.

1 Einleitung

Die Globalisierung des Handels und des Transports führte während der letzten Jahrzehnte zu einem immer intensiveren weltweiten Transport von Waren und Menschen. Über diese internationalen Handels- und Transportnetzwerke werden aber auch Tier- und Pflanzenarten transportiert (Hulme, 2009). Häufig werden diese Arten – bewusst oder unbewusst – in Gebiete gebracht, die sie ohne die Hilfe des Menschen nie erreicht hätten. Manche dieser sogenannten gebietsfremden Arten (oder auch „Neobiota“ genannt) können sich in den neuen Regionen etablieren und Populationen außerhalb ihres ursprünglichen Verbreitungsgebiets aufbauen. Sie werden damit fester Bestandteil der lokalen Flora und Fauna.

Die Etablierung von Neobiota kann in manchen Fällen schwerwiegende Konsequenzen haben. Die Einführung und Etablierung einiger Arten führte z.B. zur Verdrängung und Verlust einheimischer Arten, zur Beschädigung der Infrastruktur, zu erhöhten Pestizideinsätzen in der Land- und Forstwirtschaft und zu Beeinträchtigungen der menschlichen Gesundheit (Simberloff et al., 2013). Neobiota, deren Etablierung derart negative Folgen hat, werden als invasive Arten bezeichnet. Invasive Arten stellen daher einen Teil der Neobiota dar. Die Unterscheidung dieser Gruppen liegt nur in der Art der Auswirkung, während die Ausbreitungswege und die Etablierung für beide Gruppen dieselben sind. Die Ausbreitung von invasiven Arten ist als eine der größten Gefahren für die einheimische Biodiversität identifiziert (Tilman et al., 2017), beeinträchtigt die menschliche Gesundheit (Pyšek & Richardson, 2010) und führt zu Kosten von mehreren Milliarden Euro pro Jahr innerhalb der EU (Kettunen et al., 2009). Eine effiziente Vermeidung der Einwanderung weiterer invasiver Arten und Reduzierung der Ausbreitung bereits eingewanderter Arten ist nur bei einer genauen Kenntnis der Einfallstore, Ausbreitungswege und Risikogebiete möglich.

In den letzten Jahren gelangen einige Fortschritte in der Modellierung der Ausbreitung gebietsfremder Arten oder auch Neobiota, die zum Verständnis der Ausbreitung dieser Arten beitragen (Seebens, Essl, & Blasius, 2017). So konnte z.B. gezeigt werden, dass sich über die Intensitäten des globalen Schiffsverkehrs die Routen der marinen Bioinvasion sehr gut darstellen lassen (Seebens, Gastner, & Blasius, 2013). Dies ermöglicht eine Identifizierung der wichtigsten Ursprungsgebiete, Einfallstore und Hochrisikogebiete weltweit. Weiterhin konnte gezeigt werden, dass sich mit Hilfe vergleichsweise einfacher Modelle die Ausbreitung einzelner Arten modellieren lässt und auch die Vorhersage der nächsten Invasoren möglich ist (Seebens, Schwartz, Schupp, & Blasius, 2016). Die Modelle sind flexibel und leicht adaptierbar, so dass auch die Ausbreitung anderer Organismengruppen wie vaskuläre Pflanzen über den internationalen Handel simuliert werden konnte (Seebens et al., 2015).

Allerdings haben alle vorhandenen Modelle einen stark begrenzten Fokus, da nur einzelne Ausbreitungswege, wie z.B. Ballastwasser im internationalen Schiffsverkehr, oder einzelne Organismengruppen, wie z.B. höhere Pflanzen, betrachtet wurden. Sofern detailliertere Studien vorliegen, wurden nur einzelne Aspekte im Prozess der Ausbreitung betrachtet. So gibt es Studien zum Transport von Samen über einzelne Kraftfahrzeuge (von der Lippe & Kowarik, 2007), aber es fehlen umfassende Ansätze, bei denen der gesamte Prozess der Ausbreitung und Etablierung für verschiedene Ausbreitungswege und verschiedene Organismengruppen betrachtet wird. Gerade für Länder mit dichten und eng verzahnten Verkehrsnetzwerken wie Deutschland stellt die verkehrsträgerübergreifende Betrachtung des Prozesses der biologischen Invasion eine wichtige Komponente dar, ohne die die Verteilung von Neobiota nicht zu verstehen ist. Allerdings gibt es weltweit bislang kein Modell, welches die Ausbreitung und Etablierung von Organismen verkehrsträgerübergreifend simulieren kann.

Bestandteil dieses Forschungsprojektes ist es, die neuen Erkenntnisse der globalen Ausbreitung und die entwickelten Modelle für Deutschland zu adaptieren. Ähnlich wie in den Studien zur globalen Ausbreitung, sollen Daten zu Verkehrsflüssen mit Umweltdaten wie Habitateigenschaften verbunden werden, um eine möglichst genaue Aufenthalts- und Ausbreitungswahrscheinlichkeit gebietsfremder Arten zu ermitteln.

Die Zielsetzung des Projekts ist eine deutschlandweite Darstellung der Invasionswege und -ströme über die Verkehrswege Wasserstraße, Straße und Schiene. Dabei soll berücksichtigt werden, dass Arten verschiedene Ausbreitungswege und -vektoren nutzen können, indem Artgruppen mit ähnlichen Ausbreitungswegen definiert und simuliert werden. Hierbei standen nach Vorgabe des Auftraggebers vor allem Fragen zu Einfallstoren, Vektoren der Ausbreitung, möglichen Umschlagpunkten, den Ausbreitungsrouten, der Ausbreitungsgeschwindigkeit und der Intensität der Ausbreitung im Vordergrund.

2 Methodik

Entsprechend der Leistungsbeschreibung wurde in der ersten Phase (AP 1) des Projekts die Verfügbarkeit relevanter Daten zur Modellierung und Validierung geprüft. Die vorhandenen Daten wurden in der zweiten Phase des Projekts (AP 2) aufbereitet und harmonisiert, so dass sie als Eingangsvariablen für ein neu zu entwickelndes Modell verwendet werden konnten. Die dritte Phase des Projekts (AP 3) beinhaltete die Entwicklung dieses Modells zur Ausbreitung gebietsfremder Arten durch Verkehrsträger und dessen Kalibrierung und Validierung mit Daten zu beobachteten Ausbreitungen gut untersuchter Neobiota. Abschließend wurden die ermittelten Ergebnisse analysiert.

Die Evaluierung der Möglichkeiten der Modellentwicklung waren ebenso Bestandteil des Projekts wie auch die Modellkalibrierung und -validierung sowie die Recherche der Datenverfügbarkeit und Datenbeschaffung. Daher wurden zu Beginn des Projekts die Arbeitsziele zunächst grob gesteckt und nach der eingehenden Evaluierung der Möglichkeiten konkretisiert. Dies erfolgte in Absprache mit dem Auftraggeber und der Projektbegleitenden Arbeitsgruppe. Der Prozess der Projektentwicklung wird im Folgenden dargestellt.

2.1 Ermittlung der verfügbaren Datengrundlagen und Datenbeschaffung (AP 1)

Zur Simulation der Ausbreitung und Etablierung von Neobiota oder invasiven Arten sind drei Arten von Datensätzen notwendig. Diese umfassen Daten von i) **Prädiktoren der Ausbreitung** (Variablen, mit deren Hilfe Vorhersagen gemacht werden können, z.B. Verkehrsdaten), ii) **Prädiktoren zur Etablierung** (z.B. Habitatcharakteristiken) und iii) **tatsächliche Beobachtungen der Ausbreitung und Etablierung** von Arten zur Kalibrierung und Validierung des Modells (z.B. Verbreitungskarten einzelner Arten zu unterschiedlichen Zeitpunkten). Fehlt einer dieser Bausteine, kann die entsprechende Komponente nicht vollständig implementiert und getestet werden.

Im AP 1 wurden zunächst die möglichen Ausbreitungswege der Neobiota identifiziert, die für die Verkehrsträger relevant sind. Anschließend wurde die Relevanz jedes Ausbreitungswegs für dieses Projekt wie auch die Güte des Verständnisses der zugrundeliegenden Mechanismen sowie die Qualität und Quantität der hierzu verfügbaren Daten nach eigener Einschätzung bewertet. Die Kombination dieser Kriterien ergab eine Aufstellung der Ausbreitungswege, die sich am geeignetsten und vielversprechendsten zur Erreichung der Projektziele herausstellten. Idealerweise sind dies Ausbreitungswege, die gut verstanden sind, eine gute Datengrundlage haben, relevant für die Erreichung der Projektziele sind und eine große räumliche Abdeckung aufweisen. In der Regel wurden aber nicht alle Kriterien im gleichen Maße erfüllt. Die geeignetsten Ausbreitungswege wurden in das Modell aufgenommen (Tabelle 1).

Die **Prädiktoren der Ausbreitung** von Neobiota stellen in diesem Projekt Verkehrsnetzwerke (Straßen, Schienen und Wasserstraßen) und Verkehrs- und Güterflüsse dar. Diese Daten wurden mit einem Verkehrsmodell, welches von SSP Consult entwickelt wurde, generiert (s. Abschnitt 2.2). Als Datengrundlage für das Verkehrsmodell wurden frei zugängliche Datensätze analysiert. Hierfür wurde einerseits auf Daten des Statistischen Bundesamtes, Eurostat, des Kraftfahrtbundesamtes, des Bundesamtes für Güterverkehr und des Bundesamtes für Straßenwesen zurückgegriffen. Darüber hinaus wurde ein Datensatz analysiert, der im Rahmen der Erstellung des aktuellen Bundesverkehrswegeplans (BVWP) 2030 verwendet und erstellt wurde. Die notwendigen Daten wurden zusammengetragen, harmonisiert und zur weiteren Verwendung aufgearbeitet (s. Abschnitt 2.2).

TABELLE 1: EVALUIERUNG DER MÖGLICHEN PRÄDIKTOREN DER AUSBREITUNG VON NEOBI-OTA/INVASIVEN ARTEN DURCH VERKEHRSTRÄGER UND DEREN RELEVANZ FÜR DAS FORSCHUNGSVORHABEN.

TRANSPORT-MODUS	UNTERKATEGORIE	AUSBREITUNGSWEG	RELEVANZ	DISTANZ	DATENVERFÜGBARKEIT	VERSTÄNDNIS	MODELL
STRASSE	Fahrzeug	Anheftung am Fahrzeug	sehr hoch	lang	sehr gut	gut	x
		Transport im Fahrzeug (Fahrerkabine/Auto)	mittel	lang	sehr gut	mittel	
		Fahrtwind	hoch	kurz	gut	gut	x
	Transport	Ladung (Container und Paletten)	sehr hoch	lang	gut	mittel	x
		Direkte Ausbreitung durch den Menschen (angeheftet oder getragen)	mittel	kurz – lang	schlecht	mittel	
		Korridor	Natürliche Ausbreitung	gering	kurz	schlecht	mittel
	Mäharbeiten		hoch	sehr kurz	schlecht	mittel	
	Ausbringung von Mähgut/Erdbmaterial		gering	kurz	schlecht	gering	
	SCHIENE	Fahrzeug	Anheftung am Fahrzeug	hoch	pot. lang	mittel	gering
Fahrtwind			hoch	kurz	mittel	mittel	x
Abbruch von Ästen/Wurzeln			gering	kurz	mittel	gering	
Transport		Transport im Fahrzeug (Fahrerkabine oder Personenwaggons)	mittel	Lang	mittel	gering	
		Direkte Ausbreitung durch den Menschen (angeheftet oder getragen)	mittel	kurz – lang	schlecht	gering	
		Ladung (Container und Paletten)	sehr hoch	lang	gut	mittel	x
Korridor		Rollende Landstraße	gering	lang	gut	gering	x
		Natürliche Ausbreitung	gering	kurz	schlecht	gering	x
		Mäharbeiten	hoch	sehr kurz	schlecht	mittel	
Ausbringung von Mähgut/Erdbmaterial	gering	kurz	schlecht	gering			
WASSER-STRASSE	Fahrzeug	Äußerer Bewuchs	hoch	lang	mittel	mittel	x
		Ballastwasser	hoch	lang	gering	gut	x
		Transport im Fahrzeug (Fahrerkabine)	sehr gering	lang	gering	mittel	
	Transport	Ladung (Container und Paletten)	sehr hoch	lang	gut	mittel	x
		Direkte Ausbreitung durch den Menschen (angeheftet oder getragen)	sehr gering	kurz – lang	schlecht	mittel	
	Korridor	Natürliche Ausbreitung	gering	lang	sehr gut	gut	x

Anmerkung: Nach eigener Recherche und Einschätzung wurde die Relevanz einzelner Ausbreitungswege für die Modellierung in fünf Stufen (sehr gering bis sehr hoch), die geographische Distanz der Ausbreitung in drei Stufen (kurz bis lang), die Datenverfügbarkeit in fünf Stufen (sehr schlecht bis sehr gut) und das Verständnis des Ausbreitungswegs in drei Stufen (gering bis gut) eingestuft. Hieraus ergab sich die Einschätzung, ob der Ausbreitungsweg in das Modell aufgenommen werden soll, was durch ein „x“ in der Spalte „Modell“ angezeigt ist.

Die **Prädiktoren der Etablierung** von Neobiota stellen Faktoren zur Charakterisierung der ökologischen Nische der Arten dar. Idealerweise beschreiben sie eine Vielzahl verschiedener Faktoren, wie z.B. Habitatstruktur, Umweltbedingungen und biotische Interaktionen. Die Einbindung dieser Faktoren ist in vielen Fällen allerdings schwierig, da die zugrundeliegenden Mechanismen wenig verstanden sind oder Daten fehlen. Dies trifft in besonderem Maße auf biotische Interaktionen zu, die in der Regel nicht berücksichtigt werden können.

Die Prädiktoren der Etablierung wurden auf dieselbe Art bewertet wie die Prädiktoren der Ausbreitung. Für dieses Projekt ergab die Bewertung der Verfügbarkeit und Relevanz der verschiedenen Faktoren, dass die Etablierung neuer Arten im terrestrischen Raum am besten mit dem Parameter Landbedeckung abgeschätzt werden kann (Tabelle 2). Dieser ermöglicht eine gute Abschätzung der ökologischen Nische einer Art (Chytrý et al., 2008). Daten zu Umweltparametern wie Temperatur oder Niederschlag sind zwar in hoher Auflösung vorhanden, allerdings eignen sie sich nicht für die Vorhersage von Aufhalten der Arten, da sie eine sehr geringe Variation innerhalb Deutschlands aufweisen. Diese Parameter enthalten daher nur ein geringes Maß an zusätzlicher Information und werden daher in dieser Analyse nicht berücksichtigt.

In aquatischen Habitaten werden Informationen zu Wassertemperatur und elektrischer Leitfähigkeit als Maß für den Salzgehalt herangezogen. Mit Hilfe dieser Parameter können wichtige aquatische Habitate der großen Wasserstraßen in Deutschland unterschieden werden. Andere Faktoren wie z.B. die Strukturgüte der Fließgewässer stellten sich nach Diskussion mit Experten als von geringer Relevanz heraus, da sie so gut wie keine Variation entlang der zu betrachteten Gewässer aufweisen. So ist z.B. die Habitatstruktur der Bundeswasserstraßen nahezu ausschließlich durch Blocksteinschüttungen charakterisiert, wodurch keine Unterscheidung zwischen Habitaten möglich war. Auf die Einbindung von solchen einheitlichen Faktoren wurde verzichtet, um das Modell möglichst einfach und verständlich zu gestalten.

TABELLE 2: EVALUIERUNG DER PRÄDIKTOREN DER ETABLIERUNG UND DEREN RELEVANZ FÜR DAS FORSCHUNGSVORHABEN.

PARAMETER	TREIBER	ARTGRUPPEN	RELEVANZ	DATENVERFÜGBARKEIT	VERSTÄNDNIS	MODELL
HABITATE	Landbedeckung	Terrestrische Arten	sehr hoch	gering - sehr gut	mittel	x
	Uferstruktur	Limnische Arten	hoch	gering	mittel	
	Küstenstruktur	Marine Arten	gering	schlecht	mittel	
	Umweltbedingungen	Alle Arten	mittel - hoch	sehr gut	mittel	x
BIOTISCHE INTERAKTIONEN	Konkurrenz, Räuber, Pathogene	Alle Arten	hoch	schlecht	sehr gering	

Anmerkung: Nach eigener Recherche und Einschätzung wurde die Relevanz einzelner Ausbreitungswege für die Modellierung in fünf Stufen (sehr gering bis sehr hoch), die geographische Distanz der Ausbreitung in drei Stufen (kurz bis lang), die Datenverfügbarkeit in fünf Stufen (sehr schlecht bis sehr gut) und das Verständnis des Ausbreitungswegs in drei Stufen (gering bis gut) eingestuft. Hieraus ergab sich die Einschätzung, ob der Ausbreitungsweg in das Modell aufgenommen werden soll, was durch ein „x“ in der Spalte „Modell“ gekennzeichnet ist. Die Einschätzung variiert zwischen Arten.

Aufgrund der Vielzahl der möglichen Kombinationen an Faktoren zur Charakterisierung der möglichen Etablierung von Neobiota bzw. invasiven Arten erschien eine umfassende Einbindung aller relevanten Faktoren in diesem Projekt als nicht realistisch. Es wurde daher bewusst die Entscheidung getroffen, nur

wenige allgemeingültige Faktoren zu berücksichtigen. Hinzu kommt, dass es sehr weit entwickelte alternative Modelle gibt, die eine umfassende Simulation der Habitate erlauben. Diese sogenannten Habitatmodelle („species distribution models“) decken eine große Vielfalt an Faktoren ab. Sie sind frei verfügbar und lassen sich leicht mit dem in diesem Projekt entwickeltem Modell verknüpfen. Dies wird in der Beschreibung zum AP 3 näher ausgeführt (s. Abschnitt 2.3.2.3).

Um die Güte eines Modells zu bewerten, empfiehlt es sich, ein neu entwickeltes Computermodell mit **unabhängigen Datensätzen zur Ausbreitung von Neobiota** zu testen. Idealerweise werden hierzu Daten von Arten verwendet, die nur über den zu testenden Ausbreitungsweg verbreitet werden, und die sich nur nahe der Verkehrsträger etablieren können. Weiterhin müssen die Daten in ausreichender Auflösung, Quantität, Qualität und zugänglich vorliegen. Zusätzlich müssen die Daten die Ausbreitung einer Art über eine gewisse Zeitperiode zeigen, die sich für die Kalibrierung des Modells eignet. Eine umfassende Literaturrecherche, Durchsicht von Onlinedatenbanken und persönliche Gespräche und Treffen mit verschiedenen Experten (z.B. Prof. Frank Schurr, Dr. Moritz von der Lippe, Prof. Jürgen Böhmer, Prof. Ingolf Kühn, Dr. Tina Heger, Dr. Franz Schöll, Dr. Nicole Heibeck) ergaben, dass nur sehr wenige Datensätze in Deutschland vorliegen, die zumindest einen Teil dieser Kriterien erfüllen. Daten aus der Literatur sind in der Regel nicht frei zugänglich und Onlinedatenbanken sind häufig unvollständig, räumlich und zeitlich limitiert und/oder nicht frei zugänglich. Um die bestmöglichen Projektergebnisse zu erzielen, wurde daher entschieden, gute Datensätze für wenige Arten anstatt vieler Datensätze mit geringer Qualität zu verwenden. Hier musste vom ursprünglich geplanten Ansatz der Verwendung vieler Datensätze abgewichen werden. Weiterhin wurden anstatt frei zugänglicher Daten alle Datensätze durch persönlichen Kontakt zu Einzelpersonen ermittelt, die für dieses Projekt zur Verfügung gestellt wurden.

2.2 Datenaufbereitung (AP 2)

In diesem Projekt kamen Datensätze zur Anwendung, die in einem unterschiedlichen Maße aufbereitet werden mussten, um sie für die Modellierung der Ausbreitung und Etablierung von Neobiota verwendbar zu machen. Im einfachsten Fall konnten Daten ohne aufwendige Aufbereitung verwendet werden. So konnten z.B. Verkehrsdichten der Schiffe auf Bundeswasserstraßen deutschlandweit direkt bezogen und verwendet werden. Ebenso lagen bereits Informationen zur Landbedeckung als aufbereitete Satellitenprodukte vor, die ohne weitere Aufbereitung verwendet werden konnten. In den meisten Fällen musste Daten aber weiter aufbereitet werden, wobei unterschiedliche Methoden zur Anwendung kamen, die von einer Extrapolation einzelner Messpunkte in die Fläche (z.B. Temperatur der Wasserkörper) bis zu komplexer Modellierung des Straßenverkehrs reichten. Zur Aufbereitung der Datensätze wurden bestehende Methoden und Modelle angewandt.

Im Folgenden wird beschrieben, welche Datensätze in welcher Form aufbereitet wurden. Die so aufbereiteten Datensätze stellen die Basis der Entwicklung des Modells zur Ausbreitung von Neobiota (CASPIAN) dar, welches im Abschnitt 2.3 im Detail beschrieben wird.

2.2.1 Verkehrsmodellierung

Allgemeine Beschreibung

In Deutschland existieren keine offiziellen Daten zu den Fahrzeugströmen (Pkw-, Lkw-, Züge, Binnenschiffe). Daten zu Verkehrsströmen wurden daher mit dem Verkehrsmodell „SSP-D-Modell“ oder auch „SSP-Deutschlandmodell“ generiert. Hierbei handelt es sich um einen integrierten, d. h. Personen- und Güterverkehr gemeinsam betrachtenden, und verkehrsträgerübergreifenden Modellansatz. Das Modell ist eine Eigenentwicklung von SSP Consult, die ständig weiterentwickelt und vielfach für Kurz-, Mittel- und Langfristprognosen im Personen- und Güterverkehr genutzt wurde und wird (u. a. im Auftrag des BMVI). Es ist für die besondere Anwendung und Anforderungen bei der Fortschreibung und Evaluierung integrierter Fragestellungen angepasst und konzipiert. Das „SSP-Deutschlandmodell“ hat folgende grundsätzliche Eigenschaften, die für dieses Forschungsvorhaben besonders geeignet sind:

- Das Modell arbeitet mit bundesweiten Verkehrsströmen.
- Das Modell bestimmt für definierte Nachfragesegmente (Fahrtzwecke im Personenverkehr bzw. Gütergruppen im Güterverkehr) die regionalen Verkehrsaufkommen, die Verkehrsverteilung zwischen den Verkehrszellen und den Wechsel von Transportmodi im Güterverkehr.
- Fahrzeugmodelle wandeln die Verkehrsströme in tägliche Fahrzeugbewegungen für die Landverkehrsträger um, so dass die Personen- und Güterverkehre auf Netzmodelle umgelegt werden können.
- Zu den Verkehrsmodellen existieren passende Netzmodelle für Schiene, Straße und Wasserstraße auf Basis der Netze des Bundesverkehrswegeplans, so dass die Fahrzeugströme in Netzbelastungen umgeformt werden können.
- Das Modell liefert als Ergebnis Verkehrsmengen, d.h. neben Güterströmen auch Verkehrsstärken der Verkehrsnetze.

Das „SSP-Deutschlandmodell“ ist nach einem klassischen 4-stufigen Modellablauf aufgebaut, dass 1. Verkehrserzeugung (Ermittlung der Quell- und Zielverkehrsaufkommen der Verkehrsbezirke), 2. Verkehrsverteilung (Ermittlung der Verkehrsströme zwischen den Verkehrsbezirken), 3. Verkehrsaufteilung (Aufteilung des Verkehrs auf verschiedene Verkehrsmittel) und 4. Verkehrsumlegung (Bestimmung der Verkehrsstärken der Strecken der Verkehrsnetze) berechnet. Um den Input der hier benötigten Verkehrsdaten zu generieren, wurden aus dem „SSP-Deutschlandmodell“ für die hier relevanten Fragestellungen folgende Teilmodule genutzt:

- *Fahrzeugmodell*: Das Modell ermöglicht die Umrechnung der Verkehrsströme auf Lkw- und Pkw-Fahrten über z. B. Fahrzeugbesetzungsgrade oder segment-spezifische durchschnittliche Zuladungen im Güterverkehr.
- *Umlagemodell*: Das Modell berechnet die Aufteilung der Verkehre auf das Verkehrszellensystem, z. B. entlang der für das Vorhaben ausgewählten Korridore. Die berechneten Fahrzeugströme werden auf die Feinzellen verteilt. Anschließend werden die Belastungen der Streckensegmente („Verkehrsstärken“) ermittelt.
- *Zugbildungsmodell*: Ein Extra-Modul kommt für den Bereich des Schienengüterverkehrs zum Einsatz, das den betrieblichen und operativen Ablauf bei der Bildung von Zügen simuliert und damit wertvolle Hinweise auf die tatsächlichen Verlagerungsmöglichkeiten von Güterverkehren auf die Schiene gibt.
- *Paletten- und Containermodell*: Ein Extra-Modul, das für jede Gutart entfernungsabhängige Anteile von palettisierten und containerisierbaren Gütern beinhaltet.

Anwendung für das vorliegende Forschungsvorhaben

Die für das Forschungsvorhaben erstellten Verkehrsdaten sind das Ergebnis einer Abstimmung der Projektpartner über die Anforderungen für die Modellierung der Einfuhr und Ausbreitung der invasiven Arten auf der einen Seite und der Möglichkeiten der Verkehrsmodelle unter Verwendung vorhandener Daten auf der anderen Seite.

Basis für die Berechnung des Güter- und Personenverkehrs bildeten die bei der Clearingstelle des DLR¹ verfügbaren Daten der Verkehrsverflechtungsprognose der Bundesverkehrswegeplanung 2030 (VP 2030) (Abbildung 1). Diese Daten sind in Deutschland auf Kreisbasis formuliert. Güterströme sind zusätzlich nach rund 20 verschiedenen Güterarten differenziert. Diese Datengrundlage lässt allerdings keine Berechnung der Verkehrsströme innerhalb der Kreise zu. Für eine feinere Darstellung des Verkehrs wird die räumliche Struktur sogenannter Verkehrszellen verwendet, die auf einem vom BMVI entwickelten Raumstruktursystem basiert. Unter Verwendung feinräumiger Strukturdaten zu Bevölkerung und Beschäftigung können somit die Ströme des Personen- und Güterverkehrs für einzelne Straßen- und Schienenabschnitte berechnet werden.

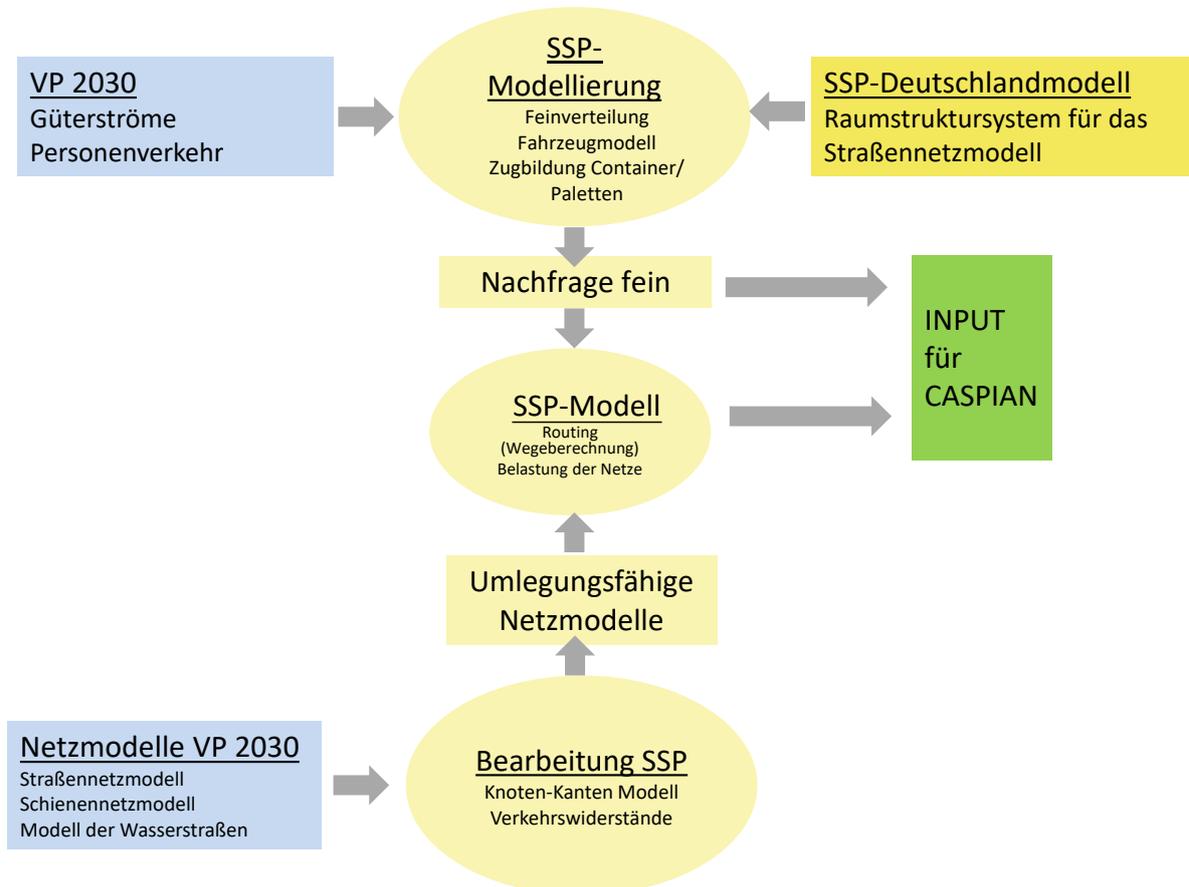


Abbildung 1: Übersicht über die Arbeiten von SSP Consult. Aufbauend auf der Verkehrsverflechtungsprognose der Bundesverkehrswegeplanung 2030 (VP 2030) werden mit Hilfe des SSP-Deutschlandmodells Verkehrs- und Güterströme berechnet, die als Eingangsvariablen für das Invasionsmodell CASPIAN verwendet wurden.

¹ <https://www.dlr.de/cs/Desktopdefault.aspx/tabid-701/>

In die Berechnung der Fahrzeugströme gehen zusätzlich Belegungsgrade im Personenverkehr sowie Beladungskennziffern und Leerfahrtenwahrscheinlichkeiten im Güterverkehr mit ein. Im Güterzugverkehr werden weiterhin verschiedene Produktionssysteme (Ganzzug, Einzelwagenlauf, kombinierter Verkehr) unterschieden. Hierfür wird ein von SSP Consult entwickeltes Zugbildungsmodell angewandt, welches aus Güterströmen einen Güterzugfahrplan erstellt. Im Schienenpersonenverkehr existiert ein fester Fahrplan, für den bei SSP Consult die Belastungszahlen der Modellstrecken vorhanden sind.

Differenziertere Auswertungen zu den Containerströmen und Palettenströmen werden mit speziellen Modellrechnungen von SSP Consult getätigt. In der VP 2030 sind Containerströme für Binnenschiff- und Bahnverkehre auf Kreisbasis ausgewiesen. Informationen zu Palettenströmen liegen öffentlich nicht vor. Die Modellberechnung für Container- und Palettenströme geschieht auf bei SSP Consult vorhandenen Auswertungen zu entfernungsabhängigen Anteilen für Paletten und Container je ausgewiesener Gutart.

2.2.1.1 Straßenverkehr

Basis für die Modellierung des Straßennetzes im hier bearbeiteten Projekt liefert das Netzmodell für die Bundesfernstraßenplanung (NEMOBFStr). NEMOBFStr ist ein GIS-Datensatz, der Strecken des Straßennetzes und Knoten (gemeinsame Anfangs- und Endpunkte der Strecken) unterscheidet. NEMOBFStr bildet Autobahnen, Bundes-, Landes- und Kreisstraßen sowie wichtige kommunale Straßen ab. Die Knoten (z.B. Kreuzungen) und Links (Straßensegmente) des Modells sind mit Informationen und Eigenschaften zu Nummer, Bezeichnung, Topographie, Lage, Länge, Querschnittstyp, Straßenbezeichnung und Länderkennung versehen. Das Netz beinhaltet dabei den Ausbauzustand im Jahr 2015, d.h. alle Strecken, die bis zu diesem Zeitraum für den Verkehr freigegeben worden sind.

Um die im Straßennetz zurückgelegten Wege der Verkehrsströme zu modellieren und daraus die Verkehrsbelastungen an den Strecken zu generieren, wurde das Netzmodell in ein Verkehrsnetz überführt, welches die Simulation der Bewegungen einzelner Verkehrsteilnehmer erlaubt. Hierzu wurden weitere Informationen, wie z.B. Kapazitäten und Durchschnittsgeschwindigkeiten basierend auf Daten der Bundesverkehrswegeplanung für alle Straßenabschnitte ergänzt. Als Ergebnis stand für den Verkehrsträger Straße ein geografisch referenziertes und routingfähiges Netzmodell zur Verfügung, das aus 113.956 Knoten und 334.819 Streckensegmenten zwischen den Knoten besteht.

Dieses Netzmodell bildete die Grundlage für die Berechnung der Verkehrs- und Güterströme im Straßenverkehr mit Hilfe des Verkehrsmodells „SSP-Deutschlandmodell“ (s. Einleitung zu Abschnitt 2.2.1). Weitere Datensätze wie z.B. empirische Daten und Verkehrsmodelldaten (Dauerzählstellen, Straßenverkehrszählungen 2015, Verkehrsverflechtungsmatrizen für den Straßenpersonenverkehr und Straßengüterverkehr) gingen in die Berechnung mit ein.

Das Verkehrsmodell lieferte dann Berechnung für die durchschnittlichen täglichen Verkehrsstärken (DTV) differenziert nach Lkw und Pkw (Abbildung 2). Zusätzlich wurden Informationen zu den Routen der Ströme und zu Quelle und Ziel der einzelnen Fahrzeuge generiert, aufbereitet und dargestellt.

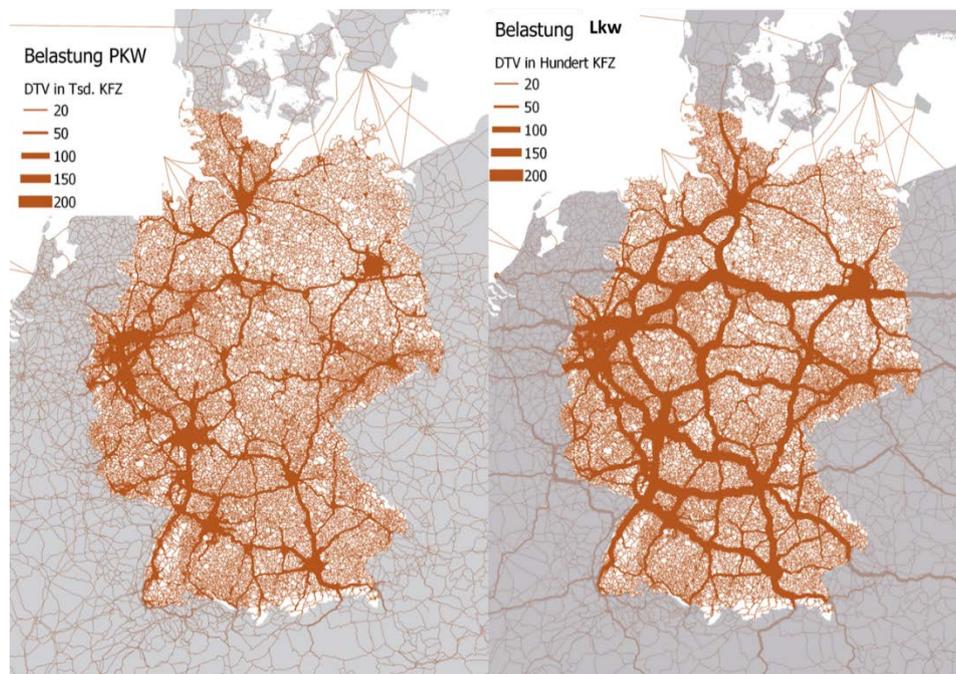


Abbildung 2: Verkehrsbelastungen Pkw und Lkw (DTV: Durchschnittlich täglicher Verkehr).

2.2.1.2 Schienenverkehr

Analog zum Vorgehen im Straßenverkehr, wurde auch für den Schienenverkehr ein Netzmodell generiert und verwendet, welches über die Clearingstelle Verkehr² des DLR frei zugänglich ist. Basis des Schienennetzes ist hierbei das im Rahmen der Prognose zur Bundesverkehrswegeplanung 2030 (VP 2030) verwendete Netzmodell mit ca. 20.000 Knoten und 47.000 Streckenabschnitten (INTRAPLAN et al., 2014). Der Ausbauzustand spiegelt das Analysejahr 2010 wider, musste aber dennoch teilweise aufbereitet und routingfähig gemacht werden. Hierzu mussten Attribute und Eigenschaften ergänzt werden, die eine Anwendung der Verkehrsmodellierungsmodule ermöglichen. Die Knoten- und Links beinhalten Informationen zu Streckentyp, Streckennutzung, Streckenlänge, Gleisigkeit, Traktion, Kapazitäten etc.

Auf Grundlage dieses Netzmodells wurden auf Basis von empirischen und Verkehrsmodelldaten (Zugbelastungen, Verkehrsverflechtungsmatrizen für den Schienenpersonenverkehr und Daten zum Schienengüterverkehr des Statistischen Bundesamts) mit Hilfe des Verkehrsmodells (s. Einleitung zu Abschnitt 2.2.1) die durchschnittlichen täglichen Zugbelastungen, differenziert nach Personenverkehrs- und Güterverkehrszügen, modelliert und dargestellt (Abbildung 3). Hierbei wurden für den Schienengüterverkehr auch die betrieblichen Abläufe der einzelnen Verkehrsarten (Einzelwagenlauf, Ganzzug und kombinierter Verkehr) modelliert und mitberücksichtigt.

² <https://www.dlr.de/cs/Desktopdefault.aspx/tabid-701/>

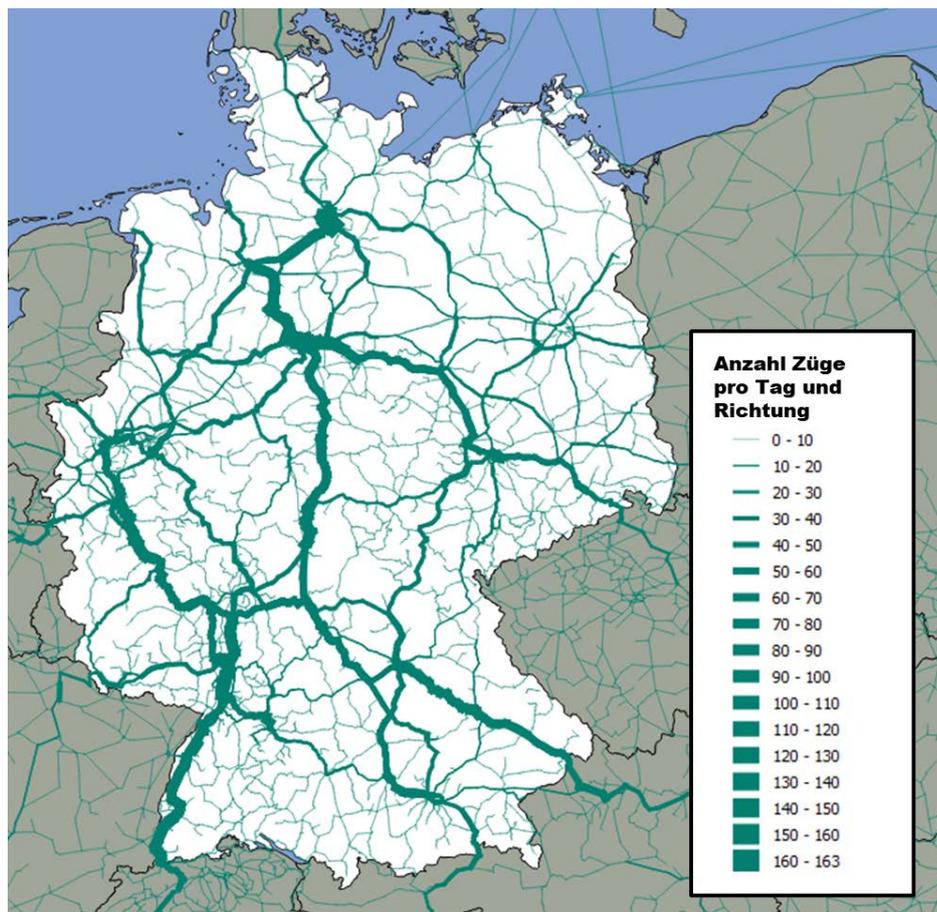


Abbildung 3: Anzahl Züge pro Tag in Deutschland.

2.2.1.3 Wasserstraßen (Binnenschifffahrt)

Das (Binnen-)Wasserstraßennetz beruht ebenfalls auf dem im Rahmen der Prognose zur Bundesverkehrswegeplanung 2030 (VP 2030) verwendeten europäischen Netzmodells, das ca. 500 Knoten und 600 Strecken enthält (Flüsse und Kanalsystem sowie relevante Fährverbindungen). Der Ausbauzustand spiegelt das Analysejahr 2010 wider. Um ein geografisch referenziertes und routingfähiges Netzmodell zu erlangen, wurden folgende Attribute und Eigenschaften ergänzt und zur Verfügung gestellt: Streckenlänge, Streckentyp, Wasserstraße, Fahrtgeschwindigkeit je Richtung (INTRAPLAN et al., 2014).

Auf Grundlage dieses Netzmodells wurde basierend auf empirischen Daten und Verkehrsmodelldaten (Verkehrsverflechtungsmatrizen für den Gütertransport per Binnenschiff) die jährliche Anzahl von Güterbinnenschiffen modelliert und dargestellt (Abbildung 4). Individualverkehr auf den Wasserstraßen wurde aufgrund fehlender Daten nicht im Modell berücksichtigt.

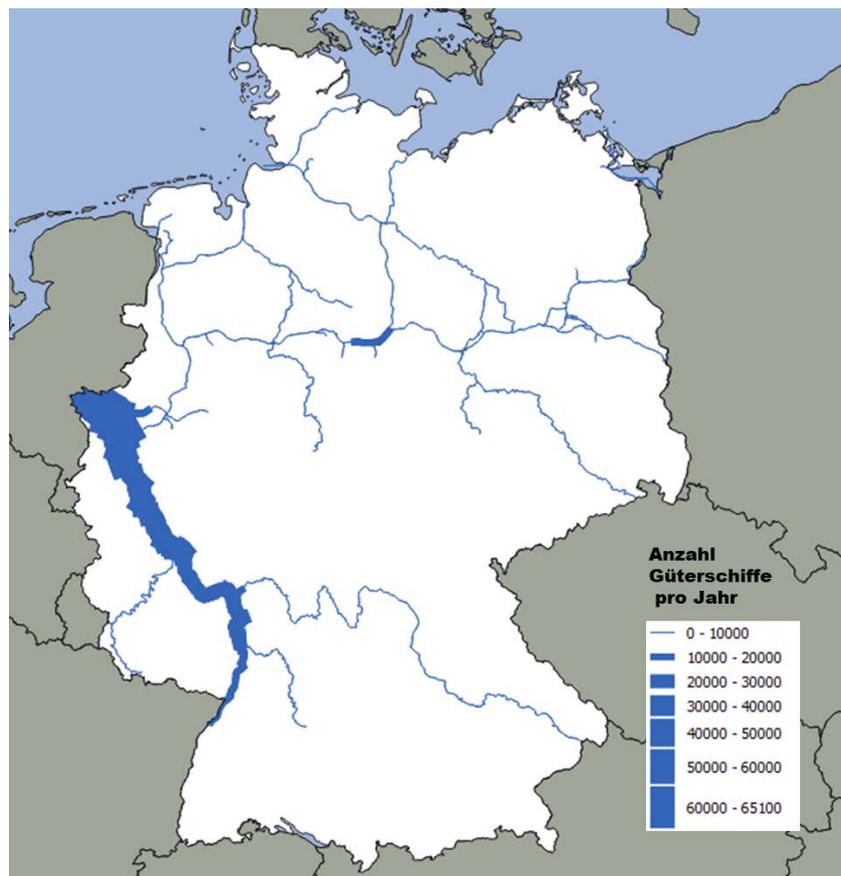


Abbildung 4: Anzahl Güterbinnenschiffe pro Jahr in Deutschland.

2.2.1.4 Verschneidung der Netzmodelle

Die Netzmodelle für Straße und Schiene wurden miteinander verschnitten, indem die Netzwerke übereinander gelegt wurden und an Schnittpunkten von Straßen- und Schienentrassen neue Knoten eingefügt wurden. Dies ermöglicht die Simulation der Ausbreitung von Neobiota über Verkehrsträger hinweg. Dadurch ergab sich ein kombiniertes Netzmodell mit Verkehrsbelastungen für Straße und Schiene mit insgesamt 260.829 Knoten und 634.817 Streckenabschnitten mit einem Median der Streckenlängen von 1km.

2.2.1.5 Güterströme

Als Grundlage für die Darstellung von Güterströmen dient hier die Analyseverkehrsverflechtungsmatrix aus der aktuellen Bundesprognose VP 2030, die im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur erstellt wurde. Diese Daten stehen bei der Clearingstelle des DLR³ frei zur Verfügung (s. auch Einleitung zu Abschnitt 2.2.1). Dabei handelt es sich um einen Datensatz der Güterverkehrsverflechtungen zwischen Ausgangs- und Endpunkt der Güterfahrten, welche nach Gütergruppen und Ver-

³ <https://www.dlr.de/cs/Desktopdefault.aspx/tabid-701/>

kehrsträger unterteilt sind. Die Daten sind auf der räumlichen Einheit von 7332 sogenannten Verkehrszellen in Deutschland sowie, bei grenzüberschreitenden Verflechtungen und Transitverkehren, in einem abgestuften System an Auslandsverkehrszellen erfasst. Einbezogen sind alle Verkehrsströme, die das Territorium Deutschlands berühren, d.h. zum einen Ströme mit Quelle und/oder Ziel in Deutschland und zum anderen die Transitverkehre, soweit sie die deutschen Landverkehrswege beanspruchen. Auf Basis dieser Ausgangsdaten wurden im Rahmen des Forschungsvorhabens Transportverflechtungen (Quelle und Ziel) von Güterströmen, differenziert nach Gütergruppen und Ladungsträger (Container, palettisierte Güter) modelliert und aufbereitet.

Für die Modellierung der **Containerströme** wurden insbesondere maritime Ströme, d.h. Güterströme des kombinierten Verkehrs, bei denen als Quelle und/oder Ziel ein Seehafen angegeben ist, berücksichtigt. Für den Transport im kontinentalen kombinierten Verkehr, der in der Regel nicht per Container, sondern Wechselbehälter oder Trailer bzw. ganzen Fahrzeugen (Rollende Landstraße) durchgeführt wird, wurden diese Informationen gesondert modelliert. Zusätzlich dazu wurde ein **Palettenmodell** als Teilmodul des gesamten Verkehrsmodells (s. auch Einleitung zu Abschnitt 2.2.1) entwickelt, das für jede Gutart entfernungsabhängige Anteile von palettisierten Gütern beinhaltet. Die Paletten- und Containerströme der verschiedenen Transportarten und Verkehrsträger wurden auf der räumlichen Struktur der Verkehrszellen berechnet. Eine Auflösung für einzelne Links ist in diesem Fall aufgrund der fehlenden Datengrundlage nicht möglich (s. auch Einleitung zu Abschnitt 2.2.1). Dieser Datensatz hat einen starken modellbasierten Ansatz. Dies betrifft v.a. die Ableitung der palettisierten Güter und die Verteilung der Güter auf die einzelnen Verkehrszellen eines Kreises.

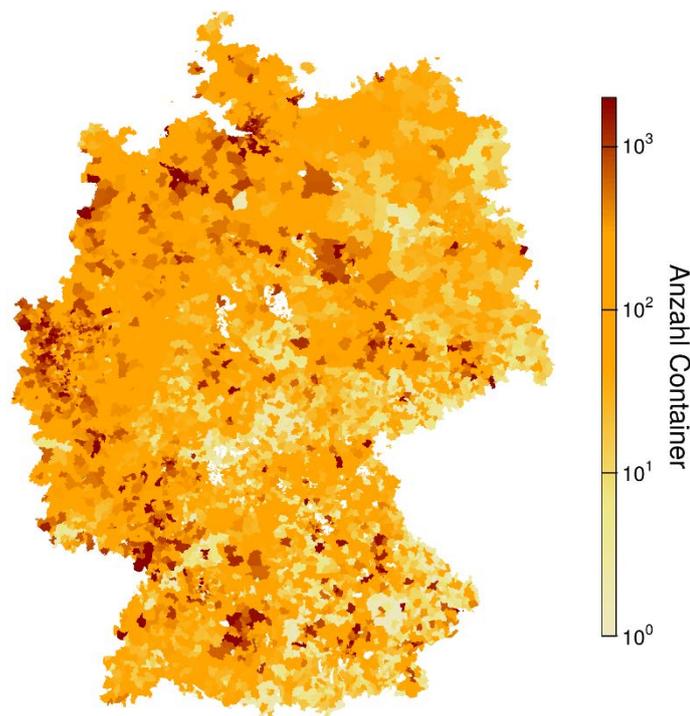


Abbildung 5: Verteilung der Güterströme in den Zielverkehrszellen am Beispiel der Container. Die Verkehrszellen sind nicht flächentreu.

2.2.2 Habitatcharakteristiken

Informationen zur Flächenbedeckung in Deutschland wurden aus Satellitenaufnahmen extrahiert, die in der „CORINE Land Cover European seamless vector“ Datenbank (version v18_5) verfügbar sind (<https://land.copernicus.eu/pan-european/corine-land-cover>). Dieses Satellitenprodukt beinhaltet eine flächendeckende Verteilung von 44 Landbedeckungstypen mit einer Auflösung von 100 m in Europa. Diese Daten wurden zur **Charakterisierung terrestrischer Habitats** entlang von Straßen und Schienen verwendet.

Auf Basis der CORINE land cover database wurden die prozentualen Anteile der Bedeckungskategorien entlang aller Streckenabschnitte des Straßen- und Schienennetzes berechnet. Die 44 Kategorien wurden hierzu in fünf Hauptklassen aggregiert (Tabelle 3), mit denen die Bedingungen eines Habitats im Modell beschrieben wurden. Somit können die Habitatansprüche einer zu simulierenden Art vom Anwender definiert und die Güte eines Habitats für diese Art berechnet werden. Beispielhaft ist die in Abbildung 12 für eine Art (*Senecio inaequidens*) zu sehen. Die Beschreibung der Berechnung findet sich in Abschnitt 2.3.2.3.

TABELLE 3: KATEGORIEN DER FLÄCHENBEDECKUNG DER CORINE DATENBANK UND DES MODELLS UND ANTEIL AN DER GESAMTBEDECKUNG IN DEUTSCHLAND.

CORINE CODE	CORINE BEZEICHNUNG	KATEGORIE DES MODELLS	PROZENTUALE BEDECKUNG IN DEUTSCHLAND
111	Durchgängig städtische Prägung	1	0.685
112	Nicht durchgängig städtische Prägung	1	34.656
121	Industrie und Gewerbeflächen, öffentliche Einrichtungen	1	4.755
122	Straßen-, Eisenbahnnetze und funktionell zugeordnete Flächen	1	0.488
123	Hafengebiete	1	0.050
124	Flughäfen	1	0.059
131	Abbauflächen	1	0.088
132	Deponien und Abraumhalden	1	0.020
133	Baustellen	1	0.037
141	Städtische Grünfläche	1	0.602
142	Sport- und Freizeitanlagen	1	0.615
211	Nicht bewässertes Ackerland	2	28.764
221	Weinbauflächen	2	0.326
222	Obst- und Beerenobstbestände	2	0.427
231	Wiesen und Weiden	3	15.634
242	Komplexe Parzellenstruktur	3	0.720
243	Landwirtschaftlich genutztes Land mit Flächen natürlicher Bodenbedeckung von signifikanter Größe	3	0.615

311	Laubwälder	4	4.095
312	Nadelwälder	4	4.346
313	Mischwälder	4	1.656
321	Natürliches Grünland	3	0.124
322	Heiden und Moorheiden	3	0.054
324	Wald-Strauch-Übergangsstadien	3	0.610
331	Strände, Dünen und Sandflächen	3	0.002
332	Felsen ohne Vegetation	3	0.004
333	Flächen mit spärlicher Vegetation	3	0.016
334	Brandflächen	3	0
335	Gletscher und Dauerschneegebiete	3	0
411	Sümpfe	5	0.024
412	Torfmoore	5	0.035
421	Salzwiesen	5	0.007
423	In der Gezeitenzone liegende Flächen	5	0.009
511	Gewässerläufe	5	0.170
512	Wasserflächen	5	0.118
521	Lagunen	5	0.005
522	Mündungsgebiete	5	0.012
523	Meere und Ozeane	5	0.109

Zur **Charakterisierung aquatischer Habitate** wurden Informationen aus Onlineportalen der folgenden Flussgebietsgemeinschaften herangezogen:

- Elbe (<https://www.fgg-elbe.de>)
- Rhein (<http://iksr.bafg.de/iksr>)
- Weser (<https://www.fgg-weser.de>, <http://www.iksms-cipms.org>)
- Oder (<http://geoportal.mkoo.pl>)

Für jedes Gewässer wurden 4-11 Messstellen ausgewählt, die eine repräsentative Darstellung der Entwicklung von Umweltparametern entlang der Gewässer erlauben. Für die verbleibenden Seitengewässer und Kanäle wurden Informationen aus angrenzenden Gewässern übernommen, so dass eine flächendeckende Darstellung der Umweltbedingungen möglich war.

Es standen nicht für alle Messstellen dieselben Messwerte zur Verfügung, so dass Umweltparameter ermittelt werden musste, für die eine flächendeckende Abdeckung vorlag, wodurch sich die Auswahl an möglichen Parametern deutlich reduzierte. Vollständige Datensätze lagen für mittlere Wassertemperatur und elektrische Leitfähigkeit vor. Für diese Parameter wurden für jede Messstelle Daten über einen Zeitraum von fünf Jahren extrahiert. Sofern kontinuierliche Zeitreihen vorlagen, wurde diese auf Tagesbasis umgerechnet, um einheitliche Jahresmittelwerte zu berechnen, die wiederum zu einem Mittelwert über den 5-Jahreszeitraum zusammengefasst wurden (Abbildung 6).

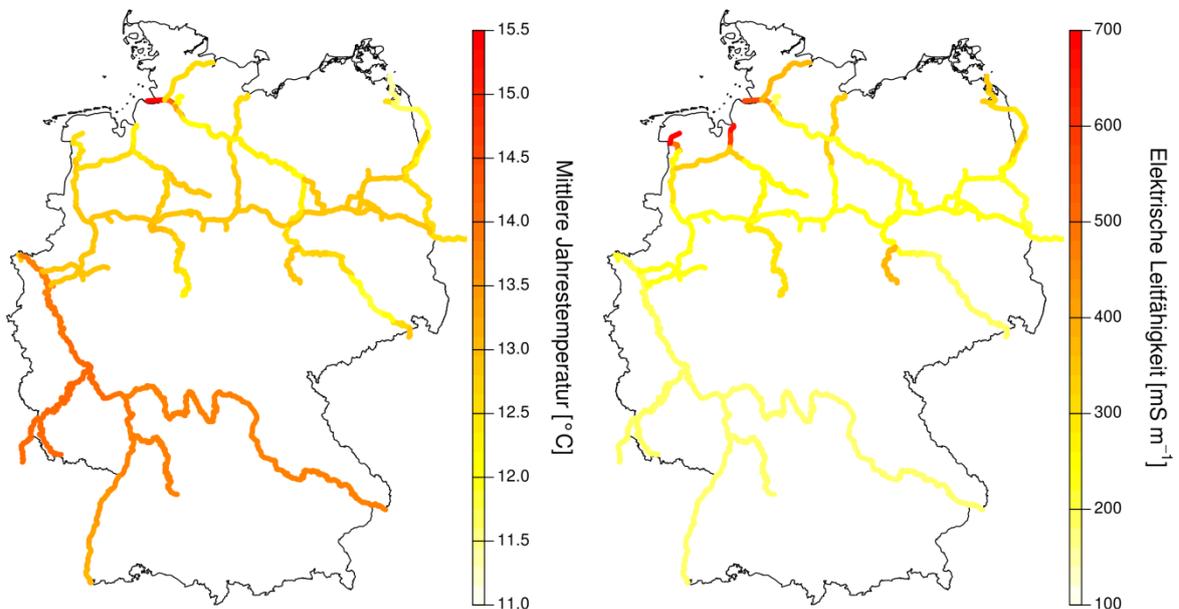


Abbildung 6: Verteilung der aquatischen Umweltparameter mittlere Wassertemperatur und elektrischer Leitfähigkeit der Bundeswasserstraßen.

2.2.3 Verbreitungsdaten

Insgesamt wurden fünf Datensätze zur Verbreitung von Arten ermittelt, die von Einzelpersonen für dieses Projekt zur Verfügung gestellt wurden. Keiner dieser Datensätze ist veröffentlicht. Die verwendeten Datensätze beinhalten Fundorte und Zeitpunkte der Erstfunde von *Senecio inaequidens* (Prof. Jürgen Böhmer, University of the South Pacific & Dr. Tina Heger, Universität Potsdam sowie ein zweiter Datensatz von Prof. Frank Schurr, Universität Hohenheim), *Ambrosia artemisiifolia* (Prof. Ingo Kowarik & Dr. Moritz von der Lippe, TU Berlin;), *Dikerogammarus villosus* und *Corbicula fluminea* (beide Dr. Franz Schöll, Bundesanstalt für Gewässerkunde). Alle Datensätze zeigten die Ausbreitung von einzelnen Arten über mehrere Jahre in Deutschland. Alle Datensätze mussten zunächst für eine weitere Verwendung in diesem Projekt entsprechend aufbereitet werden. So musste z.B. ein Datensatz von *Senecio inaequidens* zunächst von Hand digitalisiert werden (Abbildung 7).

Die Datensätze unterschieden sich in Umfang, Qualität und Art der Messwerte zum Teil erheblich. Im Laufe des Projekts stellten sich zwei der fünf Datensätze als ungenügend für die Erreichung der Projektziele heraus. Der Vergleich der beiden Datensätze von *Senecio inaequidens* ergab, dass ein Datensatz deutlich weniger Messpunkte mit größerer räumlicher Auflösung aufwies, so dass nur der Datensatz von Prof. Jürgen Böhmer und Dr. Tina Heger für die Kalibrierung verwendet werden konnte. Der Datensatz mit Vorkommen von *Ambrosia artemisiifolia* zeigt die Entwicklung von Populationen der Art entlang von Straßen innerhalb eines vergleichsweise kleinen Untersuchungsgebiets. Nach eingehender Untersuchung stellte sich heraus, dass diese Daten hauptsächlich die lokalen Populationsdynamiken und weniger die Ausbreitung entlang von Straßen und Schienen zeigte. Da Populationsdynamiken nicht im Projektvorhaben berücksichtigt werden konnten, kann das Modell die beobachteten Muster nicht reproduzieren. Dieser Datensatz konnte daher ebenfalls nicht verwendet werden.

In Abhängigkeit der Qualität und räumlichen und zeitlichen Verteilung der Daten wurde entschieden, in welcher Form die Daten verwendet werden können, um die besten Ergebnisse der Simulation zu erzielen. Dabei mussten verschiedene Aspekte gegeneinander abgewogen werden, wie z.B. die möglichst gleichmäßige Verteilung von Vorkommen („presence“) und Abwesenheiten („absence“), die Auswahl eines zeitlichen und räumlichen Bereichs, in dem tatsächliche Ausbreitung beobachtet wurde, die auch mit dem Modell reproduziert werden kann, und die mögliche Aufteilung der Daten für die Kalibrierung und Validierung. Dieses Vorgehen wird im Folgenden näher erläutert.

Senecio inaequidens

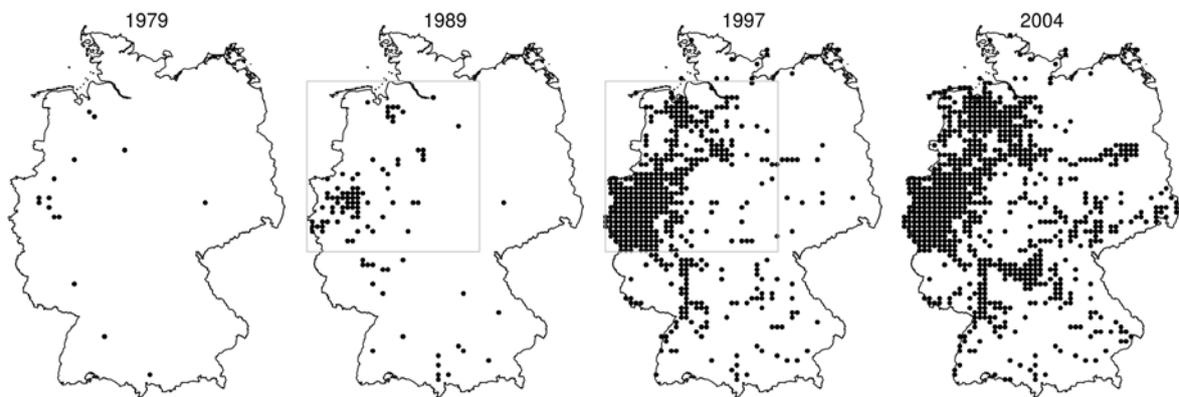


Abbildung 7: Vorkommen von *Senecio inaequidens* für vier Jahre (Heger & Böhmer, 2005). Punkte innerhalb des grauen Kastens wurde zur Kalibrierung verwendet.

Der Datensatz von Prof. Jürgen Böhmer und Dr. Tina Heger für *Senecio inaequidens* beinhaltet Vorkommen der Art auf der räumlichen Auflösung von Rasterzellen für Deutschland zu vier Zeitpunkten (Abbildung 7). Die räumliche Ausdehnung einer Rasterzelle beträgt etwas mehr als 10 km. Sofern die Art ein Vorkommen in einer Rasterzelle aufwies, wurden alle Knoten der terrestrischen Verkehrsnetzwerke in einem entsprechenden Umkreis um den zentralen Punkt der Rasterzelle als Fundort klassifiziert. Fehlende Datenpunkte wurden als Abwesenheit der Art interpretiert. Da es sich bei diesen Punkten aber in der Regel um fehlende Beobachtungen anstatt gesicherten Abwesenheiten handelt, wurde in der Kalibrierung eine Wahrscheinlichkeit von 0.1 für das Vorkommen der Art in diesen Rasterzellen angenommen. Damit geht die Unsicherheit in der Datenlage in die Kalibrierung mit ein.

Bei der Betrachtung der Ausbreitung fällt auf, dass es einen Schwerpunkt des Vorkommens zu Beginn der Ausbreitung in Nordrhein-Westfalen gab, dessen Ausbreitung gut mit Verkehr erklärt werden könnte. Daneben tauchen aber viele isolierte Vorkommen auf, die mit einem auf Verkehr basierten Modell allein höchstwahrscheinlich nicht erklärt werden können. Weiterhin erschwert die hohe Anzahl an Abwesenheiten in Süd- und Ostdeutschland die Kalibrierung des Modells. Idealerweise verwendet man eine ähnliche Anzahl an Vorkommen und Abwesenheiten, um ein einseitiges Gewicht zugunsten des Einen oder Anderen bei der Berechnung der Güte⁴ der Modellvorhersagen zu vermeiden. Daher wurde die Kalibrierung nur für den nord-westlichen Bereich des Vorkommens (grauer Kasten in Abbildung 7) für die Entwicklung von 1989 bis 1997 durchgeführt. Zur Validierung der Modellergebnisse wurde das Vorkommen der Art im gesamten Bundesgebiet im Jahr 2004 verwendet.

⁴Die Modellgüte beschreibt die Genauigkeit, mit der das Modell die Vorkommen der Arten reproduzieren kann.

Dikerogammarus villosus

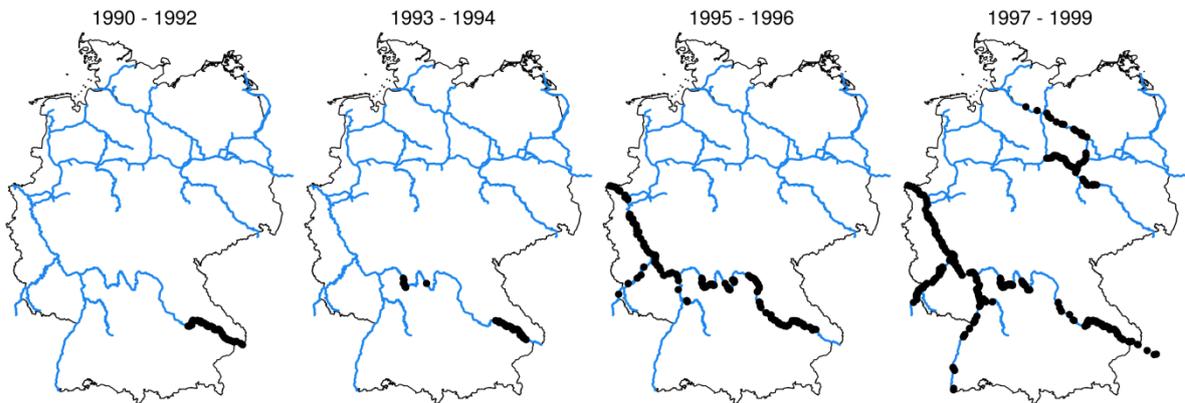


Abbildung 8: Vorkommen von *Dikerogammarus villosus* in deutschen Bundeswasserstraßen über vier Zeiträume.

Die Datensätze der beiden aquatischen Arten wiesen eine deutlich höhere räumliche und zeitliche Auflösung auf als die Datensätze der terrestrischen Arten. Abbildung 8 zeigt das Vorkommen von *Dikerogammarus villosus* in vier verschiedenen Zeitperioden. Für jede Zeitperiode liegen mehrere Zeitpunkte mit Messdaten vor. Im Zeitraum von 1990 – 1999 wurden insgesamt 96 Probenahmen zur Verfügung gestellt, wobei die räumliche Abdeckung nicht konsistent war. Die Daten liegen punktgenau mit hochauflösenden Koordinaten vor. Bei den Probenahmen handelte es sich in der Regel um vollständige Aufnahme des Makrozoobenthos. Im Gegensatz zum Datensatz von *Senecio inaequidens* kann man daher davon ausgehen, dass ein Fehlen der Art in den Proben die tatsächliche Abwesenheit der Art an diesem Ort darstellt. Abbildung 8 zeigt auch, dass die Ausbreitung von *Dikerogammarus villosus* nicht immer kontinuierlich verlief, da nicht alle Vorkommen zu späteren Zeitpunkten bestätigt werden konnten. So tauchten im unteren Donaublich Anfang der 90er Jahre Populationen auf, die in darauffolgenden Beprobungen nicht gefunden wurden. Solche augenscheinlichen Verluste an Individuen kann das Modell nicht vorhersagen. Es wurden alle Daten bis einschließlich 1996 zur Kalibrierung des Modells und Vorkommen in 1997 – 1999 zur Validierung verwendet. Es ist zu beachten, dass Probenahme nur an Flüssen, aber nicht an Kanälen durchgeführt wurde.

Corbicula fluminea

Die Qualität des Datensatzes zum Vorkommen von *Corbicula fluminea* war vergleichbar mit der von *Dikerogammarus villosus*. Die Interpretation der Datenpunkte konnte daher in der gleichen Weise durchgeführt werden wie für *Dikerogammarus villosus* weiter oben beschrieben. Abbildung 9 zeigt das Vorkommen und die zeitliche Entwicklung von *Corbicula fluminea*. Es fällt auf, dass ab 1993 eine isolierte Population in der Weser bestand. Allerdings fehlen auch für diese Art Probenahmen an Kanälen, so dass dieses plötzliche Auftauchen eher durch eine unvollständige Datengrundlage als durch tatsächliche Ausbreitung erklärt werden kann. Für ein rein auf Verkehr beruhendes Modell sind solche anscheinend sprunghaften Dynamiken allerdings nur schwer zu erfassen. Wie im obigen Fall werden die Jahre 1990 – 1996 zur Modellkalibrierung und 1997 – 1999 zur Validierung der Modellergebnisse herangezogen.

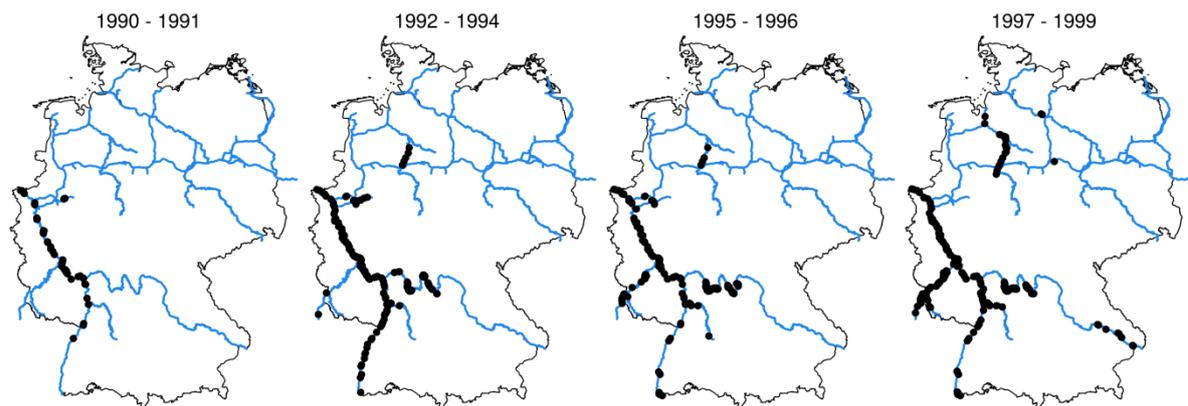


Abbildung 9: Vorkommen von *Corbicula fluminea* in deutschen Bundeswasserstraßen über vier Zeiträume.

2.3 Modellentwicklung (AP 3)

Ein Computermodell, das die Ausbreitung von Neobiota über verschiedene Verkehrsträger flächendeckend für ein ganzes Land simuliert, existiert weltweit nicht, so dass es für dieses Projektvorhaben neu entwickelt wurde. Ziel war es, ein Modell zu entwickeln, das sowohl die Ausbreitung einzelner Arten über verschiedene Kombinationen an Verkehrsträgern und Verkehrsflüssen simulieren kann, als auch generelle Aussagen zur Ausbreitung innerhalb Deutschlands zulässt.

2.3.1 Überblick

Das Modell basiert auf den Verkehrsnetzwerken für Straße, Schiene und Binnengewässer. Es können daher nur Vorhersagen für diese Netzwerke und die entsprechenden Knoten und Verbindungen gemacht werden. Das Modell ist modular aufgebaut, so dass einzelne Komponenten ein- bzw. ausgeschaltet werden können. Komponenten stellen die verschiedenen Ausbreitungswege und Verkehrsnetzwerke dar (Tabelle 4). So kann z.B. die natürliche Ausbreitung entlang der Wasserstraßen mit der Ausbreitung über den Schiffsverkehr verglichen werden. Die aquatische Komponente liegt getrennt von der terrestrischen Komponente vor, da es unwahrscheinlich erschien, dass aquatische Organismen über Straßen- oder Schienenverkehr oder terrestrische Arten im Gewässer transportiert werden. Ein Wechsel zwischen aquatischen und terrestrischen Systemen ist daher im Modell ausgeschlossen. Im terrestrischen Bereich sind die entsprechenden Verkehrsnetzwerke miteinander verschnitten, so dass die Ausbreitung von Arten verkehrsträgerübergreifend (Schiene und Straße) simuliert wird.

Informationen zum Transport über Container und palettierte Güter liegen auf der räumlichen Skala der 7332 Verkehrszellen vor, so dass die Ausbreitung über diese Wege separat von den Verkehrsträgern implementiert werden musste. Hinzu kommt, dass diese Güter verkehrsträgerübergreifend transportiert werden, so dass auch die Ausbreitung von Neobiota über die verschiedenen Verkehrsträger stattfindet. Für Container und palettierte Güter wurden nur Neobiota berücksichtigt, die sich innerhalb der Container und in den Paletten (z.B. holzbohrende Käfer) befinden. Daher spielt der Weg des Transports innerhalb Deutschlands keine wesentliche Rolle, so dass nur die Start- und Zielpunkte für palettierte Güter verwendet wurden, unabhängig von der tatsächlich genommenen Route. Container stellen Seefracht dar, deren Ursprung in diesem Projekt aufgrund des Fokus auf Deutschland nicht berücksichtigt

werden konnte. Somit wurden für Container nur die Zielpunkte aufgenommen in der Annahme, dass die Container an diesen Punkten geöffnet werden, welches potentiell die Gefahr der Freilassung von Neobiota birgt.

TABELLE 4: ÜBERSICHT DER AUSBREITUNGSWEGE, DIE IN DEM MODELL BERÜCKSICHTIGT WURDEN.

NETZWERKE	ANHEFTUNG	FAHRTWIND	INNERHALB TRANSPORTEIN- HEIT	KORRI- DOR ⁵
STRAÙE	x	x		x
SCHIENE	x	x		x
BINNENWASSER- STRAÙE	x ⁶		x ⁷	x
PALLETISIERT GÜTER			x	
CONTAINER			x	

Diese einzelnen Modellkomponenten werden von einem allgemeinen Rahmen des Modells eingefasst, das die verschiedenen Datensätze und Komponenten miteinander kombiniert, die Initialisierung und Berechnungen der Modellkomponenten vornimmt und die Ausgabe in Text und Karte regelt. Die zugrundeliegenden Skripte sind in der Sprache R geschrieben. Alle Skripte und Datensätze sind in einem R Paket mit dem Namen „CASPIAN“ zusammengefasst und stehen online zur Verfügung (s. Abschnitt Anlage 8.2. für eine detaillierte Beschreibung des Pakets).

2.3.2 Modellstruktur

Das Modell berechnet die Wahrscheinlichkeit einer Ausbreitung und Etablierung von gebietsfremden Arten. Die Berechnung kann hierbei in einzelne Teilschritte unterteilt werden (Abbildung 10). Im ersten Schritt werden die Datensätze, die zur Berechnung der Wahrscheinlichkeiten notwendig sind, geladen. Diese werden verwendet, um in einem zweiten Schritt die Wahrscheinlichkeit der Ausbreitung und der Etablierung zu berechnen. Zur Berechnung der Ausbreitungswahrscheinlichkeit werden für jeden Ausbreitungsweg (Tabelle 4) sogenannte „dispersal kernels“ verwendet, die die Wahrscheinlichkeit der Ausbreitung über die geographische Distanz beschreiben. Abhängig vom Ausbreitungsweg gehen weitere Faktoren wie Verkehrsintensität in die Berechnung mit ein. Die ermittelten Wahrscheinlichkeiten für jeden Ausbreitungsweg werden dann zusammengefasst und als „Probability of introduction“ $P(I)$ ausgegeben (Abbildung 10). Die Wahrscheinlichkeit der Etablierung einer Art berechnet sich aus der Abweichung der lokalen Umwelt- und Habitateigenschaften von den Präferenzen der jeweiligen Art und wird in einem zweiten Schritt berechnet („Probability of establishment“ $P(E)$). Im letzten Schritt werden die Wahrscheinlichkeiten der Ausbreitung und Etablierung zu einer gesamten Invasionswahrscheinlichkeit

⁵ „Korridor“ beschreibt die Ausbreitung entlang eines Infrastruktursegments ohne die Hilfe des Verkehrs.

⁶ In Form von Schiffsaufwuchs

⁷ Im Ballastwasser

kombiniert („Probability of invasion“ $P(Inv)$). Der Begriff „Invasion“ bezieht sich hierbei nur auf die Phasen der Ausbreitung und Etablierung. Mögliche Konsequenzen, die dem Begriff der „invasiven Arten“ zugrunde liegen, stehen außerhalb des Fokus dieses Forschungsvorhabens. Die Modellergebnisse werden anschließend als Tabellen und Karten ausgegeben. Im Folgenden werden die einzelnen Schritte näher erläutert.

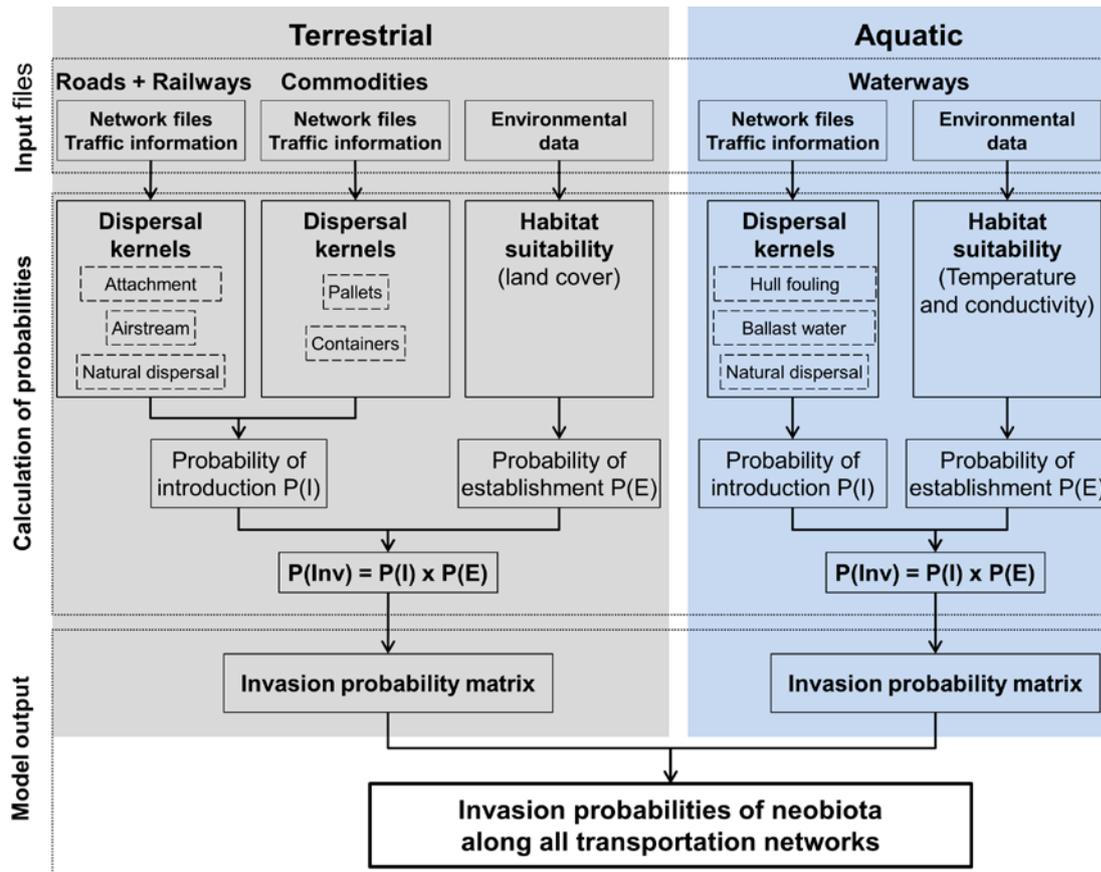


Abbildung 10: Schematische Darstellung der Modellstruktur. Die Eingangsvariablen umfassen Informationen zu Verkehrsnetzwerken und Verkehrsbelastungen und Umweltvariablen („input files“). Die Wahrscheinlichkeiten der Ausbreitung werden mit Funktionen („dispersal kernels“) für jeden Ausbreitungsweg separat berechnet und anschließend zusammengefasst. Die Wahrscheinlichkeiten der Ausbreitung ($P(I)$) und Etablierung ($P(E)$) werden für terrestrische und aquatische Habitate separat berechnet. Die Gesamtwahrscheinlichkeit des Auftretens einer Art ($P(Inv)$) ergibt sich aus dem Produkt beider Einzelwahrscheinlichkeiten. Am Ende stellt das Modell Ausbreitungs- und Etablierungswahrscheinlichkeiten für die entsprechenden Transportnetzwerke zur Verfügung.

2.3.2.1 Eingangsdatensätze

Die Eingangsdatensätze müssen entsprechend den Vorgaben des Modells aufbereitet werden. Die Transportnetzwerke müssen als Tabellen vorliegen, in der jede Zeile einen Link des Netzwerks (z.B. ein Straßensegment) beschreibt. Zu jedem Link werden in derselben Zeile Informationen zu Charakteristiken des Links beigefügt wie z.B. Länge und Verkehrsintensität. Zusätzlich benötigt das Modell für jeden Link ein Maß für die Güte des Habitats entlang dieses Links, welches als weitere Spalte eingefügt ist. Container- und Palettenflüsse können nicht für einzelne Links angegeben werden, so dass ein weiteres Format

für diese Ausbreitungswege verwendet werden muss. Die Flüsse der Paletten werden zwischen räumlichen Einheiten der Verkehrszellen simuliert. Der Eingangsdatensatz besteht daher aus einer Tabelle mit der Anzahl der transportierten Paletten zwischen zwei Verkehrszellen in jeder Zeile. Für Container wird in jeder Zeile die Ziel-Verkehrszelle aufgeführt, da Startgebiete außerhalb Deutschlands liegen und daher nicht berücksichtigt werden können. Die Formate dieser Eingangsdatensätze werden im Abschnitt Anhang 8.2 im Detail beschrieben.

Das Einlesen der Datensätze wurde möglichst benutzerfreundlich gestaltet. Die Formate der Eingangsdatensätze sind daher shapefiles und einfache Tabellenformate. Allerdings verlängert sich dadurch die Zeit des Einlesens erheblich. Das Einlesen kann beschleunigt werden, indem komprimierte Formate verwendet werden, die allerdings eine Transformation der ursprünglichen Datensätze erfordern. Eine mögliche Vorgehensweise hierfür ist im Anhang 8.2 Abschnitt „Anwendung – Starten von CASPIAN“ aufgeführt.

2.3.2.2 Ausbreitung

Die Berechnung der Ausbreitungswahrscheinlichkeit wird auf der Basis der Transportnetzwerke durchgeführt. Der Bezug zu einem Netzwerk bestehend aus Knoten und verbindenden Segmenten bedingt, dass das Modell in einem Zeitschritt die Ausbreitung immer von einem Knoten zum nächsten Knoten berechnet. Eine direkte Berechnung zu allen anderen Knoten des Netzwerks ist aufgrund der hohen Anzahl der möglichen Routen nicht möglich. Die Wahrscheinlichkeit der Ausbreitung wird durch verschiedene Parameter des Streckenabschnitts wie z.B. Streckenlänge oder Verkehrsdichte bestimmt. Somit ist die Wahrscheinlichkeit einer Ausbreitung über große Streckenlängen mit wenig Verkehr geringer als über kurze Streckensegmente mit hohem Verkehrsaufkommen. Ein Zeitschritt des Modells beschreibt die Ausbreitung innerhalb 1/12 des Jahres, also etwa eines Monats, von einem Knoten zum nächsten. Die Simulation der Ausbreitung findet hochaufgelöst bis auf die Ebene der Kreisstraßen deutschlandweit statt. Dies kann vom Benutzer im Vorfeld angepasst werden. Für das Straßennetzwerk können unterschiedliche Kategorien an Straßen in Abhängigkeit ihrer Größe ausgewählt werden (z.B. Autobahnen, Landstraßen oder Kreisstraßen), wodurch die Berechnung der Simulation beschleunigt werden kann.

Die Berechnung der Ausbreitungswahrscheinlichkeiten basiert auf „dispersal kernels“ oder Ausbreitungsfunktionen f , die sofern vorhanden aus der Literatur entnommen wurden (Tabelle 5, Abbildung 11). Soweit möglich wurden gängige Funktionen der Ausbreitung verwendet, die die betrachteten Prozesse möglichst einfach beschreiben. Damit erhält man eine Auswahl an Funktionen, die weit verbreitet sind und damit häufig getestet wurden, und gleichzeitig eine vergleichsweise niedrige Anzahl an Parametern aufweisen. Es wurde versucht, die veröffentlichten Funktionen soweit wie möglich zu übernehmen. Da aber die veröffentlichten Funktionen nicht immer genau die Dynamiken beschreiben, welche für das Modell relevant sind, mussten in einigen Fällen Anpassungen nach eigener Einschätzung vorgenommen werden. Die Funktionen beschreiben in der Regel die Ausbreitung von Individuen für ein Fahrzeug oder Schiff. Die Wahrscheinlichkeit der Ausbreitung $P(I)$ über alle Verkehrsteilnehmer entlang einer Distanz D zwischen Knoten i und j wurde dann entsprechend der Verkehrsintensität T hochgerechnet (Leung, Drake, & Lodge, 2004; Liebhold, Brockerhoff, & Kimberley, 2017):

$$P(I)_{ij} = 1 - \left(1 - f(D_{ij})\right)^{\alpha T_{ij}}. \quad \text{Formel (1)}$$

α beschreibt die Krümmung der Funktion, mit der Verkehr die Ausbreitungswahrscheinlichkeit über Anheftung an Fahrzeugen beeinflusst, und ist für Fahrtwind des Straßen- (α_{FS}) und Schienenverkehrs (α_{FR}) und Anheftung (α_{AS} und α_{AR}) separat implementiert. Die verschiedenen Ausbreitungsfunktionen f sind in

Tabelle 5 aufgeführt. Die Wahrscheinlichkeit der Einfuhr von Individuen an Knoten j ergab sich demnach aus Ausbreitungswahrscheinlichkeiten aller Links nach j :

$$P(I)_j = 1 - \prod_i 1 - P(I)_{ij}. \quad \text{Formel (2)}$$

In der oben beschriebenen Form lassen sich alle Prozesse beschreiben, die einen Start- und Zielpunkt innehaben. Die Ausbreitung über Container und Ballastwasser musste davon abweichend implementiert werden, da die Startpunkte unbekannt waren. Im Falle von Containertransport war nur bekannt, wo die

TABELLE 5: ÜBERSICHT ÜBER FUNKTIONEN DER AUSBREITUNG.

NR.	AUSBREITUNG	FORMEL f	PARAMETER	REFERENZ
1	Natürliche Ausbreitung (terrestrisch)	$\frac{b}{2\pi a^2 \Gamma(\frac{2}{b})} e^{\left(\frac{-D^b}{a^b}\right)}$	D: Distanz a: Formparameter b: Formparameter	Nathan et al. 2012
2	Anheftung an Fahrzeug	$e^{(ce^{(dD^g)})}$	D: Distanz c: Formparameter d: Formparameter g: Formparameter	Taylor et al. 2012
3	Fahrtwind	$\frac{1}{hD\sqrt{2\pi}} e^{\left(\frac{-(\log(D)-k)^2}{2h^2}\right)}$	D: Distanz h: Formparameter k: Formparameter	von der Lippe et al. 2012
4A	Aufwuchs	e^{-lD^m}	D: Distanz l: Formparameter m: Formparameter	Sylvester et al. 2012
4B	Aufwuchs	$(1 - e^{-wP^x})(1 - e^{-yS^z})e^{-lD^m}$	D: Distanz l: Formparameter m: Formparameter P: Zeit seit letztem Anstrich w: Formparameter x: Formparameter S: Aufenthalt im Hafen y: Formparameter z: Formparameter	Sylvester et al. 2011
5	Ballastwasser	$\frac{G}{n + G}$	G: Volumen Güterumschlag n: Formparameter	
6	Natürliche Ausbreitung (aquatisch)	qD^r	D: Distanz q: Formparameter r: Formparameter	Elliot et al. 2003
7	Paletten	$\frac{P}{v + P}$	P: Anzahl Paletten v: Formparameter	
8	Container	$\frac{C}{u + C}$	C: Anzahl Container u: Formparameter	

Container die deutsche Grenze passierten. Die Herkunftsländer konnten aufgrund des Fokus des Forschungsvorhabens auf Deutschland nicht berücksichtigt werden. Im Fall von Ballastwasser waren die Herkunftsgebiete gänzlich unbekannt. Für den Transport mittels Containern und Ballastwasser wurde angenommen, dass sich eine höhere Wahrscheinlichkeit der Einfuhr von Neobiota an Orten ergab, an denen Container geöffnet bzw. Ballastwasser abgelassen wurde. Anders formuliert bedeutet dies, dass der Transport über Ballastwasser und Container nur für $P(I)_j$ aber nicht für $P(I)_{ij}$ ermittelt werden konnte. Im Folgenden werden die einzelnen Ausbreitungswege vorgestellt:

- **Natürliche Ausbreitung:** Die natürliche Ausbreitung beschreibt den Vorgang, dass sich Individuen entlang von Verkehrsinfrastruktur ohne die Hilfe von Verkehr oder Verkehrsteilnehmern ausbreiten. Dies wurde für terrestrische und aquatische Transportnetzwerke separat implementiert. Für die terrestrische Ausbreitung wurde eine exponentielle „Power“ Funktion (Nr. 1 in Tabelle 5) verwendet, die von Nathan et al. (2012) als gut geeignet für Samenausbreitung beschrieben wurde. Für Wasserstraßen lieferte eine einfachere Form (Nr. 6) bei einem Test für 10 aquatischen Invertebraten sehr gute Ergebnisse (Elliott, 2003). Diese bezieht sich nur auf die natürliche Ausbreitung von aquatischen Organismen (keine Ufervegetation).
- **Anheftung:** Anheftung beschreibt den Prozess der Ausbreitung über die Anheftung an die Karosserie oder Räder von Autos, Lastwagen oder Zügen. Die entsprechende Ausbreitungsfunktion (Nr. 2) wurde von Taylor et al. (2012) als am besten geeignete Funktion ermittelt und übernommen.
- **Fahrtwind:** Die Ausbreitung über Fahrtwind beschreibt den Prozess, dass Individuen oder Samen mit dem Fahrtwind entlang der Verkehrsträger verbreitet werden. Eine log-normale Funktion (Nr. 3) hat sich als geeignetste Beschreibung der Ausbreitung ergeben (von der Lippe, Bullock, Kowarik, Knopp & Wichmann, 2013). Diese Funktion wurde für Straßen- und Schienenverkehr verwendet, wobei der Einfluss einzelner Fahrzeuge über die Formel (1) für beide Verkehrsträger separat kalibriert wurde (unterschiedliches α für Straßen- und Schienenverkehr), um die Unterschiede des Fahrtwindes für Schienen- und Straßenverkehr berücksichtigen zu können.
- **Aufwuchs:** Aufwuchs beschreibt den Prozess des Transports von Individuen, die sich an im Wasser befindlichen Oberflächen von Schiffen wie z.B. Rumpf oder Schiffsschraube anheften können. Dies können sessile Organismen wie Muscheln oder Algen aber auch Kleinkrebse sein, die sich im Aufwuchs festhalten können. Die Ausbreitungsfunktion (Nr. 4) wurde von Sylvester et al. (2011) übernommen. Die ursprüngliche Funktion (Nr. 4B) beinhaltet die Komponenten „Aufenthaltsdauer im Hafen“ und „Zeit seit letztem Anstrich“. Diese Komponenten sind im Modell implementiert, gehen aber nicht die Simulation ein, da die entsprechenden Daten fehlen. Sofern Daten zur Verfügung stehen, können sie zur Simulation verwendet werden. In der aktuellen Form wird Transport über Aufwuchs als Funktion der Distanz berechnet. Die ursprüngliche Funktion reduziert sich daher auf die Funktion Nr. 4A.
- **Ballastwasser:** Dies beschreibt den Prozess, dass Individuen im Ballastwasser von Frachtschiffen transportiert und an anderer Stelle abgelassen werden. Da keine Informationen über die Verfrachtung von Ballastwasser oder die Bewegung von einzelnen Schiffen vorlagen, wurde angenommen, dass Ballastwasser in Häfen ausgetauscht wird. Weiterhin wurde angenommen, dass die Gesamtmenge an ausgetauschtem Ballastwasser mit der Hafengröße zunimmt. Somit erhalten größere Häfen mehr Ballastwasser als kleinere Häfen. Im Modell wurde daher die Einfuhr über Ballastwasser nur an den Knoten des Schiffsnetzwerks, die Häfen darstellen, und entsprechend des Güterumschlags des jeweiligen Hafens berücksichtigt. Der Güterumschlag der Häfen wurde aus der Güterverkehrsstatistik der Binnenschifffahrt des Statistischen Bundesamts von 2016 entnommen. Da keine geeignete Funktion zu diesem Ausbreitungsweg publiziert vorlag, wurde eine einfache saturierende Wahrscheinlichkeitsfunktion angenommen (Nr. 5).
- **Container:** Die Ausbreitung über Container beschreibt den Prozess, dass Individuen in Containern befördert werden und bei Öffnung des Containers die Möglichkeit einer Etablierung am Ort der Öffnung haben. Der Transportweg des Containers und die Art des Transports (Schiene, Straße oder Schiff) spielen für diese Ausbreitung keine Rolle. Daher wurde nur der Zielpunkt mit einer erhöhten Wahrscheinlichkeit der Einfuhr entsprechend der Anzahl an Containern berücksichtigt. Es lagen

keine Untersuchungen zur Wahrscheinlichkeit der Einfuhr über Container vor, so dass eine einfache saturierende Wahrscheinlichkeitsfunktion (Nr. 8) verwendet wurde. Der Parameter dieser Funktion konnte aufgrund mangelnder Daten nicht kalibriert werden. Die Wahrscheinlichkeiten über diesen Ausbreitungsweg können daher nicht getestet und interpretiert werden. Dies kann erfolgen, sobald geeignete Daten zur Kalibrierung oder Ergebnisse aus anderen Studien über z.B. die Häufigkeit des Eintrags von Neobiota über Container vorliegen.

- **Paletten:** Der Transport von Paletten beschreibt den Prozess der Einfuhr von Organismen auf oder in Paletten. Dies stellen in der Regel holzbohrende Arten wie einige Forstschädlinge dar. Da für Paletten Informationen zu Start- und Zielgebiet vorlagen, wurden diese Transporte ebenfalls auf Basis eines Transportnetzwerks simuliert. Da die Auflösung für die Ströme palettisierter Güter aber größer als die Auflösung für Verkehr war, wurde der Transport über Paletten separat von den Transportnetzwerken vorgenommen. Die berechneten Ausbreitungswahrscheinlichkeiten wurden in jedem Zeitschritt der Simulation zu den Knoten der Transportnetzwerke hinzugefügt, um eine einheitliche Auflösung zu erreichen. Die Ausbreitungsfunktion (Nr. 7) stellt ebenfalls eine einfache saturierende Wahrscheinlichkeitsfunktion dar, die ähnlich dem Containertransport aufgrund mangelnder Daten nicht kalibriert werden konnte. Auch für diesen Weg gilt, dass die Wahrscheinlichkeiten der Ausbreitung nicht getestet und interpretiert werden konnte. Dies kann erfolgen, sobald geeignete Daten zur Kalibrierung oder Ergebnisse aus anderen Studien über z.B. die Häufigkeit des Eintrags von Neobiota über Paletten vorliegen.

Die Ausbreitungswahrscheinlichkeit wird für jeden Ausbreitungsweg k (z.B. Anheftung oder Fahrtwind) individuell berechnet und zu einer Gesamt-Ausbreitungswahrscheinlichkeit $P(I)_{ij}$ für jeden Link der Transportnetzwerke zusammengefasst (Abbildung 11): $P(I)_{ij} = 1 - \prod_k 1 - P(I)_{ijk}$.

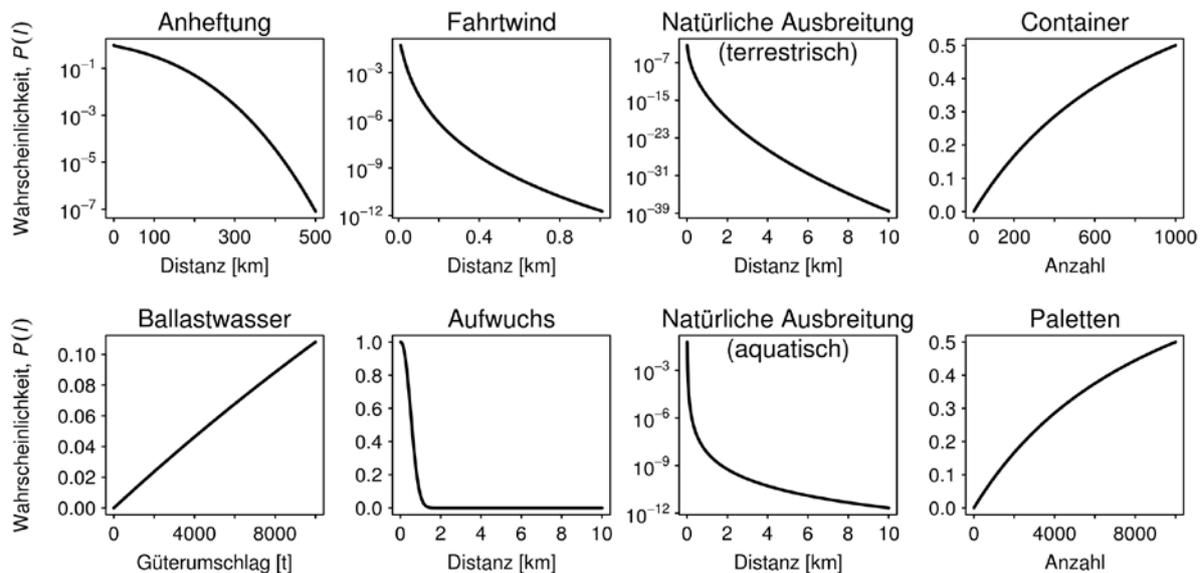


Abbildung 12: Darstellung der Funktionen f aus Tabelle 5 unter Verwendung der kalibrierten Parameter für *Senecio inaequidens* und *Dikerogammarus villosus*.

2.3.2.3 Etablierung

Die Berechnung der Wahrscheinlichkeit einer Etablierung einer neuen Population geschieht auf Basis der lokalen Habitatcharakterisierung und der Präferenzen der jeweiligen Art. Die Eignung eines Habitats

ergibt sich aus den Abweichungen der Habitatcharakteristiken und Artpräferenzen. Die Habitatcharakteristiken sind durch das Modell vorgegeben, während die Artpräferenzen von der zu simulierenden Art abhängen und durch den Nutzer festgesetzt werden können. Dies wird in den folgenden Abschnitten und in Anhang 8.2 Abschnitt „Eingangsdatensätze“ im Detail beschrieben. Aufgrund der Vielzahl der möglichen Berechnungen der Habitateignung wurde in diesem Forschungsvorhaben ein genereller Ansatz gewählt, der für möglichst viele Arten Anwendung finden kann. Der Aufbau des Modells erlaubt eine einfache Kopplung mit deutlich komplexeren Habitatmodellen („species distribution models“), die eine große Vielfalt an Umweltbedingungen berücksichtigen können. Hierzu müssten zuerst Habitatmodelle für die zu simulierende Art aufgesetzt werden, was in der Softwareumgebung R möglich ist, und die Ergebnisse der Habitatmodelle könnten als Eingangsvariable für das vorliegende Invasionsmodell fungieren. Habitatmodelle finden häufig Anwendung und stehen daher in vielen Varianten und frei zugänglich zur Verfügung.

In diesem Forschungsvorhaben wurde die Habitateignung für terrestrische Habitate entlang von Schienen und Straßen über Landbedeckungen ermittelt. Wie im Abschnitt 2.2.2 beschrieben, liegen für jeden Link prozentuale Anteile von Landbedeckungsklassen vor. Die Habitateignung H entlang eines Links s berechnet sich aus den prozentualen Anteilen p der Landbedeckungsklassen N und der Präferenzen I der jeweiligen Art für die entsprechende Landbedeckungsklasse: $H_s = \sum_p (I_{ps} N_{ps})$. Bei einer Präferenz von $I = 1$ für eine Landbedeckungsklasse, die 100% der Landbedeckung entlang des Links s ausmacht, ergibt sich ein Maximalwert der Habitateignung von $H_s = 1$. Die Wahrscheinlichkeit der Etablierung $P(E)$ ergibt sich dann aus dem Produkt der Grundwahrscheinlichkeit einer Etablierung β_1 und H_s : $P(E)_s = \beta_1 H_s$. β_1 und beschreibt die Wahrscheinlichkeit einer Etablierung unter perfekten Bedingungen in terrestrischen Habitaten. Dies ist beispielhaft für *Senecio inaequidens*, als Vertreter für eine terrestrische Art, und eine Süßwasserart in Abbildung 12 gezeigt.

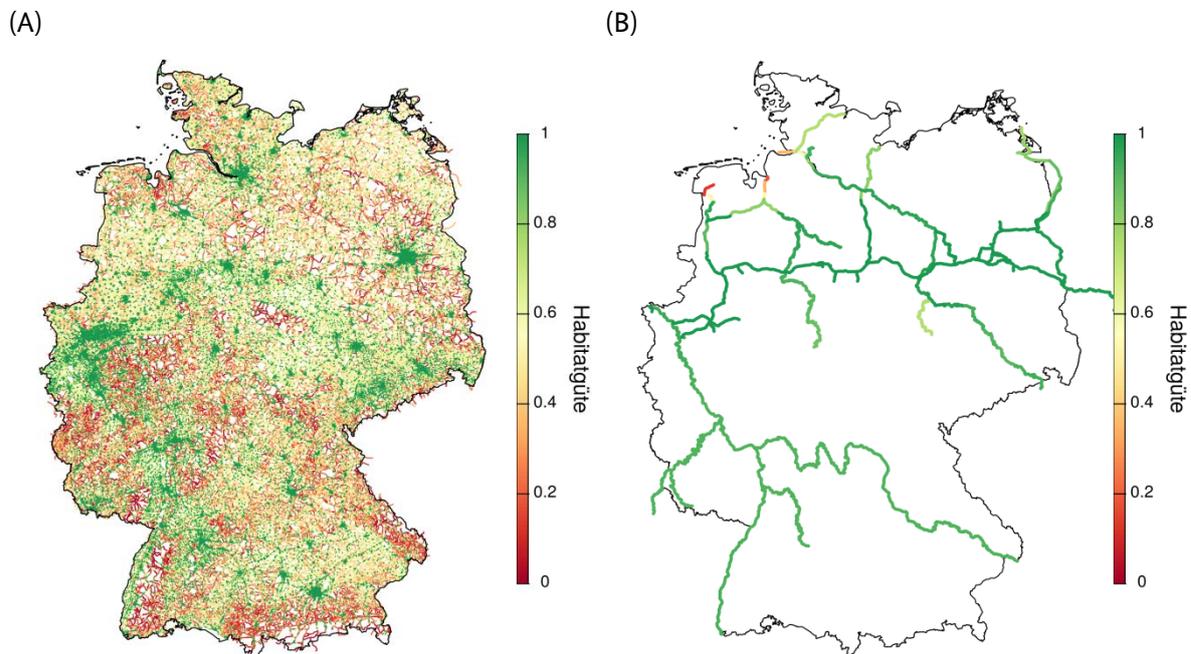


Abbildung 13: Exemplarische Darstellung der Habitatgüte. Dargestellt ist die Habitatgüte (A) für eine Pflanze mit guten Wuchsbedingungen in städtischen und ruderalen Bereichen wie *Senecio inaequidens* und (B) für aquatische Arten, die an mittlere klimatische Bedingungen in deutschen Binnengewässern angepasst sind (rot=schlecht geeignet ($H_s = 0$); grün=gut geeignet ($H_s = 1$)).

Für aquatische Habitate wurde die Habitateignung über die jährlichen Mittelwerte der Wassertemperatur und der elektrischen Leitfähigkeit abgeschätzt. Die Auswahl dieser Faktoren ist limitiert durch die

Verfügbarkeit der Messungen für verschiedene Gewässer (s. Absatz 2.2.2). Die Habitataignung wurde als Euklidische Distanz zwischen den Werten der Umweltparameter und der Artpräferenz berechnet:

$P(E)_s = \beta_2 \sqrt{(T_s - \bar{T})^2 + (L_s + \bar{L})^2}$, wobei T die Temperatur und L die elektrische Leitfähigkeit für Link s und \bar{T} und \bar{L} die jeweilige Artpräferenz darstellen. β_2 beschreibt die Wahrscheinlichkeit einer Etablierung unter perfekten Bedingungen in aquatischen Habitaten.

Die Gesamtwahrscheinlichkeit der Einführung und Etablierung $P(Inv)_s$, ergibt sich aus dem Produkt von $P(I)_s$ und $P(E)_s$.

2.3.2.4 Simulation der Ausbreitung und Etablierung

Das Modell simuliert den Verlauf einer Ausbreitung ausgehend von vorab definierten Startpunkten. Hierzu müssen Koordinaten des Vorkommens von Arten eingelesen werden. Anhand dieser Koordinaten werden die Startknoten ausgewählt, die als „invasiert“ angesehen werden. Dies stellt den Ausgangspunkt (oder die Initialisierung) der Simulation dar. Alle Knoten im Netzwerk erhalten zu Beginn der Simulation einen Wert N , der den Status des Vorkommens beschreibt. Dieser kann als Aufenthaltswahrscheinlichkeit interpretiert werden. Startknoten erhalten den Statuswert 1, während alle anderen Knoten zu Beginn den Status 0 haben.

Beim Start der Simulation berechnet das Modell nun anhand der in den Abschnitten 2.3.2.2 beschriebenen Abläufe die Wahrscheinlichkeit der Ausbreitung $P(I)$ zu allen benachbarten Knoten j . Die berechnete Wahrscheinlichkeit, einen Knoten j zu erreichen, ergibt sich nun aus dem Status des Startknotens i und der Ausbreitungswahrscheinlichkeit $P(I)_{ij}$: Je höher der Status des Startknotens und je höher die Ausbreitungswahrscheinlichkeit ist, desto höher ist die Wahrscheinlichkeit, den Knoten j zu erreichen. Bei einer Verbindung von nur zwei Knoten ergibt sich der Status des Zielknotens aus $N_j = N_i * P(I)_{ij}$. Falls mehrere Links zum Zielknoten existieren, werden der Status und die Ausbreitungswahrscheinlichkeiten der Links aller i zu j zusammengefasst:

$$N_j = 1 - \prod_i [1 - (N_i P(I)_{ij})]. \quad \text{Formel (3)}$$

In dieser Form wird in jedem Zeitschritt der Knotenstatus aller Knoten aktualisiert. Da die Ausbreitungswahrscheinlichkeit $P(I)$ abhängig von den Charakteristiken des Links, wie z.B. Streckenlänge oder Verkehrsdichte, ist, bleibt $P(I)$ bei der Simulation konstant für jeden Link, während N für jeden Knoten stetig aktualisiert wird. Dies beschreibt den Ausbreitungsprozess und ergibt die Wahrscheinlichkeit, dass ein Knoten erreicht wird. Zur Berechnung der Invasionswahrscheinlichkeit $P_f(Inv)$, wird der Status N_j eines Knotens mit der Wahrscheinlichkeit der Etablierung $P_j(E)$ an Knoten j verrechnet: $P_f(Inv) = N_j P_j(E)$. Die Etablierung einer Art an einem Ort hat demnach keinen Einfluss auf die weitere Ausbreitung einer Art. Dies kann so interpretiert werden, dass das Modell zuerst die Ausbreitung über z.B. Anheftung an Fahrzeuge über alle Knoten hinweg simuliert und dann die Etablierung nach Erreichen der Knoten berechnet. In der Natur können die Etablierung und der Aufbau größerer lokaler Populationen zu einer erhöhten Wahrscheinlichkeit der Ausbreitung führen, aber dies kann nur über die Einbindung von Populationsdynamiken in das Modell wiedergegeben werden, was nicht Bestandteil des Projektvorhabens war.

Die berechneten Wahrscheinlichkeiten der Ausbreitung und Etablierung werden in Form von Tabellen und als Karte (shapefile) ausgegeben. Ein beispielhaftes Modellergebnis zur Ausbreitung ist in Abbildung 13 dargestellt. Dies stellt die simulierte Ausbreitung von *Senecio inaequidens* dar. Die Originaldaten zur beobachteten Verbreitung dieser Art sind in Abbildung 7 dargestellt..

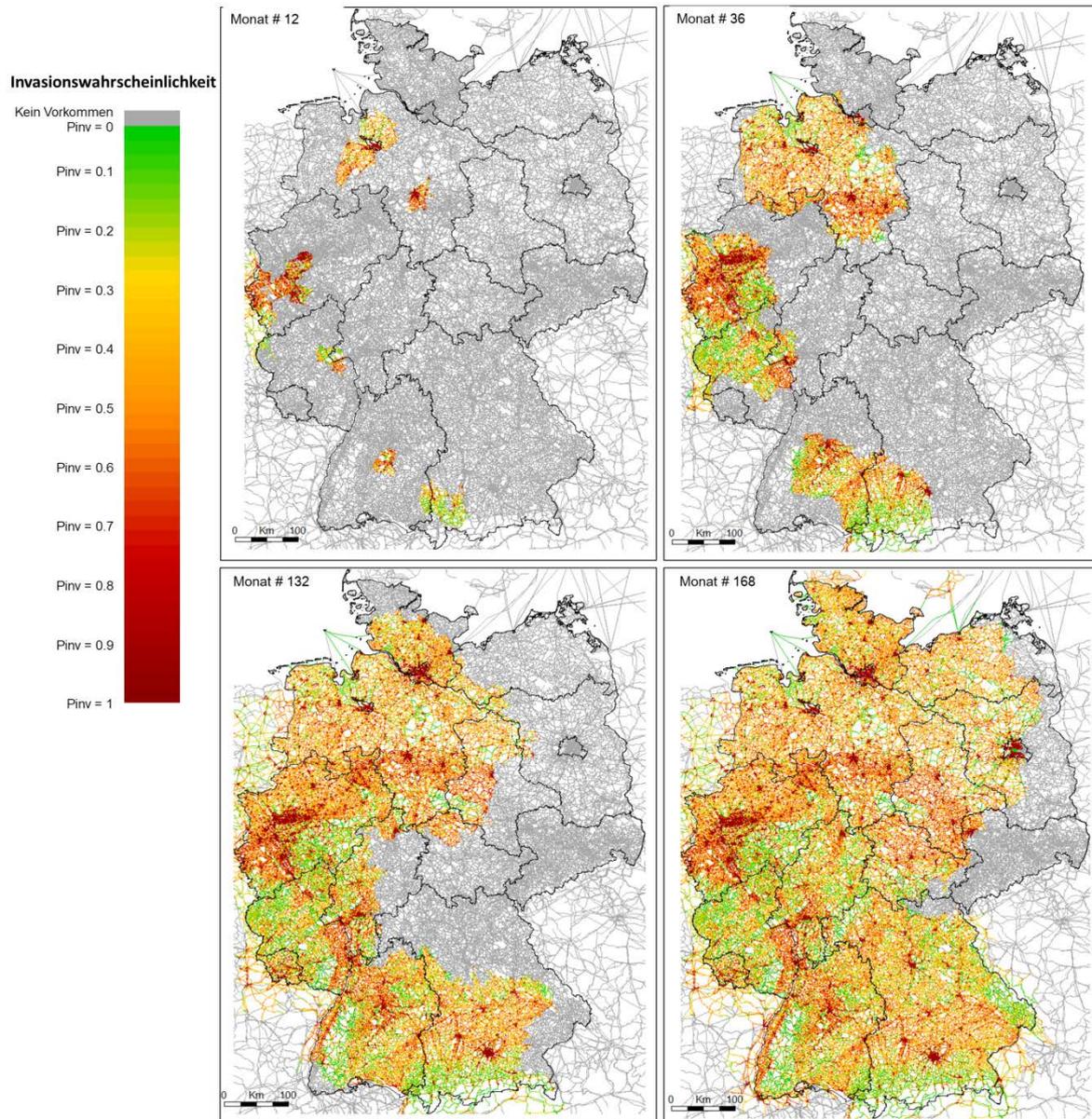


Abbildung 14: Beispielhafte Simulation der Ausbreitung von *Senecio inaequidens* durch den Straßen- und Schienenverkehr. Dargestellt ist die Wahrscheinlichkeit der Ausbreitung und Etablierung der Art, $P(Inv)$, ausgehend von Initialpopulationen.

2.3.3 Kalibrierung und Validierung des Modells

Bei der Kalibrierung des Modells werden in einem iterativen Verfahren die Werte der Modellparameter ermittelt. Hierzu werden Modellergebnisse mit Beobachtungen der Ausbreitung einer Art verglichen und eine Kombination an Parameterwerten ermittelt, mit denen die Modellergebnisse die Beobachtungen am besten widerspiegeln. Aufgrund der Datenverfügbarkeit und der Transportmechanismen der Beispielarten werden für Arten die entsprechenden Ausbreitungswege vorab gewählt. Bei *Senecio inaequidens* werden die Ausbreitungswege natürliche Ausbreitung, Anheftung und Fahrtwind für Straßen- und Schienenverkehr berücksichtigt. Für die aquatischen Organismen werden Aufwuchs, Ballastwasser und natürliche Ausbreitung kalibriert.

Die Modellparameter werden mittels Bayesscher Kalibrierung ermittelt. Die Bayessche Kalibrierung stellt eine weit entwickelte und technisch ausgereifte aber auch anspruchsvolle Art der Modellkalibrierung dar (Bagnara et al., 2018; Clark, 2004). Ein großer Vorteil des Bayesschen Ansatzes ist, dass Verteilungen der Parameterwerte betrachtet werden, die eine Abschätzung der Unsicherheit der Kalibrierung erlauben. Zur Initialisierung der Kalibrierung werden Verteilungen der zu ermittelnden Parameterwerte, die sogenannten „priors“, gewählt. Im vorliegenden Fall wurden nach eingehenden Tests Normalverteilungen für die Parameterwerte gewählt mit den Literaturwerten als Mittelwert sofern diese vorhanden waren. Damit wurde den Parameterwerten aus der Literatur Vorrang gegeben, aber gleichzeitig auch Abweichungen von diesen Werten durch die Kalibrierung zugelassen. Anhand dieser Anfangsverteilungen werden zufällige Startwerte der einzelnen Parameter gewählt und in einem iterativen Verfahren („Markov Chain Monte Carlo (MCMC)“) angepasst, so dass die Modellergebnisse die bestmögliche Übereinstimmung mit den beobachteten Werten ergeben. In der Regel sind 30.000 - 100.000 Iterationen nötig, um eine zufriedenstellende Konvergenz des Algorithmus zu erhalten, wobei in jedem Schritt eine unterschiedliche Kombination an Parameterwerten getestet werden kann. Ein weiterer Vorteil der Bayesschen Kalibrierung ist, dass auch für die ermittelten Parameter Verteilungsfunktionen (sogenannte „posteriors“) errechnet werden, wodurch die Unsicherheiten bzw. Variationen der Parameterschätzungen dargestellt werden können. Ein Nachteil dieser Art der Kalibrierung stellt die rechenintensive Berechnung dar, die teilweise durch Parallelisierung ausgeglichen werden kann. Um eine möglichst große Abdeckung des Parameterraums und robuste Ergebnisse zu erzielen, wurden neun MCMC Läufe parallel gestartet, wobei jeweils drei Läufe nach einer Anfangsphase („burn-in“) zu Einheiten zusammengefasst wurden, die dann miteinander kommunizieren, um eine schnellere Konvergenz zu erreichen. Die drei Gruppen liefen unabhängig voneinander, um robuste Ergebnisse zu erhalten. Zur Beschleunigung der Laufzeit wurden die Modellskripte auf Rechenzeit optimiert. Hiernach dauerte eine Kalibrierung in der Regel 1-5 Tage auf den leistungsstarken Workstations des Senckenberg Biodiversität und Klima Forschungszentrums.

Der Erfolg einer Kalibrierung wird durch eine Konvergenz des Algorithmus zu einem möglichst eindeutigen Satz an Parameterwerten mit geringer Unsicherheit dargestellt. Dies hängt von vielen Faktoren ab, unter anderem auch von den Datensätzen, die zur Kalibrierung verwendet werden, dem Modell selbst und deren Interaktionen. Häufig müssen Lösungen erarbeitet werden, die spezifisch für die jeweilige Kombination von Modell und Datensätzen sind. In diesem Forschungsvorhaben wurde ein konsistenter Ansatz für alle aquatischen Datensätze ermittelt. Dies garantiert die Vergleichbarkeit der Ergebnisse zwischen den Arten. Allerdings könnten die Ergebnisse möglicherweise verbessert werden, wenn für jeden Datensatz bzw. für jede Art individuelle Lösungen gefunden werden. Da dieses Forschungsvorhaben aber auf die Modellentwicklung abzielt, wurde der Ansatz der guten Vergleichbarkeit gewählt. Für die terrestrische Ausbreitung lag nur ein Datensatz zur Ausbreitung vor, so dass hier ein individueller Ansatz erarbeitet wurde, der sich an denen der aquatischen Arten orientiert.

Die verwendeten Datensätze weisen Unsicherheiten zum tatsächlichen Vorkommen der Arten auf. So liegen teilweise nur grobe räumliche Auflösungen vor, die Probenahmen erfolgen unregelmäßig und

nicht flächendeckend. In Abhängigkeit der Datensätze sind die Vorkommen und vor allem die Abwesenheiten einer Art daher nicht eindeutig zu bestimmen. Um diese Unsicherheit in die Kalibrierung mit aufzunehmen und zu berücksichtigen, wurde den Vorkommen einer Art in den Daten eine Wahrscheinlichkeit von 0,9 und einer Abwesenheit eine Wahrscheinlichkeit von 0,1 zugewiesen. Damit werden Unsicherheiten in den Daten bei der Abschätzung der Parameter berücksichtigt.

Um einen Vergleich der Modellvorhersagen und der Daten zu ermöglichen, mussten die kontinuierlichen Wahrscheinlichkeiten in binäre (Vorkommen/Abwesenheit) Daten übersetzt werden. Hierzu wurde ein Grenzwert von 0.5 angesetzt, der in der logistischen Regression üblich ist. Wahrscheinlichkeiten unter 0.5 wurden als Abwesenheit und über 0.5 als Vorkommen klassifiziert. In der logistischen Regression gibt es verschiedene Möglichkeiten die Güte eines Modells zu beschreiben. Das am weitesten verbreitete Maß stellt die sogenannte „area under the curve“ (AUC) dar, die häufig als „receiver operating characteristic“ (ROC) – Kurve angezeigt wird. Diese konnte aber im vorliegenden Fall nicht für die aquatischen Datensätze angewandt werden, da nicht in allen Fällen die benötigten Kombinationen von beobachteten und simulierten Vorkommen und Abwesenheiten vorhanden waren. Daher wurde ein anderes Standardmaß, das „percentage correctly classified“ (PCC), für alle Datensätze verwendet (Mouton, De Baets, & Goethals, 2010). Zur Vergleichbarkeit wurde PCC auch für die terrestrische Art verwendet. Basis für die Berechnung von PCC liefert eine Tabelle der korrekt und falsch vorhergesagten Vorkommen und Abwesenheiten. PCC berechnet sich aus dem Anteil der korrekt vorgesagten Abwesenheiten und Vorkommen und liefert damit ein leicht zu interpretierendes Maß zur Güte der Simulation. Ein PCC = 0.5 gilt als zufällige Verteilung und ein PCC = 1 als perfekter Fit (Mouton et al., 2010).

Die Möglichkeit der Validierung der Modellergebnisse hängt sehr stark von den vorhandenen Datensätzen ab. Idealerweise liegen unabhängige Datensätze vor, die für Kalibrierung und Validierung verwendet werden können. Da dies für dieses Projekt nicht der Fall war, mussten die Datensätze zeitlich getrennt werden für Kalibrierung und Validierung. Dies bedeutet, dass das Modell mit den ersten Zeitschritten der beobachteten Ausbreitung kalibriert wurde und mit den letzten Zeitschritten validiert (s. Abschnitt 2.2.3). So wurde beispielsweise das Modell mit der Ausbreitung von *Senecio inaequidens* über den Zeitraum bis 1997 kalibriert und anschließend mit Vorhersagen von 1998 bis 2004 validiert.

Aufgrund fehlender Daten zu den Ausbreitungswegen über Paletten und Container konnten die Modellergebnisse nicht mit Daten zur tatsächlichen Ausbreitung verglichen werden. Daher konnten die Ausbreitungswege über Paletten- und Containertransport nicht kalibriert und validiert werden. Die Parameter für diese Ausbreitungswege wurden manuell auf $v = 10.000$ Paletten und $u = 1.000$ Container gesetzt. Das Modell liefert damit Ergebnisse zur Ausbreitung über Paletten- und Containertransport, aber die Aussagekraft der Modellergebnisse konnte aufgrund fehlender Datensätze zur Ausbreitung von Neobiota über Container und Paletten nicht bewertet werden. Daher dienen die Ergebnisse für diese Ausbreitungswege nur der Anschauung, was potenziell möglich ist, sofern geeignete Daten zur Kalibrierung vorliegen.

2.4 Modellannahmen

Wie bei allen Computermodellen üblich, wurde auch im vorliegenden Fall eine Reihe von Annahmen getroffen, die die zugrunde liegenden Prozesse vereinfachen und somit für ein Modell verfügbar machen. Einige dieser Annahmen können Modellergebnisse beeinflussen und sollten daher bei der Interpretation der Ergebnisse berücksichtigt werden. Im Folgenden werden die wichtigsten Annahmen vorgestellt und diskutiert, die für den vorliegenden Ansatz getroffen wurden.

- **Annahme 1:** Die Ausbreitung von Neobiota findet nur entlang von Verkehrsträgern und durch den Verkehr statt. **Diskussion:** Neobiota werden über eine Vielzahl von Wegen ausgebreitet und können sich auch unabhängig von Verkehrsträgern selbstständig ausbreiten. Daher ist diese Annahme sehr vereinfachend, aber aufgrund der Fragestellung des Projektgebers vorgegeben. Die berechnete Wahrscheinlichkeit einer Einwanderung ist nur in Bezug auf Verkehrsträger zu sehen und kann daher nicht direkt in die Realität übertragen werden. Dies macht einen Vergleich von Modellergebnissen mit Artaufnahmen schwierig und könnte zu einer Überschätzung der Ausbreitungswahrscheinlichkeit durch die Verkehrsträger führen, da alternative Ausbreitungswege, die ebenfalls das Vorkommen einer Art erklären könnten, nicht ins Modell aufgenommen werden konnten. Die Bedeutung des Verkehrs zur Erklärung der Vorkommen einer Art ist daher vermutlich überhöht. Eine Überschätzung könnte vor allem im terrestrischen Bereich auftreten, da hier alternative Ausbreitungswege bestehen, die nicht berücksichtigt werden konnten. Im Gegensatz dazu sind die Ausbreitungen im aquatischen Bereich sowohl über Schifffahrt als auch über natürliche Ausbreitungswege sehr gut abgedeckt, so dass es für den aquatischen Bereich weder zu einer systematischen Über- noch Unterschätzung kommen sollte. Die Art und das Ausmaß einer möglichen Abweichung kann allerdings nur ermittelt werden, wenn die entsprechenden alternativen Ausbreitungswege ebenfalls implementiert werden, so dass die relativen Bedeutungen aller Wege verglichen werden können.
- **Annahme 2:** Lokale Populationsdynamiken und biotische Interaktionen beeinflussen die Ausbreitung nicht. **Diskussion:** Sobald sich Individuen einer Art in einer neuen Region ansiedeln können, findet häufig eine lokale Ausbreitung über Nachkommen statt, die hauptsächlich über Populationsdynamiken und das lokale Habitat bestimmt werden. Diese Aspekte und die jeweilige Bedeutung der wirkenden Faktoren sind jedoch sehr spezifisch für jede einzelne Art und können daher nicht mit einem so allgemeinen Ansatz wie dem vorliegenden berücksichtigt werden. Auf großer räumlicher Skala, wie z.B. auf der Ebene des Bundes, sollten Populationsdynamiken die Muster der Ausbreitung nur wenig beeinflussen. Für kleinräumige Vorhersagen ist aber zu empfehlen, Populationsdynamiken in die Modellierung mit aufzunehmen, wodurch allerdings die Generalität des Ansatzes verloren geht.
- **Annahme 3:** Biotische Interaktionen wie Konkurrenz, Fraß oder Pathogene beeinflussen die Ausbreitung nicht. **Diskussion:** Biotische Interaktionen können wichtige Faktoren vor allem für die Etablierung von neuen Populationen darstellen. Selbst wenn Individuen potenziell gute Habitate erreichen, können biotische Interaktionen die Etablierung verhindern (Mitchell et al., 2006). Diese Mechanismen sind stark Kontext abhängig, was bedeutet, dass sie von einer hohen Variabilität zwischen Arten, zwischen Individuen derselben Art und deren Aktionspartnern sowie den örtlichen Gegebenheiten abhängen. Damit ist es nahezu unmöglich, biotische Interaktionen in einen generellen Modellierungsansatz wie in dieser Studie mit aufzunehmen. Da biotische Interaktionen vor allem kleinräumig eine Rolle spielen, sollten die großräumigen Ausbreitungsmuster wenig beeinflusst sein. Dennoch stellt dies eine Limitierung des Modellansatzes dar, wobei hervorzuheben ist, dass es aktuell keinen Forschungsansatz gibt, der diese Limitierung zufriedenstellend aufheben könnte.
- **Annahme 4:** Die Daten zum Vorkommen der Arten spiegeln akkurat das Vorkommen der jeweiligen Art zu diesem Zeitpunkt dar. **Diskussion:** Eine feldökologische Aufnahme ist in der Regel nicht vollständig, da nur in den seltensten Fällen alle Organismen aufgenommen werden können, so dass immer ein gewisser Grad an Unsicherheit bestehen bleibt. Ein Teil dieser Unsicherheit kann mit dem hier angewandten Ansatz der Bayesschen Modellkalibrierung aufgefangen werden, da Vorkommen und Abwesenheiten einer Art mit Wahrscheinlichkeiten größer null bzw. kleiner eins mit einer definierten Streuung festgelegt wurden. Damit geht die Unsicherheit einer fehlerhaften Darstellung der Vorkommen der Arten mit in die Analyse ein. Dennoch bleibt ein großes Maß an Unsicherheit in den Daten, das nur schwer oder gar nicht abzuschätzen ist. So können z.B. Angaben auf Basis der Rasterzellen nur mit Unsicherheiten in punktgenaue Koordinaten übertragen werden.
- **Annahme 5:** Die Ausbreitung einer Art findet in einem Simulationsschritt (ca. ein Monat) nur von einem Knoten zum nächsten statt. **Diskussion:** In der Regel breiten sich Arten nicht kontinuierlich in Zeit und Raum aus, da jede Art bestimmte Zeitpunkte im Lebens- und Jahreszyklus hat, die für

eine Ausbreitung besser geeignet ist. Pflanzen werden meistens über Samen ausgebreitet, die in der Regel nur in bestimmten Jahreszeiten gebildet werden. Weiterhin können sich Arten auch über den nächsten Knoten hinaus innerhalb eines Zeitschritts ausbreiten. Daher ist die Annahme der kontinuierlichen Ausbreitung von Knoten zu Knoten eine Vereinfachung, die durch die vorhandene Datengrundlage gegeben ist. Eine Berechnung zu allen Knoten anstatt der benachbarten Knoten ist aufgrund der Vielzahl aller möglichen Verbindungen nicht möglich. Simulationsläufe haben allerdings gezeigt, dass die simulierte Ausbreitung mit dem erstellten Modell sehr schnell vonstattengehen kann, so dass eine Limitierung der Ausbreitung nur aufgrund des gewählten Ansatzes unwahrscheinlich ist.

3 Ergebnisse und Diskussion

Die hier präsentierten Ergebnisse resultieren aus einer Vielzahl an Zwischenschritten mit Bewertungen der Zwischenergebnisse, vielen Testläufen und kontinuierlichen Anpassungen und Verbesserungen der Methodik, die sich über die gesamte Projektlaufzeit hinzog. Das Modell und die Simulationsergebnisse wurden umfassend getestet und stellen nach Einschätzung der Auftragnehmer die beste und geeignetste Form zur Beantwortung der Fragestellungen des Forschungsvorhabens dar. Die folgenden Ergebnisse zeigen das Endprodukt dieses Prozesses.

3.1 *Senecio inaequidens*

Die visuelle Inspektion der Vorkommensdaten von *Senecio inaequidens* verdeutlicht, dass es einen Schwerpunkt des Vorkommens zu Beginn der Simulation im Jahr 1989 im Bereich von Nordrhein-Westfalen gab (Abbildung 7). In späteren Jahren weitete sich das Vorkommen in alle Richtungen aus und es kamen einige isolierte Vorkommen hinzu. Da das Modell auf Verkehr basiert, kann es nur eine gleichmäßige Ausbreitung von Arten simulieren. Das spontane Auftauchen einer Art an einem isolierten Ort ist nur schwer mit Verkehrsströmen zu erklären, sondern basiert vermutlich auf der Eintragung in Einzelfällen. Möglicherweise spielen hier weitere Faktoren der Ausbreitung eine Rolle, die nicht berücksichtigt werden konnten. Eine weitere Erklärung könnte in der Unvollständigkeit der Daten liegen, da möglicherweise nicht alle Vorkommen der Art erfasst sind.

Die Kalibrierung führte im Verlauf von über 33.000 Iterationsschritten zur Konvergenz, was mit dem Gelman-Rubin Konvergenzmaß nahe eins (1,014) angezeigt wurde (Tabelle 6). Dies bedeutet, dass ein eindeutiger Satz an Parameterwerten gefunden werden konnte, der die beobachteten Dynamiken von *Senecio inaequidens* am besten beschreibt. Der Verlauf der Kalibrierung kann für die einzelnen Parameter in Abbildung S1 und Abbildung S2 gesehen werden (siehe Abschnitt 8.1.1). Für alle Parameter zeigen die Verläufe eine weitgehend konstante Streuung um einen zentralen Wert. Dies bedeutet, dass der Algorithmus in der Lage war, einen großen Bereich des Parameterraums zu testen. Gleichzeitig konnte aber keine deutliche Verbesserung der durch die Literatur vorgegebenen Werte erzielt werden. Dies zeigt aber auch, dass die vorgegebenen Parameterwerte bereits zufriedenstellende Ergebnisse liefern. Die resultierende Verteilung der Parameterwerte und damit die Unsicherheit der Parameterschätzung ist ebenfalls in den Abbildungen S1 und S2 dargestellt. Die Parameterwerte der „dispersal kernels“, die sich aus der Kalibrierung ergaben, sind in Tabelle 6 aufgeführt und stellen die derzeitige Einstellung des Modells für terrestrische Ausbreitung und Etablierung dar.

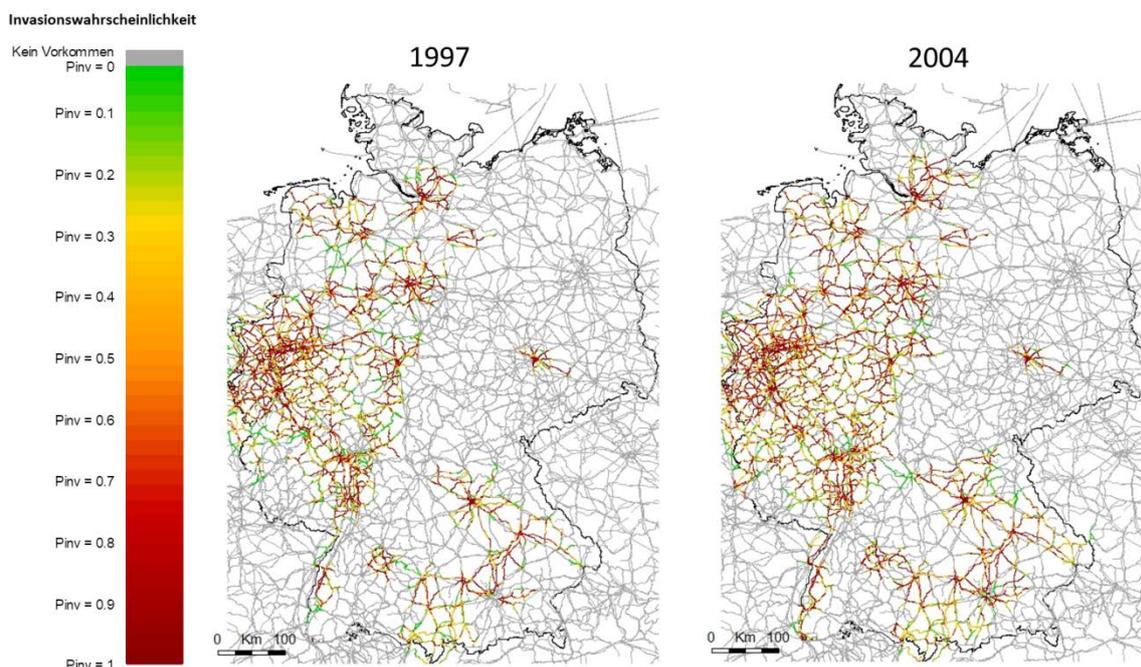


Abbildung 15: Ergebnisse der Kalibrierung (A) und Validierung (B) für die simulierte Ausbreitung von *Senecio inaequidens* über die Verkehrsträger Straße und Schiene. Die Startpunkte der Simulation stellten die Vorkommen der Art im Jahr 1989 dar (Abbildung 7). Die Karten zeigen die simulierte Ausbreitung der Art im Jahr 1997, die zur Kalibrierung verwendet wurde, (A) und 2004 für die Validierung der Modellergebnisse (B) unter Verwendung der kalibrierten Parameter (Tabelle 6).

Insgesamt ist die Güte der Modellvorhersagen der Kalibrierung für *Senecio inaequidens* mit $PCC = 0.666$ als moderat zu bewerten (Abbildung 14). Ein PCC Wert von 0.5 beschreibt eine zufällige Verteilung der Beobachtungen. Dies ist aber nicht unbedingt mit einer geringen Qualität der Modellvorhersagen gleichzusetzen, da die tatsächlichen Vorkommen der Art nicht ausschließlich durch Verkehr bedingt sind. Daher wird es immer eine Diskrepanz zwischen Modellvorhersagen und tatsächlichen Vorkommen geben, selbst wenn das Modell die Ausbreitung über Verkehr perfekt beschreiben könnte. Es konnte im Rahmen dieser Studie nicht abschließend geklärt werden, welche Faktoren (z.B. die Qualität der zugrundeliegenden Daten oder das Fehlen weiterer Ausbreitungswege) oder Modellannahmen in erster Linie zu einer Verbesserung des PCC Wertes führen würden. Um dies zu klären, wären weitere Forschungsaktivitäten notwendig.

Anders als die Kalibrierung erfolgte die Validierung der Modellergebnisse für das Vorkommen der Art im gesamten Bundesgebiet, um den vollen Umfang an vorhandenen Daten auszuschöpfen. Die Modellgüte fiel dabei nur leicht auf einen Wert von $PCC = 0,661$. Die Abnahme der Modellgüte für die Validierung war zu erwarten, da dieser Test außerhalb des für die Kalibrierung genutzten Bereichs durchgeführt wurde. Die geringe Abnahme des PCC Wertes zeigt jedoch, dass die moderate Güte der Vorhersage auch für Vorhersagen von zukünftigen Vorkommen gilt. Eine Verbesserung sowohl der Kalibrierung als auch der Validierung ist zu erwarten, wenn zusätzliche Ausbreitungswege unabhängig vom Verkehrsträger berücksichtigt werden können.

TABELLE 6: ERGEBNISSE DER MODELLKALIBRIERUNG FÜR *SENECIO INAEQUIDENS*

PARAMETER	VALUE	5%	MEDIAN	95%
A_{KS} (ANHEFTUNG STRAÙE)	$8,31 \cdot 10^{-7}$	$2,72 \cdot 10^{-7}$	$1,14 \cdot 10^{-6}$	$2,53 \cdot 10^{-6}$
A_{KR} (ANHEFTUNG SCHIENE)	$9,40 \cdot 10^{-6}$	$1,31 \cdot 10^{-6}$	$9,99 \cdot 10^{-6}$	$2,21 \cdot 10^{-5}$
D (ANHEFTUNG)	0,663	0,489	0,651	0,810
C (ANHEFTUNG)	-0,037	-0,062	-0,035	-0,01
G (ANHEFTUNG)	0,357	0,173	0,332	0,489
A_{FS} (FAHRTWIND STRAÙE)	0,006	0,001	0,007	0,020
A_{FR} (FAHRTWIND SCHIENE)	0,002	0,001	0,009	0,022
K (FAHRTWIND)	2,288	2,106	2,258	2,341
H (FAHRTWIND)	0,742	0,521	0,683	0,786
A (NATÜRLICHE AUSBREITUNG)	0,481	0,036	0,329	0,939
B (NATÜRLICHE AUSBREITUNG)	0,481	0,340	0,503	0,655
B_1 (ETABLIERUNG)	0,941	0,770	0,904	0,980
ITERATIONSSCHRITTE	33.334			
RECHENZEIT (STD.)	306			
GELMAN-RUBIN	1,014			
PCC (KALIBRIERUNG)	0,666			
PCC (VALIDIERUNG)	0,611			

3.2 *Corbicula fluminea*

Bei der Kalibrierung für die aquatische Art *Corbicula fluminea* stellte sich heraus, dass bei Berücksichtigung der Schiffsintensität keine Konvergenz des Algorithmus erreicht werden konnte (Gelman-Rubin >> 1). Die Schiffsintensität geht nur bei dem Ausbreitungsweg Aufwuchs in die Berechnung mit ein, da Ballastwasser über die Größe der Häfen abgeschätzt wird (s. Abschnitt 2.3.2.2). Der Grund für die fehlende Konvergenz besteht in der Diskrepanz zwischen Intensitäten des Schiffverkehrs und der Ausbreitungsbewegung von *Corbicula fluminea*. Beim Vergleich der Schiffsintensitäten (Abbildung 4) mit der Ausbreitung von *Corbicula fluminea* (Abbildung 9) fällt auf, dass die Schiffsintensitäten vor allem im unteren Rhein sehr hoch sind, während in den ersten Jahren *Corbicula fluminea* in die Oberläufe von Mosel, Main und Rhein einwanderte. Diese Ausbreitung führt von den hohen Schiffsintensitäten im unteren Rhein weg, was darauf hinweist, dass die Ausbreitung dieser Art innerhalb Deutschlands nicht allein mit hohem Schiffverkehr erklärt werden kann. Das hohe Aufkommen im Rhein legt die Vermutung nahe, dass die Art mit Schiffen eingebracht wurde (Schöll, 2007), während die weitere Ausbreitung innerhalb Deutschlands nicht nur durch Schiffverkehr bestimmt wurde. Diese Diskrepanz zwischen Daten und Modellvorhersagen hat die Konvergenz des Algorithmus bei der Kalibrierung erschwert. Es wurde daher entschieden, dass die Schiffsintensität nicht in die Kalibrierung mit eingeht, wobei der Ausbreitungsweg Aufwuchs berücksichtigt wurde. Anders formuliert bedeutet dies, dass die Formel (1) nicht auf Aufwuchs angewandt wurde. Dies bedarf weiterer Untersuchung und wird über das Projekt hinausgehend weiterhin analysiert und angepasst. Es wird vorgeschlagen, dies in späteren Modellversionen anzupassen.

Mit der nun getätigten Auswahl an Ausbreitungswegen konnte die Kalibrierung für *Corbicula fluminea* nach etwa 33.000 Iterationsschritten beendet werden, da der Optimierungsalgorithmus dann eine deutliche Konvergenz aufwies, was durch einen Gelman-Rubin Wert nahe eins ausgedrückt wird (Tabelle 7). Dies spiegelt sich auch in den Verläufen der Kalibrierung für die einzelnen Parameter wider (Abbildung S3): Obwohl in den ersten Iterationsschritten ein großer Bereich des Parameterraums beprobt wurde, wurde in den meisten Fällen schon nach vergleichsweise wenigen Schritten eine optimale Kombination an Parameterwerten gefunden. Dies drückt sich auch in der häufig engen Dichteverteilung der Parameterwerte aus. Einzig die Parameterwerte von q zeigt eine vergleichsweise hohe Streuung. Dieser Parameter bestimmt die Ausbreitungsfunktion für die natürliche Ausbreitung. Die einzelnen Parameterwerte mit der jeweiligen Variation sind in Tabelle 7 dargestellt.

Mit einem PCC = 0.84 ergab sich ein vergleichsweise guter Wert der Modellvorhersage, was darauf schließen lässt, dass das Modell die beobachtete Ausbreitung von *Corbicula fluminea* über den Zeitraum der Kalibrierung (1990 – 1996) gut wiedergeben kann. Im Vergleich dazu fällt die Güte der Modellvorhersage bei der Validierung der Modellergebnisse auf PCC = 0.55 stark ab, was auf eine geringe Kapazität der Vorhersage neuer Vorkommen außerhalb des kalibrierten Bereichs schließen lässt. Dies kann so interpretiert werden, dass beobachtete Ausbreitungen der Art in der Retrospektive gut wiedergegeben werden können, wobei Aussagen über zukünftige Vorkommen nur mit großen Unsicherheiten möglich sind. Dies stellt ein weitverbreitetes und generelles Problem vieler Modelle und wissenschaftlicher Studien über Neobiota dar und konnte auch für eine Reihe anderer und häufig verwendeter statistischer Modelle gezeigt werden (Capinha, Essl, Seebens, Pereira, & Kühn, 2018). Die deutliche Abnahme der Modellgüte kann damit erklärt werden, dass ab 1997 neue Vorkommen mit einem großen Abstand zu den bisherigen Beobachtungen gefunden wurden, die nur schwer mit einer kontinuierlichen Ausbreitung erklärt werden können (Abbildung 9). Diese anscheinend „sprunghafte“ Ausbreitung der Art könnte mit Lücken im Datensatz erklärt werden, da es keine Beprobungen an den Kanälen gab, die die verschiedenen Flusssysteme miteinander verbinden. Möglicherweise spielen aber auch Einzelereignisse der Eintragung eine Rolle, die nur sehr schwer mit Modellen erfasst werden können. Hinzu kommt, dass bei der Verwendung von Daten aus Zeitreihen eine technische Schwierigkeit in der Berücksichtigung bereits simulierter Vorkommen für den Bereich der Kalibrierung liegt. Im vorliegenden Fall wurden Vorkommen

vor 1997 nicht für die Validierung verwendet. Hierdurch sollte vermieden werden, dass Daten für Kalibrierung und Validierung verwendet werden. Hierdurch kann ebenfalls die starke Abnahme der Güte der Modellvorhersage erklärt werden.

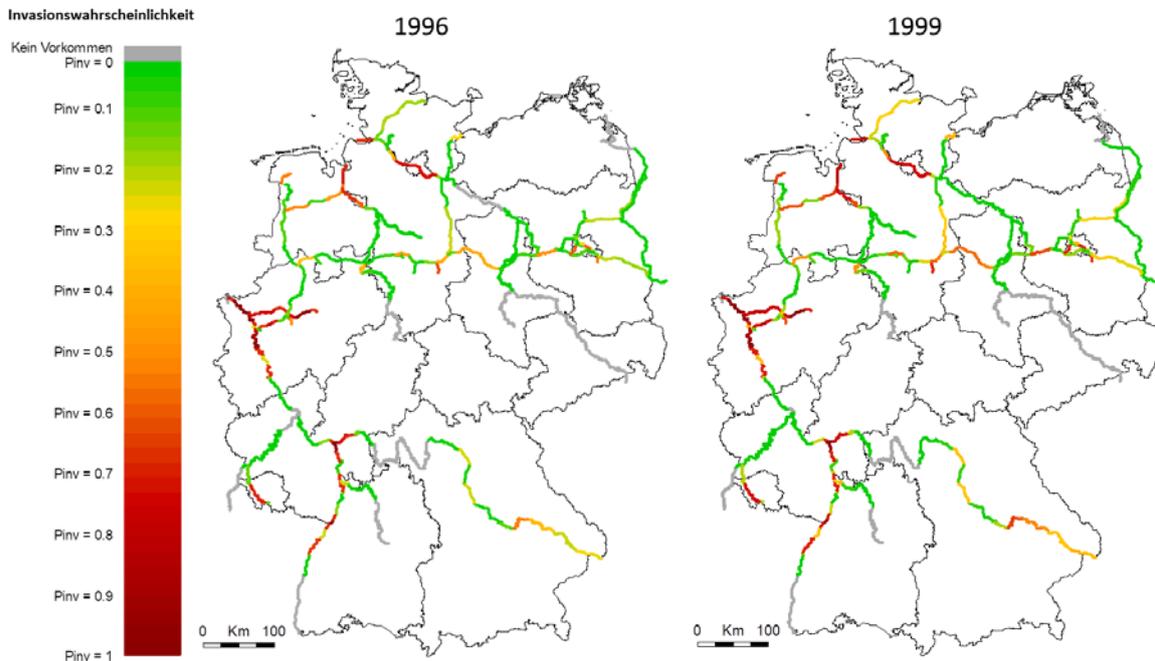


Abbildung 16: Ergebnisse der Kalibrierung (A) und Validierung (B) für die simulierte Ausbreitung von *Corbicula fluminea*. Die Startpunkte der Simulation stellten die Vorkommen der Art im Jahr 1990 dar (Abbildung 9). Die Karten zeigen die simulierte Ausbreitung der Art im Jahr 1996, die zur Kalibrierung verwendet wurde, (A) und 1999 zur Validierung des Modells (B) unter Verwendung der kalibrierten Parameter (Tabelle 7).

Insgesamt zeigen die Ergebnisse für *Corbicula fluminea*, dass die Ausbreitung dieser Art in den Wasserstraßen gut mit dem Modell zu erfassen sind (s. Abbildung 9 und Abbildung 15). Allein der Eintrag der ersten Individuen, die den Startpunkt der Ausbreitung vorgeben, lässt sich nur mit hoher Unsicherheit vorhersagen. Es ist anzunehmen, dass eine deutliche Verbesserung erreicht werden kann, wenn detailliertere Informationen zu Schiffsbewegungen, Herkunftsgebieten und länderübergreifenden Ausbreitungen berücksichtigt werden können.

TABELLE 7: ERGEBNISSE DER MODELLKALIBRIERUNG FÜR *CORBICULA FLUMINEA*

PARAMETER	VALUE	5%	MEDIAN	95%
Q (NATÜRLICHE AUSBREITUNG)	9,441	6,190	7,941	10,967
R (NATÜRLICHE AUSBREITUNG)	3,471	3,365	3,457	3,496
N (BALLASTWASSER)	73.977	73.034	76.450	82.822
L (AUFWUCHS)	2,748	2,639	2,788	2,936
M (AUFWUCHS)	2,302	2,304	2,396	2,646
B_2 (ETABLIERUNG)	0,700	0,700	0,705	0,729
ITERATIONSSCHRITTE	33.334			
RECHENZEIT [STUNDEN]	127			
GELMAN-RUBIN	1,055			
PCC (KALIBRIERUNG)	0,838			
PCC (VALIDIERUNG)	0,557			

3.3 *Dikerogammarus villosus*

Die Kalibrierung für *Dikerogammarus villosus* wurde in gleicher Weise durchgeführt wie für *Corbicula fluminea* (siehe Kap. 3.2). Ähnlich wie bei *Corbicula fluminea* ergab die Kalibrierung für *Dikerogammarus villosus* eine ähnlich gute Konvergenz (Gelman-Rubin Wert = 1.016), was wiederum auf einen eindeutigen Satz an erklärenden Parameterwerten hindeutet. Auch hier zeigen die Verläufe der einzelnen Optimierungen eine schnelle Konvergenz für nahezu alle Parameter (Abbildung S4). Die Verteilung der Parameterwerte war hierbei schmaler als bei *Corbicula fluminea*, was auf eine noch geringere Unsicherheit der kalibrierten Parameter hindeutet. Auch die bei der vorherigen Art noch etwas unsichere Abschätzung der Parameters q zeigte eine enge Bandbreite und damit eine geringe Unsicherheit. Dies drückt sich auch in der leicht verbesserten Güte der Modellvorhersage von PCC = 0,84 aus (Tabelle 8). Im Vergleich zur Kalibrierung von *Corbicula fluminea* ergeben sich nur geringe Unterschiede, was wiederum ein Hinweis auf die Robustheit der Parameterwerte ist.

Ähnlich wie für *Corbicula fluminea* fällt die Validierung mit einem PCC = 0,42 deutlich schlechter aus als der Wert für die Kalibrierung. Die Interpretation folgt derselben Argumentation wie für *Corbicula fluminea* bereits dargelegt (Abschnitt 3.2). Auch hier können die Ausbreitungen grundsätzlich gut vorhergesagt werden, wobei Aussagen zur zukünftigen Verteilung nur mit großer Unsicherheit möglich sind. Hinzu kommt, dass im Zeitraum der Validierung (1997-1999) eine neue Population in der Elbe auftauchte, die deutlich abseits der bisherigen Population im Rhein-Donau Gebiet lag (Abbildung 8). Diese sprunghafte Ausbreitung weicht deutlich von der zuvor beobachteten Ausbreitung ab, die zur Kalibrierung verwendet wurde, so dass das Modell nicht in der Lage ist, solche unerwarteten Ausbreitungssprünge zu erfassen.

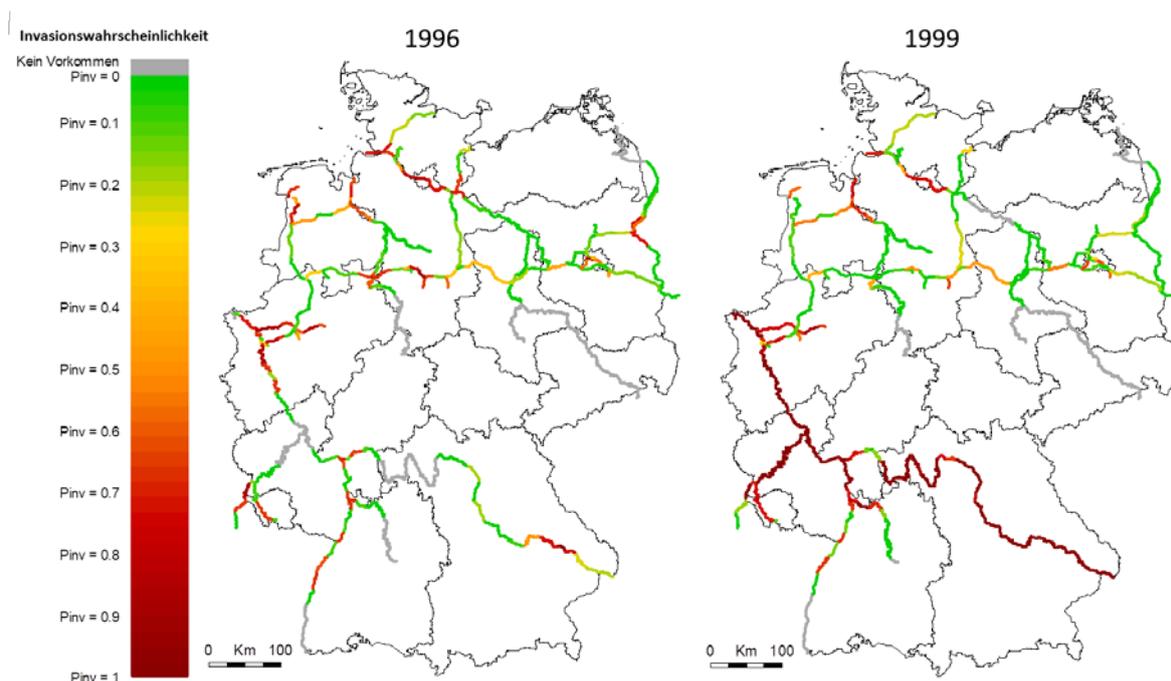


Abbildung 17: Ergebnisse der Kalibrierung (A) und Validierung (B) für die simulierte Ausbreitung von *Dikerogammarus villosus*. Die Startpunkte der Simulation stellten die Vorkommen der Art im Jahr 1990 dar (Abbildung 8). Die Karten zeigen die simulierte Ausbreitung der Art im Jahr 1996, die zur Kalibrierung verwendet wurde, (A) und 1999 zur Validierung des Modells (B) unter Verwendung der kalibrierten Parameter (Tabelle 8).

TABELLE 8: ERGEBNISSE DER MODELLKALIBRIERUNG FÜR *DIKEROGAMMARUS VILLOSUS*

PARAMETER	VALUE	5%	MEDIAN	95%
Q (NATÜRLICHE AUSBREITUNG)	6,830	6,028	6,428	7,725
R (NATÜRLICHE AUSBREITUNG)	3,499	3,445	3,486	3,499
N (BALLASTWASSER)	98.191	98.170	99.350	99.922
L (AUFWUCHS)	2,727	2,724	2,813	2,926
M (AUFWUCHS)	2,306	2,302	2,329	2,433
B₂ (ETABLIERUNG)	0,700	0,700	0,702	0,706
ITERATIONSSCHRITTE	33.334			
RECHENZEIT (STD.)	42			
GELMAN-RUBIN	1,016			
PCC (KALIBRIERUNG)	0,846			
PCC (VALIDIERUNG)	0,427			

3.4 Spezifische Fragestellungen

Im Folgenden werden die Fragestellungen des Auftraggebers aus der Leistungsbeschreibung aufgegriffen und im Einzelnen erörtert und diskutiert. In Absprache mit dem Auftraggeber wurde die Detailtiefe, mit der die einzelnen Fragestellungen bearbeitet wurden, unter Berücksichtigung der vorhandenen und verfügbaren Daten angepasst. Bei der Darstellung und Diskussion der Ergebnisse werden auch die Schwierigkeiten in der Beantwortung der jeweiligen Fragestellung dargelegt.

3.4.1 Einfallstore der Invasion

Als Einfallstore werden geographische Punkte charakterisiert, an denen transportierte Arten das erste Mal potenziell die Möglichkeit haben, innerhalb Deutschlands in die Natur entlassen zu werden. Für den Straßen- und Schienenverkehr stellen das die Grenzübergänge dar. Für den Schiffsverkehr spielen die Häfen eine größere Rolle, da in Häfen potenziell Arten über Ballastwasser oder Aufwuchs eingebracht werden können.

Abbildung 17 zeigt die Intensitäten des Straßen-, Schienen- und Schiffsverkehrs an den Grenzübergängen und den jeweiligen Häfen. Die Intensitäten sind in der Einheit Anzahl Fahrzeuge (Pkw + Lkw, Züge) für den Landverkehr und Umschlag in Gütervolumen der Binnenschifffahrt für die Häfen dargestellt. Die wichtigsten Einfallstore für den Landverkehr sind in Süd- und Westdeutschland zu finden, während die wichtigsten Einfallstore für den Schiffsverkehr in Duisburg und Hamburg liegen. Mögliche Einfallstore über Gütertransporte ist am Beispiel der Container in Abbildung 5 dargestellt. Container und Paletten werden in sehr großer Zahl nach und in Deutschland weiter transportiert und gelangen an sehr unterschiedlichen Stellen an ihre Empfänger. Aufgrund dieser nahezu flächendeckenden Verteilung ist es daher kaum möglich, generelle Aussage über mögliche Einfallstore über Güterströme zu machen.

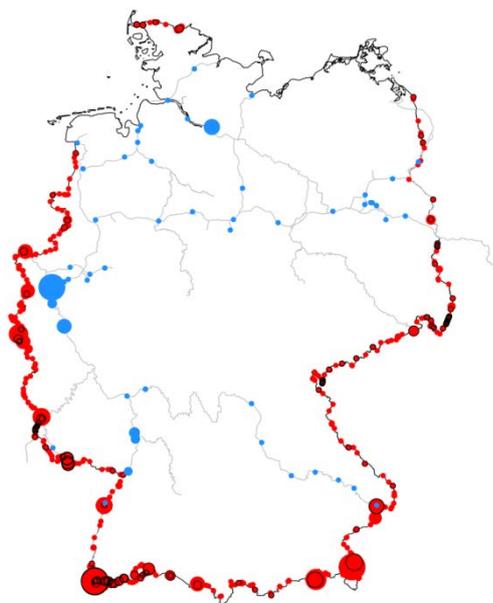


Abbildung 18: Einfallstore der Invasion für Straßen- (rot), Schienen- (schwarze Kreise) und Schiffsverkehr (blau).

3.4.2 Vektoren

Die Vektoren der Invasion stellen die verschiedenen Verkehrsträger Straße, Schiene und Wasserstraße dar. Die Vektoren konnten im vorliegenden Projekt sehr detailliert und umfassend berücksichtigt werden und stellen eine zentrale Komponente des Modells und der Projektergebnisse dar. Neben einer flächendeckenden Betrachtung des Straßen-, Schienen- und Schiffsverkehrs wurden auch Gütertransporte über Container und palettierte Güter berücksichtigt. Über die Vorgaben des Projektgebers hinausgehend, wurde auch die natürliche Ausbreitung von Arten in das Modell implementiert. Die natürliche Ausbreitung kann vor allem für die sekundäre Ausbreitung nach einer Eintragung eine große Rolle spielen. Eine detaillierte Aufstellung der berücksichtigten Vektoren ist in Tabelle 4 gegeben. Ein Vergleich der verschiedenen Verkehrsträger befindet sich im Abschnitt 3.5, in dem die Ausbreitungswege detailliert analysiert werden.

3.4.3 Mögliche Umschlagpunkte

Umschlagpunkte stellen die geographischen Orte dar, an denen Arten potenziell die Möglichkeit haben, von dem einen auf einen anderen Verkehrsträger zu wechseln. In diesem Forschungsvorhaben wurden die terrestrischen Transportnetzwerke Schiene und Straße kombiniert, sodass die Ausbreitung über Anheftung, Fahrtwind und natürliche Ausbreitung entlang beider Verkehrsträger parallel stattfinden kann. Dies stellt weltweit ein Novum im Bereich der Invasionsökologie dar. Aufgrund der Vielzahl an Schnittpunkten von Straßen und Schienen ergeben sich so insgesamt 327.806 Umschlagpunkte. Die Wahrscheinlichkeit eines Wechsels der Verkehrsträger kann nicht ohne weiteres abgeschätzt werden, da eine hohe Verkehrsdichte nicht zwangsläufig eine hohe Wahrscheinlichkeit für einen Wechsel darstellt. Im Gegenteil sind z.B. Autobahnen mit hohen Verkehrsdichten in der Regel räumlich deutlich von kreuzenden Schienen getrennt, sodass ein Wechsel von der Autobahn auf die Schiene in dieser Konstellation eher unwahrscheinlich ist. Demgegenüber sollte ein Verkehrsträgerwechsel zwischen Straße und Schiene an Bahnübergängen der Kreis- und Gemeindestraße mit deutlich höherer Wahrscheinlichkeit stattfinden, obwohl diese in der Regel ein deutlich niedrigeres Verkehrsaufkommen aufweisen als Autobahnen. Zur Berechnung der Wahrscheinlichkeit eines Verkehrsträgerwechsels wären eingehende empirische Untersuchungen und Analyse dieser Prozesse und eine entsprechende Datengrundlage notwendig. Beides ist bislang nicht vorhanden, weshalb im Rahmen dieses Projektes keine weitergehende Berücksichtigung erfolgen konnte.

Beim Gütertransport werden in diesem Forschungsvorhaben Arten berücksichtigt, die sich innerhalb von Containern oder Paletten befinden. Die Ausbreitung über diese Transportwege wird unabhängig vom Verkehrsträger simuliert, da die Arten keine direkte Berührung mit dem Transportmittel (Schiff oder Lkw) haben. Umschlagpunkte im Sinne der Transportlogistik z.B. von Zügen auf Lkw werden dadurch berücksichtigt. Sie spielen aber für diese Ausbreitungswege der Arten keine Rolle.

3.4.4 Ausbreitungsrouten

Die Ausbreitungsrouten stellen die geographischen Wege dar, entlang derer sich Neobiota ausbreiten können. Bedingt durch den Fokus des Forschungsvorhabens sind dies Straßen, Schienen und Wasserstraßen in Deutschland. Um eine detailliertere Bewertung der Ausbreitungsrouten zu ermöglichen, bedarf es einer Quantifizierung der Ausbreitung für jeden Streckenabschnitt aller drei Verkehrsträger. Da die Ausbreitung über Vektoren ein zentraler Bestandteil der Projektergebnisse ist, eignet sich das Modell ideal, um Ausbreitungsrouten zu quantifizieren.

In diesem Forschungsvorhaben werden die Ergebnisse der Simulationen der Ausbreitungswahrscheinlichkeiten genutzt, um die Ausbreitungsrouten zu quantifizieren. Hierfür wurde eine sogenannte „shortest paths“ (kürzester Pfad) Analyse durchgeführt. Diese Analyse ermittelt die „günstigsten“ Routen zwischen zwei beliebigen Punkten in den jeweiligen Transportnetzwerken auf Basis der ermittelten Ausbreitungswahrscheinlichkeiten. Ein kürzester Pfad stellt damit die wahrscheinlichste Route der Ausbreitung dar. Die Gesamtheit aller kürzesten Pfade ergibt ein flächendeckendes Bild für Deutschland.

Eine umfassende Darstellung und Analyse der Ausbreitungsrouten wird in Abschnitt 3.5 vorgestellt.

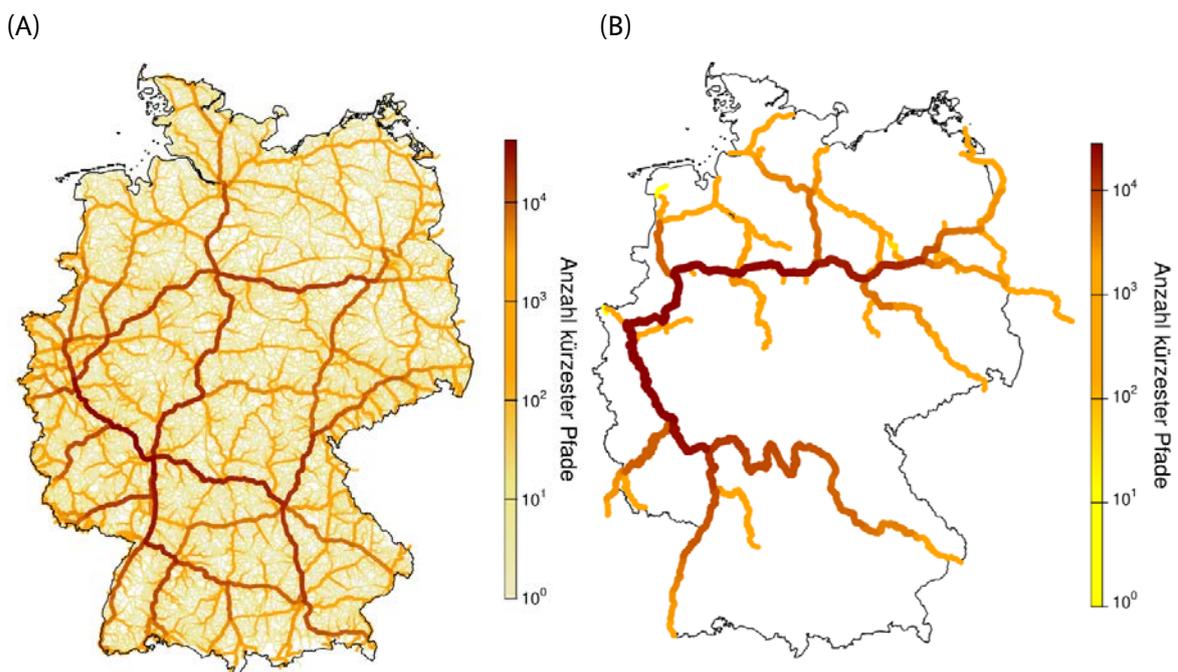


Abbildung 19: Routen der Ausbreitung entlang (A) terrestrischer und (B) aquatischer Verkehrsnetzwerke in Deutschland. Die Anzahl kürzester Pfade kann als Wahrscheinlichkeit einer Ausbreitung interpretiert werden. Siehe Abschnitt 3.5 für eine detailliertere Analyse.

3.4.5 Geschwindigkeiten der Ausbreitungen

Die Geschwindigkeit der Ausbreitung einzelner Arten hängt von einer Vielzahl von Faktoren ab. So spielen Ausbreitungsstrategien wie z.B. Windausbreitung bei Pflanzen oder Mobilität bei Tieren, aber auch die Größe der Arten, die Umweltbedingungen und Populationsdynamiken entscheidende Rollen. Da sich diese Faktoren stark von Art zu Art unterscheiden, sind generelle Aussagen über die Geschwindigkeit der Ausbreitung nicht möglich. Selbst für die gleiche Art sind solche Aussagen schwierig, da es von Habitaten, saisonalen und interannuellen Schwankungen abhängt.

Das Modell, welches in diesem Forschungsvorhaben entwickelt wurde, ermöglicht die Berechnung von Aufenthaltswahrscheinlichkeiten, die nicht direkt in Geschwindigkeiten übersetzt werden können. Die „Ausbreitungsfrent“, wie sie in Abbildung 13 zu sehen ist, spiegelt nicht zwangsläufig die tatsächliche Ausbreitung wider, da sie nur die Wahrscheinlichkeit eines Vorkommens beschreibt, welches sehr gering sein kann. Man würde die tatsächliche Front der Ausbreitung nicht mit dem ersten Auftauchen dieser niedrigen Wahrscheinlichkeiten gleichsetzen. Hinzu kommt, dass die „Ausbreitungsfrent“ aus der Abbil-

dung hauptsächlich durch die Struktur der Transportnetzwerke vorgegeben ist: In einem dichten Netzwerk sind mehr Schritte zur Berechnung der Wahrscheinlichkeiten nötig als in einem dünnen Netzwerk, obwohl die berechneten Wahrscheinlichkeiten ähnlich sein können.

Um eine Ausbreitungsgeschwindigkeit aus den Modellergebnissen zu ermitteln, müsste zuerst ein Grenzwert der Aufenthaltswahrscheinlichkeit (z.B. $P(Inv) = 0.5$) festgelegt werden. Mit Überschreiten dieses Grenzwerts gilt die Art als vor Ort etabliert. Nun kann berechnet werden, wann die Invasionswahrscheinlichkeiten den Grenzwert überschreiten, woraus sich die Geschwindigkeit berechnen ließe. Allerdings ist dieses Vorgehen von der Definition der Geschwindigkeit und des Modells abhängig. Es stellt daher ein wenig intuitives und nicht robustes Maß dar. Es wird daher empfohlen, die Geschwindigkeiten der Ausbreitung aus Beobachtungen des Aufenthalts der Arten aus den Daten direkt zu extrahieren, ohne den Umweg über die Modellsimulationen zu nehmen. Dies wird deutlich realistischere Ergebnisse liefern.

3.4.6 „Loads“ und Schwerpunkte der Invasion

„Loads“ der Invasion beschreibt die Quantifizierung der Ausbreitung und des Vorkommens von Neobiota. Die Quantifizierung der Ausbreitung und des Vorkommens erfolgt in diesem Forschungsvorhaben über die Berechnung der Wahrscheinlichkeiten der Ausbreitung und Etablierung. Die Wahrscheinlichkeit der Ausbreitung wird für jedes Segment in den Transportnetzwerken ausgerechnet. Es beschreibt die Wahrscheinlichkeit der Ausbreitung $P(I)$ von einem Knoten des Netzwerks zum nächsten. Aufgrund der hohen räumlichen Auflösung der Transportnetzwerke wird eine hohe Zahl von Ausbreitungswahrscheinlichkeiten errechnet (>600.000), womit eine detaillierte Simulation der Ausbreitung möglich ist. Allerdings eignen sich die einzelnen Wahrscheinlichkeiten $P(I)_{ij}$ nur bedingt für eine einfache Darstellung und Interpretation der Ausbreitung. Dies hat zwei Gründe: 1. Die hohe Zahl an Wahrscheinlichkeit erschwert die Darstellung und Interpretation und 2. die Ausbreitungswahrscheinlichkeiten beschreiben nur die Ausbreitung entlang eines Segments, welches häufig unter 1 km lang ist. Um die Aussagekraft der Simulationen zu erhöhen und generellere Aussagen zu ermöglichen, wurde der Ansatz der „shortest paths“ gewählt, bei dem weitreichende Verbindungen zwischen Knoten analysiert werden (Abbildung 18). Dies wird in einem eigenem Abschnitt (Abschnitt 3.5) im Detail vorgestellt und besprochen.

Die Wahrscheinlichkeit einer Etablierung ergibt sich aus den jeweiligen lokalen Umweltbedingungen und den Präferenzen der zu betrachtenden Arten. Dies wird für die Beispielarten in den Abbildung 12 und Abbildung 6 dargestellt. Die Gesamtwahrscheinlichkeit einer Invasion – und damit der Invasionschwerpunkt – ist für *Senecio inaequidens* in Abbildung 13 zu sehen.

3.5 Analyse der Modellergebnisse

Die bisher präsentierten Ergebnisse zeigen Simulationen zur Ausbreitung einzelner Arten. Einzelne Simulationen sind allerdings kaum zu generalisieren, da sie von der betrachteten Art, der Simulationszeit und den Anfangsbedingungen abhängen. Die einzelnen Simulationen dienen hauptsächlich der Anschauung und der Kalibrierung der Parameter. Mit Hilfe der kalibrierten Parameter können die Wahrscheinlichkeiten der Ausbreitung und Etablierung für jeden Link und Knoten berechnet werden. In einem weiteren Schritt können diese individuellen Wahrscheinlichkeiten genutzt werden, um generelle Aussagen zur Ausbreitung zu machen. Dies soll im Folgenden vorgestellt werden.

Die berechneten Wahrscheinlichkeiten bilden die Grundlage für weitere Analysen zu Ausbreitungen und Verbreitungsschwerpunkten. Hierzu werden auf Ansätze der Forschungsdisziplin der komplexen Netzwerke zurückgegriffen (Albert & Barabasi, 2002). In diesem Forschungsbereich wurden sehr effiziente Algorithmen zur Berechnung von kürzesten Pfaden („shortest paths“) entwickelt (s. auch Glossar). Diese erlauben die schnelle Suche nach der kürzesten Verbindung zwischen zwei beliebigen Punkten, wobei die Distanz zwischen zwei Punkten nicht die geographische Distanz sein muss. In diesem Projekt bilden die Ausbreitungswahrscheinlichkeiten das Maß der „Distanz“, die dann besonders „kurz“ ist, wenn die Gesamtheit der Ausbreitungswahrscheinlichkeiten aller Segmente zwischen zwei Punkten besonders hoch ist (Abbildung 19). Dies stellt die Route mit der höchsten Ausbreitungswahrscheinlichkeit zwischen zwei Punkten dar.

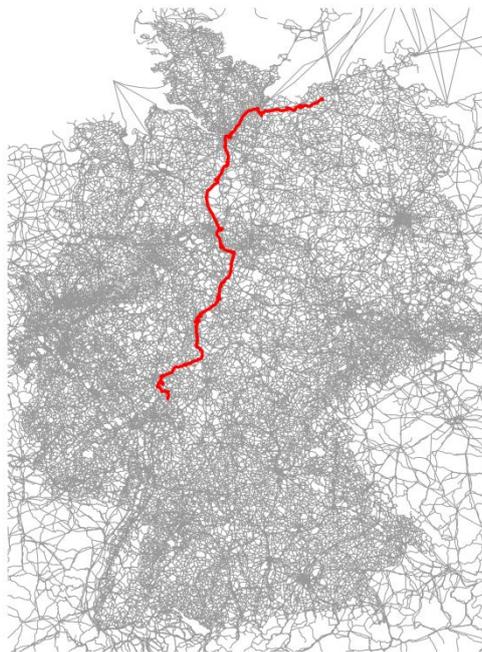


Abbildung 20: Beispielhafte Route der höchsten Wahrscheinlichkeit der Ausbreitung zwischen zwei beliebigen Punkten.

Die Berechnung der kürzesten Pfade zwischen allen Punkten des Netzwerks beschreibt dann, wie häufig ein bestimmter Link Teil eines kürzesten Pfads ist. Links, die sehr häufig „gewählt“ wurden, können als die wahrscheinlichsten Ausbreitungsrouten interpretiert werden. Die Berechnung der kürzesten Pfade kann für jedes Netzwerk, Teile des Netzwerks oder bestimmte Knoten ermittelt werden. Im Folgenden werden Überblicke über die Netzwerke und ein paar Beispiele mit engerem Fokus dargestellt.

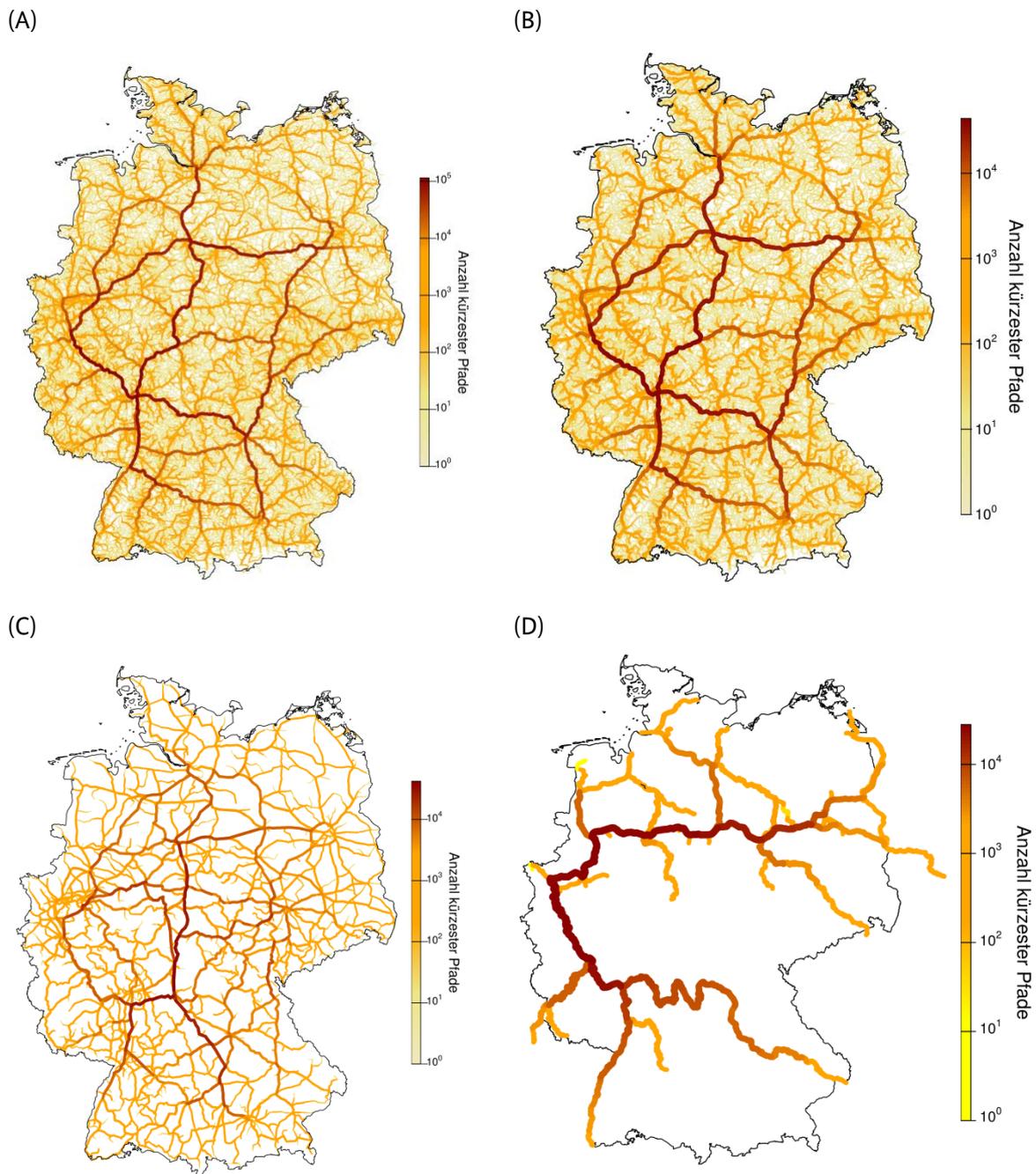


Abbildung 21: Ausbreitungsrouten innerhalb Deutschlands entlang (A) des Straßen- und Schienennetzes, (B) entlang des Straßennetzes, (C) entlang des Schienennetzes und (D) entlang des Wasserstraßennetzes.

Bei Betrachtung der Ausbreitungsrouten fällt auf, dass die Ausbreitungsrouten des Straßenverkehrs gleichmäßig über Deutschland verteilt liegen, während sich für den Schienenverkehr eine zentrale vertikale Achse als Hauptroute ergibt (Abbildung 20). Die Routen mit der höchsten Wahrscheinlichkeit der Ausbreitung stellen in der Regel die Hauptverkehrsachsen in Deutschland dar, während auf Links mit einem geringen Verkehrsaufkommen nur eine geringe Bedeutung für die Ausbreitung von Neobiota zukommt (Abbildung 21). Umgekehrt sind aber die Routen mit der höchsten Verkehrsdichte nicht zwangsläufig die Hauptrouten der Ausbreitung. So gibt es viele Links mit einer Verkehrsdichte von 80.000 Fahr-

zeugen pro Tag, die die Hauptrouten der Ausbreitung darstellen neben Links mit gleichem Verkehrsaufkommen, die nur eine unbedeutende Rolle für die Ausbreitung spielen (Abbildung 21). Allein auf Verkehrsdichten basierend lassen sich die Hauptrouten der Ausbreitung nicht erschließen.

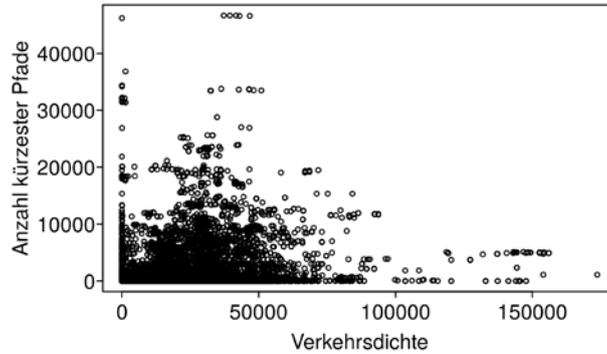


Abbildung 22: Zusammenhang zwischen Verkehrsdichte und Anzahl kürzester Pfade aller Links.

Beim Vergleich der verschiedenen terrestrische Netzwerke und Kategorien der Links kann man erkennen, dass Landesstraßen in ihrer Gesamtheit eine größere Rolle für die Ausbreitung von Neobiota spielen im Vergleich zu anderen Straßentypen (Abbildung 22A). Bei Betrachtung der mittleren Anzahl an kürzesten Pfaden ergibt sich ein abweichendes Bild, bei dem Autobahnen die mit Abstand größte Rolle spielen (Abbildung 22B). Die Unterschiede sind vor allem mit der unterschiedlichen Gesamtstreckenlänge der Straßentypen in Deutschland zu erklären: Pro Streckenabschnitt stellen Autobahn im Mittel die häufigsten Ausbreitungsrouten, wobei aufgrund der vergleichsweise niedrigeren Streckenlänge das Netzwerk der Autobahnen insgesamt gegenüber z.B. Landes- und Bundesstraßen nur eine untergeordnete Rolle spielt. Der Schienenverkehr nimmt bei beiden Betrachtungsweisen für die Ausbreitung einen mittleren Platz ein.

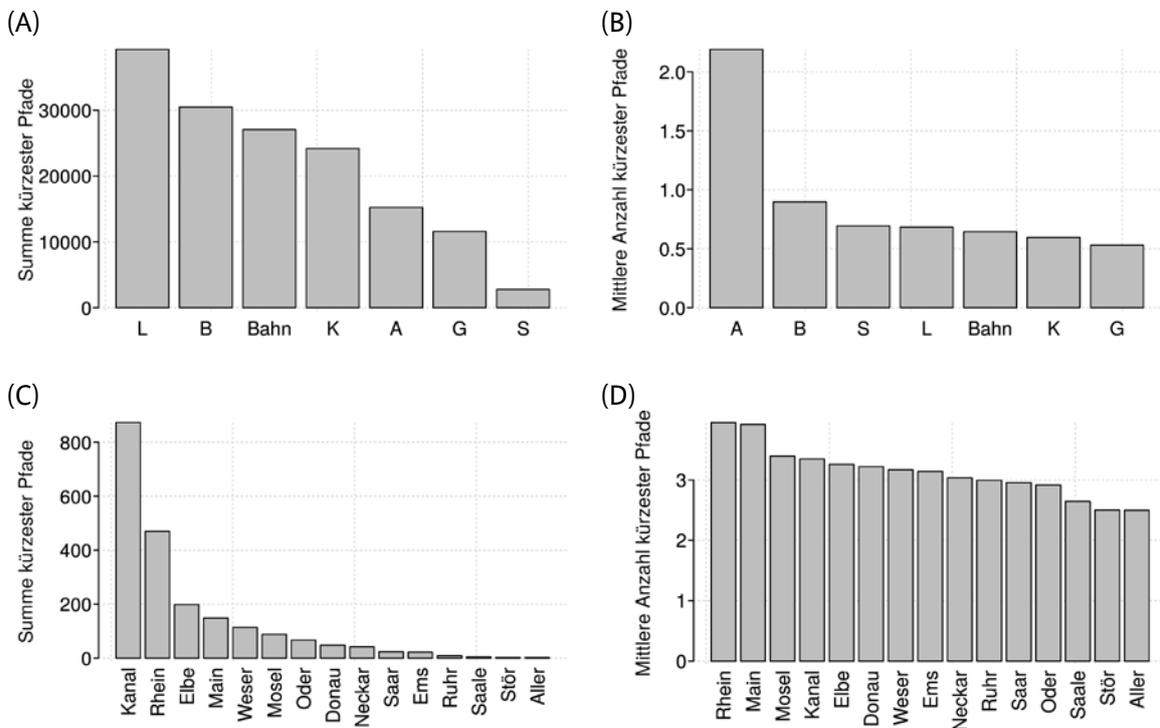


Abbildung 23: Vergleich der Hauptrouten der Ausbreitung zwischen Verkehrsträgern, Straßentypen und Wasserstraßen. Die Hauptrouten sind für das kombinierte terrestrische Netzwerk (A, B) und das aquatische Netzwerk (C, D) als Mittelwerte pro Segment (A, C) und als Summe über alle Segmente (B, D) dargestellt. A: Autobahn, B: Bundesstraße, L: Landesstraße, Bahn: Schienenverkehr, K: Kreisstraße, G: Gemeindestraße, S: Staatsstraßen.

Bei den Wasserstraßen fällt auf, dass im Mittel pro Link die Ausbreitung entlang aller Links ähnlich häufig stattfindet (Abbildung 22D). Dies liegt vor allem an dem vergleichsweise kleinen Netzwerk der Wasserstraßen in Deutschland, da es kaum Alternativrouten bei der Verbindung von zwei Knoten gibt. Bei Betrachtung der Summe über alle Links pro Gewässer zeigen die Kanäle die höchste „Nutzung“ zur Ausbreitung, was hauptsächlich mit der Gesamtlänge aller Kanäle im Vergleich zu einzelnen Flüssen erklärt werden kann (Abbildung 22C). Der Rhein stellt aufgrund seiner Länge und hohen Schiffsintensität mit Abstand den Fluss mit der häufigsten Ausbreitungsrouten dar, während Elbe, Main und Weser folgen.

Die Ausbreitungsrouten lassen sich auch für einzelne Startgebiete und Netzwerktypen ermitteln. In Abbildung 23 wurde exemplarisch Hamburg als Startpunkt gewählt, um die Ausbreitung entlang der verschiedenen Netzwerktypen darzustellen. Beim Vergleich der Ausbreitungsrouten über beide terrestrischen Verkehrsnetzwerke Schiene und Straße (Abbildung 23A) mit der Ausbreitung über das Straßennetzwerk (Abbildung 23B) fällt deren hohe Ähnlichkeit auf. Dies zeigt, dass die Ausbreitung für terrestrische Arten durch den Straßenverkehr dominiert wird, was die Gegenüberstellung der relativen Bedeutung der Verkehrsträger widerspiegelt (Abbildung 22). Im Folgenden werden weitere Beispiele für Ausbreitungsrouten ausgehend von Großstädten (Abbildung 24) und Grenzübergängen (Abbildung 25) dargestellt.

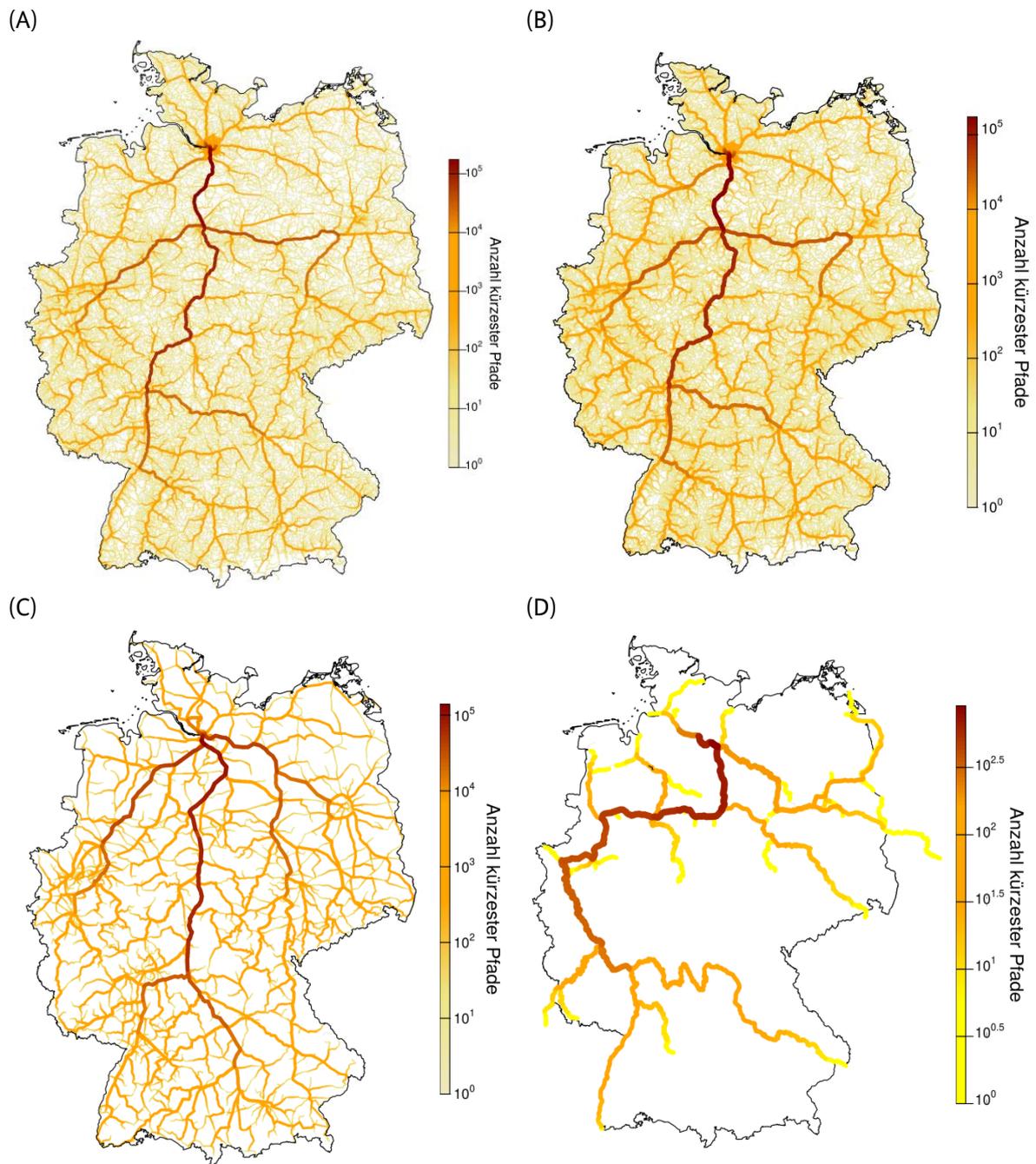


Abbildung 24: Ausbreitungsrouten mit Startpunkt Hamburg über (A) alle terrestrischen Netzwerke, (B) das Straßennetzwerk, (C) das Schienennetzwerk und (D) das Wasserstraßennetzwerk.

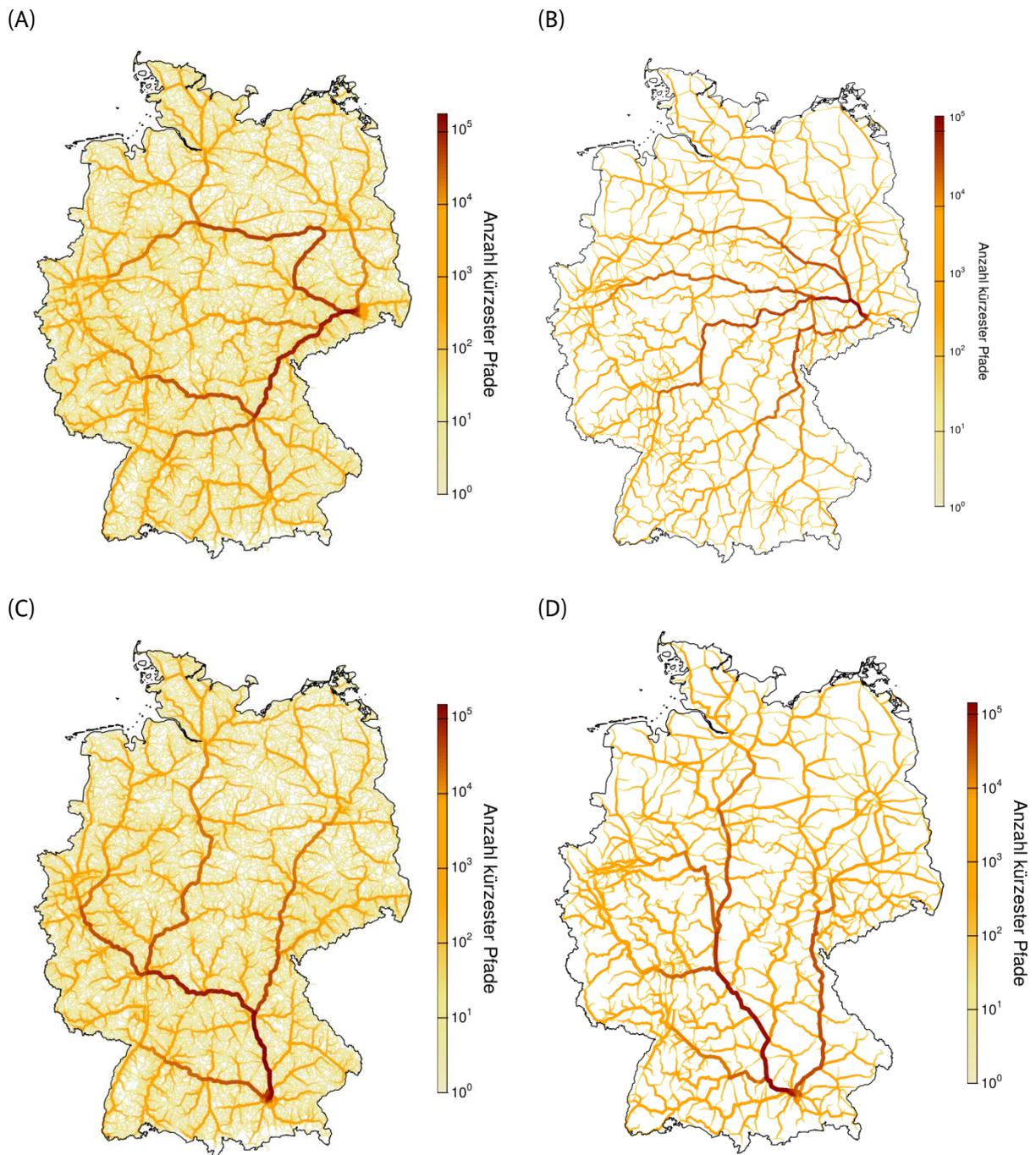


Abbildung 25: Beispiele für Ausbreitungsrouten mit Startpunkt Dresden (A,B) und München (C,D) über alle terrestrischen Netzwerke (A,C) und das Schienennetzwerk (B,D).

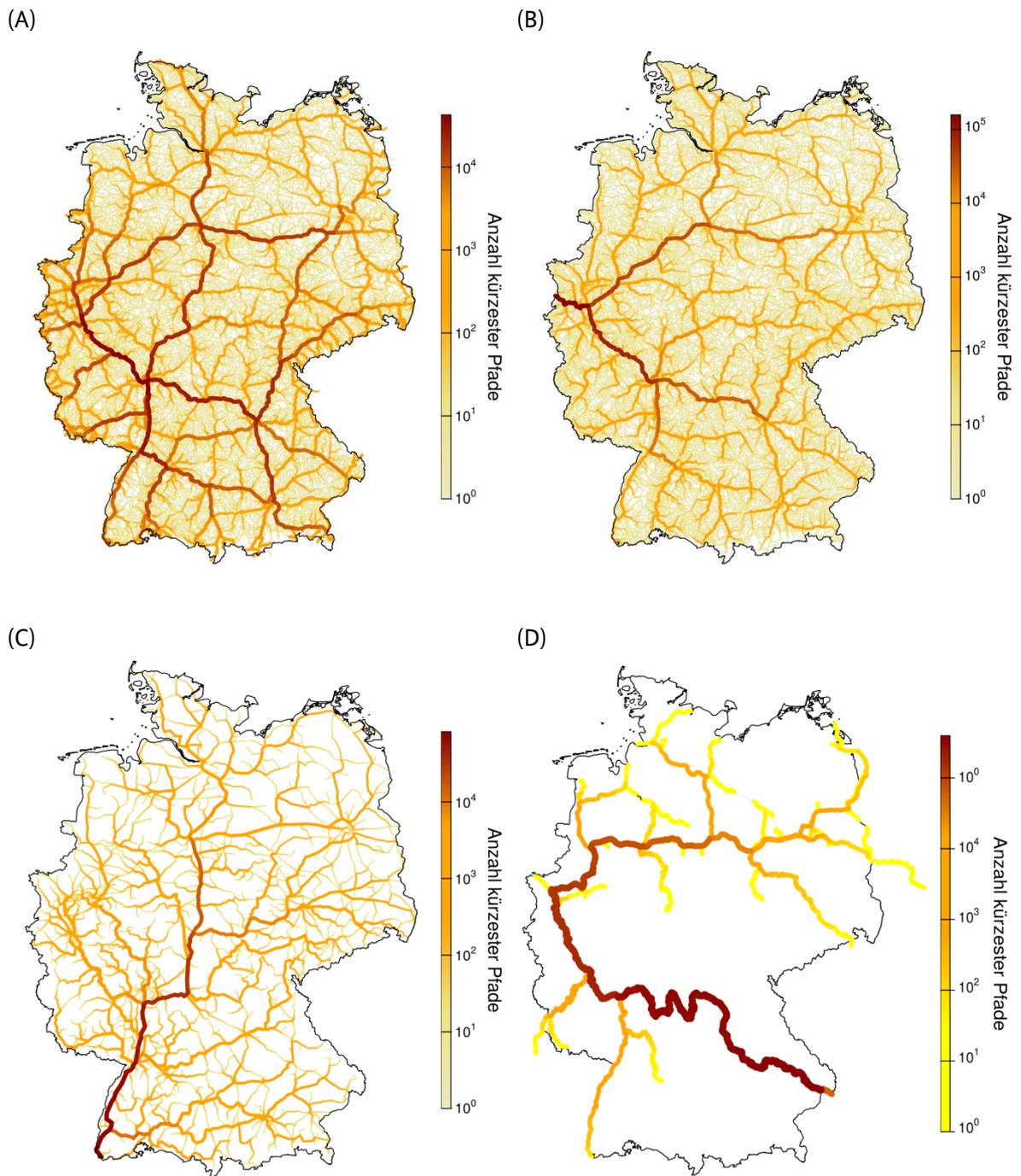


Abbildung 26: Beispiele für Ausbreitungsrouten mit Startpunkten an Grenzübergängen. Start der Routen bei allen Grenzübergängen des Straßen- und Schienenverkehrs (A), ausgewählter Grenzübergänge der Straße (B), Schiene (C) und Wasserstraße (Donau) (D).

4 Weitere Schritte

4.1 Wissens- und Datenlücken

Die Evaluierung der Datengrundlage (AP1) hat deutlich gezeigt, dass es ein großes Gefälle der Datenverfügbarkeit zwischen Daten der Verkehrsflüsse, der Umweltparameter und biologischer Datensätze gibt. Datensätze zu Verkehr und Infrastruktur liegen vor allem im Landverkehr in sehr hoher zeitlicher und räumlicher Auflösung vor. Die Daten besitzen eine hohe Genauigkeit und bilden daher eine sehr gute Basis, um die Verkehrsströme in Deutschland zu simulieren. Informationen zur Binnenschifffahrt liegen dagegen mit vergleichsweise geringem Informationsgehalt nur als Intensitäten pro Segment vor. Zur genaueren Ermittlung der Wahrscheinlichkeit der Ausbreitung von Arten über Schiffe werden Informationen zur Herkunft und Route einzelner Schiffe benötigt, die nicht verfügbar sind. Weiterhin wären Informationen über Ballastwasseraustausch, Liegezeiten, Güterumschlag und Schiffscharakteristika wie Größe, Ladung und Dockzeiten hilfreich.

Für Umwelt- und Habitatparameter gibt es deutlich unterschiedliche Grundlagen. Manche Datensätze wie der Bedeckungsgrad oder Umweltparameter in terrestrischen Habitaten liegen als Satellitenprodukte in hoher Auflösung vor und können gut genutzt werden. Für aquatische Habitate werden Informationen nur für Einzugsgebiete einzelner Flüsse bereitgestellt. Umfang und Genauigkeit dieser Daten variieren sehr stark zwischen den online Portalen der einzelnen Flussgebietsgemeinschaften und können daher schlecht verglichen und kombiniert werden. Dies hatte für das Forschungsprojekt zur Folge, dass die teilweise in hoher Güte vorliegenden Daten für einzelne Gewässer nicht verwendet werden konnten, da diese für andere Bereiche nicht in ähnlicher Güte vorlagen. **Hier wäre eine zentrale oder vereinheitlichte Datendarstellung hilfreich und wünschenswert.**

Das mit Abstand größte Problem für die Umsetzung dieses Forschungsvorhabens lag in der Verfügbarkeit an Daten zum Vorkommen einzelner Arten im terrestrischen Bereich. Öffentliche Portale wie Flora-Web.de oder Florkart (deutschlandflora.de) erwiesen sich als unvollständig und/oder nicht zugänglich, da Daten nicht bezogen werden konnten. Anfragen bei zuständigen Personen erwiesen sich als wenig hilfreich. So scheint es, dass derartige Daten, wenn überhaupt nur über langwierige Prozesse bereitgestellt werden können und zudem der starken Nachbesserung bedürfen, was aus im Rahmen eines Forschungsvorhabens wie diesem nicht zu bewältigen ist. Der ursprünglich für dieses Forschungsvorhaben geplante Ansatz, die Verbreitungskarten möglichst vieler Arten zu verwenden, um Artenzahlen anstatt Vorkommen einzelner Arten vorherzusagen, musste aufgrund der geringen Datenverfügbarkeit verworfen werden. **Hier besteht dringender Handlungsbedarf, um Veränderung von Vorkommen wie bei der Einwanderung neuer Arten schnell erfassen und bewerten zu können.**

Zur Erreichung der Projektziele mussten daher Datensätze einzelner Arten und deren Anbieter ermittelt, die Anbieter kontaktiert und überzeugt sowie die Datensätze aufbereitet werden. Leider stellten sich auch diese vergleichsweise guten Datensätze in zwei von fünf Fällen als unzureichend für die Erreichung der Projektziele heraus (s. Abschnitt 2.2.3), so dass sie für diese Analyse verworfen werden mussten. Im Gegensatz zum terrestrischen Bereich lagen im aquatischen Bereich umfangreiche Datensätze vor, die sehr gut für das Projektvorhaben genutzt werden konnten. Umfang und Genauigkeit der aquatischen Arten lag deutlich über dem für terrestrische Arten.

Als ein weiteres großes Defizit stellten sich das fehlende Verständnis und die fehlende Datengrundlage der Ausbreitung von Arten über Verkehr und Verkehrsträger dar. Für viele Ausbreitungswege ist das Verständnis der zu Grunde liegenden Mechanismen nur unzureichend. So ist zwar generell bekannt, dass Arten über z.B. Anheftung an Fahrzeuge verbreitet werden können, aber es gibt so gut wie keine

Studien über die Anzahl und Reichweite angehefteter Organismen. Die vorhandenen Studien stellen Einzelstudien dar, deren Ergebnisse häufig nur schwer zu verallgemeinern sind. Für viele Ausbreitungswege wie z.B. Schienenverkehr oder die Schifffahrt liegen bislang kaum Studien vor und das Verständnis des Transports von Neobiota durch diese Verkehrsträger ist nur gering. Für die Schifffahrt wurden bislang hauptsächlich Untersuchungen im marinen Bereich durchgeführt, aber auch hier ist kaum bekannt, wie viele Arten ein Schiff über Aufwuchs von einem Ort an den nächsten transportiert. Für Deutschland sind keine Studien bekannt, die die Ausbreitung von Arten über Schiffe innerhalb Deutschlands untersuchen. Das gleiche gilt für den Schienenverkehr. Informationen zur Einwanderung über Container und Paletten oder andere Gütertransportwege liegen weltweit kaum vor. Für Deutschland sind keine Studien bekannt, die z.B. das Potential der Einwanderung über Palettentransport untersuchen. **Auch hier besteht Nachholbedarf, da Container und Palettentransport wichtige Wege der Einbringung von z.B. Forstschädlingen darstellen.**

Einige weitere mit dem Verkehr assoziierte Ausbreitungswege konnten nicht im Modell berücksichtigt werden, da weder für Deutschland als Bezugsraum noch weltweit ausreichende Informationen zur Ausbreitung der Arten vorliegen trotz guter Daten zum Verkehr. So ist z.B. anekdotisch bekannt, dass Insekten wie die Tigermücke in Fahrerkabinen von Autos oder Lkws oder in Personenzügen mitreisen können, aber es gibt keine Daten, die eine Quantifizierung dieses Ausbreitungswegs zulassen. **Hier besteht ebenfalls Handlungsbedarf, Daten zur Anzahl der transportierten Organismen für die nicht berücksichtigten Ausbreitungswege zu erheben und die zugrundeliegenden Mechanismen näher zu untersuchen.**

4.2 Ausblick

Als Ausblick soll hier ein Überblick über mögliche weitere Schritte gegeben werden, die zu einer Verbesserung des Modells aber auch zu einer verbesserten Grundlage von Wissen und Daten führen können. Es gibt eine große Vielfalt an Möglichkeiten, wobei im Folgenden nur eine Auswahl an sinnvoll erscheinenden und realistisch umsetzbaren Schritten vorgestellt werden soll:

- **Weitere Anwendungen des Modells:** Die Qualität und Anwendbarkeit eines Modells stellt sich häufig erst bei eingehender Verwendung heraus. Daher wäre ein sinnvoller nächster Schritt, das Modell in anderen Bereichen zu testen. So könnten weitere Datensätze zu Vorkommen hinzugenommen, komplexere Habitat- und Umweltvariationen, andere Verkehrsnetzwerke oder regionale Studien vorgenommen werden. Dies würde helfen, die Möglichkeiten des Modells zur Vorhersage weiter auszuleuchten und zu bewerten.
- **Verbesserung der Datengrundlage für Vorkommen:** Wie im vorherigen Abschnitt beschrieben, stellt das Fehlen von adäquaten Verbreitungskarten mit zeitlichen Veränderungen das größte Defizit für die Erstellung von Vorhersagen zur Ausbreitung und Etablierung von Neobiota dar. Die Qualität von Modellvorhersagen hängt direkt von der Qualität und Quantität der zugrundeliegenden Daten zur Artverbreitung ab. Vor allem im terrestrischen Bereich wurde ein eklatanter Mangel an geeigneten Datensätzen festgestellt, wodurch eine gründliche Validierung der Ergebnisse nur eingeschränkt möglich war. Eine Verbesserung der Datengrundlage würde daher zu einer deutlich verbesserten Einschätzung der Modellvorhersagen führen. Es könnten weitere Arten getestet werden, um die Robustheit der Ergebnisse zu überprüfen und ggf. Verbesserungen vorzunehmen. In Abhängigkeit der jeweiligen Arten könnten auch bestimmte Ausbreitungswege im Detail untersucht und ergänzt werden wie z.B. die Ausbreitung entlang von Uferböschungen. Informationen zu holzbohrenden Käfern in Paletten würden helfen, den Eintrag über Paletten besser abschätzen zu können. In vielen Fällen ist die Verbesserung der Datengrundlage durchaus mit einem vergleichsweise geringen finanziellen Aufwand möglich, da viele

Daten bereits vorhanden sind. Diese müssten bei Vorliegen „nur“ identifiziert, zusammengetragen und harmonisiert werden.

- **Erweiterung des Modells um weitere Ausbreitungswege:** Eine verbesserte Datengrundlage der Vorkommen würde auch die Möglichkeit eröffnen, weitere bislang noch nicht berücksichtigte Ausbreitungswege in das Modell zu implementieren. So könnten mit einem Vorhandensein eines Datensatzes zur Ausbreitung von Ufervegetation die Ausbreitung entlang von Uferböschung implementiert werden. Weiterhin könnten Ausbreitungen in Fahrerksabinen und Personenzüge von z.B. Mücken implementiert werden. Hierfür wären Funddaten von Mücken oder anderen Arten und gute Kenntnisse der Mechanismen und Habitate nötig. Ebenso ist die Ausbreitung über den Individualverkehr zu Wasser nicht berücksichtigt, obwohl Sportboote wesentlich zu Ausbreitung von Neobiota beitragen können (Ulman et al., 2019). Hierzu wären Informationen zum Individualverkehr notwendig. Die modulare Struktur des in diesem Projekt erstellten Modells CASPIAN ermöglicht eine einfache Implementierung weiterer Ausbreitungswege.
- **Implementierung von Populationsdynamiken:** Grundsätzlich kann das Modell auch für eine kleinräumige Vorhersage der Vorkommen verwendet werden. Allerdings spielen hier weitere Faktoren eine entscheidende Rolle, die unabhängig von den Verkehrsträgern sind. Zumindest für einzelne Arten könnten solche Faktoren, wie Populationsdynamiken und Saisonalität der Umweltbedingungen implementiert werden.
- **Berücksichtigung von Hegemaßnahmen:** Um die Effizienz bestimmter Hegemaßnahmen abschätzen zu können, könnten diese in das Modell implementiert werden. Hierzu müssten diese Maßnahmen aber zunächst entwickelt und getestet werden, um den Einfluss auf die Ausbreitung und Etablierung quantifizieren zu können. Sobald bekannt ist, in welchem Maße eine Hegemaßnahme die Ausbreitung einer Art einschränkt, könnte diese Maßnahme im Modell in verschiedenen Szenarien getestet werden.
- **Zukunftsszenarien:** Die Einwanderung neuer Arten wie auch die Entwicklung des Verkehrs und der Infrastruktur sind stetig verändernde Prozesse. Für eine Abschätzung der zukünftigen Entwicklung könnten Szenarien zur Entwicklung der Verkehrssituation in Deutschland verwendet werden, um die Auswirkung auf die Ausbreitung und Etablierung von Neobiota zu untersuchen. Auch die Interaktion mit klimatischen Veränderungen könnten so abgeschätzt werden.
- **Erweiterung auf größere räumliche Skalen:** Die Ausbreitung von Neobiota ist nicht auf Deutschland beschränkt und auch die Ausbreitung innerhalb Deutschlands hängt von der Verteilung und Bewegung von Art und Individuen außerhalb Deutschlands ab. Eine wesentliche Rolle spielen hier Vorkommen in Anrainerstaaten, aber viele Arten kommen auch aus weit entfernten Gebieten. Dies spielt vor allem für den marinen Bereich eine große Rolle, da die internationale Schifffahrt die wichtigste Antriebsfeder für die Verbreitung von Arten darstellt. Ein Fokus auf Deutschland allein greift daher häufig zu kurz, um die eine umfassende Darstellung zur Ausbreitung und Etablierung von Neobiota in Deutschland zu ermöglichen. Daher würde eine Betrachtung der Ausbreitung über die Grenzen Deutschlands hinaus die Vorhersagen auch für das Bundesgebiet verbessern. Die Transportnetzwerke, die die Basis für dieses Forschungsvorhaben bilden, liegen auch für andere Staaten innerhalb und außerhalb der EU vor, so dass das Modell leicht auf Anrainerstaaten oder die gesamte EU ausgedehnt werden kann. Dadurch lassen sich genauere Aussagen für Deutschland machen, die die Basis für eine effektive Prävention weiterer Einfuhren und Ausbreitung von Neobiota darstellen könnten.
- **Erweiterung um weitere Netzwerke:** Das Modell könnte mit bereits vorhandenen Modellen für den globalen Transport gekoppelt werden. So könnte z.B. ein Modell zur Ausbreitung über den globalen Schiffsverkehr verwendet werden, um die Einfuhr nach Deutschland und die anschließende weitere regionale Ausbreitung zu simulieren. Hierdurch würden auch die Seewege in deutschen Hoheitsgewässern in die Untersuchung mit eingehen, wodurch Vorhersagen für den marinen Bereich überhaupt erst möglich wären. Zur Priorisierung von Einfuhrwegen und Prävention weiterer Einfuhren ist eine Erweiterung des räumlichen Fokus essentiell.

Mit der Ausschreibung des Forschungsvorhabens „Modellierung der Einfuhr und Verbreitung von invasiven Arten durch Verkehrsträger“ hat das BMVI-Expertennetzwerk eine weitsichtige Entscheidung getroffen und das Fundament für wichtige weitere Schritte zum Verständnis der Ausbreitung und Etablierung gebietsfremder Arten gelegt. Ein solches Vorhaben ist weltweit einzigartig und es ist davon auszugehen, dass dieses Engagement über die Grenzen Deutschlands hinaus Beachtung finden wird. Die freie Verfügbarkeit des Modells bietet die große Möglichkeit, dass andere Arbeitsgruppen, das Modell anwenden, testen und verbessern. Dies würde auch die Vorhersagekraft des Modells für Deutschland verbessern.

5 Abbildungsverzeichnis

Abbildung 1: Übersicht über die Arbeiten von SSP Consult.	19
Abbildung 2: Verkehrsbelastungen Pkw und Lkw (DTV: Durchschnittlich täglicher Verkehr).	21
Abbildung 3: Anzahl Züge pro Tag in Deutschland.....	22
Abbildung 4: Anzahl Güterbinnenschiffe pro Jahr in Deutschland.	23
Abbildung 5: Verteilung der Güterströme in den Zielverkehrszellen am Beispiel der Container.	24
Abbildung 6: Verteilung der aquatischen Umweltparameter mittlere Wassertemperatur und elektrischer Leitfähigkeit der Bundeswasserstraßen.....	27
Abbildung 7: Vorkommen von <i>Senecio inaequidens</i> für vier Jahre (Heger & Böhmer, 2005). Punkte innerhalb des grauen Kastens wurde zur Kalibrierung verwendet.....	28
Abbildung 8: Vorkommen von <i>Dikerogammarus villosus</i> in deutschen Bundeswasserstraßen über vier Zeiträume.	29
Abbildung 9: Vorkommen von <i>Corbicula fluminea</i> in deutschen Bundeswasserstraßen über vier Zeiträume.	30
Abbildung 10: Schematische Darstellung der Modellstruktur.	32
Abbildung 11: Darstellung der Funktionen f aus Tabelle 5 unter Verwendung der kalibrierten Parameter für <i>Senecio inaequidens</i> und <i>Dikerogammarus villosus</i>	36
Abbildung 12: Exemplarische Darstellung der Habitatgüte.....	37
Abbildung 13: Beispielhafte Simulation der Ausbreitung von <i>Senecio inaequidens</i> durch den Straßen- und Schienenverkehr. Dargestellt ist die Wahrscheinlichkeit der Ausbreitung und Etablierung der Art, $P(Inv)$, ausgehend von Initialpopulationen.....	39
Abbildung 14: Ergebnisse der Kalibrierung (A) und Validierung (B) für die simulierte Ausbreitung von <i>Senecio inaequidens</i> über die Verkehrsträger Straße und Schiene.	45
Abbildung 15: Ergebnisse der Kalibrierung (A) und Validierung (B) für die simulierte Ausbreitung von <i>Corbicula fluminea</i>	48
Abbildung 16: Ergebnisse der Kalibrierung (A) und Validierung (B) für die simulierte Ausbreitung von <i>Dikerogammarus villosus</i>	50
Abbildung 17: Einfallstore der Invasion für Straßen- (rot), Schienen- (schwarze Kreise) und Schiffsverkehr (blau).....	51
Abbildung 18: Routen der Ausbreitung entlang (A) terrestrischer und (B) aquatischer Verkehrsnetzwerke in Deutschland. Die Anzahl kürzester Pfade kann als Wahrscheinlichkeit einer Ausbreitung interpretiert werden. Siehe Abschnitt 3.5 für eine detailliertere Analyse.....	53
Abbildung 19: Beispielhafte Route der höchsten Wahrscheinlichkeit der Ausbreitung zwischen zwei beliebigen Punkten.	55
Abbildung 20: Ausbreitungsrouten innerhalb Deutschlands entlang (A) des Straßen- und Schienennetzes, (B) entlang des Straßennetzes, (C) entlang des Schienennetzes und (D) entlang des Wasserstraßennetzes.....	56
Abbildung 21: Zusammenhang zwischen Verkehrsdichte und Anzahl kürzester Pfade aller Links... 57	
Abbildung 22: Vergleich der Hauptrouten der Ausbreitung zwischen Verkehrsträgern, Straßentypen und Wasserstraßen.	58

Abbildung 23: Ausbreitungsrouten mit Startpunkt Hamburg über (A) alle terrestrischen Netzwerke, (B) das Straßennetzwerk, (C) das Schienennetzwerk und (D) das Wasserstraßennetzwerk.	59
Abbildung 24: Beispiele für Ausbreitungsrouten mit Startpunkt Dresden (A,B) und München (C,D) über alle terrestrischen Netzwerke (A,C) und das Schienennetzwerk (B,D).	60
Abbildung 25: Beispiele für Ausbreitungsrouten mit Startpunkten an Grenzübergängen.	61

6 Tabellenverzeichnis

Tabelle 1: Evaluierung der möglichen Prädiktoren der Ausbreitung von Neobiota/invasiven Arten durch Verkehrsträger und deren Relevanz für das Forschungsvorhaben.	15
Tabelle 2: Evaluierung der Prädiktoren der Etablierung und deren Relevanz für das Forschungsvorhaben.	16
Tabelle 3: Kategorien der Flächenbedeckung der corine datenbank und des modells und Anteil an der gesamtbedeckung in deutschland.	25
Tabelle 4: Übersicht der Ausbreitungswege, die in dem Modell berücksichtigt wurden.....	31
Tabelle 5: Übersicht über funktionen der ausbreitung.	34
Tabelle 6: Ergebnisse der Modellkalibrierung für <i>Senecio inaequidens</i>	46
Tabelle 7: Ergebnisse der Modellkalibrierung für <i>Corbicula fluminea</i>	49
Tabelle 8: Ergebnisse der modellkalibrierung für <i>Dikerogammarus villosus</i>	50

7 Quellenverzeichnis

- Albert, R., & Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews Of Modern Physics*, 74(1), 47–97.
- Bagnara, M., Van Oijen, M., Cameron, D., Gianelle, D., Magnani, F., & Sottocornola, M. (2018). Bayesian calibration of simple forest models with multiplicative mathematical structure: A case study with two Light Use Efficiency models in an alpine forest. *Ecological Modelling*, 371, 90–100. <https://doi.org/10.1016/j.ecolmodel.2018.01.014>
- Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., ... Richardson, D. M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology and Evolution*, 26(7), 333–339. <https://doi.org/10.1016/j.tree.2011.03.023>
- Capinha, C., Essl, F., Seebens, H., Pereira, H. M., & Kühn, I. (2018). Models of alien species richness show moderate predictive accuracy and poor transferability. *NeoBiota*, 38, 77–96. <https://doi.org/10.3897/neobiota.38.23518>
- Chytrý, M., Maskell, L. C., Pino, J., Pyšek, P., Vilà, M., Font, X., & Smart, S. M. (2008). Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. *Journal of Applied Ecology*, 45(2), 448–458. <https://doi.org/10.1111/j.1365-2664.2007.01398.x>
- Clark, J. S. (2004). Why environmental scientists are becoming Bayesians. *Ecology Letters*, 8(1), 2–14. <https://doi.org/10.1111/j.1461-0248.2004.00702.x>
- Elliott, J. M. (2003). A comparative study of the dispersal of 10 species of stream invertebrates. *Freshwater Biology*. <https://doi.org/10.1046/j.1365-2427.2003.01117.x>
- Heger, T., & Böhmer, H. J. (2005). The Invasion of Central Europe by *Senecio Inaequidens* DC. - A Complex Biogeographical Problem. *Erdkunde*, 59(1), 34–49. <https://doi.org/10.3112/erdkunde.2005.01.03>
- Hulme, P. E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology*, 46(1), 10–18. <https://doi.org/10.1111/j.1365-2664.2008.01600.x>
- INTRAPLAN et al. (2014). *Verkehrsverflechtungsprognose 2030*. München.
- Kettunen, M., Genovesi, P., Gollasch, S., Pagad, S., Starfinger, U., ten Brink, P., & Shine, C. (2009). Technical support to EU strategy on invasive species (IAS) – Assessment of the impacts of IAS in Europe and the EU: Final report for the European Commission, Institute for European Environmental Policy (IEEP). Brussel, Belgium.
- Leung, B., Drake, J. M., & Lodge, D. M. (2004). Predicting invasions: Propagule pressure and the gravity of allee effects. *Ecology*, 85(6), 1651–1660.
- Liebold, A. M., Brockerhoff, E. G., & Kimberley, M. (2017). Depletion of heterogeneous source species pools predicts future invasion rates. *Journal of Applied Ecology*, 54(6), 1968–1977. <https://doi.org/10.1111/1365-2664.12895>
- Mitchell, C. E., Agrawal, A. A., Bever, J. D., Gilbert, G. S., Hufbauer, R. A., Klironomos, J. N., ... Vázquez, D. P. (2006). Biotic interactions and plant invasions. *Ecology Letters*, 9(6), 726–740. <https://doi.org/10.1111/j.1461-0248.2006.00908.x>
- Mouton, A. M., De Baets, B., & Goethals, P. L. M. (2010). Ecological relevance of performance criteria for

- species distribution models. *Ecological Modelling*, 221(16), 1995–2002.
<https://doi.org/10.1016/j.ecolmodel.2010.04.017>
- Nathan, R., Klein, E., Robledo-Arnuncio, J. J., & Revilla, E. (2012). Dispersal kernels: review. In J. Clobert, M. Baguette, T. G. Benton, & J. M. Bullock (Eds.), *Dispersal Ecology and Evolution*, 187–210. Oxford University Press.
- Pyšek, P., & Richardson, D. M. (2010). Invasive Species, Environmental Change and Management, and Health. *Annual Review of Environment and Resources*, 35(1), 25–55. <https://doi.org/10.1146/annurev-environ-033009-095548>
- Schöll, F. (2007). Der Main-Donau-Kanal als Drehscheibe wandernder Tierarten. *Wasser- Und Schifffahrtsdirektion Süd Jahresbericht*, 69–70.
- Seebens, H., Essl, F., & Blasius, B. (2017). The intermediate distance hypothesis of biological invasions. *Ecology Letters*, 20(2), 158–165. <https://doi.org/10.1111/ele.12715>
- Seebens, H., Essl, F., Dawson, W., Fuentes, N., Moser, D., Pergl, J., ... Blasius, B. (2015). Global trade will accelerate plant invasions in emerging economies under climate change. *Global Change Biology*, 21(11), 4128–4140. <https://doi.org/10.1111/gcb.13021>
- Seebens, H., Gastner, M. T., & Blasius, B. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters*, 16(6), 782–790. <https://doi.org/10.1111/ele.12111>
- Seebens, H., Schwartz, N., Schupp, P. J., & Blasius, B. (2016). Predicting the spread of marine species introduced by global shipping. *Proceedings of the National Academy of Science USA*, 113(20), 5646–5651. <https://doi.org/10.1073/pnas.1524427113>
- Simberloff, D., Martin, J.-L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., ... Vilà, M. (2013). Impacts of biological invasions: what's what and the way forward. *Trends in Ecology and Evolution*, 28(1), 58–66. <https://doi.org/10.1016/j.tree.2012.07.013>
- Sylvester, F., Kalaci, O., Leung, B., Lacoursière-Roussel, A., Murray, C. C., Choi, F. M., ... MacIsaac, H. J. (2011). Hull fouling as an invasion vector: can simple models explain a complex problem? *Journal of Applied Ecology*, 48(2), 415–423. <https://doi.org/10.1111/j.1365-2664.2011.01957.x>
- Taylor, K., Brummer, T., Taper, M. L., Wing, A., & Rew, L. J. (2012). Human-mediated long-distance dispersal: an empirical evaluation of seed dispersal by vehicles. *Diversity and Distributions*, 18(9), 942–951. <https://doi.org/10.1111/j.1472-4642.2012.00926.x>
- Tilman, D., Clark, M., Williams, D. R., Kimmel, K., Polasky, S., & Packer, C. (2017). Future threats to biodiversity and pathways to their prevention. *Nature*, 546(7656), 73–81. <https://doi.org/10.1038/nature22900>
- Ulman, A., Ferrario, J., Forcada, A., Seebens, H., Marchini, A., Arvanitidis, C., & Ambrogi, A. O. (2019). Alien species spreading via biofouling on recreational vessels in the Mediterranean Sea. *Journal of Applied Ecology*, (August), 1–10. <https://doi.org/10.1111/1365-2664.13502>
- von der Lippe, M., Bullock, J. M., Kowarik, I., Knopp, T., & Wichmann, M. (2013). Human-Mediated Dispersal of Seeds by the Airflow of Vehicles. *PLoS ONE*, 8(1), 1–10. <https://doi.org/10.1371/journal.pone.0052733>
- von der Lippe, M., & Kowarik, I. (2007). Long-distance dispersal of plants by vehicles as a driver of plant invasions. *Conservation Biology*, 21(4), 986–996. <https://doi.org/10.1111/j.1523-1739.2007.00722.x>

8 Anhänge

8.1 Detailergebnisse der Kalibrierung

Im Folgenden werden Ergebnisse der Kalibrierung für die drei Beispielarten präsentiert. Den Verlauf der Parameterwerte während der Kalibrierung wird in den sogenannten traceplots dargestellt. Die traceplots zeigen die Parameterwerte in jedem Iterationsschritt der Kalibrierung für alle neun parallel verlaufenden Läufe der Optimierung und dienen der visuellen Bewertung der Konvergenz. Idealerweise starten die einzelnen Läufe mit sehr unterschiedlichen Parameterwerten und konvergieren dann zu einem ähnlichen Wert. Im Laufe der Kalibrierung werden auch andere Parameterkombinationen getestet, die wieder verworfen werden können. In den traceplots ist dies als plötzlicher Sprung in einen anderen Parameterraum zu erkennen.

Die Startwerte der Parameter sind aus der Literatur entnommen und können daher schon den optimalen Werten entsprechen. Dies zeigt sich darin, dass die Kalibrierung keine deutlichen Abweichungen von den ursprünglich gewählten Parameterwerten hervorbringt. Die traceplots erscheinen dann als Streuung um den Idealstart, da immer wieder neue Parameterkombinationen getestet werden, diese aber keine Verbesserung des Modellfits ergeben.

Die Verteilung der getesteten Parameterwerte der letzten 10.000 Iterationsschritte ist den Dichtegrafiken zu erkennen. Diese densityplots zeigen die sogenannten „posteriors“, die die endgültige Verteilung und Unsicherheit für jeden Parameter angibt. So schmaler die Verteilung ist, desto eindeutiger ist die Wahl dieses Parameterwerts.

8.1.1 Senecio inaequidens

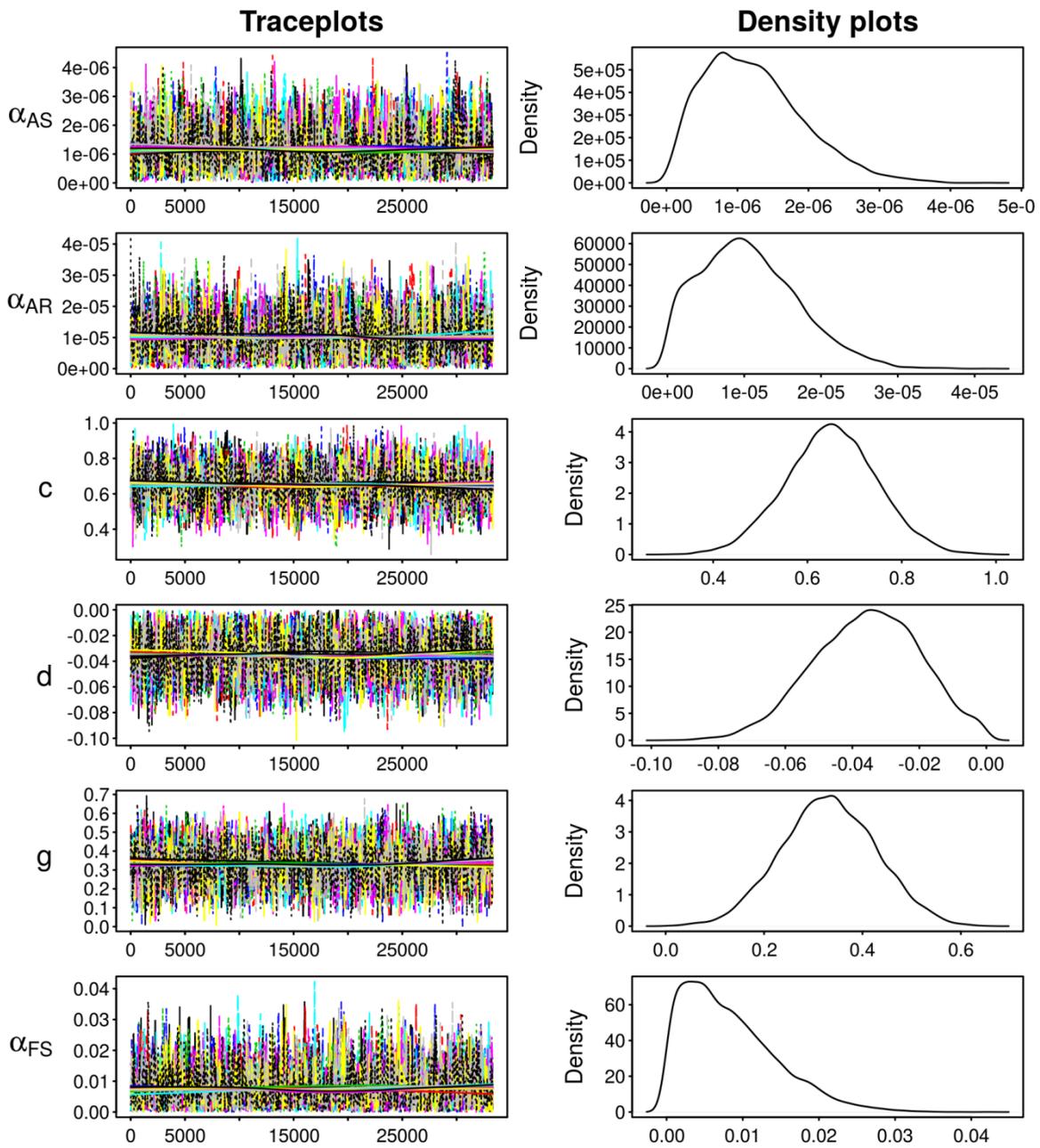


Abbildung S1: Ergebnisse der Kalibrierung für *Senecio inaequidens* (erster Teil)

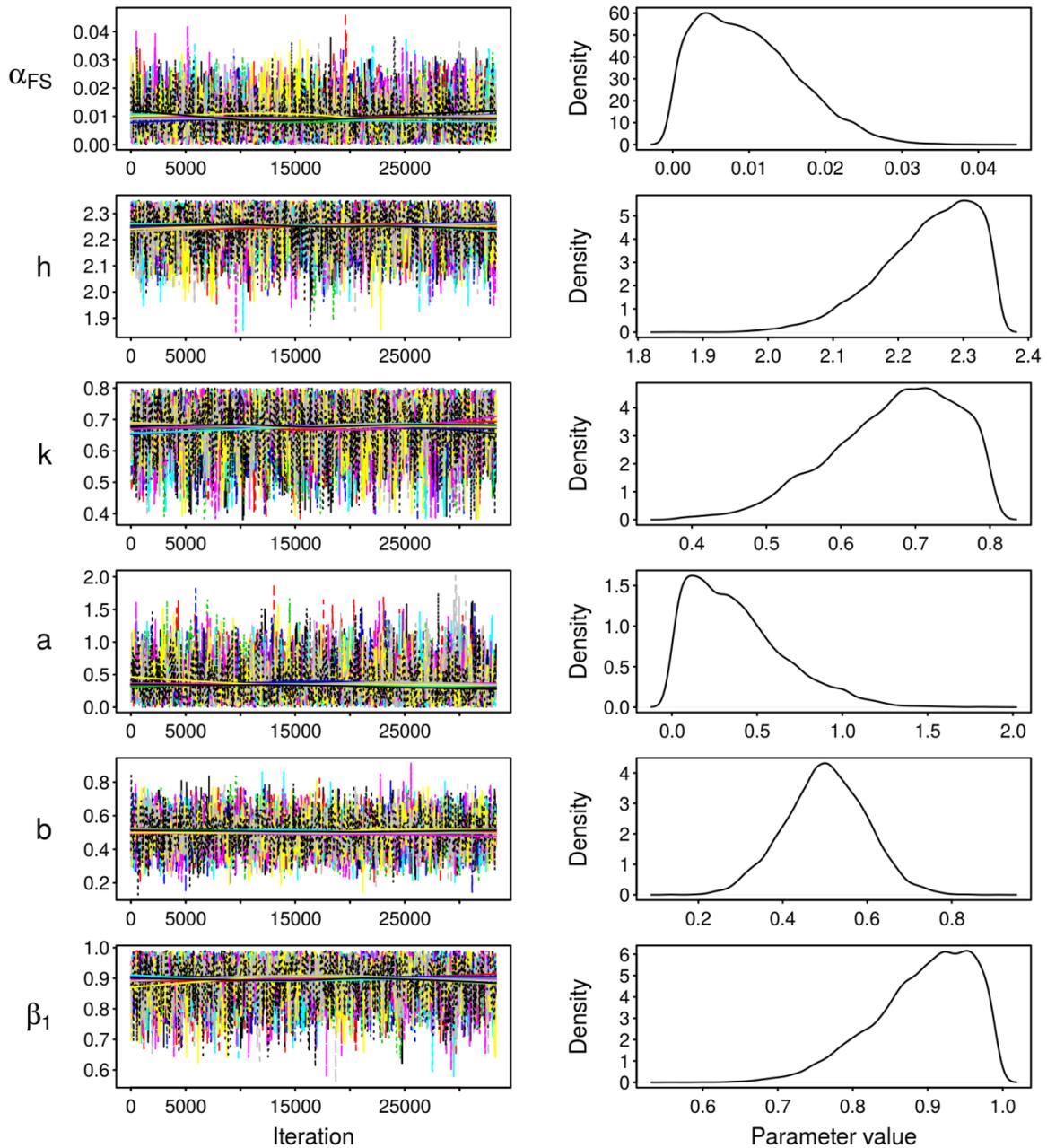


Abbildung S2: Ergebnisse der Kalibrierung für *Senecio inaequidens* (zweiter Teil)

Die ermittelten Parameterwerte für die Ausbreitung von *Senecio inaequidens* weisen eine relativ große Streuung am Ende der Kalibrierung auf. Dies könnte daraufhin deuten, dass es verschiedene Kombinationen an Parameterwerten gibt, die die Daten ähnlich gut beschreiben können. Eine andere Erklärung könnte sein, dass die Daten nicht aussagekräftig genug sind, um die Parameter mit einer engeren Verteilung zu ermitteln. Beides könnte unter Verwendung weiterer Datensätze für unterschiedliche Kombinationen an Ausbreitungswegen näher untersucht werden.

8.1.2 Corbicula fluminea

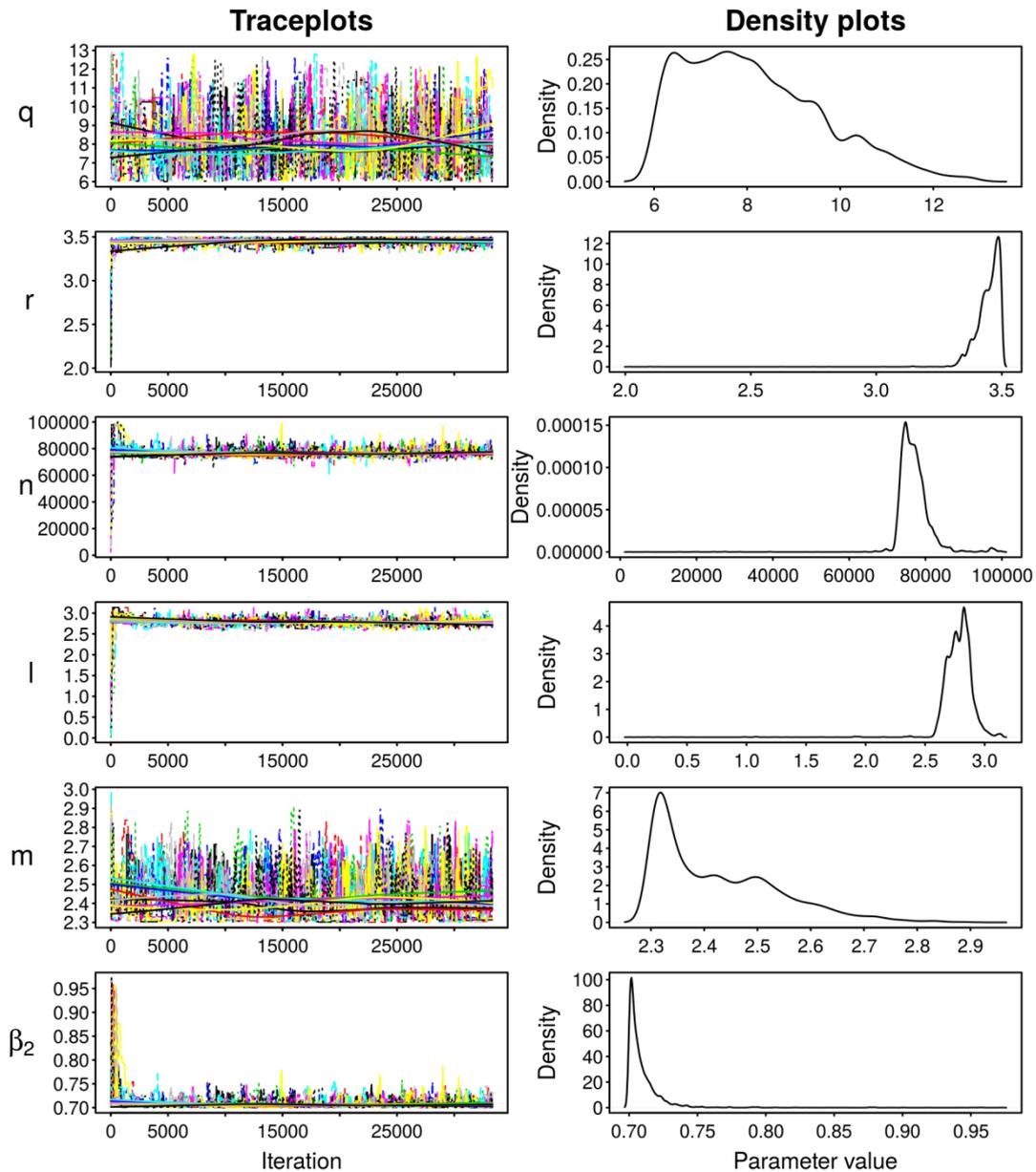
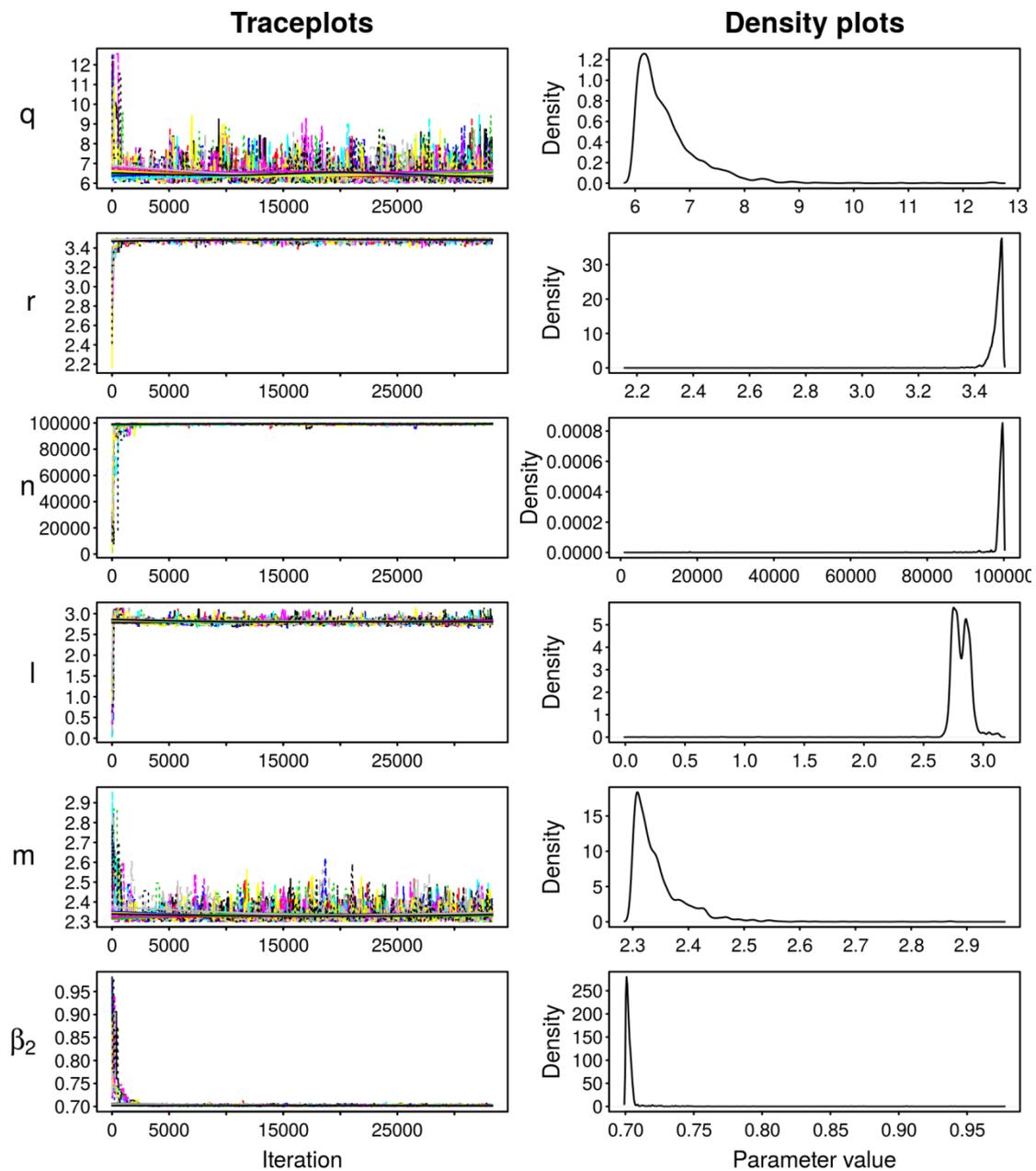


Abbildung S3: Ergebnisse der Kalibrierung für *Corbicula fluminea*

Die ermittelten Parameterwerte für die Ausbreitung von *Corbicula fluminea* weisen in den meisten Fällen eine vergleichsweise geringe Bandbreite an möglichen Werten am Ende der Kalibrierung auf. Dies deutet daraufhin, dass eine eindeutige Kombination an Parameterwerten gefunden wurde, die die beobachteten Vorkommensdaten gut beschreiben kann. Es gibt nur eine geringe Unsicherheit bei der Wahl der ermittelten Parameterwerte. Einzig die Parameterwerte von q zeigen noch eine hohe Streuung.

8.1.3 *Dikerogammarus villosus*Abbildung S4: Ergebnisse der Kalibrierung für *Dikerogammarus villosus*

Ähnlich wie für die andere aquatische Art weisen die ermittelten Parameterwerte für die Ausbreitung von *Dikerogammarus villosus* in den meisten Fällen eine vergleichsweise geringe Bandbreite an möglichen Werten am Ende der Kalibrierung auf. Dies deutet daraufhin, dass eine eindeutige Kombination an Parameterwerten gefunden wurde, die die beobachteten Vorkommensdaten gut beschreiben kann. Es gibt nur eine geringe Unsicherheit bei der Wahl der ermittelten Parameterwerte.

8.2 Handbuch CASPIAN

CASPIAN Handbuch

Hanno Seebens, Maurizio Bagnara
Senckenberg Biodiversität und Klima Forschungszentrum, Frankfurt

24.05.2019

Contents

1	Installation	2
1.1	Schritt 1: Verbindung zum Server	2
1.2	Schritt 2: Installation von CASPIAN	2
1.3	Schritt 3: Ergänzung der Datensätze	3
1.4	Schritt 4: Laden von CASPIAN	3
2	Anwendung	3
2.1	Starten von CASPIAN	3
2.2	Erweiterte Einstellungen	4
2.3	Eingangsdatensätze	4
2.4	Konfiguration der Simulation	6
2.4.1	Struktur des Modells	7
2.4.2	Simulation und Ausgabe	8
2.4.3	Startbedingungen	8

CASPIAN (Calculating Alien Species Probability of Invasion Along Networks) ist ein Computermodell zur Simulation der Ausbreitung und Etablierung von Arten entlang von Verkehrsträgern wie Straßen, Schienen und Wasserstraßen. Das Modell liegt als R Paket ‘CASPIAN’ vor. R und CASPIAN sind ‘open source’ Anwendungen, so dass alle Funktionen frei zugänglich und editierbar sind. Im Folgenden werden die Installation und Anwendung von CASPIAN vorgestellt.

CASPIAN ist im Detail im Abschlussberichts des Projekts “Modellierung der Einfuhr und Verbreitung von invasiven Arten durch Verkehrsträger” beschrieben.

1 Installation

Um CASPIAN anwenden zu können, müssen einige Softwarekomponenten vorliegen, die - sofern nicht vorhanden - erst installiert werden müssen. Zunächst wird die freie Programmiersprache R benötigt, welche unter <http://www.r-project.org> heruntergeladen werden kann. Es empfiehlt sich zusätzlich die Entwicklungsumgebung RStudio zu installieren, welches die Anwendung von R und CASPIAN deutlich erleichtert. Es kann kostenfrei unter <https://www.rstudio.com> bezogen werden. Für die weiteren Ausführungen wird angenommen, dass R und RStudio vorliegen. Sämtliche Befehle werden in RStudio ausgeführt.

Nach der Installation von R kann das R Paket ‘CASPIAN’ installiert werden, welches das Modell enthält. Dies erfolgt in drei Schritten:

1.1 Schritt 1: Verbindung zum Server

Zunächst muss eine Verbindung zur Programmier Plattform GitHub <https://github.com/> hergestellt werden. Hierzu muss das R Paket “devtools” installiert werden, welches die notwendigen Funktionen enthält. Kopieren Sie dazu den folgenden Befehl in die Konsole (“Console” befindet sich in der Regel am unteren Bildrand) von RStudio (Zeilen beginnen mit “>”) und drücken Sie ‘Enter’ zum Ausführen:

```
install.packages("devtools")
```

1.2 Schritt 2: Installation von CASPIAN

Nun muss das Paket ‘devtools’ in R geladen werden, wodurch CASPIAN heruntergeladen und installiert werden kann. Kopieren Sie dafür den folgenden Befehl in die Konsole und führen ihn mit ‘Enter’ aus:

```
library("devtools")
install_url("https://github.com/mauriziobagnara/CASPIAN/releases/download/1.0/
CASPIAN_1.0.tar.gz", dependencies = T)
```

Alternativ kann das Paket aus dem Verzeichnis “<https://github.com/mauriziobagnara/CASPIAN/releases/tag/1.0>” heruntergeladen und manuell in R installiert werden. Hierzu muss der folgende Befehl mit entsprechendem Pfad zum Paket ausgeführt werden:

```
install.packages("/path/to/package/CASPIAN_1.0.tar.gz", repos = NULL, type = "source")
```

In Abhängigkeit der vorliegenden Installation von R müssen verschiedenen Pakete zusätzlich installiert werden (‘data.table’, ‘rgdal’, ‘SearchTrees’, ‘geosphere’, ‘sp’, ‘raster’). Diese Pakete sollten während der Installation von CASPIAN automatisch mit installiert werden, so dass keine zusätzlichen Befehle notwendig sind.

Abhängig vom verwendeten Betriebssystem kann es notwendig sein, weitere Software zum Bearbeiten von Karten zu installieren. Unter Linux Distributionen muss z.B. GDAL außerhalb von R installiert werden. Unter Windows erfolgt dies mit der Installation der R Pakete, so dass dieser Schritte in Windows entfällt.

1.3 Schritt 3: Ergänzung der Datensätze

Um CASPIAN im vollen Umfang nutzen zu können, müssen verschiedene Datensätze zur Verfügung gestellt werden. Diese sind im Detail im Abschnitt 2.3 “Eingangsdatensätze” beschreiben. Die Pfade zu den Datensätzen müssen im Konfigurationsskript configFile.R angegeben werden (s. Abschnitt 2.4).

1.4 Schritt 4: Laden von CASPIAN

Nach der erfolgreichen Installation von CASPIAN kann es mit dem folgenden Befehl in R geladen und angewandt werden:

```
library("CASPIAN")
```

2 Anwendung

2.1 Starten von CASPIAN

Nach einer Neuinstallation von CASPIAN sind Standardeinstellungen vorgegeben, die einen direkten Start des Modells erlauben. Es muss nur der Pfad zum Verzeichnis angegeben werden, in dem die Eingangsdatensätze (s. 2.3 Eingangsdatensätze) vorliegen und die Ergebnisse gespeichert werden sollen. Die Eingangsdatensätze müssen in einem Unterordner names “data” gespeichert sein (s. Punkte 3 unten). Um CASPIAN zu starten, öffnen Sie in RStudio ein leeres R Skript (File->New File->R Script) und kopieren Sie die folgenden Zeilen in das leere Skript:

```
## 1. clear environment and close figures
rm(list=ls())
graphics.off()

## 2. load package CASPIAN
library(CASPIAN)

## 3. set path to directory to store model output
mainDir <- "C:/Pfad/zum/Verzeichnis/Modellergebnisse"
setwd(mainDir)

## 4. set path to script configFile
getConfigFile(outputDir = mainDir)
configFile <- file.path(mainDir,"ConfigFile.R")
# example to load a modified version of ConfigFile.R:
# configFile <- "C:/Pfad/zum/Verzeichnis/NewConfigFile.R"

## 5. run CASPIAN
modelResults <- runCASPIAN(configFile=configFile)
```

Die folgenden Schritte werden in diesem Beispielskript ausgeführt:

1. Bereinigung der R Umgebung und schließen aller Fenster: Die Bereinigung der R Umgebung führt dazu, dass vorherige Operationen in R gelöscht werden, so dass ein sauberer Start des Modells gewährleistet ist. Ansonsten kann es zu ungewünschten Interaktionen mit vorherigen Operationen kommen.
2. Laden des Modells: Das Modell muss in R geladen werden, um es anwenden zu können. Weitere Pakete werden automatisch geladen.

3. Angabe des Verzeichnispfads zum Laden und Speichern der Dateien: Hier muss ein Verzeichnispfad zum Ordner des Modells angegeben werden. In diesem Ordner werden die Ergebnisse der Simulationen gespeichert. Der Ordner muss einen Unterordner mit dem Namen "data" beinhalten, in dem die für die Simulationen notwendigen Dateien (s. 2.3 Eingangsdatensätze) gespeichert sind. Es wird eine Angabe in der Form "Pfad/zum/Verzeichnis/" erwartet.
4. Angabe des Verzeichnispfads zum Konfigurationsskript: In diesem Schritt wird der Pfad zum Konfigurationsskript "ConfigFile.R" angegeben, welches alle Einstellungen für das Modell beinhaltet. Bei Ausführung dieser zwei Zeilen wird das Konfigurationsskript aus dem Paket CASPIAN in das Verzeichnis des Modells aus Schritt 3 kopiert. Dieses Skript enthält die Standardkonfiguration des Pakets. Bei erneuter Ausführung dieser Zeilen wird das Skript überschrieben. Falls der Benutzer die Konfigurationen ändern möchte, sollte das modifizierte Skript "ConfigFile.R" unter einem anderen Namen abgespeichert und der Pfad in diesem Schritt entsprechend angepasst werden. Eine Beispiel ist im Skript oben hinter "#" angegeben.
5. Starten des Modells: Mit dieser Zeile wird das Modell gestartet. Es werden die Einstellungen aus dem Objekt "configFile" eingelesen. Falls die Konfigurationen in einem anderen Skript gespeichert werden, muss dies in Zeile 4 entsprechend angepasst werden.

Einträge, die mit '#' beginnen, gelten als Kommentare und werden bei der Ausführung ignoriert.

Speichern Sie das Skript unter einem aussagekräftigen Namen (z.B. "StartCASPIAN.R") ab und führen Sie das Skript aus. Dies können Sie auf unterschiedliche Weise tun. Wählen Sie das Skript in der Tab-Leiste aus, so dass es auf dem Bildschirm zu sehen ist. Dann können Sie z.B. in RStudio oberhalb des Skripts auf "Source" klicken oder Strg+Umschalt+S drücken oder Code->Source auswählen. Das Modell wird dann gestartet und entsprechend der gewählten Konfigurationen ausgeführt. Modellergebnisse werden im vorgegebenen Verzeichnis gespeichert.

Die Standardeinstellungen berechnen die Ausbreitung einer terrestrischen und einer aquatischen Beispielart ausgehend von einzelnen Startpunkten über einen bestimmten Zeitraum. Das Einlesen der Eingangsdatensätze kann ein paar Minuten dauern, da die Verkehrsnetzwerke sehr umfangreich sind und man aufgrund der Nutzerfreundlichkeit die Originalformate (Tabellen und shapefiles) anstatt komprimierter Datensätze verwenden wollte. Dieser Schritt kann deutlich beschleunigt werden, in dem die Datensätze einmalig in R geladen und als komprimiertes R Format gespeichert werden (z.B. mit Hilfe der Funktion "readRDS(Datensatz,"Pfad/zum/Datensatz/Datensatz.rds)"). In allen folgenden Simulationen kann nun auf die komprimierte Dateien zurückgegriffen werden.

Die Modellergebnisse bestehen aus 1.) einem shapefile mit den Verkehrsnetzwerken und den ermittelten Wahrscheinlichkeiten der Invasion und 2.) einem .txt Dokument mit den Wahrscheinlichkeiten in tabellarischer Ausführung. Die Ergebnisse werden in einem Unterordner

2.2 Erweiterte Einstellungen

Für eine Simulation benötigt CASPIAN eine Liste der Modellparameter und der Startbedingungen. Standardmäßig werden diese in dem R Skript 'ConfigFile.r' aufgeführt, welches im Paket CASPIAN enthalten ist. Die in diesem Skript aufgeführten Einstellungen gelten als Standardeinstellungen und können ohne Änderungen übernommen werden. Wenn Änderungen in den Standardeinstellungen vorgenommen werden, beachten Sie die notwendigen Anpassungen im Startskript (s. oben).

2.3 Eingangsdatensätze

CASPIAN benötigt Eingangsdatensätze, auf deren Basis die Ausbreitung und Etablierung einer Art simuliert wird. Die Eingangsdatensätze umfassen zwei Kategorien: Verkehrsintensitäten und Habitatgüte. Im Folgenden wird beschrieben, wie die Datensätze in der ausgelieferten Version aussehen. Wenn neue Netzwerke verwendet werden sollen, müssen diese den Formaten entsprechen.

FromNode	ToNode	Type	Length	cargo	passengers	ID
R200003	R203854	Rail	1.46	0.51	10.5	R1
R200012	R202484	Rail	0.96	94.75	65.5	R2
R200013	R200682	Rail	1.68	31.74	17.0	R3
R200013	R206800	Rail	1.95	0.00	26.0	R4
R200016	R205880	Rail	2.10	11.48	0.0	R5
R200043	R203119	Rail	5.60	19.18	42.5	R6

Figure 1: Beispielhafter Datensatz eines terrestrischen Verkehrsnetzwerks

Die Verkehrsintensitäten werden als shapefile eingelesen mit einer eingebetteten Tabelle der Attribute. In R werde diese Objekte als “SpatialLinesDataFrame” eingelesen, wobei jedes Segment dieses Objekt eine Verbindungen zwischen zwei Knoten darstellt.

Verkehrsnetzwerke bestehen aus Knoten (z.B. Kreuzungen) und Verbindungen (z.B. Straßen) sowie Verkehrsintensitäten (z.B. Anzahl KFZ pro Tag). Aufbauend auf diesen Informationen berechnet CASPIAN die Wahrscheinlichkeit einer Ausbreitung von Knoten zu Knoten.

In der ausgelieferten Version folgende Datensätze verwendet:

- terrestrisches Verkehrsnetzwerk (Straße + Schiene)
- aquatisches Verkehrsnetzwerk (Wasserstraßen)
- Güternetzwerk (Transport von Containern und Paletten)

Jede Zeile der eingebetteten Tabelle beschreibt ein Segment des Netzwerks. Die Spalten liefern Informationen zu jedem Segment. Die Reihenfolge der Spalten muss genau den unten beschriebenen Formaten entsprechen, wobei die Spaltenüberschrift irrelevant ist.

Das **terrestrische Verkehrsnetzwerk** bestehend aus dem Straßen- und Schienennetzwerk muss folgende Angaben beinhalten. Die Spaltenüberschriften sind nicht relevant und dienen nur der Vergleichbarkeit mit den Beispieldatensätzen (in den)Abbildungen 1-4):

1. Spalte: Name des Startknotens des Links (“FromNode”)
2. Spalte: Name des Zielknotens des Links (“ToNode”)
3. Spalte: Typ des Links z.B. (“Type”): Der Typ beschreibt die Art des Verkehrsträgers. Mit “Rail” wird zwischen Straßen- und Schienennetzwerken unterschieden. Für Straßen sind in der ausgelieferten Version die Typen “A”utobahn, “B”undesstraße, “L”andstraße, “K”reisstraße und “G”emeindestraße unterschieden. Daneben gibt es noch weitere Kategorien, die aber nur räumlich begrenzt Verwendung finden wie z.B. “S”taatsstraßen in Sachsen.
4. Spalte: Länge des Links in Kilometer (“Length”)
5. Spalte: Verkehrsintensität für Güterverkehr in Anzahl Fahrzeugen/Zügen pro Tag (“cargo”)
6. Spalte: Verkehrsintensität für Personenverkehr in Anzahl Fahrzeugen/Zügen pro Tag (“passenger”)
7. Spalte: Eindeutige Bezeichnung des Links (“ID”)
8. Spalte (optional): Habitatgüte entlang des Links zwischen 0 (ungeeignet) bis 1 (perfekt geeignet). Sofern die Daten zu Habitatgüte vorliegen, muss diese Spalte mit “Env_suit” überschrieben sein. In diesem Fall werden die Daten aus dieser Spalte zur Abschätzung der Habitatgüte verwendet. Wenn diese Spalte fehlt oder anders überschrieben ist, wird die Habitatgüte entsprechend der Grundeinstellungen des Modells berechnet (s. Abschlussbericht des Projekts).

FromNode	ToNode	Motorized	Non-motorized	Length	ID	CargoToNode
325200001	325200002	404	765	10.220854	W10	0
325200003	325201000	870	651	4.767631	W11	0
325201000	325200003	870	651	4.767631	W12	0
305100001	305111000	5044	1179	38.633023	W13	251
305111000	305100001	5044	1179	38.633023	W14	0
303100001	304206000	3272	116	8.375770	W15	1182

Figure 2: Beispielhafter Datensatz eines aquatischen Verkehrsnetzwerks

Im Netzwerk der *Wasserstraßen* werden folgende Spalte in dieser Reihenfolge erwartet:

1. Spalte: Name des Startknotens des Links (“FromNode”)
2. Spalte: Name des Zielknotens des Links (“ToNode”)
3. Spalte: Intensität des motorisierten Schiffsverkehrs in Anzahl Schiffen pro Monat (“Motorized”)
4. Spalte: Intensität des motorisierten Schiffsverkehrs in Anzahl Schiffen pro Monat (“Non-Motorized”)
5. Spalte: Länge des Links in Kilometer (“Length”)
6. Spalte: Eindeutige Bezeichnung des Links (“ID”)
7. Spalte: Güterumschlag des Zielknotens sofern dieser einen Hafen repräsentiert (“CargoToNode”). Diese Information wird benötigt, um die Menge an abgelassenen Ballastwassers abzuschätzen.
8. Spalte (optional): Habitatgüte entlang des Links zwischen 0 (ungeeignet) bis 1 (perfekt geeignet). Sofern die Daten zu Habitatgüte vorliegen, muss diese Spalte mit “Env_suit” überschrieben sein. In diesem Fall werden die Daten aus dieser Spalte zur Abschätzung der Habitatgüte verwendet. Wenn diese Daten nicht vorliegen, müssen die Parameter Temperatur in °C und elektrische Leitfähigkeit in mS/m in zwei Spalten getrennt vorliegen (“Temperature” und “Conductivity”) (s. Abschlussbericht des Projekts).

Container- und Palettentransporte werden in leicht abgewandelter Weise behandelt. Die räumliche Auflösung liegt nicht auf der Basis von Knoten vor sondern auf der von Verkehrszellen und dem Austausch zwischen diesen. Es werden insgesamt drei Datensätze benötigt. Container- und Palettentransporte werden getrennt als jeweils dreispaltige Datensätze eingelesen (Abbildungen 3,4):

1. Spalte: Startgebiet (“FromArea”)
2. Spalte: Zielgebiet (“ToArea”)
3. Spalte: Anzahl transportierter Container pro Jahr (“numContainers”) bzw. Paletten (“numPallets”)

Das shapefile enthält die Lage der Verkehrszellen als Polygone und die IDs der Verkehrszellen für Container (1. Spalte) und Paletten (2. Paletten).

2.4 Konfiguration der Simulation

Alle Konfigurationen von CASPIAN sind in einem separatem Skript mit dem Namen “configFile.R” gespeichert. Hier werden alle Modellparameter und Einstellungen zur Simulation aufgeführt. Falls Modifikationen der

FromArea	ToArea	numContainers
2185001	4821042	1.07
2185001	4321006	0.93
2185001	41991	0.26
2185001	53993	0.04
2185001	593	0.53
2185001	73041	2.43

Figure 3: Beispielhafter Datensatz für den Containertransport

FromArea	ToArea	numPallets
2185001	2185001	1500.9
2185001	2185002	3988.2
2185001	2185003	2065.8
2185001	2185006	1856.7
2185001	2185007	1613.8
2185001	2187001	567.5

Figure 4: Beispielhafter Datensatz für den Containertransport

Standardeinstellungen vorgenommen werden, ist zu empfehlen, das Skript unter einem aussagekräftigen Namen abzuspeichern, so dass die ursprünglichen Einstellungen weiterhin vorliegen. In diesem Fall muss im Initialisationsskript (s. Abschnitt 2.1) der Pfad zum configFile.R angepasst werden.

Grundsätzlich können alle Einstellungen verändert werden. Allerdings sind manche Einstellungen sensibler als andere und sollten nur mit soliden Kenntnissen der Dynamiken verändert werden. Es wird grundsätzlich unterschieden zwischen Modellparametern und Einstellungen zur Simulation. Die Modellparameter beschreiben die Ausbreitung der Art entlang eines entsprechenden Verkehrsträgers. Diese Modellparameter wurden größtenteils durch Kalibrierung des Modells an Daten zu Beobachtungen ermittelt. Dies ist ein zeitintensiver und aufwändiger Prozess, da eine große Zahl an Kombinationen von Parameterwerten getestet wurde, um das beste Ergebnis zu erzielen. Es empfiehlt sich daher, die Werte der Modellparameter nicht ohne gute Begründung zu ändern.

Die Einstellungen zur Simulation dienen dazu, die Modellläufe an die Ansprüche des Benutzers anzupassen. Es können drei Komponenten eingestellt werden: 1. Struktur des Modells, 2. Simulation und Ausgabe und 3. Startbedingungen:

2.4.1 Struktur des Modells

CASPIAN ist modular aufgebaut und verschiedene Mechanismen der Ausbreitung können ein- und ausgeschaltet werden. Jeder Ausbreitungsweg wie z.B. Anheftung an KFZ oder natürliche Ausbreitung kann separat ausgewählt werden. Somit kann CASPIAN individuell an die Ansprüche des Benutzers angepasst werden. Die Auswahl erfolgt im Konfigurationsskript configFile.R, in dem die Variable des jeweiligen Ausbreitungswegs auf "TRUE" (an) oder "FALSE" (aus) gesetzt wird. Alle Variablen, die einen Ausbreitungsweg ein- und ausschalten beginnen mit "incl_" (z.B. "incl_attachment<-TRUE").

In der Simulation werden nur diejenigen Ausbreitungswege berücksichtigt, die im Konfigurationsskript auf "TRUE" gesetzt sind.

Es gibt zwei übergeordnete Einstellungen, die die Auswahl der Typen der Verkehrsnetzwerke bestimmen. Die entsprechenden Variablen beginnen mit "run" (z.B. "runTerrestrialModel<-TRUE"). Mit diesen Variablen kann zwischen aquatischen und terrestrischen Ausbreitungen unterschieden werden. Wenn die entsprechende Variable nicht ausgewählt ist (z.B. "runAquaticModel<-FALSE"), werden die Ausbreitungswege des entsprechenden Verkehrsnetzes nicht berücksichtigt unabhängig, ob diese auf "TRUE" oder "FALSE" stehen. Die Güterströme fallen in die Kategorie der terrestrischen Verkehrsnetzwerke.

In Abhängigkeit der Verkehrsnetzwerke können zusätzliche Einstellungen gewählt werden. So können für das terrestrische Netzwerk (Straße + Schiene) verschiedene Kategorien wie Autobahn oder Bundesstraße ausgewählt werden. Dies geht aus den Beschreibungen im configFile.R hervor.

2.4.2 Simulation und Ausgabe

Im Folgenden wird beschrieben, wie man mit CASPIAN die Ausbreitung einer Art simulieren kann. Eine Simulation beschreibt einen Modellauf von CASPIAN und einen bestimmten Zeitraum und mit vorgegebenen Startbedingungen. Die Simulation wird durch verschiedene Einstellungen gesteuert, die im Skript configFile.R in der Rubrik "General model settings" enthalten sind. Hier kann entschieden werden, ob z.B. Zwischenergebnisse graphisch dargestellt ("makeplot<-TRUE") und ob diese Grafik gespeichert werden sollen ("save_plot<-TRUE").

Spezifischere Angaben über die Länge der Simulation und den Export können zu den jeweiligen Verkehrsnetzwerken individuell vorgenommen werden. Dies wurde separat implementiert, da Simulationen über das deutlich kleinere Netzwerk der Wasserstraßen deutlich schneller ablaufen können als dies für terrestrische Netzwerke der Fall ist. Daher kann die Anzahl der simulierten Zeitschritt separat eingestellt werden.

Die Ausgabe der Simulationen kann in drei verschiedenen Formaten erfolgen. Es können zwei tabellarische Formate ("csv" und "txt") ausgewählt. Die ausgegebene Tabelle enthält für jedes Segment eines Netzwerks die Simulationsergebnisse (Wahrscheinlichkeiten für Invasion, Einführung und Etablierung) und die zugrunde liegenden Parameter (z.B. Länge, Verkehrsbelastung und Habitatgüte). Im dritten Format ("shp") wird die Tabelle in ein shapefile eingebettet, welches die geographische Information der Verkehrsnetzwerke enthält. Mit dem shapefile können die Ergebnisse als Karte dargestellt werden. In Abhängigkeit der verwendeten Einstellungen und Netzwerke kann die Erstellung der shapefile einige Minuten in Anspruch nehmen. Zusätzlich können Ergebnisse zwischengespeichert werden.

Für die jeweiligen Verkehrsnetzwerke können weitere Einstellungen zur Berechnung der Ausbreitung und Etablierung vorgenommen werden. Dies umfasst Modellparameter, die die Ausbreitung und Etablierung einer Art bestimmen, und Schwellwerte, mit denen eine Mindestanzahl an Containern und Paletten festgesetzt werden kann. Da die Parameterwerte Ergebnis einer umfangreichen Kalibrierung sind, wird empfohlen, diese Werte nicht zu ändern.

2.4.3 Startbedingungen

Zur Simulation der Ausbreitung und Etablierung werden Startbedingungen benötigt, die die Verbreitung und die Charakteristiken einer Art beschreiben. Die Verbreitung einer Art wird als Tabelle von Koordinaten (Längen- und Breitengrad) eingelesen. Jeder Koordinate beschreibt ein Vorkommen der Art. Diese Vorkommen werden während der Simulation als stabil angesehen und dienen als Startpunkte der Ausbreitung. Die einzulesende Tabelle muss zwei Spalten mit den Namen "Long" und "Lat" enthalten. Bisher werden die Koordinaten im configFile.R eingefügt. Umfangreichere Datensätze können aber auch eingelesen werden. Hierfür muss die Zeile zur Festlegung der Variablen "init_coords_*" im configFile.R entsprechend angepasst werden.

Die Vorkommen einer Art sind über die Koordinaten punktweise angegeben. Da aber die Ausbreitung auf einem Netzwerk mit Liniensegmenten simuliert wird, müssen in einem ersten Schritt die punktwisen Vorkommen in Beobachtungen entlang der Netzwerksegmente wie Straßen übersetzt werden. Dafür wird angenommen, dass alle Segmente in einem bestimmten Radius um die punktweise Beobachtung ebenfalls Vorkommen aufweisen.

Der Radius kann vom Nutzer im Konfigurationsskript festgelegt werden. Über den Radius können auch Unsicherheiten der Beobachtungen in die Simulation mit einfließen, da Koordinatenangaben ungenau (z.B. bei Angaben für Rasterzellen) und unvollständig sein können. Der Radius kann individuell für aquatische und terrestrische Organismen festgesetzt werden.