



Abschätzung des Denitrifikationspotenzials mit Random-Forest-Modellen unter Verwendung hydrochemischer Kovariablen

Marc Ohmer¹  · Tanja Liesch¹ · Julian Xanke²

Eingegangen: 7. Dezember 2024 / Überarbeitet: 11. März 2025 / Angenommen: 11. März 2025
© The Author(s) 2025

Zusammenfassung

Aufgrund hoher Stickstoffeinträge überschreiten Nitratkonzentrationen im Grundwasser in Deutschland teils Trinkwassergrenzwerte, was präzisere Vorhersagemodelle und Managementstrategien erfordert. Diese Studie kombiniert Random-Forest-Modelle mit Explainable AI (SHAP-Werten), um die räumliche Verteilung der Nitratkonzentration im Grundwasser von Baden-Württemberg und Niedersachsen vorherzusagen und den Einfluss des Denitrifikationspotenzials zu analysieren. Die Ergebnisse zeigen, dass geodatenbasierte Modelle für Baden-Württemberg gute, für Niedersachsen jedoch fehlerhafte Vorhersagen liefern.

Die Diskrepanzen sind auf nicht erfasstes Denitrifikationspotenzial zurückzuführen. Durch SHAP-Werte wird der Einfluss des Denitrifikationspotenzials quantifiziert und sichtbar gemacht. Die Einbeziehung chemischer Parameter wie Eisen und Kalium, die eng mit Denitrifikationsprozessen verknüpft sind, erhöht die Modellgüte in Niedersachsen signifikant ($R^2 = 0,06$ auf 0,72) und bestätigt die zentrale Rolle anoxischer Bedingungen und spezifischer Elektronenakzeptoren für den Nitratabbau. SHAP zeigt zudem, dass ohne chemische Parameter räumliche Prädiktoren in Niedersachsen zufällige Muster widerspiegeln, statt kausale Zusammenhänge abzubilden. Random-Forest-Modelle kombiniert mit SHAP-Analysen liefern wertvolle Einblicke in die komplexe Wechselwirkung zwischen Nitratdynamik und Denitrifikation im Grundwasser.

Schlüsselwörter Nitrat · Monitoring · Deutschland · Explainable AI

✉ Marc Ohmer
marc.ohmer@kit.edu

Tanja Liesch
tanja.liesch@kit.edu

Julian Xanke
julian.xanke@tzw.de

¹ Karlsruher Institut für Technologie (KIT), Institut für Angewandte Geowissenschaften (AGW), Adenauerring 20b, 7613 Karlsruhe, Deutschland

² TZW: DVGW-Technologiezentrum Wasser, Karlsruher Straße 84, 76139 Karlsruhe, Deutschland

Assessment of denitrification potential using Random Forest and hydrochemical covariates

Abstract

Due to significant nitrogen inputs, nitrate concentrations in groundwater in Germany often exceed drinking water thresholds, necessitating better prediction models and management strategies. This study uses Random Forest models with Explainable AI (SHAP values) to predict nitrate concentrations in groundwater across Baden-Württemberg and Lower Saxony and evaluate the role of denitrification potential. Results show accurate predictions in Baden-Württemberg but notable errors in Lower Saxony. These discrepancies are attributed to unaccounted spatial variability of denitrification potential. SHAP values quantify and visualize its influence. Incorporating chemical parameters such as iron and potassium, closely associated with denitrification, significantly enhances model performance in Lower Saxony (R^2 rising from 0.06 to 0.72). This highlights the role of anoxic conditions and electron acceptors in nitrate reduction. Without chemical parameters, SHAP reveals how spatial predictors in Lower Saxony reflect random patterns rather than causal relationships. Combining Random Forest models and SHAP offers valuable insights into nitrate dynamics and denitrification.

Keywords Nitrate · Monitoring · Germany · Explainable AI

Einleitung

Obwohl in den letzten Jahren ein leichter Rückgang der Nitratbelastung im Grundwasser zu verzeichnen war, bleiben die Nitratkonzentrationen in vielen landwirtschaftlich genutzten Gebieten Deutschlands weiterhin problematisch hoch (BMEL 2020). Infolgedessen werden in bestimmten Fällen die Trinkwassergrenzwerte überschritten, sodass ohne zusätzliche Maßnahmen zur Reduktion der Nitratkonzentration gesundheitliche Risiken bestehen. Besonders betroffen sind Säuglinge und Kleinkinder, bei denen eine erhöhte Nitrataufnahme zu gesundheitlichen Beeinträchtigungen führen (WHO 2017). Trotz der geringfügigen Verbesserungen zeigt die Auswertung der Messwerte, dass ein erheblicher Anteil der Messstellen im EU-Nitratmessnetz weiterhin Konzentrationen über dem Grenzwert von 50 mg/l aufweist. So lagen im Zeitraum 2016–2018 an 26,7 % der Messstellen die durchschnittlichen Nitratkonzentrationen über diesem Schwellenwert, verglichen mit 28,2 % im vorangegangenen Berichtszeitraum 2012–2015 (BMEL 2020). Regionale Analysen der Nitrateinträge verdeutlichen die komplexe Natur der Nitratbelastung und zeigen erhebliche regionale Unterschiede (Bechtel 2008), die eine pauschale Betrachtung erschweren. Während einige Gebiete Fortschritte verzeichnen, stagnieren die Werte in anderen Regionen oder zeigen sogar eine Verschlechterung. Dies unterstreicht die Notwendigkeit einer differenzierten Analyse der Entwicklung der Nitratbelastung unter Berücksichtigung regionaler Besonderheiten. Zudem gibt es Hinweise darauf, dass die bisher ergriffenen Maßnahmen zur Reduzierung der Nitratbelastung nur begrenzt wirksam sind (BMEL 2020).

Gleichzeitig nimmt die Nitratabbaukapazität in einigen Grundwasserkörpern ab, was die langfristige Wirksamkeit von Maßnahmen zur Reduzierung der Nitrateinträge erschweren könnte (Bergmann und Dietrich 2013). Um

die Entwicklung geeigneter Maßnahmen zur Reduzierung der Nitrateinträge voranzutreiben und eine effektive Bewirtschaftung der Grundwasserressourcen zu ermöglichen, spielt die räumliche Modellierung der Nitratkonzentration im Grundwasser eine entscheidende Rolle. Dabei sind auch die Prozesse der Denitrifikation im Grundwasser von großer Bedeutung, da diese die Nitratkonzentrationen massiv beeinflussen können. Aus diesem Grund müssen diese Prozesse beispielsweise auch bei der Ausweisung von sogenannten „Roten (Nitrat-)Gebieten“ im Rahmen der Düngerordnung berücksichtigt werden. Um eine effektive Bewirtschaftung der Grundwasserressourcen zu ermöglichen, ist die Entwicklung präziser Methoden zur zeitlichen und auch räumlichen Vorhersage der Nitratkonzentration unerlässlich.

Traditionell werden für die räumliche Vorhersage der Nitratkonzentration im Grundwasser, vor allem von Behörden, deterministische Interpolationsmethoden wie Inverse-Distanzgewichtung (Ohlert et al. 2022) und einfache Ansätze wie die Verbindung von Datenpunkten zu Polygonen wie Voronoi- bzw. Thiessen-Polygone oder Delaney-Dreiecke verwendet. Darüber hinaus werden auch geostatistische Interpolationsmethoden der Kriging-Familie (u. a. Arslan et al. 2017; Balacco et al. 2023; Ghadermazi et al. 2011; Rostami et al. 2020; Wriedt et al. 2019; Zaresefat et al. 2023) verwendet. Diese Methoden basieren ausschließlich auf den räumlichen Beziehungen der Messdaten zueinander. Einige Ansätze berücksichtigen zusätzliche gemessene Parameter wie Arsen, Eisen und Ammonium (Dalla Libera et al. 2017) sowie Geologie- und Landnutzungsdaten (Bárdossy et al. 2003; Pebesma und De Kwaadsteniet 1997), um weitere Informationen über die räumliche Struktur und Zusammenhänge zwischen den Variablen zu erfassen. Weiterhin werden auch statistische Modelle wie die Multiple Linear Regression (Boy-Roura et al. 2013; Mfumu Kihumba et al. 2016; Ouedraogo und Vanclooster 2016) oder auch

numerische Modelle verwendet (Graf und Schäfer 2002; Karlović et al. 2022; Matiatis et al. 2019; Nguyen und Dietrich 2018; Sullivan et al. 2019).

In den letzten Jahren haben maschinelle Lernverfahren (ML), wie z. B. Feedforward-Netzwerke (Ehteshami et al. 2016; Stylianoudaki et al. 2022), Convolutional Neural Networks (CNNs) (Karimanzira et al. 2023; Nadiri et al. 2022), Recurrent Neural Networks (RNNs); (Alajas et al. 2021; Nair und Vijaya 2023) und Long Short-term Memory Netze (LSTMs); (Hrnjica et al. 2021; Saha et al. 2023), sowie Ensemble-Methoden wie AdaBoost (Sakizadeh et al. 2024), Gradient Boosting, Extreme Gradient Boosting (Gholami und Booij 2022; Ransom et al. 2022) und vor allem Random Forests (Knoll et al. 2019; Nolan et al. 2015, 2014; Ouedraogo und Vanclooster 2016; Ransom et al. 2022) an Bedeutung gewonnen, um die Modellierung der Nitratkonzentration zu verbessern. Random Forest ist ein Ensemble-Lernverfahren, das auf der Kombination mehrerer Entscheidungsbäume basiert. Jeder Baum trifft eine eigene Prognose, die anschließend zu einer Gesamtvorhersage aggregiert wird, wodurch das Modell robust gegenüber Überanpassung ist und komplexe nichtlineare Zusammenhänge zuverlässig abbilden kann (Breiman 2001).

Diese ML-Methoden erfassen die komplexen nichtlinearen Beziehungen zwischen räumlich flächenhaft vorhandenen hydrogeologischen und landwirtschaftlichen Parametern (z. B. hydrogeologische Einheit, Landnutzung, Vegetationsindizes) und der Verteilung von Nitrat im Grundwasser. Um jedoch mögliche Denitrifikationsprozesse zu berücksichtigen, ist die Einbindung von Eingangsparametern, welche diese Prozesse steuern, erforderlich. Diese sind jedoch selten flächenhaft vorhanden, was eine flächenhafte Modellierung von Denitrifikationsprozessen, insbesondere mit maschinellen Lernmodellen, zu einer Herausforderung macht.

Diese Herausforderung, das Denitrifikationspotenzial in Modellen angemessen abzubilden, wird durch bestehende Studien deutlich hervorgehoben. Wolters et al. (2022) entwickelten eine Methode zur regionalen Bestimmung der Denitrifikationsbedingungen im Grundwasser Deutschlands. Dabei kombinierten sie die Interpolation von redox-sensitiven Parametern wie Sauerstoff, Nitrat, Eisen, Mangan und gelöstem organischen Kohlenstoff (DOC) mit einem 2D-hydrodynamischen Modell zur Quantifizierung der Denitrifikation. Ihre Ergebnisse zeigen, dass in der Norddeutschen Tiefebene hohe Redoxklassen dominieren, die durch erhöhte Konzentrationen von Eisen, Mangan, DOC und Sauerstoff gekennzeichnet sind. Diese Bedingungen begünstigen reduzierende Prozesse und fördern den mikrobiellen Nitratabbau. In den Übergangsbereichen zwischen der Norddeutschen Tiefebene und den Mittelgebirgen wurden intermediäre Reduktionsraten beobachtet, während in den konsolidierten Festgesteinen Süddeutschlands oxidie-

rende Bedingungen vorherrschen, die nur eine geringe Denitrifikation zulassen. Dies resultiert in einem ausgeprägten Nord-Süd-Gefälle in Deutschland. Dabei ist zu berücksichtigen, dass großräumige Modellierungen zwar überregionale Muster abbilden, jedoch die kleinräumige Variabilität der Nitratabbaukapazität oft nur eingeschränkt erfassen können. Redoxbedingungen können bereits auf kurzen räumlichen Skalen stark schwanken, beeinflusst durch Faktoren wie Unterschiede in der Sedimentzusammensetzung, den Grundwasserfluss und die Verfügbarkeit organischer Substrate (Tesoriero et al. 2000). Knoll et al. (2020) entwickelten einen bundesweit harmonisierten Ansatz zur Schätzung von Redoxbedingungen und Nitratkonzentrationen im Grundwasser, basierend auf Stickstoffüberschüssen, Wasserflüssen und Nitratmesswerten. Mithilfe von maschinellen Lernmodellen, insbesondere Random Forest (RF) und Quantile Random Forest (QRF), wurden regionale Unterschiede im Denitrifikationspotenzial und in der Nitratreduktion modelliert. Die Ergebnisse zeigen hier auch, dass die höchsten Reduktionsraten in Norddeutschland auftreten, insbesondere in porösen Grundwasserleitern mit anaeroben Bedingungen, die eine effektive Denitrifikation ermöglichen. Im Gegensatz dazu sind die Reduktionsraten in Süddeutschland deutlich niedriger, was vor allem auf die dort dominierenden geklüfteten, konsolidierten Aquifere mit aeroben Bedingungen zurückzuführen ist. Diese hydrogeologischen Unterschiede beeinflussen die Persistenz und Mobilität von Nitrat im Grundwasser erheblich.

In der vorliegenden Studie wird die Nitratkonzentration des Grundwassers in den Bundesländern Baden-Württemberg und Niedersachsen mit einem Random-Forest-Modell modelliert und das Denitrifikationspotenzial mit Explainable AI analysiert.

Explainable-AI(XAI)-Methoden zielen darauf ab, Entscheidungen künstlicher Intelligenz nachvollziehbar und interpretierbar zu machen. Sie ermöglichen eine transparente Analyse des Einflusses einzelner Eingangsparameter auf die Modellvorhersagen und liefern Erklärungen für die zugrunde liegenden Entscheidungsprozesse. Dadurch können Modellfehler besser identifiziert, Unsicherheiten reduziert und gezielte Optimierungen vorgenommen werden.

SHAP-Werte (SHapley Additive exPlanations) sind eine besonders leistungsfähige XAI-Methode zur Interpretation komplexer maschineller Lernmodelle (Lundberg und Lee 2017). Sie quantifizieren den Beitrag jedes Eingangsparameters zur Modellvorhersage und berücksichtigen dabei auch Wechselwirkungen zwischen Variablen. SHAP-Werte ermöglichen sowohl eine globale Analyse der wichtigsten Einflussgrößen im Modell als auch eine lokale Erklärung einzelner Vorhersagen. Im Gegensatz zu herkömmlichen Feature-Importance-Metriken liefern sie präzisere und intuitivere Einblicke, insbesondere für nichtlineare Modelle wie Random Forests, Gradient Boosting oder neuronale

Netze. Beispielsweise zeigte die Studie von Ransom et al. (2022), dass landwirtschaftliche Stickstoffquellen, atmosphärische Stickstoffdeposition, Bodeneigenschaften und Klimafaktoren signifikant zur Vorhersage der Nitratkonzentration im Grundwasser in den USA beitragen, während der Niederschlag einen negativen Einfluss hatte. Yang et al. (2024) nutzten SHAP-Werte in Kombination mit einem Light-Gradient-Boosting-Machine-Modell, einem Gradienten basierten Ensemble-Algorithmus, um den Einfluss von Faktoren wie u.a. Nitrat, Sulfat und Magnesium auf die Grundwasserqualität in einer Region in Nordchina zu analysieren. Ebenso zeigten Scanlon et al. (2022), dass Umweltfaktoren wie jährliche Niederschlagsmenge, Landbedeckung und Bewässerungsflächen in Teilen der USA einen signifikanten Einfluss auf die Nitratkonzentration im Grundwasser haben.

Die Verwendung von SHAP-Werten in der vorliegenden Studie zielt darauf ab, die Beiträge der gemessenen chemischen Wasserparameter und der räumlichen Geodaten zur Nitratverteilung im Grundwasser zu quantifizieren. Dadurch können der Einfluss einzelner Prozesse wie Denitrifikation, Nitrifikation, Assimilation und andere Mechanismen besser verstanden werden. In dieser Studie sollen daher folgende Forschungsfragen beantwortet werden:

1. Wie stark beeinflusst das Denitrifikationspotenzial die Vorhersagegenauigkeit eines räumlichen Random-Forest-Modells für Nitratkonzentrationen im Grundwasser?
2. Welche spezifischen chemischen Wasserparameter und Informationen aus Geodaten haben den stärksten Einfluss auf die Modellgüte, insbesondere im Hinblick auf das Denitrifikationspotenzial?
3. Können Methoden der erklärbaren künstlichen Intelligenz (XAI) verwendet werden, um die Rolle des Denitrifikationspotenzials und anderer Einflussfaktoren im Modell hinreichend transparent zu machen?

Geografischer Überblick

Die Grundwasservorkommen in Deutschland sind stark regionalgeologisch geprägt und lassen sich verschiedenen hydrogeologischen Großräumen zuordnen. Diese unterscheiden sich im Wesentlichen durch ihre geologische Entstehungsgeschichte und ihren tektonischen Bau und werden u.a. durch den Schichtenaufbau, die Morphologie und die Grundwasservertiefungen definiert. Kleinere Teile räume wiederum ergeben sich aus einer stärkeren regionalen und naturräumlichen Gliederung und werden unter anderem durch Schichtgrenzen, Faziesgrenzen, Erosionsränder oder Störungen untergliedert (Ad-hoc-AG Hydrogeologie 2016). Die hydrogeologische Gliederung nach quantitativen Merkmalen wie Speicherkoeficient oder Transmissivität erlaubt eine grobe Einteilung der Grundwasserleiter hinsichtlich

ihrer Ergiebigkeit und somit auch hinsichtlich ihrer Bedeutung für die regionale Wasserversorgung. So reicht die Ergiebigkeit bei Porengrundwasserleitern von sehr hoch bis sehr gering, ebenso bei Kluft- und Karstgrundwasserleitern. Eine grobe Gliederung in diese Teile räume ist in der Karte der Hydrogeologie Deutschlands 1:1.000.000 (BGR und SGD 2013) enthalten und in Abb. 1a, b für die beiden Bundesländer Niedersachsen und Baden-Württemberg dargestellt.

Für Niedersachsen zeigt sich, dass ein Großteil der Landesfläche von ergiebigen Porengrundwasserleitern geprägt ist, die teilweise von Geringleitern durchzogen sind. Im Süden schließen sich Kluft- und Karstgrundwasserleiter bzw. Grundwassergeringleiter an.

In Baden-Württemberg erstreckt sich im Westen der große ergiebige Porengrundwasserleiter des Oberrheingrabens, an den sich der Schwarzwald mit seinen geringmächtigen Poren- und Kluftgrundwasserleitern anschließt. Weiter östlich erstrecken sich die ausgedehnten und ergiebigen Karst- und Kluftgrundwasserleiter der Schwäbischen Alb, die teilweise von Geringleitern überlagert werden. Weiter südöstlich geht das Gebiet in das Molassebecken über, das von untergeordneten Porengrundwasserleitern geprägt ist (Abb. 1a, b).

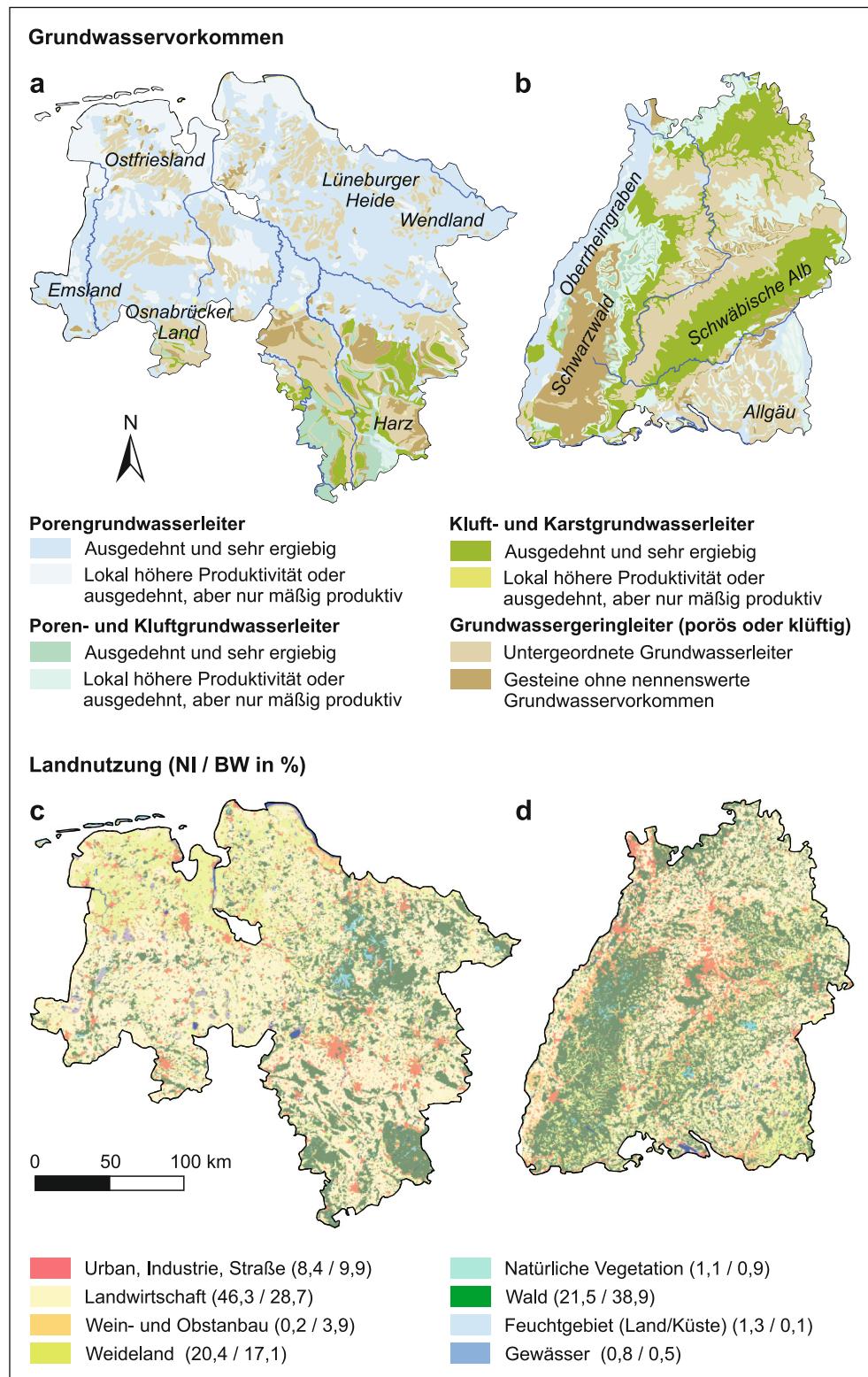
Die Landnutzung in den beiden Untersuchungsgebieten unterscheidet sich vor allem hinsichtlich des nicht bewässerten Ackerlandes und des Waldanteils (Abb. 1c, d). In Niedersachsen werden etwa 46 % der Fläche ackerbaulich genutzt, ergänzt durch etwa 20 % Weideland, das ebenfalls potenziell gedüngt wird. In Baden-Württemberg beträgt der Anteil des Ackerlandes etwa 28 %, während das Weideland rund 17 % der Fläche ausmacht. Baden-Württemberg weist mit 3,9 % zudem einen höheren Anteil an Wein- und Obstbau auf (Niedersachsen: 0,2 %), eine landwirtschaftliche Nutzung, die oft mit intensiver Düngung verbunden ist und lokal zur Nitratbelastung beitragen kann. Der deutlich höhere Waldanteil in Baden-Württemberg (38,9 % vs. 21,5 % in Niedersachsen) könnte hingegen als natürlicher Schutzfaktor wirken, da Waldflächen in der Regel nicht gedüngt werden und somit weniger zur Nitratbelastung des Grundwassers beitragen (Bechtel 2008). Insbesondere in Gebieten mit ergiebigen (durchlässigen) Grundwasserleitern und landwirtschaftlicher Nutzung ist daher das Vorkommen von Nitrat im Grundwasser wahrscheinlicher.

Nitrat im Grundwasser

Die Nitratbelastung des Grundwassers resultiert hauptsächlich aus dem Nitrateintrag durch landwirtschaftliche Aktivitäten, insbesondere aus Nährstoffüberschüssen auf bewirtschafteten Flächen. Die Art der landwirtschaftlichen Fläche und deren Bewirtschaftung beeinflussen den Nitrateintrag in das Grundwasser auf vielfältige Weise

Abb. 1 Übersicht über die Grundwasservorkommen nach Hohlraumtyp, Ausdehnung und Ergiebigkeit in **a** Niedersachsen und **b** Baden-Württemberg sowie die Landnutzungsverteilung in **c** Niedersachsen und **d** Baden-Württemberg

Fig. 1 Overview of groundwater resources by cavity type, extent, and yield in **a** Lower Saxony and **b** Baden-Württemberg, as well as land use distribution in **c** Lower Saxony and **d** Baden-Württemberg



(Bechtel 2008). Ackerland und Sonderkulturen wie Gemüse und Spargel weisen aufgrund der häufigen Düngung tendenziell höhere Einträge auf als Grünland, Obstplantagen und Streuobstwiesen, die geringere Stickstoffüberschüsse aufweisen. Rebland zeigt im Vergleich zu Gemüse- und Spargelkulturen niedrigere Eintragswerte, jedoch mit starker Variabilität je nach Standort und Bewirtschaftung. Auch die Art und Menge des Düngers (mineralisch oder organisch) sowie Düngzeitpunkt und -technik wirken sich auf den Nitrateintrag aus. Zudem beeinflussen landwirtschaftliche Maßnahmen wie Zwischenfruchtanbau, Bodenbearbeitung und Fruchtfolge die Stickstoffauswachung. Weiterhin spielt die Bodenart eine zentrale Rolle: Sandböden haben tendenziell höhere Austräge als lehmige Böden. Darüber hinaus beeinflussen Klima und Topografie die Stickstoffverlagerung und -verfügbarkeit, sodass die Kombination dieser Faktoren entscheidend für die Höhe des Nitrateintrags ins Grundwasser ist.

Neben der landwirtschaftlichen Düngung tragen teils auch nicht-landwirtschaftliche Quellen erheblich zur Nitratbelastung des Grundwassers bei (Wakida und Lerner 2005). Zu den diffusen Eintragsquellen zählen atmosphärische Deposition, geogene Freisetzung und die Infiltration belasteter Oberflächengewässer. Darüber hinaus spielen auch punktuelle Einträge, insbesondere in urbanen Gebieten, eine bedeutende Rolle. Dazu gehören Leckagen in Abwasserleitungen, Sickerwasser von Deponien, Stickstoffeinträge über Straßenabflüsse, dezentrale Abwasserentsorgung (z. B. Klärgruben), übermäßige Düngung in Gärten sowie kontaminierte Flächen. Infolge dieser Einträge können die Nitratkonzentrationen im Grundwasser städtischer Gebiete ähnlich hoch oder sogar höher sein als in landwirtschaftlich bewirtschafteten Regionen (Bechtel 2008; Shukla und Saxena 2020; Wakida und Lerner 2005). Nitrat gelangt überwiegend über das Sickerwasser in den Grundwasserleiter, kann jedoch auch durch direkte Infiltration aus Oberflächengewässern oder punktuelle Einträge in tiefere Bodenschichten vordringen (Kludt 2021).

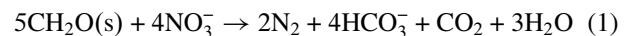
Denitrifikation als zentraler Abbauprozess

Ein zentraler Prozess für den Nitratabbau im Grundwasser ist die Denitrifikation, ein mikrobiell katalysierter Prozess, bei dem Nitrat (NO_3^-) unter anoxischen Bedingungen zu gasförmigem Stickstoff (N_2) oder Distickstoffoxid (N_2O) reduziert wird (Huno et al. 2018; Korom 1992; Soares 2000). Die Denitrifikation lässt sich in zwei Haupttypen unterteilen:

1. Heterotrophe Denitrifikation:

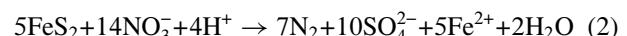
Hierbei nutzen Mikroorganismen organischen Kohlenstoff als Elektronendonator. Die Verfügbarkeit von leicht abbaubarem organischem Material, wie z. B. Huminstof-

fen, ist entscheidend für die Effizienz dieses Prozesses. In oberflächennahen Bodenschichten ist organischer Kohlenstoff meist ausreichend vorhanden, nimmt jedoch mit zunehmender Tiefe ab, was die Denitrifikationsrate beeinflussen kann:



2. Autotrophe Denitrifikation:

In diesem Prozess dienen anorganische Verbindungen, insbesondere Sulfide wie Pyrit (FeS_2), als Elektronendonator. Bei der Oxidation von Pyrit wird Nitrat reduziert, wobei Sulfat und zweiwertiges Eisen (Fe^{2+}) freigesetzt werden. Dieser Mechanismus ist besonders in tiefen Grundwasserleitern von Bedeutung, wo organischer Kohlenstoff limitiert ist. Die Verfügbarkeit von Pyrit ist jedoch endlich, sodass das Denitrifikationspotenzial im Laufe der Zeit erschöpft sein kann (Bergmann und Dietrich 2013).



Mehrere Schlüsselfaktoren beeinflussen die Denitrifikationsrate im Grundwasser:

- **Sauerstoffkonzentration:** Denitrifikation findet bevorzugt unter anoxischen Bedingungen statt, da denitrifizierende Mikroorganismen bei ausreichender Sauerstoffverfügbarkeit diesen als bevorzugten Elektronenakzeptor nutzen. Erst bei niedrigen Sauerstoffkonzentrationen ($< 2 \text{ mg O}_2/\text{l}$) oder unter anoxischen Bedingungen erfolgt die Reduktion von Nitrat (NO_3^-) zu gasförmigem Stickstoff (N_2) durch denitrifizierende Mikroorganismen (Bechtel 2008; Kludt 2021; Rissmann 2011; Rivett et al. 2008).
- **pH-Wert:** Der pH-Wert spielt eine zentrale Rolle, da er sowohl die Enzymaktivität als auch das Mikroklima der denitrifizierenden Mikroorganismen direkt beeinflusst. Die Denitrifikation erfolgt am effizientesten im pH-Bereich von 6 bis 8. Unter sauren Bedingungen ($\text{pH} < 6$) wird die Aktivität der N_2O -Reduktase gehemmt, was zu einer vermehrten Freisetzung von N_2O führt. Darüber hinaus bewirkt die Protonierung von Nitrit (NO_2^-) die Bildung von HNO_2 , was eine chemische Disproportionierung in Stickstoffmonoxid (NO) und Distickstoffmonoxid (N_2O) zur Folge hat. Bei alkalischen Bedingungen ($\text{pH} > 8$) kann die Denitrifikationsrate ebenfalls abnehmen, da die Aktivität der Nitritreduktasen sinkt und sich die Zusammensetzung der mikrobiellen Gemeinschaft verändert (Kludt 2021).
- **Temperatur:** Die Temperatur beeinflusst die Denitrifikationsrate maßgeblich, da sie sowohl die mikrobielle Stoffwechselaktivität als auch die Enzymkinetik steuert. Mit

steigender Temperatur nehmen enzymatische Reaktionen und mikrobielles Wachstum zu, wodurch die Denitrifikation beschleunigt wird. Unter 10 °C verlangsamt sich der Prozess deutlich, und unter 5 °C sind Enzymaktivität und mikrobieller Stoffwechsel so stark reduziert, dass die Denitrifikation nahezu zum Stillstand kommt (Brauer et al. 2015; Kludt 2021; Zhang et al. 2022).

- *Verfügbarkeit von Substraten als Elektronendonoren:* Die Art und Menge der Substrate bestimmen maßgeblich die Effizienz der Denitrifikation. Organische Substrate wie Huminstoffe werden von denitrifizierenden Mikroorganismen als Elektronendonoren genutzt und sind in oberflächennahen Bodenschichten meist in ausreichen- der Menge vorhanden. Mit zunehmender Tiefe nimmt die Verfügbarkeit von organischem Kohlenstoff jedoch ab, wodurch die Denitrifikationsrate begrenzt werden kann. Hohe Nitratauswaschung kann zudem den Substratver- brauch beschleunigen und das Denitrifikationspotenzial erschöpfen (Rohmann und Sontheimer 1985). In tiefen Grundwasserleitern, wo organischer Kohlenstoff limi- tiert ist, gewinnt die chemolithotrophe Denitrifikation an Bedeutung. Hier werden anorganische Substrate wie Pyrit (FeS_2) als Elektronendonoren genutzt, wobei Sulfat (SO_4^{2-}) und Fe^2 freigesetzt werden. Dieser Prozess ist besonders in sulfidreichen Aquiferen entscheidend für die Nitratreduktion (Juncher-Jørgensen et al. 2009).
- *Einfluss konkurrierender Elektronenakzeptoren:* Sobald der verfügbare Sauerstoff erschöpft ist, übernehmen an- dere Elektronenakzeptoren wie Eisen(III) (Fe^3), Man- gan(IV) (Mn^4) und Sulfat (SO_4^{2-}) die Rolle des Elek- tronenakzeptors, wodurch die Denitrifikation reduziert oder sogar verdrängt werden kann.
 - *Eisen (Fe^3/Fe^2):* Fe^3 wirkt als bevorzugter Elektronen- akzeptor und kann so die Denitrifikation hemmen. In reduzierenden Milieus wird Fe^3 zu Fe^2 reduziert, das wiederum mit Sulfid (H_2S) zur Pyritbildung (FeS_2) beitragen kann. Gleichzeitig kann Fe^2 mit Nitrat che- misch zu Ammonium (NH_4^+) reagieren (chemische Denitrifikation), wodurch kein gasförmiger Stickstoff freigesetzt wird (Di Capua et al. 2019; Ward et al. 2011).
 - *Mangan (Mn^4) und Sulfat:* Können als alternative Elektronenakzeptoren bevorzugt reduziert werden und damit die Denitrifikation unterdrücken. Sulfat kann in stark reduzierten Milieus durch sulfatreduzierende Bakterien (SRB) zu Sulfid (H_2S) reduziert werden, das wiederum mit Eisen zu Pyrit (FeS_2) reagiert (Di Capua et al. 2019).

Die Denitrifikation ist der zentrale Prozess für den Nitratabbau im Grundwasser, kann jedoch unter ungünstigen Bedingungen unvollständig ablaufen. Durch die Beeinflus- sung der oben genannten Faktoren kann verstärkt Distick- stoffoxid (N_2O) anstelle von molekularem Stickstoff (N_2)

freigesetzt werden – ein Treibhausgas mit einem 300-fach höheren Treibhauspotenzial als CO_2 (Butterbach-Bahl et al. 2013; Firestone et al. 1989). Dadurch hat die unvollständige Denitrifikation nicht nur Auswirkungen auf die Grundwas- serqualität, sondern auch auf die Atmosphäre.

Nitrat im Grundwasser stellt in Baden-Württemberg und Niedersachsen ein großes Umweltproblem dar. In beiden Bundesländern führt die intensive landwirtschaftliche Nutzung, insbesondere der Einsatz von Düngemitteln, regional zu erhöhten Nitratkonzentrationen im Grundwasser. In Baden-Württemberg waren es im Jahr 2021 etwa 8 % der Messstellen im Landesmessnetz, die den Grenzwert von 50 mg/l überschritten haben, was einen leichten Rückgang seit 2012 (10,0 %) bedeutet. Der Warnwert von 37,5 mg/l ging in diesem Zeitraum von 19,1 % auf 18,4 % der Mess- stellen zurück (LUBW 2021). In Niedersachsen waren es 21,5 % der Messstellen, die den Grenzwert, und 16,9 %, die den Warnwert überschritten haben.

Material und Methoden

Datenbasis

Die Nitratdaten sowie ergänzende chemische Parameter (Tab. 1) der Jahre 2016–2019 wurden vom Landesamt für Umwelt Baden-Württemberg (LUBW) (Abb. 2) und vom Niedersächsischen Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN) (Abb. 3) bereitge- stellt.

Messstellen, bei denen der Standardfehler des Mittel- werts (Standard Error of the Mean, SEM) aller Messungen im genannten Zeitraum mehr als 10 mg/l betrug, wurden aus der Analyse ausgeschlossen, da nur Datenpunkte mit einer hinreichend präzisen und stabilen Schätzung des Mittelwerts in die Analyse einbezogen werden sollten, ebenso wurden Ausreißer, die den 1,5-fachen Interquartilsabstand über dem 75. Perzentil überschritten, entfernt, um extrem hohe Werte, die potenziell auf Messfehler oder ungewöhnli- che lokale Einflüsse zurückzuführen sind, systematisch aus- zuschließen.

Messstellen, die gemäß Stammdaten nicht im obersten Grundwasserleiter liegen, blieben ebenfalls unberücksichtigt. Anschließend wurden für die verbleibenden Messstel- len die mittleren Nitratkonzentrationen berechnet. Diese Kriterien führten zu einer Auswahl von 1526 Messstellen in Baden-Württemberg und 1305 Messstellen in Niedersachsen. Zusätzlich wurden flächenhaft verfügbare Geodaten aus verschiedenen Quellen genutzt, die potenziell Einfluss auf die Nitratkonzentrationen im Grundwasser haben könnten (Tab. 2). Die ausgewählten Parameter umfassen geolo- gische, bodenkundliche und hydrogeologische Eigenschaf- ten sowie Informationen zur Landnutzung und Vegetations-

Tab. 1 Übersicht der verwendeten chemischen Parameter und deren jeweilige Rolle bzw. Einfluss auf den Nitratgehalt, die Denitrifikation und verwandte Nitratabbauprozesse.
Table 1 Overview of the chemical parameters used and their respective roles or impacts on nitrate content, denitrification, and related nitrate degradation processes.

Parametername	Elementsymbol/ Abkürzung	BW	NI	Indikator	Einfluss	Quelle
Ammonium	NH ₄	✓	✓	✓	Direkt	Hohe Konzentrationen weisen auf anaerobe Bedingungen oder unvollständige Nitifikation hin Wisotzky (2021)
Basenkapazität	KB _{8,2}	–	✓	Indirekt	Indikator für die Pufferkapazität des Wassers. Ein stabiler pH-Wert beginntigt die Denitrifikation	
Calcium	Ca ²⁺	✓	✓	Indirekt	Trägt zur Pufferung des pH-Werts bei, beeinflusst Ionengleichgewicht und Wechselwirkungen mit anderen Ionen	
Chlorid	Cl [–]	✓	✓	Indirekt	Hohe Werte könnten auf landwirtschaftliche Abflüsse hinweisen, die Nitrat enthalten	
Eisen	Fe ²⁺ /Fe ³⁺	✓	✓	Direkt	Fe ³ als Elektronenakzeptor anaerob; Fe ²⁺ mobil unter Reduktion, hemmt durch Elektronenkonkurrenz mit Nitrat	
Elektr. Leitfähigkeit	EL	✓	✓	Indirekt	Anzeiger für hohe Ionenkonzentration, einschließlich Nitrat	
Gel. org. Kohlenstoff	DOC	–	✓	Indirekt	Elektronendonator für heterotrophe Denitrifikation, beeinflusst Denitrifikationsrate durch Verfügbarkeit und Abbaubarkeit	
Kalium	K	✓	✓	Indirekt	Indikator für landwirtschaftliche Einträge, jedoch kein direkter Einfluss auf die Denitrifikation	
Magnesium	Mg ²⁺	✓	✓	Indirekt	Ähnlich Calcium, trägt zur Pufferung des pH-Werts bei, beeinflusst Ionen-gleichgewicht und enzymatische Aktivitäten	
Mangan	Mn ²⁺	✓	✓	Direkt	Unter anoxischen Bedingungen reduziert (Mn ⁴⁺ → Mn ²⁺), kann als Elektronen-akzeptor genutzt werden	
Natrium	Na	✓	✓	Indirekt	Hinweis auf anthropogene Einträge (z. B. Abwasser), beeinflusst ionische Stärke des Wassers	
Nitrit	NO ₂ [–]	✓	✓	Direkt	Zwischenprodukt der Denitrifikation; hohe Werte deuten auf aktive Denitrifika-tion oder unvollständige Nitritifikation hin	
Sauerstoff	O ₂	✓	✓	Direkt	O ₂ > 2 mg/l verhindert Denitrifikation durch bevorzugte Nutzung als Elektro-nenakzeptor. O ₂ < 2 mg/l ermöglicht Nitratreduktion	
Sulfat	SO ₄ ^{2–}	✓	✓	Indirekt	Hohe Werte können durch Elektronenkonkurrenz, Sulfidbildung oder toxische Effekte die Denitrifikation hemmen	
Säurekapazität	KS _{4,3}	–	✓	Indirekt	Indikator für die Fähigkeit des Wassers, Säuren zu neutralisieren. > Säurekap-a-zität trägt zur pH-Stabilität	
Temperatur	T (°C)	✓	✓	Direkt	Steuert Enzymaktivität; höhere Temperaturen erhöhen Denitrifikationsrate bis zu einem Optimum (~25–30 °C)	Kludt (2021)
pH-Wert	pH	✓	✓	Direkt	Denitrifikation optimal bei pH 7–8; niedriger pH hemmt Enzyme, hoher pH fördert Hydroxidbildung und verändert Mikroben	Benjamin (2015)

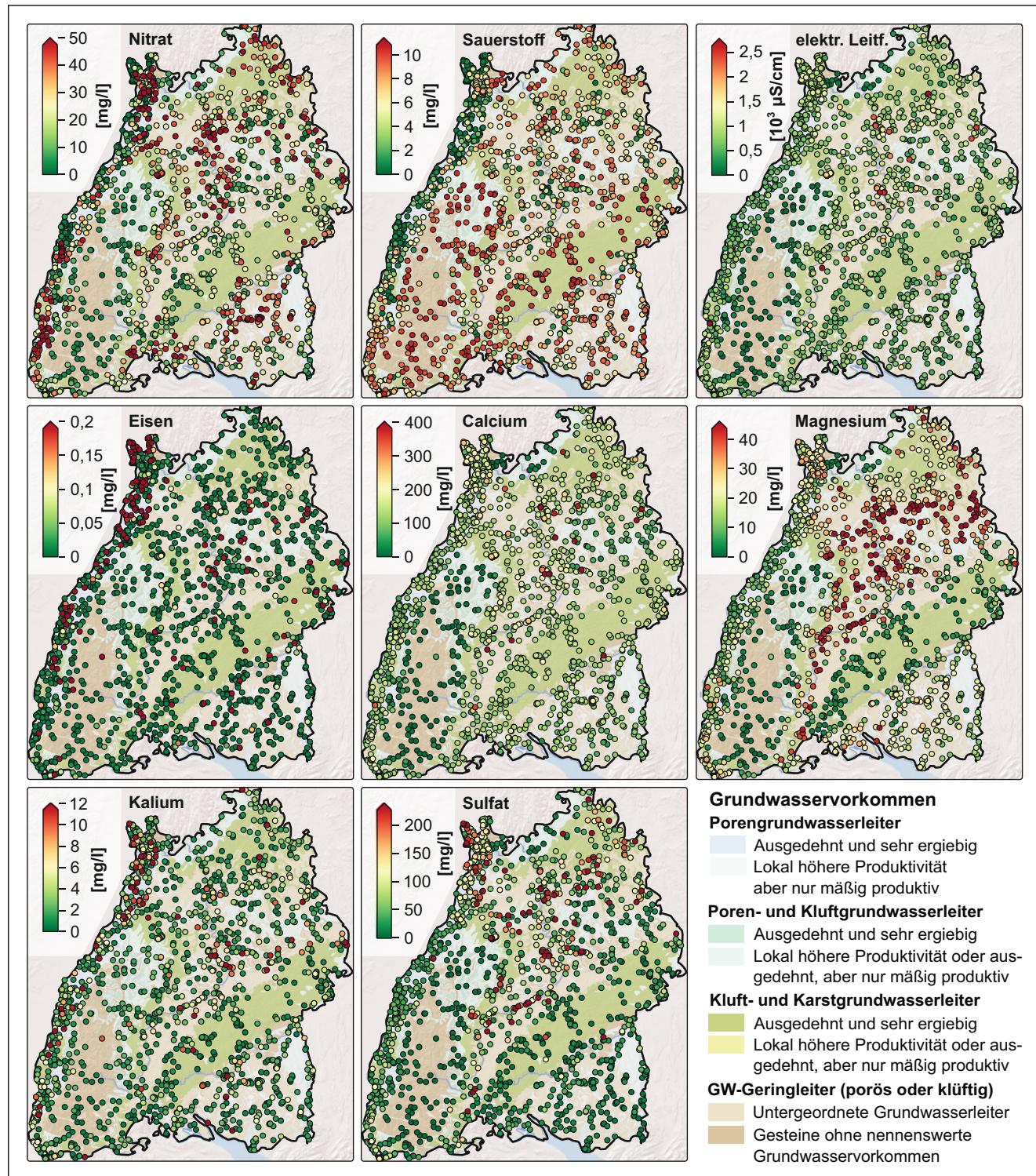


Abb. 2 Mittlere Konzentrationen ausgewählter chemischer Parameter im Grundwasser von Baden-Württemberg (2016–2019). Die Daten wurden von der Landesanstalt für Umwelt Baden-Württemberg (LUBW) bereitgestellt

Fig. 2 Average concentrations of selected chemical parameters in groundwater in Baden-Württemberg (2016–2019). The data were provided by the State Institute for the Environment Baden-Württemberg (LUBW)

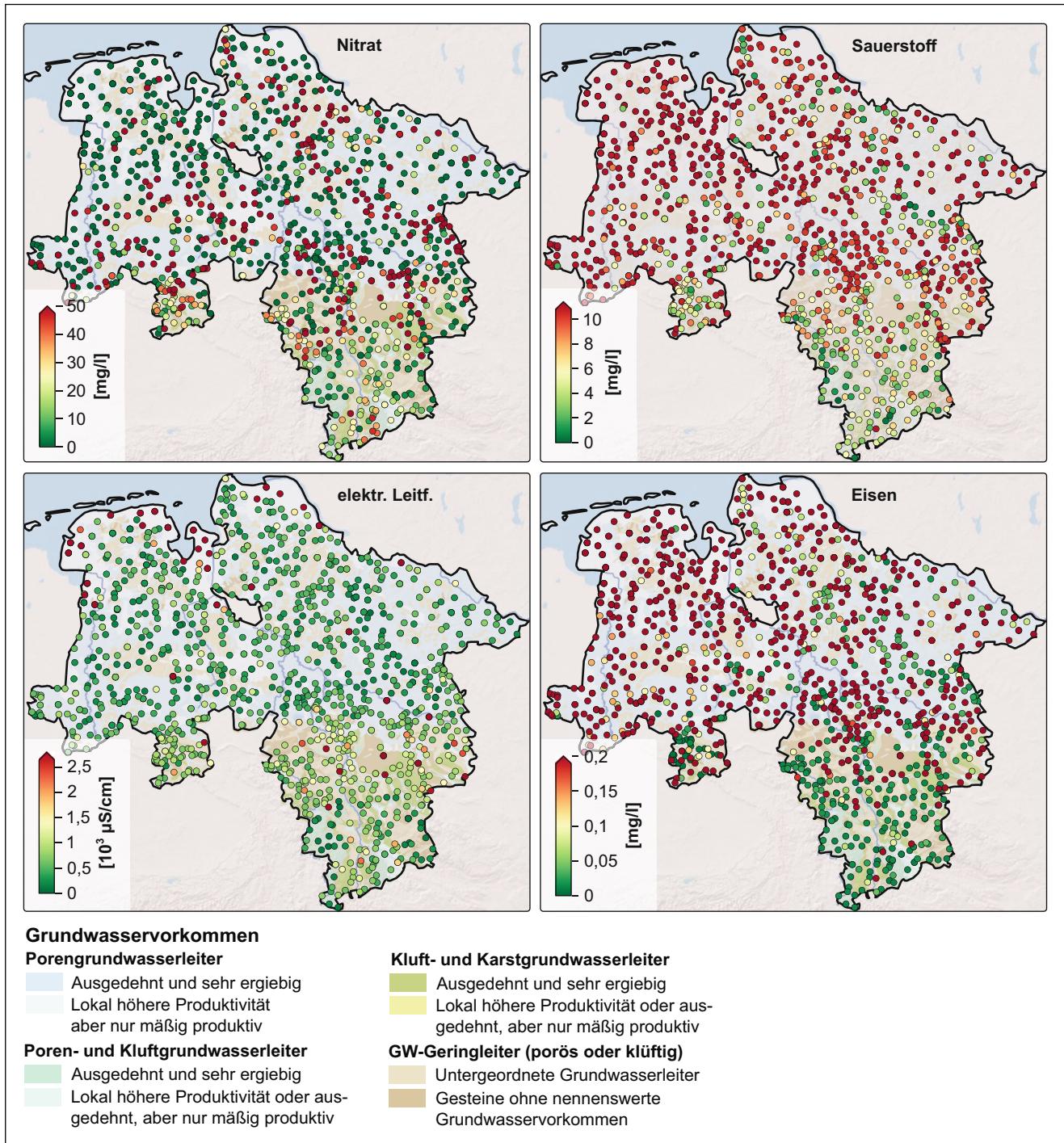


Abb. 3 Mittlere Konzentrationen ausgewählter chemischer Parameter im Grundwasser von Niedersachsen (2016–2019). Die Daten wurden von dem Niedersächsischen Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN) bereitgestellt

Fig. 3 Average concentrations of selected chemical parameters in groundwater in Lower Saxony (2016–2019). The data were provided by the Lower Saxony State Agency for Water Management, Coastal Protection, and Nature Conservation (NLWKN)

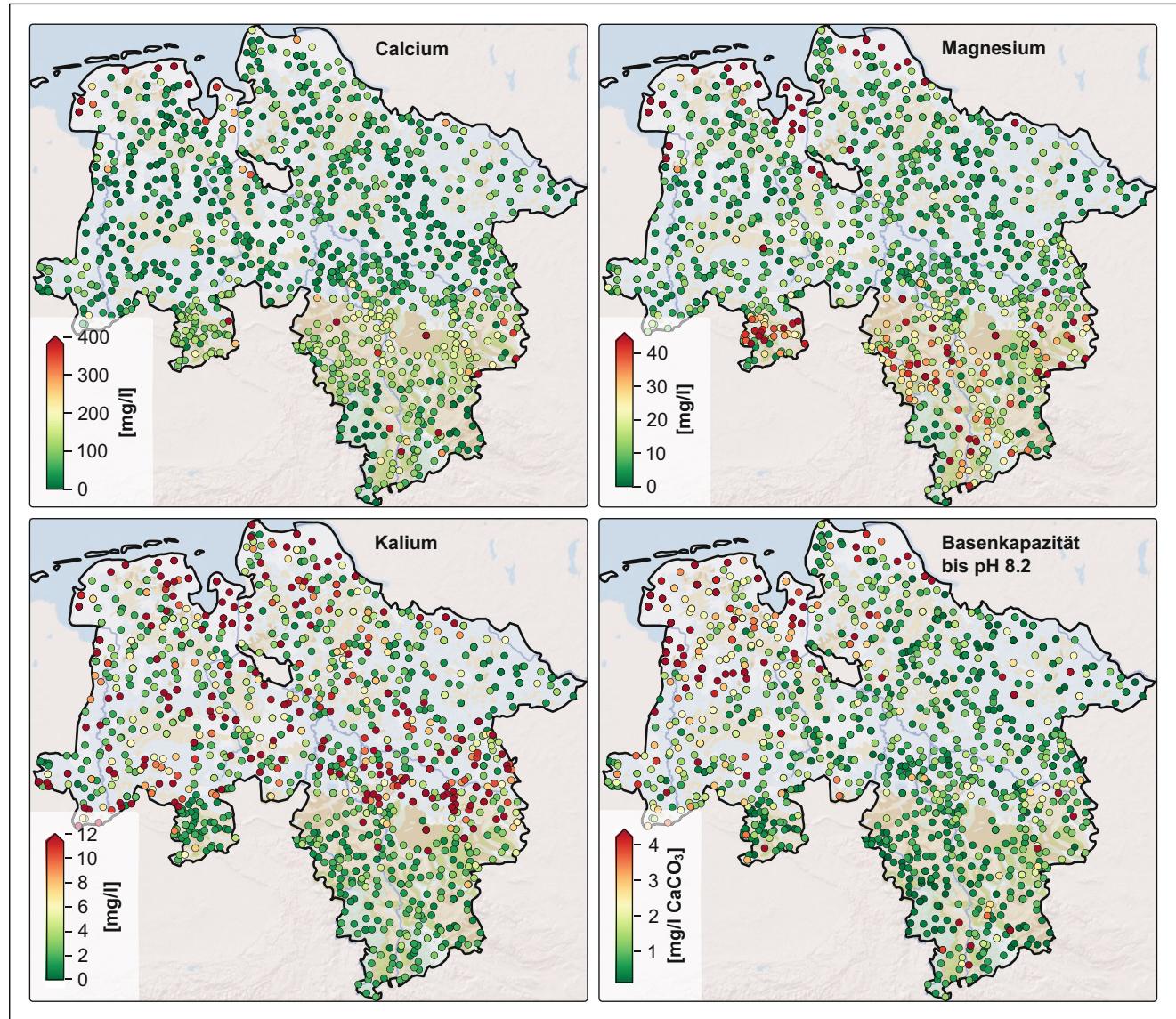


Abb. 3 (Fortsetzung)

Fig. 3 (Continuation)

bedeckung, die zusammen die Nitratdynamik im Grundwasser beeinflussen, indem sie Wasserbewegung, Bodenfiltrierkapazität und landwirtschaftliche Einträge steuern. Für jede Eingangsvariable, wie beispielsweise Bodentyp und Landnutzung, wurden die entsprechenden Rasterdaten an den Standorten der Messstellen extrahiert. Kategoriale Variablen wurden durch Mittelwert-Encoding in numerische Werte umgewandelt, um sie für die Modellierung zu nutzen.

Modellierung der Nitratkonzentration im Grundwasser mit Random Forest

Der Random-Forest-Algorithmus, entwickelt von Breiman (2001), ist eine überwachte Lernmethode für Klassifikations- oder Regressionsprobleme. Er erstellt ein Ensemble von Entscheidungsbäumen basierend auf einem Trainingsdatensatz. Der Algorithmus wählt dabei für jeden Baum zufällig eine Untergruppe von Merkmalen und Beobachtungen aus, um Vielfalt in den Entscheidungsbäumen sicherzustellen. Jeder Baum wird konstruiert, indem die Daten wiederholt anhand der ausgewählten Merkmale aufgeteilt werden, bis ein Abbruchkriterium erreicht ist. Die endgültige Vorhersage des RF wird dann als Durchschnitt der Vorhersagen

Tab. 2 Übersicht der verwendeten flächenhaft verfügbaren Geodaten zur Analyse von Einflussfaktoren auf den Nitratgehalt im Grundwasser (Typ K: kategorisch, N: numerisch).

Table 2 Overview of the spatially available geodata used for analyzing factors influencing nitrate content in groundwater (Type C: categorical, N: numerical).

Name	Parameter	Beschreibung	Typ	Quelle
HÜK250 _{he}	Hydrogeol. Einheiten	Hydrogeologische Karte von Deutschland 1:250.000 (HÜK250). Hydrogeologische Eigenschaften der oberen Grundwasserleiter	K	BGR und SGD (2019)
BÜK200bodtyp	BodenTyp	Bodenkarte von Deutschland 1:200.000 (BÜK200). Informationen über Bodentypen und die bodentypische Ausgangsgesteine in Deutschland	K	BGR und SGD (2018)
BÜK1000ba	Leitbodenassoziation	Bodenkarte von Deutschland 1:1.000.000 (BÜK1000)	K	BGR und SGD (2013)
CLC2018	Landnutzung	CORINE-Landbedeckung 2018, minimale Kartierungseinheit: 5 ha (CLC5), Deutschland	K	EEA (2019)
GK1000	Geologie	Geologische Karte von Deutschland 1:1.000.000 (GK1000)	K	BGR und SGD (2022)
HYRAUM _{TR}	Hydrolog. Teilraum	Hydrogeologische räumliche Struktur von Deutschland (HYRAUM), Regionen mit ähnlichen hydrogeologischen Eigenschaften	K	BGR und SGD (2015)
HUMUS1000	Organischer Gehalt	Organischer Gehalt im Oberboden (1:1.000.000) in Deutschland (HUMUS1000OB)	K	BGR (2007)
MUNDIALIS	Landnutzung	Deutschland 2019 – Klassifizierung der Landbedeckung basierend auf Sentinel-2-Daten	K	Riembauer et al. (2021)
NDVI	NDVI-Index	MODIS/Terra Vegetationsindex 16-Tage L3, Globales 250 m SIN-Gitter (250 m 16 Tage NDVI)	N	Didan (2019)
Preidl	Landnutzung	Landbedeckungsklassifizierungskarte der landwirtschaftlichen Fläche in Deutschland auf der Grundlage von Sentinel-2A-Daten aus dem Jahr 2016	K	Preidl et al. (2020)
THÜNEN2019	Kulturrarten	Landwirtschaftliche Flächennutzung (Vektorformat): Deutschland-weite Karten der Hauptnutzungsklassen auf Basis von Sentinel-1, Sentinel-2 und Landsat Daten (2017 bis 2021)	K	Tetteh et al. (2024)
SWR1000	Sickerwasserrate	Mittlere jährliche Sickerwasserrate aus dem Boden in Deutschland (SWR 1000)	N	BGR und SGD (2003)
Stream Distance	Entfernung Vorfluter	Entferungen zum nächstgelegenen Oberflächengewässer abgeleitet von der globalen Datenbank für Flussbreite und -tiefe Datenbank, berechnet nach Andreadis et al. (2013)	N	Andreadis et al. (2013)

aller einzelnen Entscheidungsbäume berechnet. Mathematisch lässt sich das RF-Modell folgendermaßen beschreiben:

$$\hat{y} = \frac{1}{B} \sum_{b=1}^B f_b(x; D_b)$$

Hierbei steht \hat{y} für die vorhergesagte Ausgabe für einen gegebenen Eingabevektor x . Die Vorhersage ergibt sich aus dem Durchschnitt der Vorhersagen von B individuellen Entscheidungsbäumen. Dabei repräsentiert $f_b(x; D_b)$ den b -ten Entscheidungsbäum und der Durchschnitt wird über alle Bäume gebildet. Jeder Entscheidungsbäum wird mithilfe einer Bootstrap-Stichprobe D_b erstellt, die aus zufällig ausgewählten Beobachtungen mit Wiederholung aus den ursprünglichen Trainingsdaten D erstellt wird.

Der RF-Algorithmus kombiniert dann die Vorhersagen der einzelnen Entscheidungsbäume, die jeweils auf ihrer eigenen Bootstrap-Stichprobe basieren, um eine zuverlässige und stabile Gesamtvorhersage \hat{y} für den Eingabevektor x zu erzeugen. Die Kombination der unabhängigen Entscheidungsbäume im Ensemble führt zu einer robusten Vorhersage mit verbesserter Genauigkeit und verringert das Risiko von Überanpassung.

Abschätzung des Denitrifikationspotenzials mit Explainable AI

In dieser Studie werden SHAP-Werte (Shapley Additive exPlanations) genutzt, um die Vorhersagen des Modells zu verstehen und zu interpretieren. SHAP-Werte sind ein Konzept aus der Spieltheorie, das in der künstlichen Intelligenz eingesetzt wird, um den Beitrag jedes Merkmals (Features) in einem Modell zur Vorhersage zu quantifizieren.

Um den Einfluss der verschiedenen chemischen Wasserparameter und der verfügbaren räumlichen Geodaten auf die Verteilung von Nitrat im Grundwasser zu bewerten, werden SHAP-Werte herangezogen. Jedes dieser Merkmale (Features) – also die einzelnen Eigenschaften der Daten, wie chemische Wasserparameter oder räumliche Geodaten – wird dabei als unabhängige Einheit betrachtet, die zur Vorhersage beiträgt. Die SHAP-Werte fassen den durchschnittlichen Beitrag jedes einzelnen Features über alle möglichen Kombinationen zusammen und zeigen auf, wie stark jedes Feature zum Gesamtergebnis beiträgt. Dadurch liefern sie eine individuelle Erklärung für jede Vorhersage des Modells und geben Auskunft darüber, ob das jeweilige

Feature die Modellvorhersage (also den vorhergesagten Nitratwert) positiv oder negativ beeinflusst hat.

$$SHAP_{i(f)} = \sum_{S \subseteq N \setminus \{i\}} \left[\frac{|S|! (|N| - |S| - 1)!}{|N|!} \right] \times [f(S \cup \{i\}) - f(S)]$$

Dabei ist $SHAP_{i(f)}$ der SHAP-Wert für das Merkmal i in Bezug auf die Funktion f , N die Menge aller Merkmale, S eine Teilmenge der Merkmale, und (S) der Vorhersagewert der Funktion f für die Merkmale in der Menge S .

Die Berechnung der SHAP-Werte beruht auf der Idee, dass der Beitrag eines Features zum Vorhersageergebnis auf seiner Interaktion mit anderen Features beruht. Durch die systematische Berücksichtigung aller möglichen Featuerekombinationen werden die individuellen Beiträge quantifiziert.

Die Verwendung der SHAP-Werte zielt darauf ab, die Beiträge der gemessenen Wasserparameter und der räumlichen Daten zur Verteilung von Nitraten im Grundwasser zu quantifizieren. Auf diese Weise wird ermittelt, welche Eingangsvariablen (Parameter) einen besonders großen Einfluss auf das Modellergebnis haben, um so den Einfluss einzelner Prozesse wie Denitrifikation, Nitrifikation, Assimilation besser zu verstehen. Die Abb. 4 zeigt exemplarisch einen SHAP-Plot, der die Einflüsse der Features der elektrischen Leitfähigkeit und der Sickerwasserrate auf die Modellvorhersage visualisiert.

Ergebnisse und Diskussion

Im Folgenden werden die Ergebnisse des Random-Forest-Modells zur Vorhersage der Nitratkonzentration im Grundwasser für die Bundesländer Baden-Württemberg und Niedersachsen präsentiert und diskutiert. Dabei wird das Denitrifikationspotenzial mithilfe von Explainable AI analysiert und interpretiert. Die Auswertung erfolgt getrennt nach Bundesländern sowie nach Modellen, die ausschließlich auf geodatenbasierten Parametern (Features) beruhen, und solchen, die zusätzlich in situ gemessene chemische Parameter einbeziehen.

Nitratmodellierung mit Geodaten

Die Ergebnisse der Korrelation der gemessenen und modellierten Nitratwerte für die ausschließlich geodatenbasierten Prädiktoren sind in Abb. 5a und 5b gezeigt. In Baden-Württemberg (Abb. 5a) ergibt der Vergleich zwischen gemessenen und modellierten Nitratkonzentrationen eine moderate Korrelation ($R^2=0,415$) und einen mittleren absoluten Fehler (MAE) von 9,76 mg/l. Das Modell tendiert dazu, niedrige Nitratkonzentrationen zu überschätzen und hohe

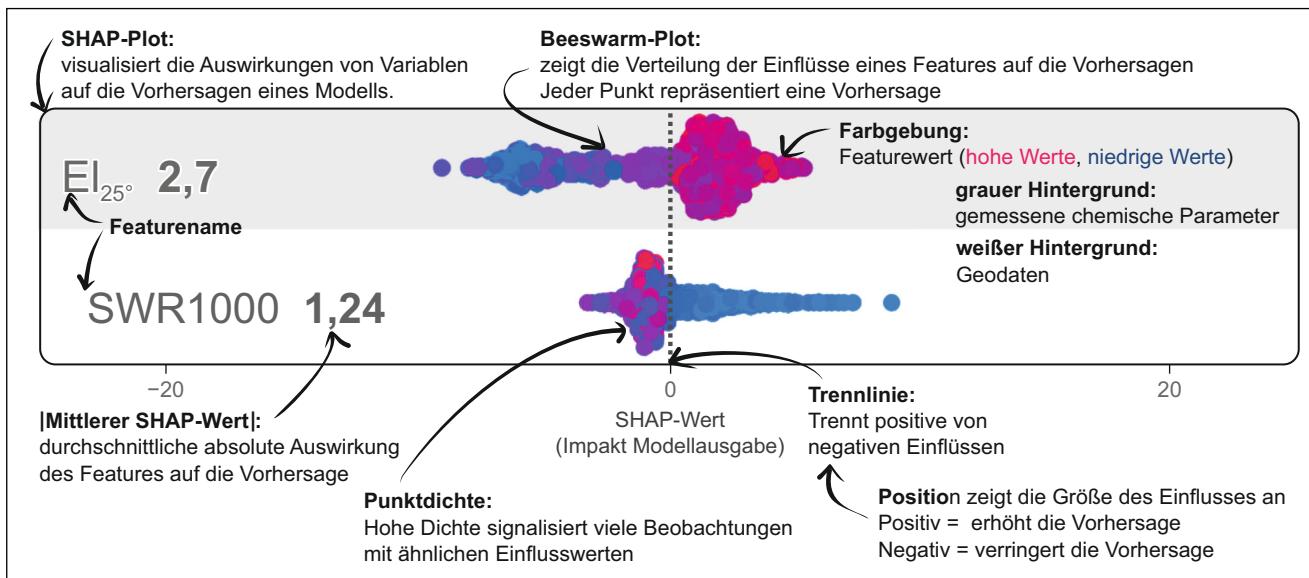


Abb. 4 SHAP-Plot zur Visualisierung der Einflüsse der Features „EI25°“ (elektrische Leitfähigkeit) und „SWR1000“ (Sickerwasserrate) auf die Modellvorhersage. Die x-Achse zeigt den SHAP-Wert, der die Stärke und Richtung des Einflusses auf die Vorhersage angibt (positiv = erhöhter Nitratwert, negativ = verringrigerter Nitratwert, der Betrag zeigt die Stärke des Einflusses). Jeder Punkt repräsentiert eine Vorhersage, farblich codiert nach dem Wert des Features (rot = hohe Werte, blau = niedrige Werte). Eine hohe Punktichte signalisiert häufig ähnliche Werte. Der graue Bereich repräsentiert wasserchemische Parameter, der weiße Bereich Geodaten. Das Beispiel zeigt, dass hohe elektrische Leitfähigkeitswerte (rot) zu hohen Nitratwerten führen, während niedrige Leitfähigkeitswerte (blau) mit niedrigeren Nitratwerten einhergehen. Für die Sickerwasserrate gilt das Gegenteil: Niedrige Sickerwasserraten führen zu höheren Nitratwerten, hohe Raten zu niedrigeren Nitratwerten

Fig. 4 SHAP plot visualizing the influence of the features “EI25°” (electrical conductivity) and “SWR1000” (seepage rate) on the model prediction. The x-axis represents the SHAP value, which indicates the strength and direction of the feature’s influence on the prediction (positive = increased nitrate level, negative = decreased nitrate level; the magnitude represents the strength of the influence). Each point represents a prediction, color-coded by the feature value (red = high values, blue = low values). A high density of points indicates frequently similar values. The grey area represents hydrochemical parameters, while the white area represents geodata. The example shows that high electrical conductivity values (red) lead to higher nitrate levels, whereas low conductivity values (blue) are associated with lower nitrate levels. For the seepage rate, the opposite is true: low seepage rates lead to higher nitrate levels, while high rates result in lower nitrate levels

Konzentration zu unterschätzen. Dies liegt an der rechtschiefen Verteilung der Messwerte: Die Häufung niedriger Nitratkonzentrationen erschwert es dem Modell, die selten auftretenden hohen Konzentrationen präzise zu lernen, da die Datenverteilung das Modell auf die häufiger auftretenden niedrigen Werte fokussiert.

Die zugehörigen SHAP-Werte (Abb. 6a) verdeutlichen die wichtigsten Einflussfaktoren auf die Modellvorhersagen für Baden-Württemberg. Dabei zeigen sich die Geologie (GK1000), die CORINE-Landbedeckung (CLC₂₀₁₈) sowie die Bodentypen (BÜK200_{bodtyp}) als die dominierenden Parameter.

Da diese kategorischen Prädiktoren durch Target Encoding in numerische Werte umgewandelt wurden, entsprechen höhere Prädiktorwerte jenen Kategorien, die in den Trainingsdaten im Durchschnitt mit höheren Nitratkonzentrationen assoziiert waren. Hohe SHAP-Werte (rote Punkte) spiegeln diesen Zusammenhang wider und zeigen, dass bestimmte Landnutzungs- und Bodenkategorien systematisch mit erhöhten Nitratkonzentrationen korrelieren.

So deutet beispielsweise die Landnutzungsklasse „CLC₂₀₁₈“ (CORINE-Landbedeckung) darauf hin, dass

landwirtschaftlich genutzte Flächen potenzielle Eintragsquellen für Nitrat darstellen. Das Modell erkennt diesen Zusammenhang aufgrund der systematischen Kopplung hoher Nitratwerte mit landwirtschaftlich dominierten Flächen in den Trainingsdaten.

Ähnlich zeigt die Klassifikation der Bodentypen (BÜK200_{bodtyp}) und Leitbodenassoziationen (BÜK1000_{iba}) eine positive Korrelation bestimmter Bodenkategorien mit erhöhten Nitratkonzentrationen, was möglicherweise auf die Durchlässigkeit oder das geringe Rückhaltevermögen dieser Böden zurückzuführen ist.

Andere geodatenbasierte Parameter wie die HÜK250_{he}, Mundialis (zusätzliche Landnutzungsinformationen), Thünen₂₀₁₉ (agrarwirtschaftliche Klassifikation) und die hydrogeologische Raumstruktur (Hyraum_{TR}) zeigen hingegen nur einen geringen Einfluss auf die Modellvorhersagen. Dies könnte auf die räumliche Auflösung oder unterschiedliche Klassifikationsmethoden dieser Parameter zurückzuführen sein. Die vergleichsweise hohe Modellgüte allein auf Basis der Geodaten legt nahe, dass Faktoren, die die Denitrifikation direkt beeinflussen, in Baden-Württemberg

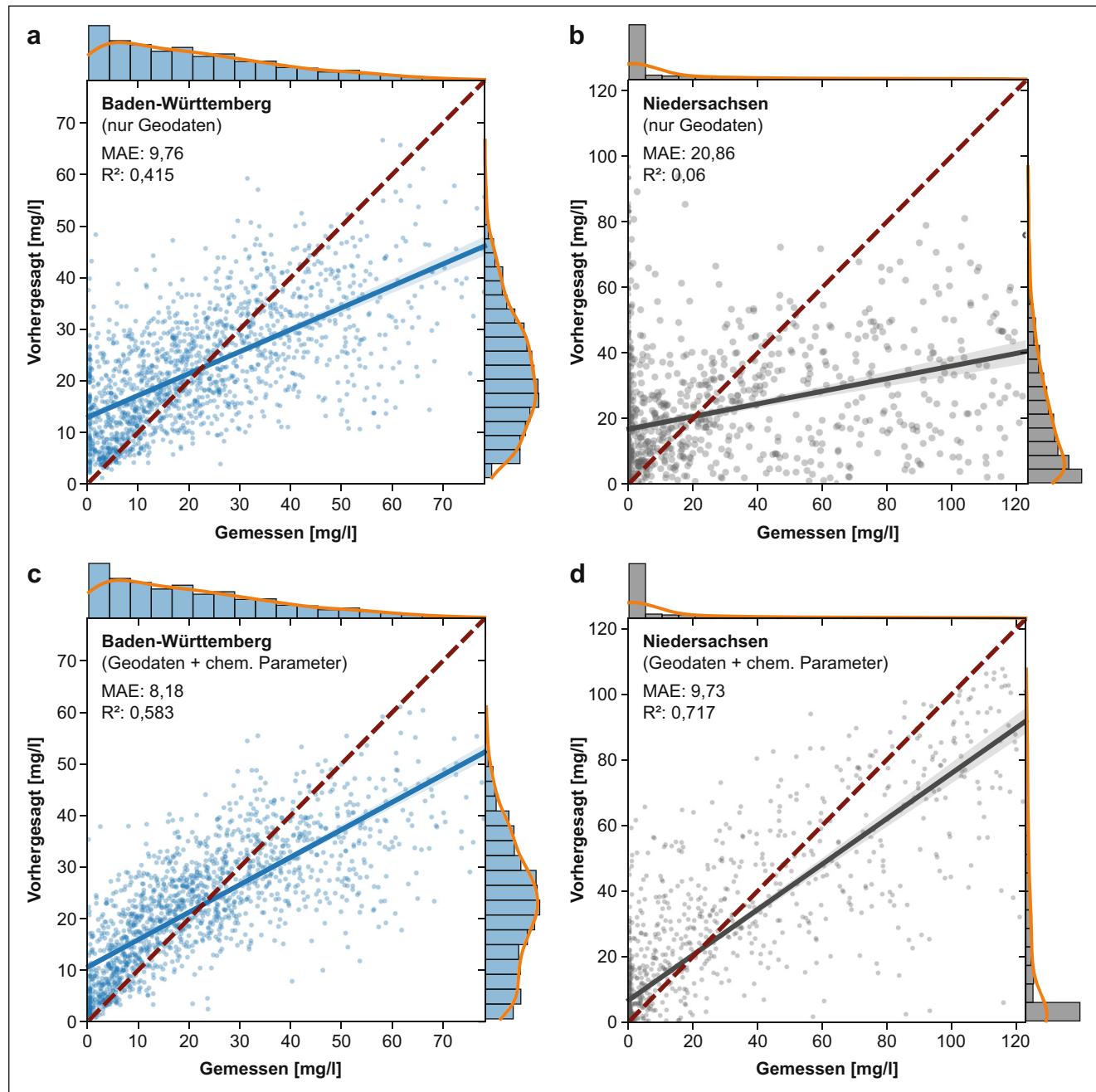


Abb. 5 Streudiagramme der gemessenen und vorhergesagten Nitratwerte für die beiden Untersuchungsgebiete Baden-Württemberg und Niedersachsen mit **a**, **b** Geodaten als Eingangsparameter und **c**, **d** Geodaten kombiniert mit chemischen Parametern

Fig. 5 Scatter plots of measured and predicted nitrate values for the two study areas, Baden-Württemberg and Lower Saxony, with **a**, **b** geodata as input parameters and **c**, **d** geodata combined with chemical parameters

eine eher untergeordnete Rolle spielen oder bereits implizit in den geodatenbasierten Prädiktoren enthalten sind.

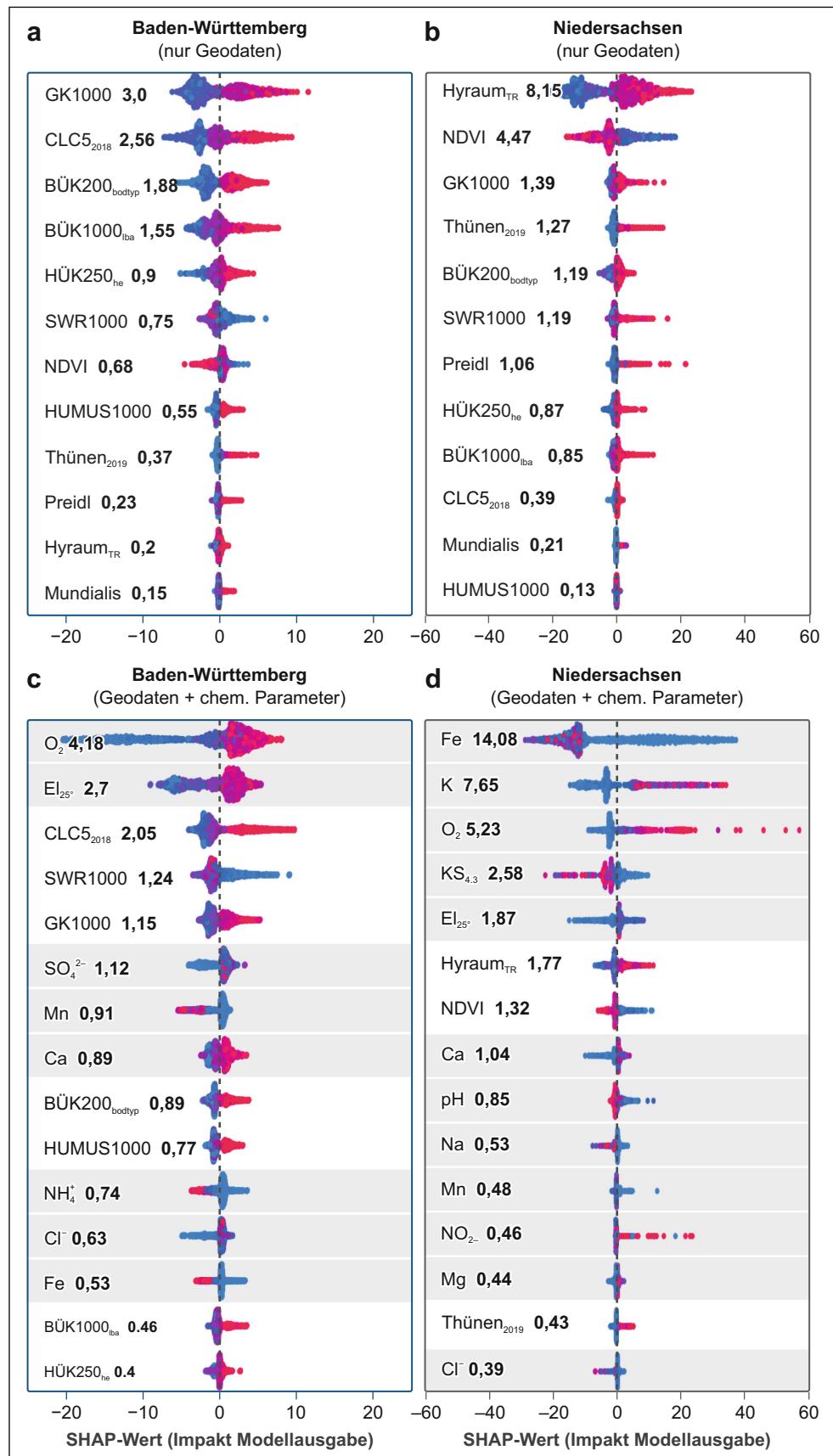
Das Modell zeigt für *Niedersachsen* ohne Berücksichtigung chemischer Parameter (Abb. 5b) eine sehr geringe Vorhersagegüte, mit einem mittleren absoluten Fehler (MAE) von 20,86 mg/l und einem R²-Wert von lediglich 0,06. Diese geringe Modellgüte verdeutlicht, dass die geodatenbasierten Parameter wie Bodentyp, Landnutzung und

Geologie allein nicht ausreichen, um die Nitratkonzentration im Grundwasser in Niedersachsen zuverlässig vorherzusagen. Diese niedrige Güte des Modells legt nahe, dass es keine kausalen Zusammenhänge erlernt, sondern stattdessen stochastische Muster in den Trainingsdaten aufgreift, die keine realen Einflussgrößen widerspiegeln.

Die aus dem Modell abgeleiteten SHAP-Werte (Abb. 6b) bilden folglich keine belastbaren Einflussfaktoren ab, son-

Abb. 6 SHAP-Werte der wichtigsten Prädiktoren für die Nitratmodellierung in Baden-Württemberg und Niedersachsen – jeweils ohne (a, b) und mit (c, d) chemischen Parametern. Rote Punkte zeigen einen positiven, blaue einen negativen Einfluss auf die Modellvorhersage

Fig. 6 SHAP values of key predictors for nitrate modelling in Baden-Württemberg and Lower Saxony—without (a, b) and with (c, d) chemical parameters. Red dots indicate positive, blue dots negative effects on the model prediction



dern reflektieren Artefakte der Modellstruktur, die aus der geringen Erklärungsleistung resultieren. Eine direkte Interpretation dieser SHAP-Werte könnte daher zu Fehlschlüssen führen, wenn diese als valide Zusammenhänge zwischen Prädiktoren und Nitratkonzentrationen verstanden werden. Dennoch zeigen die SHAP-Werte innerhalb dieser methodischen Einschränkungen die hydrogeologische Raumstruktur ($\text{Hyraum}_{\text{TR}}$) und den Vegetationsindex (NDVI) als vermeintlich wichtigste Parameter. Auffällig ist dabei, dass $\text{Hyraum}_{\text{TR}}$ sowohl positive als auch negative SHAP-Werte aufweist, was darauf hindeutet, dass dieser Parameter im Modell je nach räumlichem Kontext heterogene, jedoch nicht systematische Effekte auf die Nitratkonzentration zeigt. Weitere Parameter wie GK1000 (Geologie) und Thünen₂₀₁₉ (Landnutzung) weisen ebenfalls eine gewisse Modellrelevanz auf, was darauf hindeutet, dass spezifische geologische Eigenschaften und landwirtschaftliche Nutzungen in Niedersachsen als potenzielle Nitrateintragsquellen erkannt werden.

Die unterschiedlichen Modellergebnisse zeigen, dass geodatenbasierte Parameter den Einfluss der Denitrifikation nur begrenzt erfassen können. Während das Modell in Baden-Württemberg eine zufriedenstellende Vorhersagegüte erreicht, scheint in Niedersachsen die Denitrifikation eine größere Rolle zu spielen, sodass hier die Einbeziehung chemischer Parameter für eine belastbare Modellierung essenziell ist.

Nitratmodellierung mit Geodaten und chemischen Parametern

Die Integration chemischer Parameter, insbesondere solcher mit direktem Einfluss auf die Denitrifikation (z. B. Sauerstoff, Eisen, pH-Wert), verbessert die Modellgüte der Nitratkonzentrationsvorhersage in beiden Regionen, allerdings in unterschiedlichem Ausmaß. Während in Baden-Württemberg die Einbeziehung chemischer Parameter nur zu einer moderaten Verbesserung der Modellgüte führt (von $R^2=0,415$ auf 0,583, Abb. 5a, c), bewirkt sie in Niedersachsen eine signifikante Steigerung der Vorhersagequalität (von $R^2=0,06$ auf 0,717, Abb. 5b, d). Diese Ergebnisse bestätigen die Annahme, dass Denitrifikationsprozesse in Baden-Württemberg eine untergeordnete Rolle spielen, da die zusätzlichen chemischen Parameter dort nur eine begrenzte Verbesserung der Modellgüte bewirken.

Die leichte Verbesserung der Modellgüte ist dabei vor allem auf die Einbeziehung von Sauerstoff und elektrischer Leitfähigkeit zurückzuführen, die beide direkt mit den Redoxbedingungen im Grundwasser assoziiert sind. Ein hoher Sauerstoffkonzentration führt zu aeroben Bedingungen, unter denen Nitrat als terminaler Elektronenakzeptor nur verhindert genutzt wird. Da fakultativ anaerobe Mikroorganismen in Gegenwart von molekularem Sauerstoff (O_2) bevor-

zugt auf aerobe Atmung zurückgreifen, findet die mikrobiell vermittelte Denitrifikation nur in geringem Maße oder gar nicht statt. Die Korrelation zwischen hoher elektrischer Leitfähigkeit und erhöhten Nitratkonzentrationen kann auf vielfältige Ursachen zurückzuführen sein, darunter *landwirtschaftliche Einträge* (z. B. Düngemittel), *geogene Einflüsse* (Karbonatverwitterung, Sulfat aus Evaporiten) und *hydrogeochemische Prozesse* (Mobilisierung gelöster Ionen als *Elektronenakzeptoren*), die die Denitrifikation hemmen (Bechtel 2008). Die Geologie (GK1000) und die CORINE-Landbedeckung (CLC₂₀₁₈) bleiben weiterhin wichtige Parameter.

Im Vergleich zu Baden-Württemberg verdeutlichen die Ergebnisse für Niedersachsen, dass Denitrifikationsprozesse hier eine zentrale Rolle spielen, die durch rein geodatenbasierte Modelle nicht adäquat erfasst werden. Die SHAP-Werte (Abb. 6d) zeigen, dass Eisen ($\text{Fe}^{2+}/\text{Fe}^{3+}$) die bedeutendste Rolle spielt, gefolgt von Kalium (K^+), Sauerstoff(O_2) und der Säurekapazität ($\text{KS}_{4,3}$). Hohe Eisenkonzentrationen, vor allem in Form von Fe^2 , weisen auf reduktive Bedingungen hin, die die Denitrifikation begünstigen. Niedrige Eisenkonzentrationen korrelieren mit höheren Nitratkonzentrationen, was entweder auf unzureichend reduzierte Milieus hindeutet oder darauf, dass Fe^2 in einer früheren Redoxreaktion bereits oxidiert wurde. Eine weitere Erklärung könnte sein, dass Eisen überwiegend in der oxidierten, schwer löslichen Fe^{3+} -Form vorliegt, die nicht als Elektronendonator zur Verfügung steht. In oxischen oder nur schwach reduzierten Milieus bleibt Nitrat daher erhalten, da weder mikrobielle noch chemolithotrophe Reduktionsprozesse stattfinden (Bechtel 2008; Bergmann und Dietrich 2013).

Das Modell berechnet höhere Nitratkonzentrationen in sauerstoffreichen und niedrigere in sauerstoffarmen Bereichen, was die zentrale Rolle des Sauerstoffs als limitierenden Faktor für die Denitrifikation unterstreicht.

Die tendenziell niedrigeren Nitratwerte bei hoher Säurekapazität ($\text{KS}_{4,3}$) deuten darauf hin, dass eine stabile Pufferkapazität optimale pH-Bedingungen für die Denitrifikation schafft (Kölle 1989). Der starke Einfluss von Kalium lässt sich vermutlich auf dessen Einsatz in Kaliumnitrat (KNO_3)-Düngemitteln zurückführen und steht somit in direktem Zusammenhang mit den Nitratkonzentration im Grundwasser (Bechtel 2008).

Unsere Ergebnisse zeigen, dass die Verteilung der Denitrifikationsprozesse stark durch die hydrogeologischen Eigenschaften der Aquifere gesteuert wird, was mit den Studien von Knoll et al. (2020) und Wolters et al. (2022) übereinstimmt. In der Norddeutschen Tiefebene begünstigen mächtige, poröse Grundwasserleiter mit reduzierenden Bedingungen eine hohe Denitrifikationsrate, wodurch Nitrat effektiv abgebaut wird. Dies spiegelt sich in erhöhten Fe^{2+} - und Mn^{2+} -Konzentrationen wider und bestätigt die von

(Wolters et al. 2022) beobachteten höchsten Denitrifikationsraten Deutschlands. Im Harzvorland, einer Übergangszone mit variabler Gesteinsdurchlässigkeit, treten intermediaire Reduktionsraten auf, die teils hohe, teils niedrige Nitratwerte bedingen.

In Baden-Württemberg führen die heterogenen geologischen Strukturen zu stark variierenden Denitrifikationsraten. Im Oberrheingraben verhindern hohe Sauerstoffkonzentrationen in den durchlässigen Porengrundwasserleitern eine effiziente Denitrifikation, was mit den Erkenntnissen von Wolters et al. (2022) übereinstimmt. In der Schwäbischen Alb führt die schnelle Durchströmung in Karstquaternen zu persistierenden Nitratwerten, da keine stabilen reduzierenden Bedingungen entstehen können – ein Effekt, den Knoll et al. (2020) ebenfalls beschreiben. Der Schwarzwald, dominiert von Grundwassergleitern, zeigt zwar niedrige Nitratwerte, diese sind jedoch auf den fehlenden landwirtschaftlichen Eintrag und nicht auf Denitrifikation zurückzuführen.

Unsere Ergebnisse zeigen deutlich, dass hydrogeologische Rahmenbedingungen einen maßgeblichen Einfluss auf die Denitrifikation im Grundwasser haben. In den reduzierten Porenquaternen der Norddeutschen Tiefebene führen günstige anoxische Bedingungen zu hohen Abbauraten, während in den Karstquaternen der Schwäbischen Alb die schnellen Fließwege die Denitrifikation einschränken und in oxidierten Aquiferen des Oberrheingrabens nur eine geringe Nitratreduktion stattfindet. Diese Differenzen verdeutlichen, dass die Berücksichtigung hydrogeologischer Parameter in Nitratmodellierungen essenziell ist, um regionale Unterschiede im Nitratabbau realistisch abzubilden – auch wenn geodatenbasierte Ansätze allein nicht alle komplexen Prozesswechselwirkungen erfassen können. Zusätzlich zeigen die räumlichen Verteilungen der SHAP-Werte (siehe Anhang A1, A2 im Zusatzmaterial), dass insbesondere der Sauerstoffgehalt im Oberrheinaquifer den Nitratabbau begrenzt und die niedrige elektrische Leitfähigkeit der Kluftgrundwässer im Schwarzwald auf einen geringen Nitrateintrag hinweist. Damit wird die zentrale Rolle sowohl hydrogeologischer als auch landnutzungsbedingter Faktoren für die Nitratverteilung im Grundwasser deutlich.

Dabei ist zu berücksichtigen, dass großräumige Modellierungen zwar überregionale Muster erfassen, die kleinräumige Variabilität der Denitrifikation jedoch nur eingeschränkt abbilden. Lokale Schwankungen der Redoxbedingungen, bedingt durch Unterschiede in Sedimentzusammensetzung, Grundwasserfluss und organischer Substratverfügbarkeit (Tesoriero et al. 2000), erfordern eine feinräumige Betrachtung in heterogenen Aquiferen.

Zusammenfassung und Ausblick

Diese Studie beschreibt die Entwicklung von Random-Forest-Modellen zur Vorhersage der Nitratkonzentration im Grundwasser in Baden-Württemberg und Niedersachsen. Zunächst werden geodatenbasierte Prädiktoren wie Landnutzung, Geologie und Hydrogeologie für die Modellvorhersage verwendet. In einem zweiten Schritt auch chemische Parameter (z. B. Eisen, Sauerstoff), die eng mit Denitrifikationsprozessen verknüpft sind. Mithilfe von Explainable AI (SHAP-Werten) werden die relative Bedeutung sowie der spezifische Einfluss der einzelnen Prädiktoren auf die Modellvorhersagen analysiert und dargestellt.

Die Ergebnisse verdeutlichen, dass das höhere Denitrifikationspotenzial in Niedersachsen durch geodatenbasierte Modelle allein nicht adäquat erfasst werden kann. Ohne chemische Parameter wie Eisen und Kalium bleibt die Modellgüte dort gering ($R^2=0,06$) und es dominieren Prädiktoren wie die hydrogeologische Raumgliederung und der Vegetationsindex, die zwar hohe SHAP-Werte aufweisen, jedoch nur begrenzt zur Erklärung der Nitratkonzentration beitragen. Erst durch die Einbeziehung chemischer Parameter konnte die Modellgüte signifikant gesteigert werden (von $R^2=0,06$ auf 0,717), wodurch eine realistische Abbildung der räumlichen Nitratdynamik möglich wurde.

Diese Ergebnisse legen nahe, dass in Regionen mit hohem Denitrifikationspotenzial und stärkerer Reduktionsumgebung – wie in Niedersachsen – chemische Parameter unverzichtbar sind, um den Einfluss dieser Prozesse auf die Nitratkonzentration korrekt zu modellieren. In Regionen mit geringem Denitrifikationspotenzial wie Baden-Württemberg hingegen reichen geodatenbasierte Prädiktoren aus, da dort geologische und bodenbezogene Faktoren bereits ausreichend verlässliche Informationen zur regionalen Nitratdynamik liefern.

Die Kombination von Explainable AI und SHAP-Werten liefert wertvolle Einblicke in die räumliche Variabilität von Nitratkonzentrationen im Grundwasser und hilft, die komplexen Zusammenhänge zwischen hydrochemischen, geologischen und anthropogenen Einflussfaktoren zu entschlüsseln. Dies bietet eine fundierte Basis für gezielte Maßnahmen zur Nitratreduktion und unterstreicht die Notwendigkeit einer regelmäßigen Erfassung wasserchemischer Parameter im Grundwassermanagement.

Der verwendete Ansatz bietet Wasserbehörden und -ämtern entscheidende Vorteile. Mit öffentlich verfügbaren räumlichen Datensätzen zu Geologie, Hydrogeologie, Bodenart und Landnutzung lassen sich erste Rückschlüsse auf den Einfluss der Denitrifikation auf die Nitratkonzentration im Grundwasser ziehen. Durch die Ergänzung um chemische Parameter und die Verwendung von SHAP-Werten können zudem komplexe Zusammenhänge zwischen verschiedenen Einflussgrößen, die die Denitrifikation för-

dern oder hemmen, identifiziert werden. Random-Forest-Modelle überzeugen dabei durch ihre hohe Genauigkeit, Robustheit gegenüber Ausreißern und Skalierbarkeit, was sie für regionale und überregionale Analysen prädestiniert.

Herausforderungen ergeben sich aus dem Bedarf an ausreichend umfangreichen Datensätzen zu Nitratkonzentrationen und weiteren wasserchemischen Parametern, wenn eine kombinierte Modellierung angestrebt wird. Zudem besteht das Risiko, dass Nitratwerte unterschätzt werden, wenn das Modell vorwiegend mit niedrigen Nitratwerten trainiert wurde. Die erfolgreiche Anwendung und Interpretation der Modelle setzen daher hydrogeochemisches Fachwissen und eine kritische Bewertung der Modellvorhersagen voraus.

Gleichzeitig ist das verwendete Modell ein dynamisches, kontinuierlich kalibriertes und lernendes System. Die vorliegenden Ergebnisse sind das Produkt eines iterativen Optimierungsprozesses, der sich mit jeder neuen Datenerhebung und Validierung weiter verfeinern könnte. Zukünftige Datenerhebungen und gezielte Anpassungen würden dazu beitragen, die Vorhersagegenauigkeit weiter zu verbessern – sodass aus diesem Ansatz ein umfassenderes Systemverständnis abgeleitet werden könnte.

Zusatzmaterial online Zusätzliche Informationen sind in der Online-Version dieses Artikels (<https://doi.org/10.1007/s00767-025-00595-y>) enthalten.

Danksagung Wir danken allen Kolleginnen und Kollegen im Projekt für die wertvollen Diskussionen. Ein besonderer Dank gilt Thilo Fischer (Technologiezentrum Wasser) für seine fachkundige Unterstützung und wertvollen Anregungen, die maßgeblich zur Entwicklung dieses Papers beigetragen haben. Zudem bedanken wir uns bei den zwei anonymen Gutachtern für ihre konstruktiven Kommentare und Anregungen, die wesentlich zur Verbesserung des Manuskripts beigetragen haben.

Förderung Diese Studie ist ein Beitrag zum Projekt „Nitrat Monitoring 4.0 – Intelligente Systeme zur nachhaltigen Reduktion von Nitrat im Grundwasser (NiMo 4.0)“, gefördert vom Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (BMU) auf Grundlage eines Beschlusses des Deutschen Bundestages (Förderkennzeichen 67KI2048C).

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Ad-hoc-AG Hydrogeologie: Regionale Hydrogeologie von Deutschland – Die Grundwasserleiter: Verbreitung, Gesteine, Lagerungsverhältnisse, Schutz und Bedeutung (2016)
- Alajas, O.J., Concepcion, R., Vicerra, R.R., Bandala, A., Sybingco, E., Dadios, E., Cuello, J., Fonseca, V.: Indirect Prediction of Aquaponic Water Nitrate Concentration Using Hybrid Genetic Algorithm and Recurrent Neural Network. In: 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), S. 1–6. IEEE, Manila (2021)
- Andreadis, K.M., Schumann, G.J.-P., Pavelsky, T.: A simple global river bankfull width and depth database: data and analysis note. Water Resour. Res. **49**, 7164–7168 (2013). <https://doi.org/10.1002/wrcr.20440>
- Arslan, H., Ayyildiz Turan, N., Demir, Y., Güngör, A., Cemek, B.: Assessment of spatial and seasonal changes in groundwater nitrate pollution of agricultural lands through ordinary and indicator kriging techniques. Arch. Agron. Soil Sci. **63**, 907–917 (2017). <https://doi.org/10.1080/03650340.2016.1249472>
- Balacco, G., Fiorese, G.D., Alfio, M.R.: Assessment of groundwater nitrate pollution using the Indicator Kriging approach. Groundw. Sustain. Dev. **21**, 100920 (2023). <https://doi.org/10.1016/j.gsd.2023.100920>
- Bárdossy, A., Giese, H., Grimm-Strele, J., Barufke, K.-P.: SIMIK + – GIS implemented interpolation of groundwater parameters by means of geological and land-use data. Hydrol. Wasserbewirtsch. **47**, 13–20 (2003)
- Bechtel, A.: Abschätzung regionaler Nitrateinträge in das Grundwasser unter Verwendung von Nmin-Daten. Univ., Institutsverl. Siedlungswasserwirtschaft, Karlsruhe (2008)
- Benjamin, M. M.: Water Chemistry. 2. Auflage. Waveland Press, Long Grove (2015)
- Bergmann, D.A., Dietrich, D.P.: Konsequenzen nachlassenden Nitratbauenvermögens in Grundwasserleitern (2013)
- BoytRoura, M., Nolan, B.T., Menció, A., Mas-Pla, J.: Regression model for aquifer vulnerability assessment of nitrate pollution in the Osona region (NE Spain). J Hydrol **505**, 150–162 (2013). <https://doi.org/10.1016/j.jhydrol.2013.09.048>
- Brauer, N., Maynard, J.J., Dahlgren, R.A., O'Geen, A.T.: Fate of nitrate in seepage from a restored wetland receiving agricultural tailwater. Ecol Eng **81**, 207–217 (2015). <https://doi.org/10.1016/j.ecoleng.2015.04.003>
- Breiman, L.: Random forests. Mach Learn **45**, 5–32 (2001)
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR): Humusgehalt der Oberböden in Deutschland (HUMUS1000OB) (2007). https://www.bgr.bund.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche_Karten_Datenbanken/Themenkarten/HUMUS1000OB/humus1000ob_node.html
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR): Digitale Bodenkarte von Deutschland 1:1.000.000 (BUEK1000) (2013). <https://services.bgr.de/wms/boden/buek1000de>
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Staatliche Geologische Dienste Deutschlands (SGD): Mittlere jährliche Sickerwasserrate aus dem Boden in Deutschland (SWR 1000). Bundesanstalt für Geowissenschaften und Rohstoffe (BGR) und Staatliche Geologische Dienste Deutschlands (SGD), Hannover (2003)
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Staatliche Geologische Dienste Deutschlands (SGD): Hydrogeologie von Deutschland 1:1.000.000 (HY1000) (2013)
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Staatliche Geologische Dienste Deutschlands (SGD): Hydrogeologische Raumgliederung Deutschlands (HYRAUM) (2015). https://www.bgr.bund.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche_Karten_Datenbanken/Themenkarten/HYRAUM/hyraum_node.html

- www.bgr.bund.de/DE/Themen/Wasser/Projekte/abgeschlossen/Beratung/Hyraum/hyraum_grossraeume+raeume_karte.html
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Staatliche Geologische Dienste Deutschlands (SGD): Bodenübersichtskarte der Bundesrepublik Deutschland 1:200.000 (BÜK200) (2018). https://www.bgr.bund.de/DE/Themen/Boden/Informationsgrundlagen/Bodenkundliche_Karten_Datenbanken/BUEK200/buek200_node.html
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Staatliche Geologische Dienste Deutschlands (SGD): Hydrogeologische Übersichtskarte von Deutschland 1:250.000 (HÜK250) (2019). https://www.bgr.bund.de/DE/Themen/Wasser/Projekte/laufend/Beratung/Huek200/huek200_projektbeschreibung.html
- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Staatliche Geologische Dienste Deutschlands (SGD): Geologische Karte der Bundesrepublik Deutschland 1:1.000.000 (GK1000) (2022). https://www.bgr.bund.de/DE/Themen/Sammlungen-Grundlagen/GG_geol_Info/Karten/Deutschland/GK1000/gk1000_inhalt.html
- Bundesministerium für Ernährung und Landwirtschaft (BMEL): Nitratbericht 2020, Gemeinsamer Bericht der Bundesministerien für Umwelt, Naturschutz und nukleare Sicherheit sowie für Ernährung und Landwirtschaft. BMEL, Bonn (2020)
- Butterbach-Bahl, K., Baggs, E.M., Dannenmann, M., Kiese, R., Zechmeister-Boltenstern, S.: Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Philos. Trans. R. Soc. B Biol. Sci.* **368**, 20130122 (2013). <https://doi.org/10.1098/rstb.2013.0122>
- Dalla Libera, N., Fabbri, P., Mason, L., Piccinini, L., Pola, M.: Geostatistics as a tool to improve the natural background level definition: An application in groundwater. *Sci. Total Environ.* **598**, 330–340 (2017). <https://doi.org/10.1016/j.scitotenv.2017.04.018>
- Di Capua, F., Pirozzi, F., Lens, P.N.L., Esposito, G.: Electron donors for autotrophic denitrification. *Chem. Eng. J.* **362**, 922–937 (2019). <https://doi.org/10.1016/j.cej.2019.01.069>
- Didan, K.: MOD13C2 MODIS/Terra Vegetationsindizes Monatlich L3 Global 0.05° CMG V061. NASA EOSDIS Land Processes DAAC (Datensatz) (2019). <https://doi.org/10.5067/MODIS/MOD13C2.061>
- Didan, K.: MODIS/terra vegetation indices monthly L3 global 0.05° CMG V061 (V061), NASA EOSDIS land processes DAAC (Datensatz) (2021). <https://doi.org/10.5067/MODIS/MOD13C2.061>
- EEA: CORINE Land Cover 2018 (Raster 100 m), Europa, alle 6 Jahre – Version 2020_20u1, Mai 2020. Version 20.01. GeoTIFF. Kopenhagen: European Environment Agency. (2019). <https://doi.org/10.2909/960998C1-1870-4E82-8051-6485205EBBAC>
- Ehteshami, M., Farahani, N.D., Tavassoli, S.: Simulation of nitrate contamination in groundwater using artificial neural networks. *Model. Earth Syst. Environ.* **2**, 28 (2016). <https://doi.org/10.1007/s40808-016-0080-3>
- Firestone, M.K., Davidson, E.A., Andreae, M.O., Schimel, D.S.: Microbiological basis of NO and N₂O production and consumption in soil (1989)
- Ghadermazi, J., Sayyad, G., Mohammadi, J., Moezzi, A., Ahmadi, F., Schulin, R.: Spatial prediction of nitrate concentration in drinking water using pH as auxiliary co-kriging variable. *Procedia Environ. Sci.* **3**, 130–135 (2011). <https://doi.org/10.1016/j.proenv.2011.02.023>
- Gholami, V., Booij, M.J.: Use of machine learning and geographical information system to predict nitrate concentration in an unconfined aquifer in Iran. *J Clean Prod* **360**, 131847 (2022). <https://doi.org/10.1016/j.jclepro.2022.131847>
- Graf, H., Schäfer, W.: Simulation des Nitrattransports im Einzugsgebiet eines Wasserwerks. *Grundwasser* **7**(4), 233–242 (2002). <https://doi.org/10.1007/s007670200035>
- Hrnjica, B., Mehr, A.D., Jakupovic, E., Crnkic, A., Hasanagic, R.: Application of deep learning neural networks for nitrate prediction in the Klokoč River, Bosnia and Herzegovina. In: 2021 7th International Conference on Control, Instrumentation and Automation (ICCI), S. 1–6. IEEE, Tabriz (2021)
- Huno, S.K.M., Rene, E.R., Van Hullebusch, E.D., Annachhatre, A.P.: Nitrate removal from groundwater: a review of natural and engineered processes. *J. Water Supply Res. Technol.-Aqua.* **67**, 885–902 (2018). <https://doi.org/10.2166/aqua.2018.194>
- Juncher-Jørgensen, C., Jacobsen, O.S., Elberling, B., Aamand, J.: Microbial oxidation of pyrite coupled to nitrate reduction in anoxic groundwater sediment. *Environ. Sci. Technol.* **43**, 4851–4857 (2009). <https://doi.org/10.1021/es803417s>
- Karimanzira, D., Weis, J., Wunsch, A., Ritzau, L., Liesch, T., Ohmer, M.: Application of machine learning and deep neural networks for spatial prediction of groundwater nitrate concentration to improve land use management practices. *Front. Water.* **5**, 1193142 (2023). <https://doi.org/10.3389/fwra.2023.1193142>
- Karlovic, I., Posavec, K., Larva, O., Markovic, T.: Numerical groundwater flow and nitrate transport assessment in alluvial aquifer of Varaždin region. NW Croatia. *J. Hydrol. Reg. Stud.* **41**, 101084 (2022). <https://doi.org/10.1016/j.ejrh.2022.101084>
- Kludt, C.: Prozessorientierte Untersuchung zum Nitratabbauvermögen der Grundwasserkörper im Hessischen Ried. Technische Universität Darmstadt, Darmstadt (2021)
- Knoll, L., Breuer, L., Bach, M.: Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning. *Sci. Total Environ.* **668**, 1317–1327 (2019). <https://doi.org/10.1016/j.scitotenv.2019.03.045>
- Knoll, L., Häußermann, U., Breuer, L., Bach, M.: Spatial distribution of integrated nitrate reduction across the unsaturated zone and the groundwater body in Germany. *Water* **12**, 2456 (2020). <https://doi.org/10.3390/w12092456>
- Kölle, W.: Stickstoffverbindungen im Grund- und Rohwasser. Universität Hannover, Hannover (1989)
- Korom, S.F.: Natural denitrification in the saturated zone: A review. *Water Resour. Res.* **28**, 1657–1668 (1992). <https://doi.org/10.1029/92WR00252>
- LUBW: Grundwasser Überwachungsprogramm Ergebnisse 2021. Landesanstalt für Umwelt Baden-Württemberg Referat 42 – Grundwasse für Umwelt Baden-Württemberg Referat 42, Karlsruhe (2021)
- Lundberg, S., Lee, S.-I.: A unified approach to interpreting model predictions (2017). <http://arxiv.org/abs/1705.07874>
- Matiatos, I., Varouchakis, E.A., Papadopoulou, M.P.: Performance evaluation of multiple groundwater flow and nitrate mass transport numerical models. *Environ. Model. Assess.* **24**, 659–675 (2019). <https://doi.org/10.1007/s10666-019-9653-7>
- Mfumu Kihumba, A., Ndembo Longo, J., Vanclooster, M.: Modelling nitrate pollution pressure using a multivariate statistical approach: the case of Kinshasa groundwater body, Democratic Republic of Congo. *Hydrogeol J* **24**, 425–437 (2016). <https://doi.org/10.1007/s10040-015-1337-z>
- Nadiri, A.A., Moazamnia, M., Sadeghfam, S., Gnanachandrasamy, G., Venkatraman, S.: Formulating convolutional neural network for mapping total aquifer vulnerability to pollution. *Environ. Pollut.* **304**, 119208 (2022). <https://doi.org/10.1016/j.envpol.2022.119208>
- Nair, J.P., Vijaya, M.S.: TFT architecture and RNN variants for water quality prediction of Bharathapuzha river. *Int. J. Recent Innov. Trends Comput. Commun.* **11**(8), (2023)
- Nguyen, V.T., Dietrich, J.: Modification of the SWAT model to simulate regional groundwater flow using a multicell aquifer. *Hydrol. Process.* **32**, 939–953 (2018). <https://doi.org/10.1002/hyp.11466>
- Nolan, B.T., Gronberg, J.M., Faunt, C.C., Eberts, S.M., Belitz, K.: Modeling nitrate at domestic and public-supply well depths in the Central Valley, California. *Environ. Sci. Technol.* **48**, 5643–5651 (2014). <https://doi.org/10.1021/es405452q>

- Nolan, B.T., Fienen, M.N., Lorenz, D.L.: A statistical learning framework for groundwater nitrate models of the Central Valley, California, USA. *J Hydrol* **531**, 902–911 (2015). <https://doi.org/10.1016/j.jhydrol.2015.10.025>
- Ohlert, P.L., Bach, M., Breuer, L., Justus Liebig University Gießen: Verfahren zur Regionalisierung der Nitratkonzentration im Grundwasser – Supplement. Universitätsbibliothek Gießen (2022)
- Ouedraogo, I., Vanclooster, M.: A meta-analysis and statistical modeling of nitrates in groundwater at the African scale. *Hydrol. Earth Syst. Sci.* **20**, 2353–2381 (2016). <https://doi.org/10.5194/hess-20-2353-2016>
- Pebesma, E.J., De Kwaadsteniet, J.W.: Mapping groundwater quality in the Netherlands. *J Hydrol* **200**, 364–386 (1997). [https://doi.org/10.1016/S0022-1694\(97\)00027-9](https://doi.org/10.1016/S0022-1694(97)00027-9)
- Preidl, S., Lange, M., Doktor, D.: Landbedeckungsklassifikation der deutschen Agrarflächen basierend auf Sentinel-2A-Daten von 2016 (2020). <https://doi.org/10.1594/PANGAEA.910837>
- Ransom, K.M., Nolan, B.T., Stackelberg, P.E., Belitz, K., Fram, M.S.: Machine learning predictions of nitrate in groundwater used for drinking supply in the conterminous United States. *Sci. Total Environ.* **807**, 151065 (2022). <https://doi.org/10.1016/j.scitotenv.2021.151065>
- Riembauer, G., Weinmann, A., Xu, S., Eichfuss, S., Eberz, C., Neteiler, M.: Germany-wide Sentinel-2 based land cover classification and change detection for settlement and infrastructure monitoring. In: Proceedings of the 2021 Conference on Big Data from Space. Publications Office of the EU, (2021) <https://doi.org/10.2760/125905>
- Rissmann, C.: Regional mapping of groundwater denitrification potential and aquifer sensitivity (2011). <https://doi.org/10.13140/2.1.3412.5606>
- Rivett, M.O., Buss, S.R., Morgan, P., Smith, J.W.N., Bemment, C.D.: Nitrate attenuation in groundwater: A review of biogeochemical controlling processes. *Water Res.* **42**, 4215–4232 (2008). <https://doi.org/10.1016/j.watres.2008.07.020>
- Rohmann, U., Sontheimer, H.: Nitrat im Grundwasser: Ursachen, Bedeutung, Lösungswege. DVGW-Forschungsstelle am Engler-Bunte-Inst. d. Univ. Karlsruhe (TH) (1985)
- Rostami, A.A., Karimi, V., Khatibi, R., Pradhan, B.: An investigation on seasonal variations of groundwater nitrate by spatial modelling strategies at two levels by kriging and co-kriging models. *J. Environ. Manag.* **270**, 110843 (2020). <https://doi.org/10.1016/j.jenvman.2020.110843>
- Saha, G.K., Rahmani, F., Shen, C., Li, L., Cibin, R.: A deep learning-based novel approach to generate continuous daily stream nitrate concentration for nitrate data-sparse watersheds. *Sci. Total Environ.* **878**, 162930 (2023). <https://doi.org/10.1016/j.scitotenv.2023.162930>
- Sakizadeh, M., Zhang, C., Milewski, A.: Spatial distribution pattern and health risk of groundwater contamination by cadmium, manganese, lead and nitrate in groundwater of an arid area. *Environ. Geochem. Health* **46**, 80 (2024). <https://doi.org/10.1007/s10653-023-01845-9>
- Scanlon, B.R., Fakhreddine, S., Reedy, R.C., Yang, Q., Malito, J.G.: Drivers of spatiotemporal variability in drinking water quality in the United States. *Environ. Sci. Technol.* **56**, 12965–12974 (2022). <https://doi.org/10.1021/acs.est.1c08697>
- Shukla, S., Saxena, A.: Sources and leaching of nitrate contamination in groundwater. *Curr. Sci.* **118**, 883 (2020). <https://doi.org/10.18520/cs/v118/i6/883-891>
- Soares, M.I.M.: Biological denitrification of groundwater. In: Belkin, S. (Hrsg.) *Environmental Challenges*, S. 183–193. Springer Netherlands, Dordrecht (2000)
- Stylianoudaki, C., Trichakis, I., Karatzas, G.P.: Modeling groundwater nitrate contamination using artificial neural networks. *Water* **14**, 1173 (2022). <https://doi.org/10.3390/w14071173>
- Sullivan, T.P., Gao, Y., Reimann, T.: Nitrate transport in a karst aquifer: Numerical model development and source evaluation. *J Hydrol* **573**, 432–448 (2019). <https://doi.org/10.1016/j.jhydrol.2019.03.078>
- Tesoriero, A.J., Liebscher, H., Cox, S.E.: Mechanism and rate of denitrification in an agricultural watershed: Electron and mass balance along groundwater flow paths. *Water Resour. Res.* **36**, 1545–1559 (2000). <https://doi.org/10.1029/2000WR900035>
- Tetteh, G.O., Schwieder, M., Blickensdörfer, L., Gocht, A., Erasmi, S.: Agricultural land use (vector): National-scale crop type maps for Germany from combined time series of Sentinel-1, Sentinel-2 and Landsat data (2017–2021) (2024). <https://zenodo.org/doi/10.5281/zenodo.10619783>
- Wakida, F.T., Lerner, D.N.: Non-agricultural sources of groundwater nitrate: a review and case study. *Water Res.* **39**, 3–16 (2005). <https://doi.org/10.1016/j.watres.2004.07.026>
- Ward, B.B., Klotz, M.G., Arp, D.J.: *Nitrification*. ASM Press, Washington (2011)
- Wisotzky, F., Cremer, N., Lenk, S. (Hrsg.): *Angewandte Grundwasserchemie, Hydrogeologie und hydrogeochemische Modellierung – Grundlagen, Anwendungen und Problemlösungen*. 3. Auflage. Springer Spektrum, Berlin/Heidelberg (2021). <https://doi.org/10.1007/978-3-662-62755-6>
- WHO: Guidelines for drinking-water quality: fourth edition incorporating first addendum. World Health Organization, Geneva (2017)
- Wolters, T., Bach, T., Eisele, M., Eschenbach, W., Kunkel, R., McNamara, I., Well, R., Wendland, F.: The derivation of denitrification conditions in groundwater: Combined method approach and application for Germany. *Ecol Indic* **144**, 109564 (2022). <https://doi.org/10.1016/j.ecolind.2022.109564>
- Wriedt, G., De Vries, D., Eden, T., Federolf, C.: Regionalisierte Darstellung der Nitratbelastung im Grundwasser Niedersachsens. *Grundwasser - Zeitschrift der Fachsektion Hydrogeologie* **24**(2), 27–41 (2019). <https://doi.org/10.1007/s00767-019-00415-0>
- Yang, S., Luo, D., Tan, J., Li, S., Song, X., Xiong, R., Wang, J., Ma, C., Xiong, H.: Spatial mapping and prediction of groundwater quality using ensemble learning models and SHapley additive exPla-nations with spatial uncertainty analysis. *Water* **16**, 2375 (2024). <https://doi.org/10.3390/w16172375>
- Zaresefat, M., Hosseini, S., Ahrari Roudi, M.: Addressing nitrate contamination in groundwater: the importance of spatial and temporal understandings and interpolation methods. *Water* **15**, 4220 (2023). <https://doi.org/10.3390/w15244220>
- Zhang, D., Wang, P., Cui, R., Yang, H., Li, G., Chen, A., Wang, H.: Electrical conductivity and dissolved oxygen as predictors of nitrate concentrations in shallow groundwater in Erhai Lake region. *Sci. Total Environ.* **802**, 149879 (2022). <https://doi.org/10.1016/j.scitotenv.2021.149879>

Hinweis des Verlags Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.