Österreichweite Regionalisierung bodenhydraulischer Eigenschaften

Hanna Zeitfogel · Moritz Feigl · Karsten Schulz

Angenommen: 20. Dezember 2021 / Online publiziert: 25. Januar 2022
 © Der/die Autor(en) 2022

Zusammenfassung Angesichts der Klimawandelproblematik gewinnt auch die flächige Verfügbarkeit von bodenhydraulischen Informationen an Bedeutung. Diese Bodeninformationen bilden die Grundlage zur Modellierung hydrologischer Prozesse, speziell bei aktuellen Problemfeldern wie der Gefahrenausweisung von pluvialen Hochwasser. In Österreich gibt es derzeit kein Produkt, das die Bodendaten bundesweit flächig abbildet und gleichzeitig die hohe natürliche Variabilität der Bodeneigenschaften widerspiegelt. Ziel der vorliegenden Studie war es, auf Basis verfügbarer Daten relevante bodenhydraulische Parameter für die Gesamtfläche Österreichs abzuleiten und auch entsprechende Unsicherheiten anzugeben. Hierzu wurden zwei gängige Verfahren aus dem „Machine Learning“ (ML), XGBoost und FNN, getestet, um Zusammenhänge zwischen leicht messbaren bzw. flächig verfügbaren physiogeografischen sowie zusätzlichen Informationen aus Satellitenfernerkundung und den relevanten Bodenparametern zu entwickeln. Auf Basis der ML-Verfahren wurden die Bodenparameter Sand, Schluff, Ton und Humus flächig für ganz Österreich und für drei verschiedene Tiefenstufen auf einer Rasterbasis von $1 \times 1 \text{ km}^2$ abgeleitet. Die Ergebnisse stellen im direkten Vergleich mit dem derzeit einzig österreichweit flächig verfügbaren Bodeninformationssystem eine deutliche Verbesserung dar. Die Regionalisierung der gesättigten hydraulischen Leitfähigkeit (ks) wurde indirekt – auf Grundlage der regionalisierten Bodenparameter und mithilfe von existierenden Pedotransfer Funk-

tionen (PTFs) – und direkt – auf Basis vorhandener bodenhydraulischer Datensätze – getestet. Die Ableitung von ks ist nur mit großen Unsicherheiten möglich. Die erstellten Bodenkarten leisten einen wichtigen Beitrag zur Reduktion der vorhandenen Bodendatenlücken in Österreich und sollen als Grundlage für weitere Arbeiten zur Abschätzung der pluvialen Hochwassergefahr dienen.

Schlüsselwörter Gesättigte hydraulische Leitfähigkeit · Regionalisierung · Österreich · Machine Learning · Variabilität · Digital Soil Mapping

Prediction of soil hydraulic properties for the extent of Austria

Abstract In the light of a changing climate, access to high quality soil information is gaining immense significance. Soil parameters are fundamental for modeling hydrological processes, which is crucial for solving problems like identifying zones at risk of pluvial floods. Currently, there is no single data product available which covers the whole study area and still displays the variability of local soil observations. Thus, the challenge is the combination of soil data from different sources and resolutions and, at the same time, the preservation of the characteristically high spatial variability of soil properties. For the Austrian wide spatial prediction of soil parameters two machine learning (ML) models (XGBoost and FNN) were trained with all available soil data sources and environmental raster datasets. The soil parameters sand, silt, clay and humus were predicted at three different depth levels and a resolution of $1 \times 1 \text{ km}^2$ for the area of Austria. The resulting maps are able to largely reproduce the original data variability. Two approaches were tested for deriving the saturated hydraulic conductivity (ks): Firstly, ks was determined by applying existing

pedotransfer functions (PTFs) on the previously regionalized soil parameters. Secondly, ML models were directly trained with available soil hydraulic datasets to predict ks. The prediction of ks includes high levels of uncertainties. The derived soil maps help to reduce current gaps in soil data availability for Austria and act as a basis for identifying zones at risk of pluvial floods.

Keywords Saturated hydraulic conductivity · Regionalization · Austria · Machine learning · Variability · Digital soil mapping

1 Einleitung

Infiltration, also „Wassereintritt durch die Bodenoberfläche und die nachfolgende Wasserbewegung durch das Bodenprofil“ (Brutsaert 2005), gilt als einer der Schlüsselprozesse des Wasserkreislaufs (Vereecken et al. 2019). Ihre Rate stellt während Regenereignissen den bestimmenden Faktor für die Teilung des Niederschlagswassers an der Geländeoberfläche in Infiltration und Abflussbildung dar. Übersteigt die Regenintensität die aktuelle Infiltrationsrate, kommt es zur Bildung von Oberflächenabfluss (Horton 1933).

Die Bestimmung regionaler Infiltrationskapazitäten gewinnt angesichts aktueller Extremwetterereignisse, wie lokaler Starkregen, an Bedeutung. Solche Wetterextreme werden im Hinblick auf ein sich wandelndes Klima zukünftig deutlich häufiger und intensiver erwartet (IPCC 2013; BMNT 2017). Konvektive, raumzeitlich gesehen sehr kleinskalige Niederschlagsereignisse überschreiten aufgrund ihrer hohen Intensität rasch die Infiltrationsraten des Bodens. Die damit einhergehende Bildung von großen Mengen an Oberflächenabfluss führen häufig zu pluvialen Hochwasser (Rözer et al. 2021; ÖWAV-EP 2020). Solche Wetterextreme sind nur schwer zu prognostizieren und treten oft in Gebieten auf, die aus historischer Sicht nicht als hoch-

DI H. Zeitfogel (✉) · DI M. Feigl, Bsc. · Univ.-Prof. Dipl. Geoökol. Dr. K. Schulz
 Department für
 Wasser-Atmosphäre-Umwelt
 (WAU), Institut für Hydrologie
 und Wasserwirtschaft (HyWa),
 Universität für Bodenkultur Wien,
 Muthgasse 18, 1190 Wien, Österreich
hanna.zeitfogel@boku.ac.at

wassergefährdet gelten (Rözer et al. 2016, 2021). Um eine fundierte Risikoabschätzung vornehmen zu können, ist es notwendig, großflächig Infiltrationskapazitäten darzustellen bzw. die maßgebenden Einflussfaktoren auf die Infiltrationsrate regional abzuschätzen.

Die räumliche Verteilung der Infiltrationsrate hängt in erster Linie von den Bodeneigenschaften, der Landbedeckung und Nutzung sowie der lokalen Topografie ab (Vereecken et al. 2019; Mahapatra et al. 2020). Die zeitliche Variabilität wird durch die Bodenfeuchte bestimmt, bei zunehmender Bodenfeuchte nähert sich die Endinfiltrationsrate der gesättigten hydraulischen Leitfähigkeit (k_s) an (Fohrer et al. 2016). Demzufolge kann bei Kenntnis bodenhydraulischer Eigenschaften, insbesondere von k_s , in Kombination mit Landschaftsattributen eine Abschätzung der Infiltrationskapazitäten vorgenommen werden. Während Landnutzungs- und topografische Informationen für Österreich flächig über Geoportale wie Open Data Österreich (data.gv.at) oder INSPIRE Österreich (inspire.gv.at) abgerufen werden können, sind Bodeneigenschaften für Österreich nicht flächig verfügbar und die vorhandenen Daten oft nur eingeschränkt zugänglich.

Derzeit bestehen zahlreiche Bestrebungen zur Erfassung und Erforschung bodenhydraulischer Prozesse: Das Bundesministerium für Landwirtschaft, Regionen und Tourismus (BMLRT) betreibt zur Sammlung bodenhydraulischer Informationen ein Messnetz an 11 verschiedenen Standorten (BMLRT o.J.). Das Institut für Kulturtechnik und Bodenwasserhaushalt des Bundesamts für Wasserwirtschaft (BAW-IKT) führt eine Datenbank, in der für rd. 250 Bodenprofile horizontale Proben gezogen und bodenhydraulische Parameter im Labor analysiert wurden. Des Weiteren sammelt das Bundesforschungszentrum für Wald (BFW) in seiner Datenbank die Ergebnisse von rd. 400 Berechnungsversuchen aus unterschiedlichen Gebieten (Markart et al. 2017), im Rahmen derer Abflussreaktionen bei Regenereignissen unterschiedlicher Dauer und Intensität untersucht werden. Aus den Berechnungsexperimenten können folglich auch Rückschlüsse auf Infiltrationskapazitäten gewonnen werden. Die genannten Initiativen tragen zwar maßgeblich zum besseren Verständnis von Bodenwasserbewegungen in der vadosen Zone und Abflussreaktionen bei, allerdings ist die Anzahl der

Datenpunkte nicht ausreichend, um eine flächige Ableitung der Bodenhydraulik für Österreich daraus vorzunehmen.

Im Allgemeinen ist die direkte Bestimmung von bodenhydraulischen Eigenschaften für größere Flächen aufgrund ihrer hohen räumlichen Variabilität nicht praktikabel, und direkte Messungen sind zudem teuer und zeitintensiv (Vereecken et al. 2010). Vor diesem Hintergrund wurden Pedotransferfunktionen (PTFs) entwickelt, mithilfe derer bodenhydraulische Eigenschaften über leichter verfügbare Daten – z. B. aus Bodenaufnahmen – abgeleitet werden können (Van Looy et al. 2017). In Österreich starteten im Jahr 1958 systematische Bodenkartierungen der landwirtschaftlichen Nutzfläche mit dem Ziel, die flächenmäßige Verteilung von Bodenformen, also Flächeneinheiten, die denselben Typ und einen möglichst ähnlichen Standortcharakter aufweisen, zu erfassen. Jeder dieser Bodenformen ist mindestens ein Profil zugeordnet, für das zumindest die Textur, der pH-Wert, der Humus- und der Carbonatgehalt im Labor bestimmt werden (Schneider et al. 2001). Die Ergebnisse der Bodenkartierung sind über die digitale Bodenkarte eBOD (BFW 2019), die vom BFW betrieben wird, abrufbar. Ergänzend zur landwirtschaftlichen Bodenaufnahme stellt das Bodeninformationssystem BORIS des Umweltbundesamts (Umweltbundesamt GmbH 2021) Boden- und Waldbodenzustandsinventuren zur Verfügung.

Speziell für landwirtschaftliche Nutzflächen sind Österreichs Böden also hinsichtlich der wichtigsten Eingangsparameter für PTFs gut erfasst. Die für die Bestimmung der k_s wichtigen Bodeneigenschaften – wie z. B. die Korngrößenverteilung – müssen für alle anderen Landnutzungsarten allerdings erst abgeschätzt werden. Um diese – nicht nur in Österreich bestehende – Lücke zwischen Bedarf und Verfügbarkeit an flächigen Bodeninformationen zu schließen, gibt es sowohl in Österreich als auch überregional Projektaktivitäten: Im Rahmen zweier Projekte wurden für Niederösterreich (Eder et al. 2011; Sotier et al. 2017) und die Südoststeiermark (Klebinder et al. 2017) Bodeninformationen über einen geostatistischen (NÖ) bzw. einen konzeptionellen Ansatz (SO) regionalisiert und anschließend mithilfe von am IKT entwickelten PTFs hydrologische Bodenkennwerte abgeleitet. Auf

globaler Ebene wurde das Open Data Bodeninformationssystem „SoilGrids“ entwickelt, das seit 2014 unter ständiger Weiterentwicklung Bodendaten für verschiedene Tiefenstufen auf einer Rasterbasis von 250 m zur Verfügung stellt. Die Schätzung der räumlichen Verteilung von Bodeneigenschaften erfolgt in diesem Projekt mithilfe maschineller Lernverfahren (ML) basierend auf Bodenprofilen und Fernerkundungsdaten (Hengl et al. 2017). Seit dem neuesten Projektupdate steht auch eine Abschätzung der Vorhersage-Unsicherheiten zur Verfügung (Poggio et al. 2021).

Das SoilGrids-Projekt ist ein Beispiel für „digital soil mapping“ (DSM) bzw. „predictive soil mapping“ (PSM). In diesen Verfahren werden numerische und statistische Modelle über die Zusammenhänge von Umgebungsvariablen und Bodeneigenschaften entwickelt und in weiterer Folge zur Generierung von Karten eingesetzt (Scull et al. 2003). McBratney et al. (2003) schufen mit ihrer vielzitierten Arbeit einen generellen Rahmen, um Zusammenhänge zwischen Bodenattributen und ihrer Umgebung abzuleiten. Mit dem „scorpan“-Modell formulierten sie eine Zielfunktion für die Abschätzung von Bodenattributen, in der sieben raumbegrenzende Faktoren berücksichtigt werden:

$$S_a = f(s, c, o, r, p, a, n)$$

S_a	Bodenattribut
s	gemessene Bodeneigenschaften an einem Punkt
c	klimatische Gegebenheiten
o	Organismen, Vegetation, Fauna oder anthropogene Aktivitäten
r	Topografie und Landschaftsattribute
p	Ausgangsmaterial, Lithologie
a	Alter, Zeitfaktor
n	Lage, räumliche Positionierung

In den letzten Jahren ist der Einsatz von ML-Verfahren beim DSM (siehe Hengl et al. 2017) zur Vorhersage von Bodeneigenschaften, und ganz allgemein bei vielen anderen bodenwissenschaftlichen Fragestellungen, rasant gestiegen (Padarian et al. 2020). Ein weiteres populäres ML-Anwendungsgebiet ist die Generierung von PTFs (vgl. Araya and Ghezzehei 2019; Szabó et al. 2021). In erste Linie begründet sich die große Beliebtheit von ML-Verfahren in ihrer vielfach erprobten, guten Performance. Daneben spielen aber auch die größer werdende Anzahl an digital verfügbaren

Tab. 1 Zusammenstellung und Kurzbeschreibung der verfügbaren Bodendatensätze und der einbezogenen Parameter samt Quellenangabe

Datensatz	Beschreibung	Parameter ^a	Quelle (Datenverantwortliche Stelle)
eBOD	Digitale landwirtschaftliche Karte Österreichs (Bodenform-Polygone und 11.363 Profile)	Sand, Schluff, Ton, Humus, Tiefe, 12 Bodenartklassen	Bodenkarte.at (BFW)
BZI	Bodenzustandsinventuren der einzelnen Bundesländer (4520 Profile)	Sand, Schluff, Ton, Humus, Tiefe	Umweltbundesamt.at/boris
WBZI	Österreichweite und niederösterreichische Waldbodenzustandsinventur (547 Profile)	Sand, Schluff, Ton, Humus, Tiefe	Umweltbundesamt.at/boris
BAW-IKT Datenbank	Bodenhydraulische Auswertung von 604 Bodenproben (246 Profile)	Sand, Schluff, Ton, Humus, Tiefe, ks	Institut für Kulturtechnik und Bodenwasserhaushalt
Beregnungsexperimente	Abflussbeiwerte von 241 Beregnungsexperimenten an 203 Standorten	Abflussbeiwert	BFW Innsbruck
InfCapAT – Modul 2	Studie zur Untersuchung Bodenspez. Infiltrationskapazitäten an den Monitoring Stationen des BMLRT	ks, Tiefe	Brunetti et al. 2022, in diesem Heft
Hydrobod	50 × 50 m ² Raster für NÖ und drei Tiefenstufen	ks	Data.gv.at/hydrobod2 (Abteilung Wasserwirtschaft)
SoilGrids	250 × 250 m ² Raster für Österreich und sechs Tiefenstufen	Sand, Schluff, Ton, org. Kohlenstoff	Soilgrids.org/

^aEinheiten: Sand (%), Schluff (%), Ton (%), Humus (%), Tiefe (cm), ks (L/T), Abflussbeiwert (–), org. Kohlenstoff (g/kg)

Tab. 2 Datensatzgrundlage und Kurzbeschreibung der Prädiktorvariablen, samt abgeleiteter Eingangsgrößen und Quellenangaben

Datensatz	Beschreibung	Abgeleitete Parameter	Quelle (Datenverantwortliche Stelle)
DGM	Digitales Geländemodell (10 × 10 m ² Raster)	Geländeneigung, Exposition, TWI (jeweils Mittelwert, Median, Standardabw., Q5 und Q95)	Open Data Österreich (Geoland.at)
NDVI	Vegetationsindex (250 × 250 m ² Raster)	Jährlicher mittlerer NDVI, jährliche NDVI Spanne, mittlerer NDVI der Sommermonate, NDVI Spanne der Sommermonate	lvfi-info.boku.ac.at
AT_LB_2016	Sentinel-2 Bodenbedeckungskarte (10 × 10 m ² Raster)	13 Bodenbedeckungsklassen	Open Data Österreich (Umweltbundesamt)
SPARTACUS	Tägliche Zeitreihe der Lufttemperatur und Niederschlagssumme (1961–2019, 1 × 1 km ² Raster)	Mittlerer Jahresniederschlag Spanne des mittl. Jahresniederschlags mittlere Monatstemperaturen	ZAMG
Hydrogeologie	Hydrogeologische Karte des digitalen hydrologischen Atlas Österreichs (1:1.000.000)	12 lithologische Klassen	digHAO (BMLRT, BOKU)

Bodendaten und hochaufgelösten Co-Variablen (Topografie, Klimadaten etc.), die frei zugänglichen ML-Algorithmen, die kontinuierlich weiter entwickelt werden, sowie die steigenden Rechen-Ressourcen eine wichtige Rolle (Padarian et al. 2020; Wadoux et al. 2020). ML bezeichnet nicht-lineare Algorithmen, die in erster Linie auf große Datenmengen angewendet werden (Wadoux et al. 2020) und in der Lage sind, aus Erfahrungswerten zu lernen und Entscheidungen zu treffen. ML-Algorithmen haben auch bei der Anwendung auf Fragestellungen im Bereich der Umweltwissenschaften gezeigt, dass sie versteckte Muster in großen Datenmengen erkennen, wichtige Merkmale herausfiltern bzw. Verbindungen zwischen verschiedenen Variablen identifizieren können (Tahmasebi et al. 2020). Aus diesen Gründen wurden auch in die-

ser Studie ML-Verfahren zur Ableitung bodenhydraulischer Eigenschaften eingesetzt.

Zentrales Anliegen der vorliegenden Studie war es, Bodeninformationen – insbesondere die gesättigte hydraulische Leitfähigkeit – auf Basis verfügbarer Daten für die Gesamtfläche von Österreich abzuleiten. Zu diesem Zweck wurden alle in Österreich verfügbaren Bodendaten sowie frei zugänglichen, relevanten, physio-geografischen Parameter zusammengestellt und analysiert. Die Ableitung der Zielgröße erfolgte auf Grundlage des beschriebenen DSM-Ansatzes mithilfe zweier gängiger ML-Verfahren, die basierend auf Vorarbeiten von Feigl et al. (2021) angewendet und weiterentwickelt wurden.

2 Daten und Methoden

2.1 Datengrundlagen und Aufbereitung

Zur österreichweiten flächenhaften Ableitung der Bodeneigenschaften wurden alle zur Verfügung gestellten Bodendaten sowie sämtliche frei verfügbare Prädiktorvariablen herangezogen. Die größte Datenquelle an lokalen Bodendaten stellt die digitale Bodenkarte eBOD dar, in der die Ergebnisse der Bodenkartierung von Österreichs landwirtschaftlichen Flächen abgebildet sind. Neben dem eBOD-Datensatz wurden alle verfügbaren Bodenzustandsinventuren (BZI) der einzelnen Bundesländer und Waldbodenzustandsinventuren (WBZI) in die Studie eingebunden. Für die Ableitung und Validierung bodenhydraulischer Parameter wurden die bodenhydraulische Daten-

Tab. 3 Wertebereiche für die XGBoost-Hyperparametersuche (Bezeichnung der Parameter gemäß R Paket „caret“)

Parameter	Beschreibung	Wertebereich
eta	Lernrate	0,001–0,3
nrounds	Anzahl Iterationen	300–3000
max depth	Maximale Tiefe eines Baumes	3–20
subsample	Anteil d. Teilstichprobe zur Generierung eines Baums	0,5–1
colsample by tree	Anteil Variablen zur Generierung eines Baums	0,5–1
gamma	Regularisierungsparameter	0–5
min child weight	Min. Summe der Gewichte der untergeordneten Knoten	1–15

Tab. 4 Wertebereiche für die FNN-Hyperparametersuche

Parameter	Beschreibung	Wert
Layers	Layer	1–5
Units	Neuronen	5–200
Dropout	Dropout Rate	0–0,2
Batchsize	Anzahl an Stichproben, die durchlaufen werden müssen, bevor modellinterne Parameter upgedated werden	5–150
Epochs	Anzahl der Runden, in denen FNN das gesamte Trainingsset durchläuft	100
Early Stop Patience	Training stoppt, wenn diese Anzahl an Runden ohne Verbesserung der Gütemaße durchlaufen ist	5
Activation	Aktivierungsfunktion	SELU bzw. Softmax

bank des BAW-IKT, die Beregnungsexperimente des BFW Innsbruck, eine aktuelle Studie zum Bodenwassermontoring-Programm des Bundesministeriums für Landwirtschaft, Regionen und Tourismus (Brunetti et al. 2022, in diesem Heft) sowie die Hydrobod-NÖ-Ergebnisse analysiert. Als Benchmark wurden die Rasterbodenkarten des SoilGrids-Projekts herangezogen. Eine Zusammenfassung der berücksichtigten Bodendatensätze ist in Tab. 1 ersichtlich.

Nach dem scorpan-Modell (McBratney et al. 2003) sollen über sieben Faktoren Zusammenhänge zur Zielgröße (z.B. Textur oder ks) hergestellt werden. Im Rahmen dieser Studie wurden basierend auf verfügbaren Informationen zu Topografie, Vegetation, Landbedeckung, Geologie und Klimadaten Prädiktorvariablen abgeleitet. Die Ausgangsdaten samt abgeleiteter Prädiktorvariablen sind in Tab. 2 zusammengefasst.

Für alle raumbezogenen Datensätze wurde die gemeinsame Kartenprojektion ETRS89/Austria Lambert (EPSG: 3416) gewählt. Für alle flächig verfügbaren Daten wurde auf Basis der SPARATCUS-Raster die räumliche Diskretisierung von $1 \times 1 \text{ km}^2$ gewählt. Diese Auflösung wurde in erster Linie auf Grundlage der verfügbaren Auflösung der Prädiktorvariablen gewählt, aber auch um den Rechenaufwand für die

Datenaufbereitung und das Modelltraining selbst in einem vertretbaren Rahmen zu halten. Bei der Ableitung verschiedener Prädiktorvariablen auf Basis des vergleichsweise sehr fein aufgelösten digitalen Geländemodells ($10 \times 10 \text{ m}^2$) wurde im Zuge der Aggregation auf die Zielauflösung von $1 \times 1 \text{ km}^2$ neben dem Mittelwert auch die statistischen Kennwerte Median, Standardabweichung, 5%-Quantil und 95%-Quantil berechnet, um die lokale topografische Variabilität auch im großskaligen Bereich berücksichtigen zu können.

Alle Profildaten der eBOD-, BZI- und WBZI-Datensätze wurden von fehlenden oder unrealistischen (z.B. negativen) Werten bereinigt. Die Sand-, Schluff- und Tonanteile wurden auf ihre Konsistenz (Summe der drei Anteile muss 100 % ergeben) überprüft und gegebenenfalls korrigiert. Alle relevanten Parameter wurden für sämtliche Bodenprofile auf drei einheitliche Tiefenstufen, nämlich 0–20 cm, 20–50 cm und 50–100 cm, über gewichtete Mittelwertbildung nach der Trapezregel aggregiert. Bodeninformationen für Tiefen > 100 cm wurden nicht berücksichtigt, da sie für schnelle Abflussprozesse, wenn überhaupt, nur eine geringe Rolle spielen. Die gewählten Tiefenstufen fungieren in den ML-Modellen auch als zusätzliche Prädiktorvariable, um nicht nur eine horizontale, sondern auch

eine vertikale Verteilung der Bodeneigenschaften abschätzen zu können. Zur Bereinigung von Multikollinearität der Prädiktorvariablen und generell zur Vereinfachung der großen Eingangsdatenmenge wurde für die Anwendung eines der ML-Verfahren, XGBoost (siehe Abschn. 2.2.2), im Vorfeld eine Hauptkomponentenanalyse (PCA) durchgeführt. Im Gegensatz zum XGBoost-Algorithmus können neuronale Netze gut mit Datenredundanz umgehen, und eine PCA der Eingangsdaten bringt keine Verbesserung der Ergebnisse. Alle beschriebenen abgeleiteten Variablen sind tabellarisch in Tab. 2 zusammengefasst.

2.2 Methodik

2.2.1 Methodische Ansätze

In der vorliegenden Studie wurden zwei unterschiedliche Ansätze getestet, um die gesättigte hydraulische Leitfähigkeit österreichweit zu bestimmen:

- indirekte Ableitung von ks und
- direkte Ableitung von ks.

Bei der *indirekten Ableitung* werden in einem ersten Schritt die in großer Anzahl verfügbaren Bodeneigenschaften, wie Textur (Sand-, Schluff- und Tongehalt) mittels ML-Verfahren flächig für ganz Österreich abgeleitet. Die regionalisierten Größen werden dann als Eingangsgrößen für die EU-PTF (Szabó et al. 2021; Tóth et al. 2015) herangezogen. Über die EU-PTF wird schließlich flächig für ganz Österreich auf einer Rasterauflösung von 1 km die gesättigte hydraulische Leitfähigkeit für drei verschiedene Tiefenstufen bestimmt. Die EU-PTF wurde aufgrund ihrer guten Ergebnisse im Rahmen einer Evaluierung von PTFs an ausgewählten Standorten in Österreich (vgl. Wallner 2021) gewählt. Innerhalb der verfügbaren EU-PTFs wurde die PTF02 gewählt, die als Eingangsgrößen die Korngrößenverteilung, Tiefeninformationen und den Humusgehalt benötigt. Der Vorteil des indirekten Ansatzes liegt in den großen verfügbaren Datenmengen. Dies ist eine Grundvoraussetzung, um mit ML-Verfahren valide Ergebnisse zu erzielen.

Bei der *direkten Ableitung* wird ks über ML-Verfahren direkt vorhergesagt. Hierfür werden ML-Modelle auf Basis der abgeleiteten ks-Raster des Hydrobod-Projekts trainiert und im Anschluss auf ganz Österreich transferiert. Ein Nachteil dieser direkten Ableitung ist,

Abb. 1 Vergleich der Texturanteile (Sand (*links*), Schluff (*Mitte*) und Ton (*rechts*), jeweils in Prozent) der eBOD-Profile (X-Achse) mit den Texturanteilen der SoilGrids-Raster (Y-Achse)

Abb. 2 *Links*: Vergleich der Sandwerte der BZI-Profile innerhalb von eBOD-Polygonen, Zahlenwert oberhalb der X-Achse entspricht der Anzahl der BZI-Profile im jeweiligen eBOD-Polygon; *rechts*: Variabilität von k_s , aufgeschlüsselt nach Texturklassen (BAW-IKT Datenbank)

dass es sich bei den Hydrobod-Rastern nicht um gemessene Werte, sondern um abgeleitete Größen handelt, die dementsprechend mit zusätzlichen Unsicherheiten behaftet sind.

2.2.2 ML-Modelle und Setup

Im Rahmen dieser Studie werden zwei ML-Modelle zur Abschätzung der Zielgrößen angewendet: eXtreme Gradient Boosting (XGBoost) und Feedforward Neural Networks (FNN). Chen und Guestrin (2016) entwickelten mit XGBoost einen skalierbaren Tree-Boosting-Algorithmus, der zu einem der populärsten ML-Algorithmen der letzten Jahre avancierte. XGBoost kann für Klassifikations- und Regressionsprobleme eingesetzt werden und baut, wie es schon die Terminologie verrät, auf dem Gradient Boosting Verfahren (Friedman 2001) auf. Gradient Boosting ist ein additives Vorhersagemodell, in welchem Ensembles von Entscheidungsbäumen

basierend auf den Fehlern des vorangegangenen Entscheidungsbaumes trainiert werden. XGBoost punktet gegenüber dem klassischen Gradient Boosting vor allem durch seine Skalierbarkeit, die verbesserte Effizienz und eine Reduktion des Overfittings. Bei der Anwendung des XGBoost-Algorithmus müssen zu Beginn einige Hyperparameter gewählt werden. Die Wertebereiche für die Hyperparameter samt Kurzbeschreibung sind in Tab. 3 dargestellt.

Das FNN (White and Rosenblatt 1963) ist die simpelste Form künstlicher neuronaler Netzwerke. Ein künstliches neuronales Netzwerk setzt sich aus einem Input Layer, einem Output Layer und einem oder mehreren Hidden Layers zusammen. Ein Layer wiederum besteht aus einzelnen Neuronen. Alle Neuronen eines Layers sind mit allen Neuronen des vor- und nachgeschalteten Layers verbunden. Diesen Neuronen-Verbindungen sind Gewichte zugewiesen, die im Laufe des Lern-

prozesses modifiziert werden. Jedes Neuron erhält einen Input, der mithilfe von (linearen oder nicht-linearen) Aktivierungsfunktionen transformiert und an das nächste Neuron als Output weitergesendet wird. Beim FNN müssen ebenso vorab Hyperparameter definiert werden. Zu diesen Hyperparametern gehören zum Beispiel die Aktivierungsfunktion, die Anzahl der Layer, die Anzahl der Neuronen pro Layer oder die Dropout Rate. Dropout Layer sollen durch ein zufälliges Eliminieren von Neuronen ein Overfitting unterbinden (Feigl et al. 2021). Die Wertebereiche für die Auswahl der Hyperparameter sind in Tab. 4 dargestellt.

Die Hyperparameter von XGBoost und FNN (mit Ausnahme der Aktivierungsfunktion) wurden mittels Bayes'scher Hyperparameter-Optimierung (Bergstra et al. 2013) innerhalb vorgegebener Wertebereiche ermittelt. Dabei wurden jeweils 40 Iterationen mit 20 verschiedenen zufällig gewählten Initialisierungspunkten durchgeführt. Um die Vorhersageergebnisse durch unabhängige Daten testen zu können, wurden die Eingangsdaten für beide ML-Verfahren in Trainingsdaten (90%) und Testdaten (10%) unterteilt. Zur Optimierung des XGBoost-Modells wurde auf den Trainingsdatensatz eine 10-fache Kreuzvalidierung mit fünf Wiederholungen angewendet.

Für die Hyperparameter-Optimierung des FNN wurden 75% des Trainingsdatensatzes zum Trainieren des Modells eingesetzt und 25% zum Validieren. Zusätzlich wurden die Eingangsdaten beim FNN vor der Hyperparameter-Optimierung standardisiert. Bei der FNN-Applikation wurden aus den besten Hyperparametersets fünf

Tab. 5 Gegenüberstellung der Test-Performance bezogen auf den Test-RMSE von XGBoost und FNN für die Vorhersage der Texturanteile (%) für alle drei Tiefenstufen

Tiefenstufe	XGBoost			FNN		
	0–20 cm	20–50 cm	50–100 cm	0–20 cm	20–50 cm	50–100 cm
	TEST-RMSE			TEST-RMSE		
Sand (%)	10,92	11,34	15,35	14,18	16,22	19,77
Schluff (%)	9,06	9,13	11,64	12,19	13,22	14,74
Ton (%)	6,76	7,12	8,66	8,47	9,95	10,83

unabhängige Modelle trainiert, die für die finale Vorhersage durch Mittelwertbildung zu einem Ensemble zusammengefasst wurden.

Als Modellgütemetriken und somit als zu minimierende Zielfunktionen wurden für beide ML-Modelle sowie für die Hyperparameter-Optimierung die Wurzel der mittleren quadratischen Abweichung (RMSE) sowie der mittlere absolute Fehler (MAE) herangezogen.

Die Datenaufbereitung, Analyse und Berechnung der Ergebnisse erfolgte mithilfe der Software ArcGIS (ESRI 2018) und sämtlichen frei verfügbaren Softwarepaketen der Programmierung R (R Core Team 2020).

3 Ergebnisse

3.1 Unsicherheiten und Fehlerquellen der Input-Daten

3.1.1 Skalenproblematik

Beim Einbinden unterschiedlicher Datenquellen ist es unerlässlich, den jeweiligen Aufnahme- und Betrachtungsmaßstab zu berücksichtigen. Punktinformationen, z. B. in Form der eBOD- oder BZI-Bodenprofile, liefern lokale, sehr kleinskalige Informationen. Bei SoilGrids wiederum werden Boden-daten für eine Rasterauflösung von 250 m abgeleitet, die Vorhersage geschieht allerdings zum Teil auf viel niedriger aufgelösten Datensätzen, wie z. B. der globalen lithologischen Karte (Hartmann und Moosdorf 2012) mit einem durchschnittlichen Maßstab von 1:3.750.000 oder Informationen zu Temperatur (Wan 2006) und Niederschlag (Hijmans et al. 2005) mit einer Auflösung von 1 km. Vergleicht man nun österreichweit die Texturdaten der eBOD-Profilen mit den Texturinformationen der SoilGrids-Produkte, zeigt sich nur eine geringe Korrelation und relativ hohe RMSE. In Abb. 1 wurden für die oberste Bodenschicht (0–20 cm) die Texturwerte der eBOD-Profilen (X-Ach-

se) den Texturwerten von SoilGrids (Y-Achse) gegenübergestellt.

Beim Sand entsprechen die Mittelwerte von eBOD und SoilGrids mit 31,6 % bzw. 29,7 % einander gut. Anders sieht der Vergleich der Spannweite der Sandwerte (eBOD=97 %, SoilGrids=45 %) oder der Standardabweichung ($sd_{eBOD}=17,7\%$ versus $sd_{SoilGrids}=7,2\%$) aus. Hier zeigt sich bei SoilGrids eine deutlich geringere Variabilität. Analog zu den Sandwerten verhält sich auch der Vergleich der Schluff- und Tonwerte der zwei Bodendatensätze. Die reduzierte Variabilität ist auf die unterschiedlichen Skalen der Produkte zurückzuführen. Die charakteristischerweise sehr hohe kleinskalige Variabilität, die in den eBOD-Profilen widerspiegelt wird, kann bei der Betrachtung selbiger Eigenschaften auf größerem Betrachtungsmaßstab, wie es bei SoilGrids der Fall ist, nicht wiedergegeben werden. Durch Aggregation der Eigenschaften auf größere Skalen reduziert sich die Variabilität und alle Datenpunkte nähern sich dem Mittelwert an. Dieses Phänomen ist ein sehr allgemeines Problem beim Vergleich von Daten auf unterschiedlichen räumlichen Skalen.

3.1.2 Natürliche Heterogenität der Bodeneigenschaften

Bodeneigenschaften weisen typischerweise eine sehr hohe lokale Heterogenität auf. Diese Heterogenität ist in den verfügbaren Datensätzen selbst ersichtlich. In Abb. 2 werden in der linken Grafik Sandwerte der BZI-Profilen mit dem Sandgehalt des jeweiligen eBOD-Polygons, in dem sie sich geografisch befinden, verglichen. Der Vergleich wurde für die oberste Bodenschicht (0–20 cm) durchgeführt. Die Zahlenwerte oberhalb der X-Achse geben die Anzahl der BZI-Profilen wieder, die innerhalb der jeweiligen Bodenform liegen. Manchen eBOD-Bodenformpolygonen sind zwei Leitprofile zugeordnet, dementspre-

chend sind auch teilweise zwei eBOD-Werte pro Polygon ID abgebildet.

Wie eingangs beschrieben, werden innerhalb eines Bodenform-Polygons Flächen mit möglichst gleicher Bodenbeschaffenheit und Standortcharakteristika zusammengefasst. Aus Abb. 2 kann man gut erkennen, dass die Sandwerte innerhalb einer Bodenform zum Teil bereits eine hohe Variabilität aufweisen. Dementsprechend schwierig gestaltet es sich, die natürliche Variabilität für Regionen ohne lokale Bodenaufnahmen entsprechend abzuleiten.

Die hohe natürliche Heterogenität zeigt sich auch innerhalb der BAW-IKT Datenbank. In Abb. 2 auf der rechten Seite ist die Variabilität von ks aufgeschlüsselt nach Texturklasse dargestellt. Dabei ergeben sich je nach Texturklassen Schwankungsbreiten von bis zu vier Größenordnungen.

3.1.3 Lageungenauigkeit und Datenaktualität

Die gesamte BZI- und WBZI-Datenbank wurde aus Datenschutzgründen lediglich mit einer Lagegenauigkeit von 1 km zur Verfügung gestellt. Aufgrund der hohen lokalen Variabilität der Bodeneigenschaften kann jede Lagegenauigkeit der Profildaten zu nur schwer quantifizierbaren Unsicherheiten im Regionalisierungsprozess führen. Eine weitere potenzielle Fehlerquelle stellt das teilweise hohe Alter der Bodenaufnahmen dar. Der größte Teil der eBOD-Informationen wurde zwischen 1971 und 2002 kartiert. Selbiges gilt für WBZI- und BZI-Bodenaufnahmen: Die Erhebungen fanden zwischen 1986 und 2006 statt. Welche Profilstandorte aktuell noch dieselben Bodeneigenschaften wie zum Zeitpunkt der Erhebung aufweisen, ist aufgrund kontinuierlicher Landnutzungs- und Landbedeckungsänderungen sowie allgemeiner boden-genetischer Prozesse der letzten Jahrzehnte oftmals nur mit erheblichem Aufwand feststellbar.

Abb. 3 Boxplots der Profildaten (observation – blau) und der XGBoost-Vorhersagen (prediction – rot) der Sand- (links) und Tongehalte (rechts) über die Tiefenstufen

Abb. 4 Links: Gegenüberstellung des Tongehalts der Profilpunkte (eBOD + BZI) auf der X-Achse mit der SoilGrids-Vorhersage auf der Y-Achse, rechts: Gegenüberstellung des Tongehalts der Profilpunkte (eBOD + BZI) auf der X-Achse mit der XGBoost-Vorhersage auf der Y-Achse

3.1.4 Repräsentanz und Ungleichgewicht

Für eine flächige Ableitung von Bodeneigenschaften ist es notwendig, auf eine möglichst repräsentative Datengrundlage hinsichtlich Österreichs topografischer und klimatischer Charakteristika sowie der wichtigsten Landnutzungsklassen zurückgreifen zu können. Bodeneigenschaften landwirtschaftlich genutzter Flächen können für Österreich dank der großen Anzahl an Bodenprofilen bzw. über die Bodenformpolygone vergleichsweise gut abgebildet werden. Für Österreichs Waldflächen hingegen stehen mit den WBZI lediglich rund 550 Profile zur Verfügung, wobei jeweils nur für die Tiefenstufe 30–50 cm Texturdaten erhoben wurden. Gemäß BFW (2021) beträgt die Waldfläche Österreichs aktuell rund 47,6 % der

Gesamtfläche. Mit dementsprechend größeren Unsicherheiten ist eine Ableitung von Bodeninformationen für so einen großen Flächenanteil auf Basis von nur wenigen Datenpunkten behaftet.

Auch in Bezug auf Österreichs Topografie sind die vorhandenen Bodendaten nicht gleichmäßig über alle Höhenlagen verteilt und erwartungsgemäß sind höhere Lagen unterrepräsentiert. Das Gros der Profildaten betrifft Höhen < 1100 m ü. A., obgleich rund 40 % der Gesamtfläche Österreichs höher als 1000 m ü. A. liegen.

3.1.5 Messmethodik

Vor der Regionalisierung der Zielgrößen wurden auch die Messverfahren der unterschiedlichen Datenquellen näher betrachtet. Während zum Beispiel die

pH- und Kalkwerte für BZI, eBOD und die BAW-IKT-Datenbank nach denselben Messverfahren bestimmt wurden (pH-Wert in 0,01 m CaCl₂-Lösung nach ÖNORM L 1083 bzw. Carbonatgehalt-Bestimmung nach Scheibler), wurde der organische Kohlenstoff je nach Datenquelle durch Nass- oder Trockenoxidation bestimmt. Auch bei der Korngrößenverteilung wurde nach unterschiedlichen Messverfahren (ohne/mit H₂O₂) vorgegangen. Durch verschiedene Messverfahren sind die verfügbaren Datensätze nicht eins zu eins vergleichbar und zusätzliche, nur schwer quantifizierbare Unsicherheiten können eine Folge davon sein.

3.2 Indirekte Ableitung Ks

3.2.1 Ableitung Eingangsgrößen für PTF

Bei der indirekten Ableitung von Ks wurden in einem ersten Schritt die Eingangsgrößen für die gewählte EUPTF, nämlich Sand-, Schluff-, Ton- und Humusgehalt, österreichweit regionalisiert. Erste Modellläufe mit den Texturanteilen als Zielgrößen haben schnell gezeigt, dass XGBoost auf Basis der vorhandenen Datenlage bessere Ergebnisse als FNN erzielt. Zu einer ersten Vorabschätzung stellt das FNN-Modell trotzdem ein wichtiges Instrument dar, da es vor allem hinsichtlich des Rechenaufwands und der Nachbereitung der Ergebnisse große Vorteile aufweist. Bei XGBoost muss für jede Zielgröße ein eigenes Modell trainiert werden, wobei die Rechendauer je Modell rund 11 bis 13 h beträgt. Zusätzlich müssen die XGBoost-Ergebnisse nachträglich hinsichtlich ihrer Wertebereiche normiert

Abb. 5 Ergebnisse der XGBoost-Regionalisierung der Sand-, Schluff- und Tonanteile (jeweils in %), sowie des Humusgehalts (%) für die oberste Tiefenstufe und für ganz Österreich mit Ausnahme der Bodenbedeckungsklassen „Gebäude“, „Versiegelte Flächen“, „Eis und Schnee“ und „Wasserflächen“

Abb. 6 25 %-Quantil (links), mittlere ks (Mitte) und 75 %-Quantil (rechts) der indirekten Ableitung, jeweils in $\log_{10}(\text{cm/d})$ und für die Tiefenstufe 0–20 cm

werden, damit sichergestellt ist, dass die Summe der abgeleiteten Texturanteile 100 % ergibt. Im Gegensatz dazu liefert das FNN-Modell für alle drei Texturanteile in rund 35 min Ergebnisse und es ist keine nachträgliche Skalierung der Texturwerte auf 100 % notwendig, da dieser Schritt bereits im Training durch die Wahl der Aktivierungsfunktion des Output Layers (Softmax-Aktivierung) implementiert werden kann.

Die besten Modellläufe für die Texturanteile hinsichtlich des Test-RMSE sind in Tab. 5 für alle betrachteten Tiefenstufen separat dargestellt. Der Fehler der XGBoost-Vorhersagen ist für jede Zielgröße und Tiefenstufe geringer als jener der FNN-Modelle. Aufgrund der besseren XGBoost-Ergebnisse wurde für die vierte Eingangsgröße der EU-PTF, also für den Humusgehalt, nur das XGBoost-Modell herangezogen. Hier beträgt der Test-RMSE je nach Tiefenstufen zwischen 6,12 % und 8,37 %.

Tab. 5 zeigt, dass der Fehler in beiden Modellen mit zunehmender Tiefe größer wird. Bei näherer Betrachtung der vertikalen Verteilung der Sand- und Tongehalte in Abb. 3 zeigt sich, dass die Variabilität der Beobachtungsdaten (blaue Boxplots) mit der Tiefe zunimmt. Die XGBoost-Modelle (rote Boxplots) sind nicht in der Lage, diese größer werdende Varianz abzubilden. Zusätzlich erschwert wird dies durch die geringere Anzahl an Beobachtungspunkten mit der Tiefe. Für die Tiefenstufe 50–100 cm stehen mit rd. 10.500 Beobachtungspunkten ca. 5.000 Punkte weniger zur Verfügung als für den Oberboden.

Die Test-RMSEs der XGBoost-Ergebnisse zeigen außerdem eine deutliche Verbesserung der bis dato einzigen österreichweit verfügbaren, flächigen Bodeninformation SoilGrids. Abb. 4 zeigt links eine Gegenüberstellung des Tongehalts der vorhandenen Profildaten (X-Achse) und der zugehörigen SoilGrids-Karte (Y-Achse) für die oberste Tiefenstufe. In der rechten Abbildung werden für dieselbe Tiefenstufe die Profildaten den XGBoost-Ergebnissen gegenübergestellt. Ein Vergleich der RMSE zeigt eine Fehlerreduktion von rund 42 %. Auch für den Sand- und Tongehalt werden durch XGBoost ähnlich große Verbesserungen gegenüber SoilGrids erzielt. Beim Vergleich der zwei Grafiken ist außerdem der unter Abschn. 3.1.1 beschriebene Skaleneffekt in der XGBoost-Vorhersage (rechtes Bild) nicht erkennbar. Im Vergleich zum SoilGrids-Produkt wird also die Variabilität der Bodenparameter durch das trainierte XGBoost-Modell deutlich besser vorhergesagt. Die vorherige Grafik (Abb. 3) hat aber bereits gezeigt, dass auch in den XGBoost-Vorhersagen ein gewisser Skaleneffekt zu beobachten ist und es in den Ergebnissen zur Reduktion der Variabilität kommt. Vor allem in tieferen Horizonten geht die natürliche Variabilität der Bodendaten im Regionalisierungsprozess verloren.

Basierend auf den besten XGBoost-Ergebnissen wurden die Texturanteile und der Humusgehalt für die Fläche von Österreich für die drei Tiefenstufen 0–20 cm, 20–50 cm und 50–100 cm auf einer Rasterbasis von $1 \times 1 \text{ km}^2$ abgelei-

tet. Die Ergebnistraster für die oberste Tiefenstufe sind in Abb. 5 grafisch dargestellt.

3.2.2 Ermittlung ks und Validierung

Die Anwendung der EU-PTF (Szabó et al. 2021) erfolgte für jede Tiefenstufe separat, jeweils auf Basis der pro Tiefenstufe abgeleiteten Eingangsgrößen Sand, Schluff, Ton und Humus. Zur Anwendung des R-Pakets „eupth“ (Szabó et al. 2021; Weber et al. 2020) wurden in einem ersten Schritt für alle Tiefenstufen die Eingangsgrößen Sand, Schluff und Ton ins bodenkundliche US-Klassifikationssystem mit den Korngrößenbereichen Ton ($<2 \mu\text{m}$), Schluff ($2\text{--}50 \mu\text{m}$) und Sand ($50\text{--}2000 \mu\text{m}$) transformiert. Im Anschluss wurden für alle Tiefenstufen die mittlere ks sowie das 25 %- und das 75 %-Quantil für eine Auflösung von $1 \times 1 \text{ km}^2$ ermittelt. Die EU-PTF wurde über ML auf Basis der EU-HYDI entwickelt (Szabó et al. 2021), die Vorhersageunsicherheiten wurden dabei über Quantile Regression (Meinshausen 2006) implementiert.

Wie schon unter Abschn. 3.1.3 beschrieben, entsprechen die Landnutzungsinformationen der Bodenprofile aufgrund von Lageungenauigkeiten oder alter Informationen nicht immer der herangezogenen Bodenbedeckungskarte. Deshalb wurden die Bodenbedeckungsklassen „Gebäude“, „Versiegelte Flächen“, „Eis und Schnee“ und „Wasserflächen“ bei der ks Bestimmung nicht berücksichtigt.

Abb. 7 Validierung der indirekten Ableitung von k_s mit der BAW-IKT-Datenbank (*links*), den BFW-Beregnungsexperimenten (*Mitte*) und den Hydrobod-Rastern (*rechts*), jeweils für den obersten Bodenhorizont

Abb. 8 Schematische Darstellung der direkten k_s -Ableitung (mm/d) auf Basis von Hydrobod-NÖ für die oberste Tiefenschicht *links*: NÖ-Hydrobod- k_s -Raster ($50 \times 50 \text{ m}^2$), *Mitte*: XGBoost-Ableitung ($1 \times 1 \text{ km}^2$), *rechts*: Transfer auf Österreich ($1 \times 1 \text{ km}^2$)

Wie groß die Unsicherheiten der abgeleiteten Zielgröße k_s sind, zeigt Abb. 6. Der mittlere, mittels EU-PTF abgeleitete Wert bewegt sich österreichweit zwischen $-1,0$ und $2,2 \log_{10}(\text{cm/d})$. Das 25%-Quantil liegt zwischen $-2,2$ und $1,74 \log_{10}(\text{cm/d})$, und das 75%-Quantil zwischen $-0,4$ und $3,1 \log_{10}(\text{cm/d})$. Dieser große Schwankungsbereich zeigt sehr gut, dass bei Arbeiten aufbauend auf der gesättigten hydraulischen Leitfähigkeit kein einzelner Schwellenwert, sondern vielmehr ein Wertebereich berücksichtigt werden sollte. Die mittlere Schwankungsbreite liegt bei rund zwei Größenordnungen ($2 \log_{10}(\text{cm/d})$).

Eine Validierung der österreichweiten k_s -Ableitungen ist nur bedingt möglich. Die Ergebnisse wurden der BAW-IKT-Datenbank, den Hydrobod-Rastern und den Beregnungsexperimenten der IKT gegenübergestellt. Im ersten Fall handelt es sich um punktuelle, im Labor ausgewertete Probenahmen. Ein Vergleich von Punktmessungen und einer Rasterkarte mit einer Auflösung von 1 km hat aufgrund der natürlichen, kleinskaligen Heterogenität bodenhydraulischer Parameter nur eine beding-

te Aussagekraft. Bei der Validierung mit Hydrobod ist eine Gegenüberstellung der abgeleiteten k_s -Werte nur für die Fläche von Niederösterreich möglich. Des Weiteren muss beachtet werden, dass die Hydrobod-Produkte selbst Ableitungen darstellen, die ihrerseits mit Unsicherheiten behaftet sind. Eine Validierung der Ergebniskarten mit den Beregnungsexperimenten des BFW wurde unter der Annahme getroffen, dass bei Abflusskonstanz das Produkt aus Regenintensität und Abflussbeiwert der gesättigten hydraulischen Leitfähigkeit angenähert werden kann. Da hier keine tiefendifferenzierten Daten zur Verfügung stehen, wurden die Beregnungsexperimente nur dem obersten Bodenhorizont gegenübergestellt. Abb. 7 zeigt alle drei beschriebenen Validierungsversuche für den obersten Bodenhorizont, wobei jeweils auf der Y-Achse die abgeleiteten k_s -Werte dargestellt sind und auf der X-Achse die Validierungsdaten.

In keiner der drei Validierungsversuche ist ein guter Zusammenhang zu den abgeleiteten k_s -Werten ersichtlich. Auch zwischen den drei Validierungsversuchen ist kein Muster erkennbar.

Alle drei Grafiken in Abb. 7 verbindet die Tatsache, dass die PTF-abgeleiteten k_s -Werte tendenziell jeweils niedriger sind als die Werte der drei Validierungsdatensätze. Das mittlere und das rechte Streudiagramm weisen eine ähnliche Form auf, was wohl damit zusammenhängen dürfte, dass die über PTF berechneten k_s -Werte aus dem Hydrobod-Projekt über – aus den Beregnungsversuchen gewonnenen – Modifikatoren (Sotier et al. 2017) angepasst wurden. Eine Einschätzung der Regionalisierung von k_s auf Basis dieser Validierungsversuche ist nicht möglich.

3.3 Direkte Ableitung k_s

Die Hydrobod-Ergebniskarten können über das XGBoost-Modell mit einem Test-RMSE von $1,65 \log_{10}(\text{cm/d})$, wenn man die hohe kleinskalige Variabilität von k_s bedenkt (vgl. Abb. 2), gut nachgebildet werden. Hierzu wurden die Hydrobod-NÖ-Produkte vorab von der Original-Auflösung (50 m) auf die Auflösung der Prädiktorvariablen (1 km) aggregiert. Eine Erklärung für die guten Ergebnisse dürfte einerseits darin liegen, dass die gewählten Prädiktorva-

Abb. 9 Gegenüberstellung der k_s -Ergebnisse aus der direkten und indirekten Ableitung sowie aus der Studie (Brunetti et al. 2022, in diesem Heft) für die oberste Tiefenstufe. Für die indirekte Ableitung und die Ergebnisse der Studie (modul_2) sind jeweils das 2,5-%- bzw. das 97,5-%-Quantil abgebildet

riablen den Prädiktoren im Hydrobod-Projekt sehr ähnlich sind. Andererseits handelt es sich beim Hydrobod-Produkt um modellierte Schätzwerte, bei denen die Variabilität der Zielgröße durch Aggregierungsschritte und großskalige Prädiktoren bereits reduziert worden sein dürfte. Somit gestaltet sich auch die Vorhersage für die eingesetzten ML-Verfahren einfacher. Die Ergebnisse können aufgrund der österreichweit verfügbaren Prädiktorvariablen sehr einfach von Niederösterreich auf die restlichen Bundesländer übertragen werden. Die Prozessschritte samt österreichweiter k_s -Ergebniskarte in mm/d sind für die oberste Tiefenschicht in Abb. 8 ersichtlich.

Wie valide die Ableitungen für Restösterreich sind, lässt sich allerdings nur schwer einschätzen. Analog zur Validierung der indirekten Ableitung wurde das Ergebnis der BAW-IKT-Datenbank und den BFW-Beregnungsexperimenten gegenübergestellt, allerdings kann auch anhand dieser Gegenüberstellungen kein Zusammenhang festgestellt werden oder eine Aussage über die Qualität der Ergebnisse getroffen werden.

In einer parallel zu dieser Arbeit durchgeführten Studie (Brunetti et al. 2022, in diesem Heft) wurden Infiltrationsprozesse unter dynamischen Vorfeuchte-Verhältnissen an Referenzböden der Bodenwassermonitoring-Stationen des Bundesministeriums für

Landwirtschaft, Regionen und Tourismus untersucht. Dabei wurde für jede Messtation ein Hydrus-1D-Modell aufgesetzt und die bodenhydraulischen Parameter invers kalibriert. Die resultierenden k_s -Werte dieser inversen Kalibrierung repräsentieren k_s -Bereiche auf der Plotskala typischer österreichischer Böden. Diese kleinskaligen k_s -Werte werden in Abb. 9 mit den großskaligen Ergebnissen der indirekten und direkten Ableitung von k_s verglichen. Für die invers kalibrierten k_s -Werte (modul_2) und für die indirekte Ableitung sind auch jeweils das 2,5-%- bzw. das 97,5-%-Quantil dargestellt. Die großen Vorhersage-Unsicherheiten der PTF (vgl. Abb. 9 indirekt) zeigen sehr gut, wie schwierig eine Vorhersage dieses hochvariablen Bodenparameters ist. Dementsprechend sollten bei der Nutzung der Ergebnisraaster ausreichend große Schwankungsbereiche berücksichtigt werden.

Die Gegenüberstellung der k_s -Werte zeigt, dass die Ergebnisse der direkten Ableitung mit einer Ausnahme ausschließlich über den Ergebnissen der indirekten Ableitung liegen. Dies trifft nicht nur an den dargestellten Messstationen zu, sondern auch österreichweit liegen die k_s -Werte der direkten Ableitung im Schnitt eine Größenordnung über den Werten der indirekten Ableitung. Die k_s -Werte aus der Studie (Brunetti et al. 2022, in diesem Heft) (modul_2) liegen bei 8 von 12 Stand-

orten zwischen den großskaligen Ergebnissen.

4 Diskussion und Ausblick

In der vorliegenden Studie wurden zwei ML-Modelle getestet, um anhand aller verfügbaren Bodendaten und Umweltvariablen bodenhydraulische Eigenschaften flächig für ganz Österreich abzuleiten. Besonderes Augenmerk wurde hier auf die Ableitung der gesättigten hydraulischen Leitfähigkeit (k_s) gelegt, da dieser Bodenparameter für die Abschätzung von lokalen Infiltrationsraten von großer Bedeutung ist und somit als Grundlage für die Gefahrenabschätzung von pluvialem Hochwasser herangezogen werden kann.

In einem ersten Schritt wurden alle verfügbaren Bodendatensätze analysiert und aufbereitet. Die umfassende Datenanalyse (vgl. Abschn. 3.1) hat gezeigt, mit wie vielen Unsicherheiten und potenziellen Fehlerquellen bereits die vorhandenen Eingangsdaten behaftet sind. Bodeneigenschaften besitzen charakteristischerweise eine sehr hohe kleinskalige, räumliche und zeitliche Heterogenität. Bei Betrachtung der Bodeneigenschaften auf größeren Skalen wird diese natürliche Heterogenität durch Aggregierung dementsprechend reduziert. Dieser Skaleneffekt wurde auch in den dieser Studie zugrunde liegenden Bodendaten beobachtet und muss beim Zusammenfassen von unterschiedlich hoch aufgelösten Datensätzen bedacht werden. Auch im Rahmen dieser Studie wurden teilweise feiner aufgelöste Datengrundlagen auf die Zielauflösung von $1 \times 1 \text{ km}^2$ aggregiert. Um den höher aufgelösten Eingangsdaten, wie dem Geländemodell oder dem NDVI Rechnung zu tragen, wurden bei der Aggregierung auf die Zielauflösung nicht nur der jeweilige Mittelwert, sondern auch die Standardabweichung und das 5-% bzw. 95-%-Quantil berücksichtigt. Nichtsdestotrotz kann bei einer $1 \times 1 \text{ km}^2$ Zielauflösung eine lokale Bodenheterogenität naturgemäß nicht abgebildet werden. Weitere Unsicherheiten in den Eingangsdaten stehen im Zusammenhang mit dem Alter der Daten, Lageungenauigkeiten oder unterschiedlichen Messmethoden. Zudem sind aus topografischer Sicht tiefer gelegene und vor allem landwirtschaftlich genutzte Bodenflächen deutlich besser erfasst als zum Beispiel Waldflächen oder Almregionen.

Bei der Regionalisierung der gesättigten hydraulischen Leitfähigkeit wurden zwei unterschiedliche Ansätze verfolgt: die indirekte und die direkte Ableitung von k_s . Bei der indirekten Ableitung wurden in einem ersten Schritt die Bodenparameter Sand-, Schluff-, Ton- und Humusgehalt flächig abgeleitet. Diese Bodenparameter wurden im zweiten Schritt als Eingangsgrößen für die gewählte EU-PTF zur Ermittlung von k_s benötigt. Die Bodenparameter wurden für die drei Tiefenstufen 0–20 cm, 20–50 cm und 50–100 cm abgeleitet. Da Bodeneigenschaften auch vertikal gesehen stark variieren können, wird durch eine Aggregation der Bodendaten auf einheitliche Tiefenstufen folglich auch die vertikale Variabilität reduziert. Für die Ableitung der vier genannten Bodenparameter wurden die vielfach erprobten ML-Modelle XGBoost und FNN getestet. XGBoost liefert für die gegebene Datenlage hinsichtlich des RMSE deutlich bessere Ergebnisse als der FNN-Algorithmus. Zweiter besticht vorrangig aufgrund seiner schnellen Rechenzeit und stellt somit ein hilfreiches Instrument für schnelle Testläufe dar.

Eines der wichtigsten Ergebnisse dieser Studie ist, dass die XGBoost-Ergebnisse eine relative, im Schnitt 40-prozentige Verbesserung (bezogen auf den Test-RMSE) der derzeit einzig österreichweit flächig verfügbaren Bodeninformation SoilGrids darstellen. Die Variabilität der Bodenparameter Sand, Schluff, Ton und Humus kann vor allem für die oberste Tiefenstufe gut, allerdings nicht im vollen Maße, wiedergegeben werden.

Die tiefendifferenzierten Ergebniskarten für Sand, Schluff, Ton und Humus wurden als Eingangsgrößen für die EU-PTF herangezogen. Auf Basis der EU-PTF02 wurde die gesättigte hydraulische Leitfähigkeit für ganz Österreich für drei verschiedene Tiefenstufen ermittelt. Neben der mittleren gesättigten hydraulischen Leitfähigkeit wurde auch das 25%- und das 75%-Quantil ermittelt. Die großen Schwankungsbereiche in den Ergebnissen (vgl. Abb. 6) zeigen gut, dass bei Fragestellungen betreffend die gesättigte hydraulische Leitfähigkeit nicht von einem absoluten Wert, sondern vielmehr von einer möglichen Wertespanne ausgegangen werden sollte. Die Ergebnisse für k_s sind auf Basis der verfügbaren Datensätze nur bedingt validierbar.

Die direkte Ableitung von k_s erfolgte auf Basis der Ergebnistraster der Hydrobod-NÖ-Projekte. Auf Basis der k_s -Hydrobod-Raster wurde ein XGBoost-Modell für alle drei Tiefenstufen trainiert. Dieser Datensatz ist mit den angewendeten ML-Verfahren gut nachbildbar. Das Modell wurde auf ganz Österreich transferiert. Analog zu den Endergebnissen der indirekten Ableitung war auch bei der direkten Hydrobod-Ableitung eine Validierung nur sehr eingeschränkt möglich.

Insgesamt zeigen die vorliegenden Ergebnisse, dass leicht messbare Bodeneigenschaften, wie die Texturanteile, mit ML-Verfahren gut modellierbar sind, vorausgesetzt, eine ausreichend große Datenbasis ist vorhanden. Die resultierenden RMSE sind im Vergleich zu anderen DSM-Projekten als gut zu bewerten (vgl. Hengl et al. 2017; Pásztor et al. 2020), grundsätzlich sind die Ergebnisse aber trotzdem mit einem großen Fehler behaftet. Dies zeigt einmal mehr, wie schwierig die Modellierung hoch variabler Parameter ist. Um die Ergebnisse weiter verbessern zu können, wäre eine Überprüfung der Aktualität der bestehenden Datensätze sowie eine verdichtete Bodenaufnahme von nicht landwirtschaftlich genutzten Flächen sinnvoll. Hierzu sind derzeit auch Initiativen, wie das FORSITE-Waldtypisierung-Steiermark Projekt, im Gange. Das Einbinden solcher Projektergebnisse in der vorliegenden Studie würde zusätzlich dazu beitragen, die bestehenden Datenlücken zu Österreichs Waldflächen zu füllen.

Die österreichweite Vorhersage der gesättigten hydraulischen Leitfähigkeit war nur limitiert möglich. Die Variabilität der gesättigten hydraulischen Leitfähigkeit beträgt im Extremfall bis zu vier Größenordnungen innerhalb einer Texturklasse (vgl. Abb. 2). Zusätzlich ist die vorhandene Datenlage zu k_s aufgrund der aufwendigeren Messungen dünner. Bei der indirekten Ableitung der gesättigten hydraulischen Leitfähigkeit wurden lediglich die Einflussgrößen Sand, Schluff, Ton und Humus berücksichtigt, k_s ist aber noch von sehr vielen weiteren lokalen und saisonalen Faktoren abhängig. Zukünftige Weiterentwicklungen dieser Arbeit können somit zum Beispiel das Einbinden von zusätzlichen Einflussfaktoren, wie saisonaler Bewuchs, Makroporen (durch saisonale Bearbeitung oder Regenwurmkaktivitäten), Hydrophobie oder Verschlammung bei Starkregenereignissen,

betreffen. Ein weiterer Fokus könnte auf die Ableitung der Zielvariablen für eine höhere Auflösung gelegt werden. Dies würde allerdings mit einem exponentiell wachsenden Rechenaufwand und Speicherbedarf im Modelltraining einhergehen.

Bei weiterführenden Arbeiten auf Grundlage der Ergebnistraster wird empfohlen, die Unsicherheiten der k_s -Werte der direkten Ableitung mit einer mittleren Schwankungsbreite von zwei Größenordnungen zu adressieren, und bei den k_s -Werten der indirekten Ableitung das 25%- bzw. das 75%-Quantil zu berücksichtigen. Für kleinskalige Fragestellungen bleiben lokale Bodenuntersuchungen im Vorfeld unerlässlich.

Danksagung Ein Dank gilt dem Bundesministerium für Landwirtschaft, Regionen und Tourismus (BMLRT), der Umweltbundesamt GmbH, dem Bundesforschungszentrum für Wald (BFW) sowie dem Institut für Kulturtechnik und Bodenwasserhaushalt des Bundesamts für Wasserwirtschaft (BAW-IKT) für die Bereitstellung von Daten.

Förderung Diese Studie wurde vom Bundesministerium für Landwirtschaft, Regionen und Tourismus (BMLRT) über das InfCapAT Projekt und von der Österreichischen Akademie der Wissenschaften (ÖAW) über das Projekt RechAUT finanziert.

Funding Open access funding provided by University of Natural Resources and Life Sciences Vienna (BOKU).

Interessenkonflikt

H. Zeitfogel, M. Feigl und K. Schulz geben an, dass kein Interessenkonflikt besteht.

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 unter-

liegen 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

- Araya, S. N., & Ghezzehei, T. A. (2019): Using Machine Learning for Prediction of Saturated Hydraulic Conductivity and Its Sensitivity to Soil Structural Perturbations. *Water Resources Research*, 55(7), 5715–5737. <https://doi.org/10.1029/2018WR024357>
- Bergstra, J., Yamins, D., & Cox, D. D. (2013): Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures (Bd. 28).
- BFW (2019): eBOD – digitale Bodenkarte. Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft. Abgerufen am 19. September 2021, von <https://bfw.ac.at/rz/bfwcms2.web?dok=7066>
- BFW (2021): Österreichische Waldinventur. Abgerufen am 13. August 2021, von <http://bfw.ac.at/rz/wi.auswahl?cros=1>
- BMLRT (o. J.): Hydrografie in der Erde. Abgerufen 30. Juli 2021, von <https://info.bmlrt.gv.at/themen/wasser/wasser-oesterreich/hydrografie/wasser-im-blick-das-messnetz/wasserimboden-messen.html>
- BMNT (2017): Die österreichische Strategie zur Anpassung an den Klimawandel. Aktualisierte Fassung.
- Brunetti, G., Schöbl, M., Santner, K., Stumpp, C. (2022): Sensitivitätsanalyse zu Infiltrationsprozessen in Böden. Österreichische Wasser- und Abfallwirtschaft 74(3–4). DOI: <https://doi.org/10.1007/s00506-022-00839-8>, in diesem Heft.
- Brutsaert, W. (2005): Hydrology, an Introduction. Cambridge Univ. Press. <https://doi.org/10.1017/CBO9780511808470>
- Chen, T., & Guestrin, C. (2016): XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Eder, A., Sotier, B., Klebinder, K., Sturmlechner, R., Dorner, J., Markart, G., Schmid, G., & Strauss, P. (2011): Hydrologische Bodenkenn-daten der Böden Niederösterreichs (HydroBod-NÖ). Endbericht.
- ESRI (2018): ArcGIS Desktop (Release 10.6.1). Environmental Systems Research Institute.
- Feigl, M., Lebedzinski, K., Herrnegger, M., & Schulz, K. (2021): Machine-learning methods for stream water temperature prediction. *Hydrology and Earth System Sciences*, 25(5), 2951–2977. <https://doi.org/10.5194/hess-25-2951-2021>
- Fohrer, N., Bormann, H., Miegel, K., Casper, M., Bronstert, A., Schumann, A., & Weiler, M. (2016): Hydrologie.
- Friedman, J. H. (2001): Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Hartmann, J., & Moosdorf, N. (2012): The new global lithological map database GLiM: A representation of rock properties at the Earth surface. *Geochemistry, Geophysics, Geosystems*, 13(12). <https://doi.org/10.1029/2012GC004370>
- Hengl, T., De Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017): SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*, 12(2), 1–40. <https://doi.org/10.1371/journal.pone.0169748>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005): Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Horton, R. E. (1933): The Role of infiltration in the hydrologic cycle. *Eos, Transactions American Geophysical Union*, 14(1), 446–460. <https://doi.org/10.1029/TR014I001P00446>
- IPCC (2013): Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. <https://doi.org/10.1017/CBO9781107415324>
- Klebinder, K., Sotier, B., Lechner, V., & Strauss, P. (2017): Hydrologische und hydropedologische Kenn-daten Raabgebiet und Region Südoststeiermark.
- Mahapatra, S., Jha, M. K., Biswal, S., & Senapati, D. (2020): Assessing Variability of Infiltration characteristics and Reliability of Infiltration Models in a Tropical Subhumid Region of India. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-58333-8>
- Markart, G., Sotier, B., Stepanek, L., Lechner, V., & Kohl, B. (2017): Waldwirkung auf die Abflussbildung bei unterschiedlichen Betrachtungsmaßstäben. *Wildbach- und Lawenverbau*, 180, 100–115.
- McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003): On digital soil mapping. *Geoderma*, 117(1–2), 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Meinshausen, N. (2006): Quantile Regression Forests. *Journal of Machine Learning Research*, 7, 983–999.
- ÖWAV-EP (2020): Klimawandelanpassung Wasserwirtschaft – Pluviales Hochwasser/Oberflächenabfluss. ÖWAV-ExpertInnenpapier. ÖWAV.
- Padarian, J., Minasny, B., & McBratney, A. B. (2020): Machine learning and soil sciences: A review aided by machine learning tools. *SOIL*, 6(1), 3552. <https://doi.org/10.5194/soil-6-35-2020>
- Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021): SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL*, 7(1), 217–240. <https://doi.org/10.5194/soil-7-217-2021>
- Pásztor, L., Laborcz, A., Takács, K., Illés, G., Szabó, J., & Szatmári, G. (2020): Progress in the elaboration of GSM conform DSM products and their functional utilization in Hungary. In *Geoderma Regional* (Bd. 21, S. e00269). Elsevier B.V. <https://doi.org/10.1016/j.geodrs.2020.e00269>
- R Core Team (2020): R: A Language and Environment for Statistical Computing. Abgerufen am 01. Oktober 2021, von <https://www.r-project.org/>
- Rözer, V., Müller, M., Bubeck, P., Kienzler, S., Thieken, A., Pech, I., Schröter, K., Buchholz, O., & Kreibich, H. (2016): Coping with pluvial floods by private households. *Water (Switzerland)*, 8(7). <https://doi.org/10.3390/W8070304>
- Rözer, V., Pech, A., Berkahn, S., Feng, Y., Fuchs, L., Graf, T., Haberlandt, U., Kreibich, H., Sämann, R., Sester, M., Shehu, B., Wahl, J., & Neuweiler, I. (2021): Impact-Based Forecasting for Pluvial Floods. *Earth's Future*, 9(2), 2020EF001851. <https://doi.org/10.1029/2020EF001851>
- Schneider, W., Nelhiebl, P., Aust, G., Wandl, M., & Danneberg, O. H. (2001): Die landwirtschaftliche Bodenkartierung in Österreich. *Mitt.d. Österr. Bodenkundl. Ges.*, 62, 39–67.
- Scull, P., Franklin, J., Chadwick, O. A., & Mcarthur, D. (2003): Predictive soil mapping: a review. <https://doi.org/10.1191/0309133303pp366ra>
- Sotier, B., Klebinder, K., Bauer, T., Markart, G., & Strauss, P. (2017): Hydrologische Bodenkenn-daten Niederösterreich – HYDROBOD-NÖ, zweite Projektphase, Ergänzungsbericht.
- Szabó, B., Weynants, M., & Weber, T. K. D. (2021): Updated European hydraulic pedotransfer functions with communicated uncertainties in the predicted variables (eupftv2). *Geoscientific Model Development*, 14(1), 151–175. <https://doi.org/10.5194/GMD-14-151-2021>
- Tahmasebi, P., Kamrava, S., Bai, T., & Sahimi, M. (2020): Machine learning in geo- and environmental sciences: From small to large scale. *Advances in Water Resources*, 142, 103619. <https://doi.org/10.1016/j.advwatres.2020.103619>
- Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., & Tóth, G. (2015): New generation of hydraulic pedotransfer functions for Europe. *European Journal of Soil Science*, 66(1), 226–238. <https://doi.org/10.1111/ejss.12192>
- Umweltbundesamt GmbH (2021): BORIS – Bodeninformationssystem. Abgerufen am 12. August 2021, von <https://www.umweltbundesamt.at/boris>
- Van Looy, K., Bouma, J., Herbst, M., Koestel, J., Minasny, B., Mishra, U., Montzka, C., Nemes, A., Pachepsky, Y. A., Padarian, J., Schaap, M. G., Tóth, B., Verhoef, A., Vanderborght, J., van der Ploeg, M. J., Weihermüller, L., Zacharias, S., Zhang, Y., & Vereecken, H. (2017): Pedotransfer Functions in Earth System Science: Challenges and Perspectives. In *Reviews of Geophysics* (Bd. 55, Nummer 4, S. 1199–1256). <https://doi.org/10.1002/2017RG000581>
- Vereecken, H., Weihermüller, L., Assouline, S., Šimúnek, J., Verhoef, A., Herbst, M., Archer, N., Mohanty, B., Montzka, C., Vanderborght, J., Balsamo, G., Bechtold, M., Boone, A., Chadburn, S., Cuntz, M., Decharme, B., Ducharme, A., Ek, M., Garrigues, S., ... Xue, Y. (2019): Infiltration from the Pedon to Global Grid Scales: An Overview and Outlook for Land Surface Modeling: Infiltration from the Pedon to Global Grid Scales: An Overview and Outlook for Land Surface Modeling. <https://doi.org/10.2136/vzj2018.10.0191>
- Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M. G., & van Genuchten, M.

- T. (2010):** Using Pedotransfer Functions to Estimate the van Genuchten-Mualem Soil Hydraulic Properties: A Review. *Vadose Zone Journal*, 9(4), 795–820. <https://doi.org/10.2136/vzj2010.0045>
- Wadoux, A. M. J.-C., Minasny, B., & McBratney, A. B. (2020):** Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews*, 210, 103359. <https://doi.org/10.1016/j.earscirev.2020.103359>
- Wallner, M. (2021):** Evaluation of pedotransfer-functions for estimating soil hydraulic properties and groundwater recharge. Master Thesis. University of Natural Resources and Life Sciences, Vienna.
- Wan, Z. (2006):** MODIS landsurface temperature products users' guide. ICES, University of California.
- Weber, T. K. D., Weynants, M., & Szabó, B. (2020):** R package of updated European hydraulic pedotransfer functions (eupf2). <https://doi.org/10.5281/zenodo.4281045>
- White, B. W., & Rosenblatt, F. (1963):** Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. *The American Journal of Psychology*, 76(4), 705. <https://doi.org/10.2307/1419730>

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