



Sensitivitätsstudie zur Landbe- deckungsdatenbasis

Schriftenreihe, Heft 6/2026



Sensitivitätsstudie zur Fortschreibung veränderlicher Landbedeckung anhand abgesetzter Niederschläge

Dr. Philipp Körner, iamk GmbH, Institut für angewandte Meteorologie und Klimatologie.

im Auftrag des Sächsischen Landesamtes für Umwelt, Landwirtschaft und Geologie

Inhaltsverzeichnis

Zusammenfassung	6
Abstract	6
1 Einleitung.....	7
2 Landbedeckung.....	8
2.1 Landbedeckungsprodukt basierend auf Kugler et al. (2023) und Hanelli und Barth (2023).....	9
2.2 Produkt „CLCplus Backbone“	10
2.3 Produkt „ESRI 10m Annual Land Cover“	10
2.4 Vegetationshöhe.....	10
2.5 Berechnung Nebeldeposition	11
2.6 Auswertung.....	11
3 Meteorologische Eingangsdaten	13
3.1 Stationsdaten	13
3.2 Rasterdaten	13
3.2.1 Räumliche Interpolation – Lufttemperatur, Windgeschwindigkeit, Luftfeuchte, Luftdruck, Globalstrahlung.....	13
3.2.2 Räumliche Interpolation - Niederschlag.....	14
Schlussfolgerungen	21
Ausblick	21
Literaturverzeichnis	22

Abbildungsverzeichnis

Abbildung 1: Quantil-Quantil-Plot für tägliche Nebeldepositionssummen über alle Rasterzelle für 365 Tage im Jahr 2018 in Sachsen.	12
Abbildung 2: Korrelation-Korrelation-Plot zur Darstellung der räumlichen Konsistenz der berechneten stündlichen Niederschlagszeitreihen.	18
Abbildung 3: QQ-Plot von modelliertem und gemessenem Niederschlag in stündlicher Auflösung. ..	20

Tabellenverzeichnis

Tabelle 1: Zuordnung der Vegetationshöhe zu ID und Beschreibungen der Landbedeckung im Referenzprodukt	9
Tabelle 2: Zuordnung der Vegetationshöhe zu ID und Beschreibungen der Landbedeckung im CLCplus Backbone-Produkt	10
Tabelle 3: Zuordnung der Vegetationshöhe zu ID und Beschreibungen der Landbedeckung im ESRI-Produkt.....	11
Tabelle 4: Kennwerte der Berechnung der Nebeldeposition mit dem CLC und ESRI-Produkt im Vergleich zur Referenz.	11
Tabelle 5: Kennwerte für die Kreuzvalidierung der räumlichen Interpolation und der Disaggregation	19

Abkürzungsverzeichnis

CLC	CORINE Landcover, europäischer, europaweiter Datensatz zu Landbedeckung und Landnutzung
DWD	Deutscher Wetterdienst
LfULG	Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie

Zusammenfassung

Dieser Bericht betrachtet die Sensitivität der Landbedeckungsänderung mit Blick auf abgesetzte Niederschläge und die Erstellung von hochauflösenden meteorologischen Rasterdaten für Sachsen im Zeitraum von 1961 bis 2023. Zwei Landbedeckungsprodukte, „CLCplus Backbone“ und „ESRI 10m Annual Land Cover“, werden verglichen, um dasjenige Produkt zu identifizieren, das die höchste Übereinstimmung mit dem hauseigenem Referenzprodukt aufweist. Das ESRI-Produkt zeigte eine bessere Übereinstimmung und wird für die weitere Verwendung empfohlen. Darüber hinaus wird eine Methode zur Erstellung von stündlichen Niederschlagsdaten in hoher räumlicher Auflösung entwickelt, die auf Gradient Boosting basiert. Die Methode zeigt eine hohe räumliche Konsistenz und zeitliche Korrelation mit den gemessenen Daten.

Abstract

This report presents a sensitivity analysis of land cover changes to fog deposition and the creation of high-resolution meteorological raster data for Saxony from 1961 to 2023. Two land cover products, "CLCplus Backbone" and "ESRI 10m Annual Land Cover", were compared to identify the product with the highest agreement with the in-house reference product. The ESRI product showed better agreement and was recommended for further use. Additionally, a method for creating hourly precipitation data in high spatial resolution based on Gradient Boosting was developed. The method demonstrated high spatial consistency and temporal correlation with measured data.

1 Einleitung

Die Modellierung verschiedener klima- und wasserhaushaltrelevanter Größen erfordert eine Vielzahl an Eingangsgrößen. Je höher die raum-zeitliche Auflösung des Zielproduktes ist, desto höher muss auch die Auflösung der Eingangsgrößen sein. In diesem Bericht werden hierfür zwei wichtige Punkte adressiert:

Zum einen die Fortschreibung der Landbedeckungsklassifikation, die für den Freistaat Sachsen bisher für den Zeitraum 1961-2020 in verschiedenen Zeitschritten in 100 Meter Auflösung vorliegt (KUGLER ET AL., 2023 UND HANELLI UND BARTH, 2023). Hier gibt es unterschiedliche Landbedeckungsprodukte, die jährlich fortgeschrieben werden. Ziel ist es aus zweien dieser Produkte dasjenige Produkt zu identifizieren, welches die höchste Übereinstimmung mit dem bisher vorliegenden Produkt aufweist. Durchgeführt wird diese Überprüfung durch eine praktische Anwendung für das Vergleichsjahr 2018: Anhand der Modellierung der Nebeldeposition für Sachsen wird die Übereinstimmung mit dem Referenzprodukt überprüft. Ausgewählt wird dasjenige Produkt, welches die geringsten Abweichungen aufweist.

Der zweite entscheidende und mit dem ersten in direktem Zusammenhang stehende Punkt ist die Bereitstellung von meteorologischen Eingangsgrößen in hoher zeitlicher und räumlicher Auflösung. Das angesprochene Nebeldepositionsmodell, aber auch weitere Modelle wie Modelle zur Schneehöhenmodellierung, Wasserhaushaltsmodellierung oder Niederschlags-Abfluss-Modellierung sind sowohl auf hochaufgelöste meteorologische Eingangsdaten als auch auf hochaufgelöste Landnutzungsdaten angewiesen. Weiterhin weisen meteorologische Größen eine Abhängigkeit von der Landbedeckung auf. Mit diesem Bericht wird ein stundenbasierter Rasterdatensatz meteorologischer Größen in 100 Meter Auflösung für das Gebiet des Freistaates Sachsen vorgestellt. Die hierbei genutzten Ansätze orientieren sich an der Methodik, die bei der Erstellung des sachsenweiten Referenzdatensatzes 3.0 (räumliche Auflösung von 1 km) angewandt wurden (siehe KÖRNER UND MODEROW, 2025). Jedoch weisen die genutzten Ansätze gegenüber KÖRNER UND MODEROW (2025) Änderungen auf, um der feineren räumlichen Auflösung (100 m) gerecht werden zu können. Dies gilt insbesondere für den Niederschlag.

2 Landbedeckung

Die Landbedeckung spielt in der Hydrologie und der Meteorologie eine entscheidende Rolle, da sich mit ihr der Energie- und Wasserhaushalt ändert. Aus Informationen zur Landbedeckung können weitere Größen abgeleitet werden, die für die Analyse und Modellierung hydrologischer und meteorologischer Größen, wie z. B. der Verdunstung, wichtig sind. Hierzu zählen u.a. die Vegetationsart, der Blattflächenindex und die Vegetationshöhe. Typischerweise ist die Landbedeckung zeitlich variabel. Teilweise haben sich in Sachsen in den letzten Jahrzehnten lokal drastische Änderungen der Landbedeckungen ergeben, die mit der Gewinnung von Braunkohle im Tagebau einhergingen. Ein Beispiel hierfür ist das Gebiet des ehemaligen Tagebaus Bärwalde. Wo 1970 noch Siedlungen und Wälder waren, gab es 1980 einen Tagebau, der jedoch ab Ende der 90er Jahre des 20. Jh. geflutet wurde und heute ein See ist. Um hydrologische und meteorologische Modellierungen möglichst realitätsnah vornehmen zu können, muss einer solchen dynamischen Landbedeckung Rechnung getragen werden. Anderenfalls kann nur der Zustand zu einem bestimmten Zeitpunkt oder ein gewisser mittlerer Zustand dargestellt werden. In Ergänzung zu den in KUGLER ET AL. (2023) UND HANELLI UND BARTH (2023) erstellten Landbedeckungsprodukten sollen in diesem Abschnitt zwei frei verwendbare Produkte miteinander verglichen werden, die die Landbedeckung zeitlich fortschreiben. Dies ist zum einen das Produkt „CLC plus Backbone“ und zum anderen das Produkt „ESRI 10m Annual Land Cover“. Welchen Einfluss unterschiedliche Landnutzungsdaten auf die Modellierung hydrologischer und meteorologischer Größen haben können, wird anhand eines konkreten Anwendungsfall dargestellt. Hierbei handelt es sich um die Nebeldeposition. Diese wurde bereits für Sachsen unter Verwendung der Landbedeckungsprodukte von KUGLER ET AL. (2023) BZW. HANELLI UND BARTH (2023) für die Jahre 1961-2020 berechnet (Referenzvariante). Für die zu untersuchenden Landbedeckungsprodukte „CLC plus Backbone“ und „ESRI 10m Annual Land Cover“ wurde ebenfalls für Gesamtsachsen die Nebeldeposition für das Vergleichsjahr 2018 berechnet. Auf Grundlage der berechneten Nebeldepositionen für das Jahr 2018 wird die Auswirkung unterschiedlicher Landnutzungen auf die berechnete Nebeldeposition analysiert. Hierbei dient die Variante, die die Landnutzungsdaten aus KUGLER ET AL. (2023) UND HANELLI UND BARTH (2023) verwendet, als Referenz.

2.1 Landbedeckungsprodukt basierend auf Kugler et al. (2023) und Hanelli und Barth (2023)

Dieses Landbedeckungsprodukt stellt die Referenz für die vorliegende Arbeit dar. Es liegt in einer räumlichen Auflösung von 100 m vor und es werden 29 Landbedeckungsklassen für die vorliegende Studie unterschieden (Tabelle 1).

Tabelle 1: Zuordnung der Vegetationshöhe zu ID und Beschreibungen der Landbedeckung im Referenzprodukt

ID	Höhe	Beschreibung (Referenz)
100	0	Bebaute Fläche
111	0	durchgängig städtische Prägung
112	0	nicht durchgängig städtische Prägung
121	0	Industrie/Gewerbeflächen
122	0	Straßen/Eisenbahnnetze, funktionell zugeordnete Flächen
123	0	Hafengebiete
124	0	Flughäfen
131	0	Abbauflächen
132	0	Deponien, Abraumhalden
133	0	Baustellen
141	1	Städtische Grünflächen
142	0,1	Sport/Freizeitanlagen
200	0,3	Landwirtschaft
211	0,3	Nicht bewässertes Ackerland
231	0,2	Wiesen und Weiden
310	25	Wälder
311	25	Laubwälder
312	25	Nadelwälder
320	1	Kraut/Strauchvegetation
321	0,2	Natürliches Grünland
322	0,3	Heiden und Moorheiden
324	1	Wald/Strauch
330	0,1	Offene Flächen ohne oder mit geringer Vegetation
331	0	Strände, Dünen, Sandflächen
333	0,1	Flächen mit spärlicher Vegetation
411	0,1	Sümpfe
412	0,1	Torfmoore
510	0	Wasserflächen im Landesinneren
9999	0	andere Nutzung

2.2 Produkt „CLCplus Backbone“

„CLCplus Backbone“ stellt das neueste Produkt für Landbedeckung des Copernicus-Programms für Europa dar mit dem lückenlose Informationen zur Landbedeckung zur Verfügung gestellt werden. Es ergänzt das bekannte, langjährige und konsistente Produkt CORINE Land Cover (CLC) und zeichnet sich durch eine höhere räumliche Auflösung (10 m) und kürzere Aktualisierungsintervalle aus (dreijährlich). Die in dieser Arbeit verwandten CLCplus Backbone Landbedeckungsklassen sind in Tabelle 2 aufgeführt.

2.3 Produkt „ESRI 10m Annual Land Cover“

Die jährliche Zeitreihe globaler Landnutzungs- und Landbedeckungskarten (LULC) von ESRI (KARRA ET AL. 2021) wurde auf Version 3 aktualisiert und umfasst nun globale Landbedeckungsdaten für die Jahre 2017 bis 2023 in einer räumlichen Auflösung von 10 m. Die Karten basieren auf Sentinel-2-Satellitenbildern der ESA, wobei 9 Landbedeckungsklassen unterschieden werden (Tabelle 3). Jede Karte ist ein zusammengesetztes Produkt aus diesen neun Klassen, die über das Jahr hinweg erfasst wurden, um ein repräsentatives Abbild des jeweiligen Jahres zu erhalten.

2.4 Vegetationshöhe

Die Vegetationshöhe wird aus den Daten der Landbedeckung abgeleitet. Die Landbedeckungsklassen sind für die Referenz, das CLCplus Backbone- sowie das ESRI-Produkt nicht identisch. Dennoch lassen sich den Klassen jeweils Vegetationshöhen zuweisen, so dass die Produkte letztendlich miteinander vergleichbar sind (Tabellen 1-3).

Tabelle 2: Zuordnung der Vegetationshöhe zu ID und Beschreibungen der Landbedeckung im CLCplus Backbone-Produkt

ID	Höhe in m	Beschreibung (CLCplus Backbone)
1	0	Sealed
2	25	Woody – needle leaved trees
3	25	Woody – Broadleaved deciduous trees
4	25	Woody – Broadleaved evergreen trees
5	1	Low-growing woody plants (bushes, shrubs)
6	1	Permanent herbaceous
7	1	Periodically herbaceous
8	0,1	Lichens and mosses
9	0,1	Non- and sparsely-vegetated

Tabelle 3: Zuordnung der Vegetationshöhe zu ID und Beschreibungen der Landbedeckung im ESRI-Produkt

ID	Höhe in m	Beschreibung (ESRI)
1	0	Water
2	25	Trees
4	1	Flooded vegetation
5	0.3	Crops
7	0	Built Area
8	0	Bare ground
9	0	Snow/Ice
10	0	Clouds
11	0.2	Rangeland

2.5 Berechnung Nebeldeposition

Tägliche Werte der Nebeldeposition wurden mit dem Nebeldepositionsmodell aus KÖRNER ET AL. (2023) berechnet. Dies gilt sowohl für die Referenz, die Landbedeckungsdaten basierend auf KUGLER ET AL. (2023) UND HAENELLI ET AL. (2023) nutzt, als auch für die beiden Vergleichsvarianten die das Landbedeckungsprodukt CLCplus Backbone bzw. das Landbedeckungsprodukt „ESRI 10m Annual Land Cover“ (ESRI-Produkt) nutzen. Die Berechnung für die beiden zuletzt genannten Produkte wurden für das Jahr 2018 durchgeführt (Vergleichsjahr). Es wurden dabei dieselben meteorologischen Werte verwendet, die auch für die Referenz verwendet wurden. Dadurch ist sichergestellt, dass Veränderungen im Ergebnis ausschließlich auf die geänderte Vegetationshöhe aufgrund Nutzung anderer Landbedeckungsprodukte zurückzuführen sind.

Insgesamt wurden für jeden Tag im Jahr 2018 für alle Rasterzellen von Sachsen (> 1,8 Millionen Rasterzellen) die Nebeldeposition berechnet, jeweils für die Variante mit dem CLCplus Backbone Produkt und dem ESRI-Produkt. Bei 365 Tagen im Jahr 2018 standen somit mehr als 650 Millionen Wertepaare für die nachfolgende Auswertung zur Verfügung.

2.6 Auswertung

Tabelle 4 gibt Auskunft zu statistischen Maßen die die Übereinstimmung der Nebeldeposition berechnet mit CLCplus Backbone bzw. dem ESRI-Produkt mit der Referenz beschreiben.

Tabelle 4: Kennwerte der Berechnung der Nebeldeposition mit dem CLC und ESRI-Produkt im Vergleich zur Referenz.

	CLC	ESRI
RMSE in mm/d	0,097	0,075
MAE in mm/d	0,0039	0,0025
BIAS in mm	0,0004	0
R ²	0,45	0,58
HSS	0,63	0,92

Das Ergebnis in Tabelle 4 ist dabei eindeutig: Alle Fehlermaße (RMSE, MAE und BIAS) fallen für das ESRI-Produkt geringer aus. Gleichzeitig ist sowohl das Bestimmtheitsmaß als der Heidke Skill Score als kategorischer Index für das ESRI-Produkt höher als für das CLC-Produkt. Das ESRI-Produkt schneidet damit in allen Kategorien besser ab als das CLC-Produkt.

Dieses Bild wird durch den Quantil-Quantil-Plot in Abbildung 1 bestätigt. Die Roten Punkte der Nebeldeposition berechnet mit dem ESRI-Produkt schmiegen sich näher an die schwarze 1:1 Gerade. Dies weist darauf hin, dass die Verteilung der Werte auf Basis des ESRI-Produkts deutlich näher an der Referenz liegt, als die Werte der Nebeldeposition, die das CLCplus Backbone-Produkt nutzen.

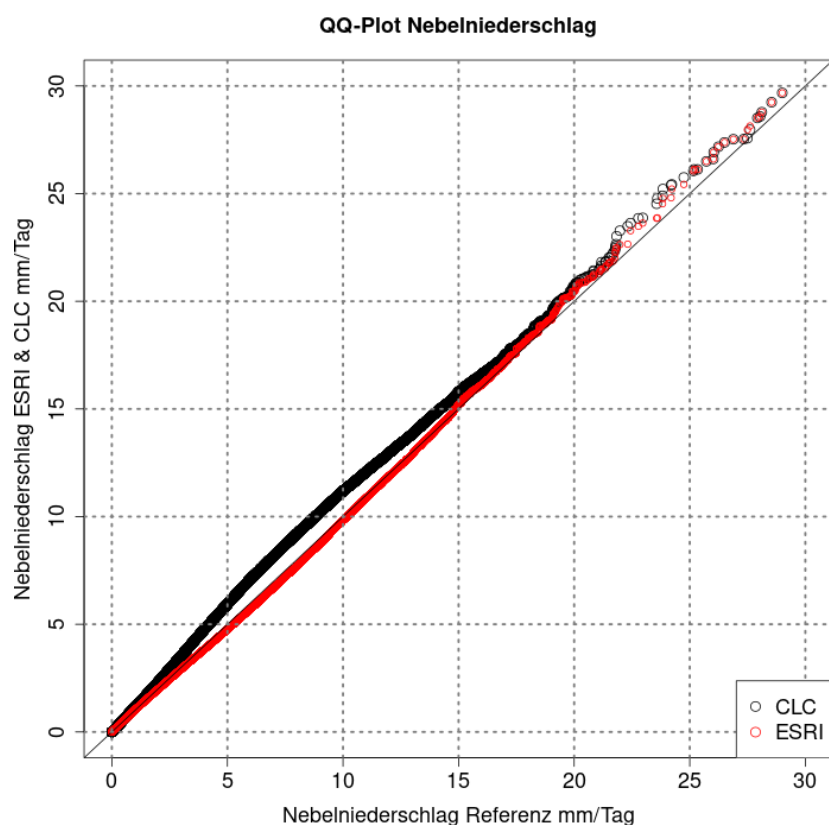


Abbildung 1: Quantil-Quantil-Plot für tägliche Nebeldepositionssummen über alle Rasterzelle für 365 Tage im Jahr 2018 in Sachsen.

Der Vergleich zeigt, dass die Nebeldeposition berechnet unter Verwendung des ESRI-Produkts objektiv besser im Vergleich zu Referenz abschneidet als die berechnete Nebeldeposition, die das CLCplus Backbone Produkt verwendet. Vor diesem Hintergrund wäre die Verwendung von Landbedeckungsdaten aus dem ESRI-Produkt gegenüber der Verwendung des Landnutzungsprodukts CLCplus Backbone zu favorisieren.

3 Meteorologische Eingangsdaten

Ziel dieses Abschnittes ist es, 100 Meter-Raster meteorologischer Eingangsgrößen für Sachsen im Zeitraum 1961 bis 2023 in stündlicher Auflösung zu erstellen. Grundlage für die Erstellung des Rasterproduktes sind die Stationsdaten des Deutschen Wetterdienstes in stündlicher Auflösung. Je nach meteorologischer Größe stehen hierfür unterschiedliche Zeiträume und Stationen zur Verfügung. Die Stationsverfügbarkeit für Stundenwerte schwankt zeitlich stark, Niederschlagsstationen sind beispielsweise erst ab September 1995 verfügbar. Neben dem Niederschlag werden noch die Größen Luftdruck, Luftfeuchte, Lufttemperatur, Windgeschwindigkeit und Globalstrahlung bearbeitet. Die Globalstrahlung wird dabei aus der stündlichen Sonnenscheindauer abgeleitet.

3.1 Stationsdaten

Zur weiteren Verarbeitung werden alle zur Verfügung stehenden Daten der unter Abschnitt 3 genannten Größen des DWD verwendet. Es erfolgt im ersten Verarbeitungsschritt kein Zuschnitt auf das Gebiet Sachsens, da für die weitere Verarbeitung, insbesondere für die Niederschlagsberechnungen, ein möglichst umfassender Datensatz benötigt wird. Die Zeitreihen der meteorologischen Größen enthalten den Zeitraum 1961 bis 2023 und umfassen Messdaten aller Stationen des DWD. Die Luftfeuchte wird einer Driftkorrektur unterzogen (KÖRNER ET AL, 2018). Anschließend wird für alle Zeitreihen eine Lückenfüllung bzw. Zeitreihenverlängerung durchgeführt. Die Methodik ist im Bericht zum Referenzdatensatz 3.0 beschrieben (KÖRNER UND MODEROW, 2025). Danach stehen für alle Größen und Stationen, die für wenigstens ein Jahr Daten messen, lückenlose Zeitreihen für den gesamten Zeitraum 1961-2023 zur Verfügung. Ausnahme ist der Niederschlag: Mangels Stützstellen vor 1995 ist eine Zeitreihenverlängerung bis zu dieser Zeit nicht möglich. Die Berechnung der stündlichen Niederschlagsdaten erfolgt räumlich und ist in Abschnitt 3.2.2 erläutert.

3.2 Rasterdaten

Aus den vollständig verlängerten und lückengefüllten Stationsdaten werden im nächsten Schritt Rasterdaten erstellt. Die Berechnung der Raster erfolgt ebenfalls auf Stundenbasis und umfasst für jede Größe somit etwas über 550.000 Raster für den Untersuchungszeitraum 1961-2023.

3.2.1 Räumliche Interpolation – Lufttemperatur, Windgeschwindigkeit, Luftfeuchte, Luftdruck, Globalstrahlung

Die meteorologischen Größen Lufttemperatur, Windgeschwindigkeit, Luftfeuchte, Luftdruck und Globalstrahlung werden mittels Thin Plate Spline Smoothing (Tps) auf das Gebiet des Freistaates Sachsen interpoliert. Details zur Methode und zur Qualität sind im Bericht zum Referenzdatensatz 3.0 zu finden (KÖRNER UND MODEROW, 2025).

3.2.2 Räumliche Interpolation - Niederschlag

Der stündliche Niederschlag lässt sich nicht für den gesamten Zeitraum 1961 bis 2023 interpolieren, weil keine gemessenen Stundenwerte der Stationen über den gesamten Zeitraum vorliegen. Die Stationsverfügbarkeit ab 1995 ist sehr gering, so dass auch eine Interpolation ab 1995 keine zufriedenstellenden Ergebnisse liefern würde. Ab wann also sollte eine räumliche Interpolation des Niederschlages durchgeführt werden? Um die Konsistenz des Produktes insgesamt zu wahren, lautet die Antwort: Gar nicht. Für den Niederschlag wird auch für Subzeiträume keine räumliche Interpolation durchgeführt, sondern die im Folgenden beschriebene Methode angewendet.

Anders als bei Stundenwerten ist die zeitliche und räumliche Abdeckung bei Tageswerten des Niederschlages sehr gut. Sowohl die zeitliche Abdeckung als auch die Stationsdichte ist bei der Niederschlagsmessung deutlich höher als bei allen anderen vom DWD gemessenen Größen. Dadurch stehen für den Niederschlag im Untersuchungszeitraum für Sachsen bereits Tagesdaten in hoher Qualität im Referenzdatensatz 3.0 zur Verfügung. Um aus Tagesdaten Stundendaten zu generieren wird eine zeitliche Disaggregation angewendet. Die disaggregierten Stundendaten müssen hierbei verschiedene Eigenschaften aufweisen:

- *Räumliche Konsistenz:* Das Zielprodukt ist ein Raster und muss als solches räumliche Konsistenz aufweisen. Zwei benachbarte Messstationen weisen untereinander typischerweise eine hohe Korrelation der Zeitreihen auf. Je weiter zwei Stationen voneinander entfernt sind, desto geringer ist die Korrelation untereinander. Das gleiche muss auch für das finale Rasterprodukt gelten.
- *Zeitliche Korrelation:* Für Anwendungen, die auf dem Produkt basieren, ist eine hohe zeitliche Korrelation mit dem tatsächlichen stündlichen Niederschlag entscheidend. Das Rasterprodukt soll nicht nur eine gute Übereinstimmung in der Verteilung zeigen, sondern gezielt die Stunden mit intensivem Niederschlag korrekt abbilden. Ebenso sollten niederschlagsfreie Stunden möglichst keinen Niederschlag im Rasterprodukt aufweisen. Daher wird ein deterministischer Ansatz verfolgt (siehe Abschnitt 3.2.2.1), der eine Nutzung der Daten, beispielsweise für Niederschlags-Abfluss-Modellierung oder Hochwassermodellierungen, ermöglicht.
- *Übereinstimmung in der Verteilung:* Zusätzlich zu genannten Punkten soll auch die Verteilung der Niederschlagsmengen pro Stunde möglichst der tatsächlichen Verteilung entsprechen. Hohe Werte sollen ähnlich oft im Rasterprodukt vorhanden sein wie sie tatsächlich auftreten. Die Anzahl der Niederschlagsstunden soll ebenso getroffen werden wie die mittlere Verteilung des Niederschlages über den Tag.

Während die Verteilung bei gängigen Disaggregationsmodellen typischerweise sehr gut getroffen wird, stellt die räumliche Konsistenz eine deutlich größere Herausforderung dar. Bei verschiedenen Formen der zur Disaggregation eingesetzten Kaskadenmodelle ist eine räumliche Konsistenz nicht vorhanden.

Ist diese aber notwendig, so wird auf ein Resampling zurückgegriffen: Die Kaskaden zur Generierung stündlicher Niederschläge werden so lange wiederholt durchlaufen, bis eine vordefinierte Korrelation zwischen den verschiedenen Zeitreihen erreicht wird. Ein solches Resampling ist auf einige wenige (weniger als 10) Stationen beschränkt, da der Zeitaufwand zur Berechnung mit jeder zusätzlichen Station exponentiell wächst. Eine Anwendung für Sachsen in 100 Meter Auflösung bedeutete mehr als 1,8 Millionen Pseudostationen, wofür Resampling gänzlich ungeeignet ist.

Abgesehen davon, sind existierende Disaggregierungsmodelle nicht in der Lage, stündliche Niederschläge so zu bestimmen, dass die Produkte mit den tatsächlich auftretenden Niederschlägen korrelieren. Da keine offen verfügbare Methode den Anforderungen entspricht, wird im Folgenden eine Eigenentwicklung vorgestellt und angewandt, die diese Herausforderungen adressiert und die genannten Kriterien erfüllt.

3.2.2.1 Methodischer Ansatz

Zur Bestimmung des stündlichen Niederschlages werden zwei Modelle erstellt. Beide Modelle sind deterministische Modelle, basierend auf Gradient Boosting (Chen und Guestin, 2016). Das erste Modell nutzt als Zielgröße den gemessenen stündlichen Niederschlag aller DWD-Stationen. Insgesamt liegen für alle Messstationen ca. 87 Millionen Werte (Stundenniederschlag an Niederschlagstagen) vor, so dass das Modell auf einer soliden Datengrundlage basiert. Unabhängige Größen bzw. Prädiktoren sind die anderen meteorologischen Größen, also Temperatur, Windgeschwindigkeit, Luftfeuchte, Luftdruck und Globalstrahlung. Diese werden mittels Tps für alle Zeitpunkte und Stationen interpoliert, so dass für jede Niederschlagsmessstation für jeden Zeitpunkt, an dem Niederschlagswerte gemessen wurden, auch die interpolierten Werte der anderen meteorologischen Größen vorliegen. Das Modell wird anschließend für den Niederschlag mit der folgenden Aufgabe trainiert: Wie hoch ist der Niederschlag für eine bestimmte Stunde, wenn für diesen Zeitpunkt die zuvor bestimmten (interpolierten) meteorologischen Größen sowie für die vorhergehenden zwei und die nachfolgenden zwei Stunden bekannt sind?

Das zweite Modell hat nicht den Stundenniederschlag, sondern die Anzahl der Niederschlagsstunden pro Tag als Zielgröße. Prädiktoren sind die Tageswerte der genannten meteorologischen Größen. Wieder wird das Modell für die gemessenen Niederschlagsstunden des DWD und die interpolierten Werte der übrigen Größen trainiert.

Gradient Boosting Ansätze basieren auf der Grundidee, viele einfache Entscheidungsbäume zu kombinieren, um eine präzise Vorhersage zu treffen. Anders als bei einem einzelnen, großen Baum, der versucht, das Problem auf einmal zu lösen, baut Gradient Boosting eine Reihe kleiner Bäume auf. Diese lernen in der Trainingsphase schrittweise aus den Fehlern der vorherigen Bäume und verbessern sich so nach und nach.

Zu Beginn startet der Algorithmus mit einer einfachen Vorhersage, die sehr rudimentär ist, beispielsweise dem Durchschnittswert der Zielvariable. Anschließend analysiert der Algorithmus, wie gut diese Vorhersage war, indem er die Fehler, auch Residuen genannt, berechnet. Diese Fehler zeigen an, wie stark sich die aktuelle Vorhersage von den tatsächlichen Werten unterscheidet und geben somit Hinweise darauf, was der Algorithmus noch lernen muss.

Um diese Fehler zu korrigieren, wird ein neuer kleiner Entscheidungsbaum erstellt. Dieser Baum konzentriert sich speziell auf die fehlerhaften Vorhersagen und versucht, diese zu verbessern. Durch das Hinzufügen der Vorhersagen dieses Baums zur ursprünglichen Schätzung entsteht eine leicht verbesserte Gesamtvorhersage.

Dieser Vorgang wird dann einige hundert Mal wiederholt. Jeder neue Baum korrigiert die verbleibenden Fehler ein kleines Stück mehr, bis am Ende ein Modell entsteht, das die Daten sehr genau beschreibt. Das Besondere an diesem Ansatz ist, dass die Fehlerreduktion systematisch erfolgt, wobei mathematische Methoden wie der Gradientenabstieg genutzt werden, um die Richtung der Optimierung zu bestimmen. So wird das Modell Schritt für Schritt immer besser.

Ist das Modell aufgebaut („angelernt“, „trainiert“), so wird es für den Zeitraum angewendet, für den für die Zielgröße an der jeweiligen Station keine Messwerte zur Verfügung stehen. Mittels der genannten unabhängigen Größen wird das Modell angetrieben, Ergebnis sind Schätzwerte der Zielgröße, also in diesem Fall Niederschlagsstunden pro Tag bzw. Niederschlagsmenge pro Stunde.

3.2.2.2 Anwendung Disaggregation

Die in 3.2.2.1 erstellten Modelle werden für die Jahre 1961-2023 für das Gebiet Sachsens angewendet. Als Prädiktoren stehen die in 3.2.1 erstellten Raster der übrigen meteorologischen Größen zur Verfügung. Für die spätere Skalierung mit den Tagessummen des Niederschlages ist zu beachten, dass durch die Messung des Tagesniederschlages um 7 Uhr es bei der Anzahl der Niederschlagsstunden ebenfalls um die Anzahl der Niederschlagsstunden von 7 bis 7 Uhr handelt. Eine entsprechende Verschiebung um sieben Stunden wurde in der Bearbeitung berücksichtigt.

Durch die Anwendung der Modelle liegen für jede Rasterzelle stündlich modellierte Niederschlagswerte sowie eine modellierte Anzahl der Niederschlagsstunden pro Tag vor. Zusätzlich ermöglicht die durchgeführte Kreuzvalidierung den Vergleich zwischen gemessenen und modellierten Werten: Für jeden beobachteten stündlichen Niederschlag existiert ein entsprechender modellierter Wert, ebenso wie für die Summe der gemessenen Niederschlagsstunden pro Tag ein modellierter Vergleichswert vorliegt.

Auf Basis dieser Daten wird ein geglättetes QQ-Mapping mittels Splines durchgeführt. Dadurch wird die Verteilung sowohl der Niederschlagsstunden als auch der stündlichen Niederschlagssummen präziser modelliert.

Das weitere Vorgehen erfolgt in mehreren Schritten:

- **Bestimmung der Niederschlagsstunden:** Für jeden Tag wird überprüft, wie viele Niederschlagsstunden das Modell für eine Rasterzelle prognostiziert hat.
- **Selektion relevanter Stundenwerte:** Wurden beispielsweise zehn Niederschlagsstunden modelliert, bleiben die zehn höchsten stündlichen Niederschlagssummen erhalten, während die verbleibenden 14 Stundenwerte auf 0 mm gesetzt werden.
- **Skalierung der Stundensummen:** Die interpolierte Tagesniederschlagssumme, die bereits bekannt ist, dient als Referenz. Falls die Summe der modellierten stündlichen Niederschlagswerte für einen Tag zunächst 5 mm, die Tagesniederschlagssumme jedoch 10 mm beträgt, werden alle stündlichen Werte mit dem Faktor 2 skaliert.

Durch dieses Verfahren wird sichergestellt, dass die stündlichen Niederschlagssummen mit dem Rasterprodukt der Tagesniederschlagssummen konsistent sind. Diese Methode wird für jede der über 1,8 Millionen Rasterzellen und jede der ca. 550.000 Zeitschritte angewendet. Ergebnis ist ein Rasterprodukt für stündliche Niederschläge für Sachsen in 100 Meter Auflösung.

3.2.2.3 Kreuzvalidierung

In diesem Abschnitt wird gezeigt, in wie weit das erstellte Rasterprodukt für den Stundenniederschlag den unter 3.2.2 formulierten Anforderungen genügt. Zur Validierung wird eine n-fache Kreuzvalidierung mit $n=20$ verwendet. Konkret bedeutet das, dass die im Abschnitt 3.2.2.1 beschriebenen Schritte für die Stationsdaten 20-mal durchgeführt wurden. Immer wurden 5% aller Stationen bei der Modellerstellung nicht verwendet. Das Modell wurde für die verbleibenden 95% der Stationen trainiert und anschließend auf die nicht zum Training verwendeten Stationen angewendet. Dadurch werden Wertepaare aus gemessenen und modellierten stündlichen Niederschlagssummen erstellt, die anschließend miteinander verglichen und ausgewertet werden können.

Zur Einordnung der Qualität der Kreuzvalidierung wurde zusätzlich eine Kreuzvalidierung der räumlichen Interpolation des stündlichen Niederschlages 1995-2023 durchgeführt. Diese dient als Referenz.

Räumliche Konsistenz

Jede Station, die zeitgleich mit einer anderen Station Messdaten erhebt, weist eine bestimmte Korrelation zu dieser Station auf. Bei der Kreuzvalidierung mit insgesamt über 1.100 Stationen gibt es so über 660.000 Stationspaare, für die sich die Korrelationen zueinander berechnen lassen. Kriterium ist, dass für jeweils beide Stationen jeweils mindestens ein Jahr (8760 Stunden) gleichzeitige Messungen vorliegen.

Diese Korrelation lassen sich für die Messwerte ebenso bestimmen wie für die modellierten Werte. Werden diese Korrelationen jeweils auf einander abgetragen, so entsteht ein Scatterplot, der eine Aussage über die räumliche Konsistenz erlaubt.

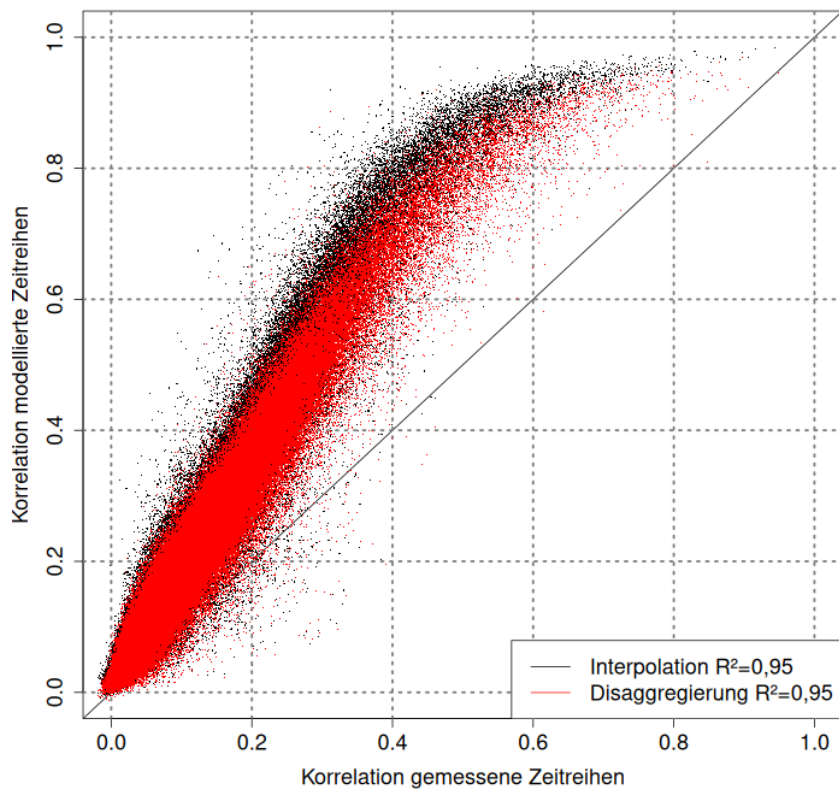


Abbildung 2: Korrelation-Korrelation-Plot zur Darstellung der räumlichen Konsistenz der berechneten stündlichen Niederschlagszeitreihen.

In Abbildung 2 ist die räumliche Konsistenz als Korrelation-Korrelation-Plot dargestellt. Bei einem perfekten Modell lägen alle Punkte auf der schwarzen 1:1 Geraden. Das heißt, die Korrelation aller Stationen zueinander in dem Modell wären exakt gleich groß wie die Korrelationen der gemessenen Zeitreihen der Stationen. Das Bestimmtheitsmaß R^2 wäre 1,0. Die räumliche Konsistenz eines realen Modells ist demnach besser, je näher die Punkte an der 1:1 Geraden liegen und je höher das Bestimmtheitsmaß von gemessenen und modellierten Korrelationen der Stationen untereinander ist. Die Bestimmtheitsmaße der räumlichen Interpolation erreichen hier ebenso wie die der Disaggregation Werte von 0,95 und sind demnach von einer ähnlich guten, sehr hohen Qualität. Zum Vergleich: Das Radarprodukt des DWD RADOLAN erreicht ein Bestimmtheitsmaß von 0,47 (Bárdossy et al., 2008). Die roten Punkte der disaggregierten Stundenwerte liegen dabei im Mittel etwas näher an der 1:1 als die schwarzen Punkte der räumlichen Interpolation.

Zeitliche Korrelation

Um die Güte der einzelnen Zeitreihen zu bestimmen, wird für jede Station der RMSE, das Bestimmtheitsmaß R^2 , die mittlere Abweichung (Bias) sowie der Heidke Skill Score (HSS) bestimmt. Zusätzlich wurde das Verhältnis von modellierten zu gemessenen Niederschlagsstunden sowie eine Trefferquote bestimmt, die angibt, zu welchem Anteil das korrekte Ereignis (Niederschlag, kein Niederschlag) modelliert wurde (Tabelle 5).

Tabelle 5: Kennwerte für die Kreuzvalidierung der räumlichen Interpolation und der Disaggregation

	Räumliche Interpolation	Disaggregation
RMSE in mm/d	0,359	0,387
Bias in mm	0,004	0,002
R^2	0,501	0,443
HSS	0,633	0,642
Verhältnis Niederschlagsstunden	1,614	0,983
Trefferquote	0,906	0,928

Der mittlere quadratische Fehler der räumlichen Interpolation ist niedriger und das Bestimmtheitsmaß höher und damit jeweils besser als bei der Disaggregation. Besser schneidet die Disaggregation beim Bias, dem Heidke Skill Score, dem Verhältnis der Niederschlagsstunden sowie bei der Trefferquote ab.

Übereinstimmung der Verteilung

Abbildung 3 zeigt den Quantil-Quantil-Plot der gemessenen und modellierten Niederschlagszeitreihen. Bei einem perfekten Modell lägen die Punkte alle auf der schwarzen 1:1 Gerade. Bei realen Modellen gilt, dass die Verteilungen besser übereinstimmen, je näher die Punkte an der 1:1 Geraden liegen. Es ist im Vergleich zu sehen, dass die Verteilung der räumlichen Interpolation und der Disaggregation nahe beieinander liegen. Die disaggregierten Werte liegen dabei fast durchgehend näher an der 1:1 Gerade als die Werte der räumlichen Interpolation.

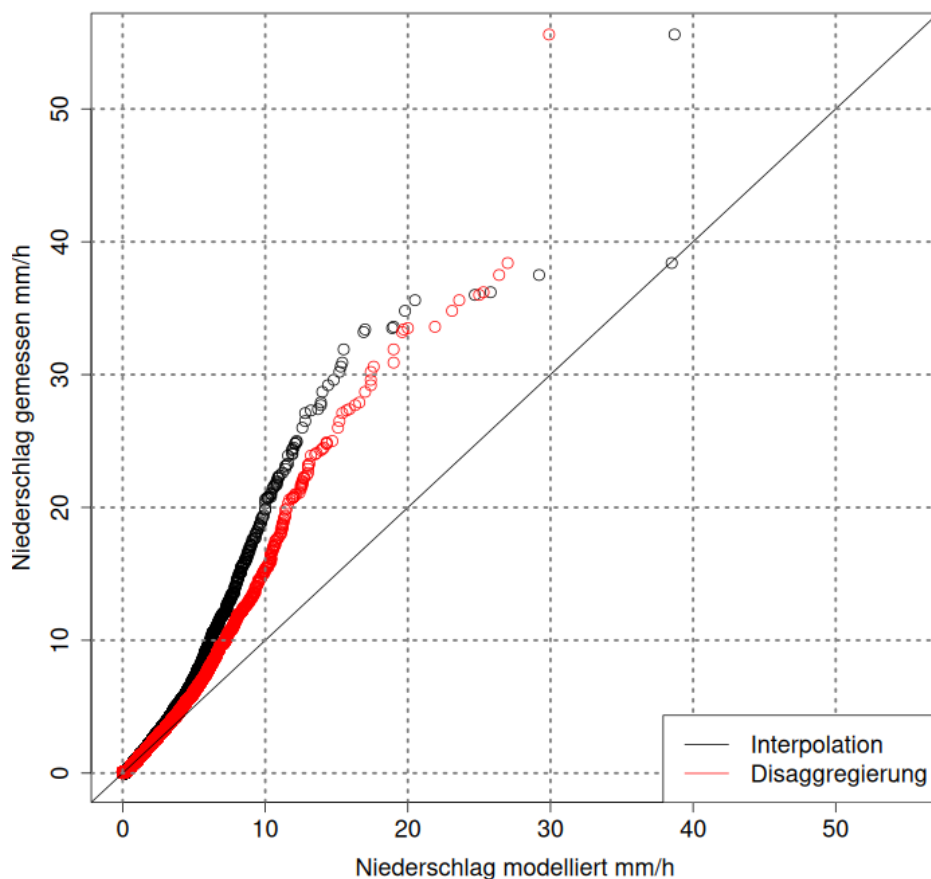


Abbildung 3: QQ-Plot von modelliertem und gemessenem Niederschlag in stündlicher Auflösung.

Fazit

Die in diesem Bericht vorgestellte Disaggregierungsmethode befindet sich etwa auf dem gleichen Niveau wie die räumliche Interpolation in Bezug auf räumliche Konsistenz, zeitliche Korrelation und statistische Verteilung. Bei einigen Kennwerten schneidet die eine, bei anderen die andere Methode besser ab. Dies unterstreicht die hohe Qualität der vorgestellten Methode zur Disaggregation. Eine definitive Entscheidung, welche Methode zu bevorzugen sei, kann hier nicht getroffen werden, da diese Entscheidung je nach konkretem Anwendungsfall anders zu beantworten wäre.

Für den hier beschriebenen Anwendungsfall beantwortet die Kreuzvalidierung die Frage, ob es sinnvoller wäre, ab den 2000er Jahren anstelle der Disaggregation die räumliche Interpolation des Niederschlags zu verwenden. Die Antwort lautet nein, da dies zu Inkonsistenzen im Produkt führen würde und die räumliche Interpolation nach der für die anderen Größen beschriebene Methode insgesamt nicht besser ist als die Disaggregation.

Schlussfolgerungen

Die Untersuchung zeigt, dass das ESRI-Produkt eine höhere Übereinstimmung mit dem Referenzprodukt aufweist und daher für die weitere Verwendung empfohlen wird. Die entwickelte Methode zur Erstellung von stündlichen Niederschlagsdaten in hoher räumlicher Auflösung basierend auf Gradient Boosting zeigt eine hohe Qualität und kann für verschiedene Anwendungen in der Meteorologie und Hydrologie genutzt werden. Die Ergebnisse unterstreichen die Bedeutung der Verwendung hochauflösender Daten für eine genaue Modellierung und Prognose.

Ausblick

Die erstellten Daten bieten eine solide Grundlage für raum-zeitlich hochauflösende Modellierungen in der Vergangenheit. Zukünftige Arbeiten sollten sich auf die Anwendung dieser Daten konzentrieren, um die Auswirkungen von Landbedeckungsänderungen auf das Klima und den Wasserhaushalt besser zu verstehen. Die Entwicklung weiterer Methoden zur Verbesserung der räumlichen und zeitlichen Auflösung der Daten könnte ebenfalls von großem Nutzen sein. Die Ergebnisse dieser Studie bieten wertvolle Einblicke und legen den Grundstein für zukünftige Forschungen in diesem Bereich.

Literaturverzeichnis

- BÁRDOSSY, A., & BROMMUNDT, J. (2008). Erzeugung simultan-synthetischer Niederschlagsreihen in hoher zeitlicher und räumlicher Auflösung für Baden Württemberg [Forschungsbericht]. Universität Stuttgart.
- KÖRNER, P., KRONENBERG, R., GENZEL, S., & BERNHOFER, C. (2018). Introducing Gradient Boosting as a universal gap filling tool for meteorological time series. *Meteorologische Zeitschrift*, 369–376. <https://doi.org/10.1127/metz/2018/0908>
- HANELLI, D., & BARTH, A. (2023). Landbedeckung in Sachsen 1985-2020 (Teil 2). Heft 07/2023.
- KARRA, K., KONTGIS, C., STATMAN-WEIL, Z., MAZZARIELLO, J. C., MATHIS, M., & BRUMBY, S. P. (2021). Global land use / land cover with Sentinel 2 and deep learning, in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 4704–4707, <https://doi.org/10.1109/IGARSS47720.2021.9553499>.
- KÖRNER, P., KRONENBERG, R., GENZEL, S., & BERNHOFER, C. (2018). Introducing Gradient Boosting as a universal gap filling tool for meteorological time series. *Meteorologische Zeitschrift*, 369–376. <https://doi.org/10.1127/metz/2018/0908>
- KÖRNER, P., & MODEROW, U. (2025). Klima- Referenzdatensatz Sachsen 1961—2023 (Schriftenreihe des LfULG).
- KÖRNER, P., VOROBESKII, I., KRONENBERG, R., & HOMOUDI, A. (2022). Klima- Referenzdatensatz Sachsen 1961—2020 (18/2022; Schriftenreihe des LfULG, Nummer 18/2022, S. 88).
- KUGLER, L., MARRS, C., KOSZOR, E., & FORKEL, M. (2023). Landbedeckung in Sachsen 1961-1979 (Teil1). Heft 06/2023.

Herausgeber

Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG)

Pillnitzer Platz 3, 01326 Dresden

Telefon: +49 351 2612-0; Telefax: +49 351 2612-1099

E-Mail: poststelle@lfulg.sachsen.de

www.lfulg.sachsen.de

Autor

Dr. Philipp Körner

iamk GmbH

Institut für Angewandte Meteorologie und Klimatologie

Dr.-Wilhelm-Külz-Straße 70, 01896 Pulsnitz

Telefon: +49 176 45747729

E-Mail: philipp.koerner@iamk-gmbh.de

Redaktion

Florian Kerl

Abteilung 5 / Referat 55

Söbrigener Str. 3a, 01326 Dresden-Pillnitz

Telefon: +49 351 2612-5502; Telefax: +49 351 2612-5099

E-Mail: florian.kerl@lfulg.sachsen.de

Bildnachweis

Helbig, Mario (Titelseite)

Redaktionsschluss

31.07.2025

ISSN

1867-2868

Bestellservice

Die Broschüre steht nicht als Printmedium zur Verfügung, kann aber als PDF-Datei heruntergeladen werden aus der Publikationsdatenbank des Freistaates Sachsen (<https://publikationen.sachsen.de>).

Hinweis

Diese Publikation wird im Rahmen der Öffentlichkeitsarbeit vom LfULG (Geschäftsbereich des SMEKUL) kostenlos herausgegeben. Sie ist nicht zum Verkauf bestimmt und darf nicht zur Wahlwerbung politischer Parteien oder Gruppen eingesetzt werden.

Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes.

Täglich für ein gutes Leben.

www.lfulg.sachsen.de