Bundesamt für
Kartographie und Geodäsie

Dokumentation

Digitaler Zwilling Deutschland Datenprodukte

DigiZ-DE



Produktstand 2022

Inhaltsverzeichnis

Inhaltsverzeichnis **2**
1 Übersicht über den Datenbestand **3**
2 Beschreibung des Datenbestandes **4**
2.1 Feedback 4
2.2 Inhalt 4
2.3 Änderungen gegenüber Vorgängerdatensatz 5
2.4 Beschreibung der Themen 6
2.4.1 Einzelbäume 6
2.4.2 Windkraftanlagen 9
3 Qualität **10**
3.1 Lage- und Höhengenauigkeit 10
3.2 Aktualität 10
4 Weiterführende Informationen **10**
5 Quellen **11**

1 Übersicht über den Datenbestand

Produkt:	Digitaler Zwilling Deutschland Datenprodukte
Inhalt:	Das Produkt enthält folgende Themen/Datensätze: - Einzelbäume (digiz-de_eb) - Windkraftanlagen (digiz-de_wka)
Gebiet:	Die Ausdehnung des Gebiets beträgt rund 8670 km² und erstreckt sich in der Metropolregion Hamburg von Helvesiek (Südwesten) über Itzehoe (Nordwesten) und Travemünde (Nordosten) bis nach Tosterglope (Südosten). Rechteck (528000,5895000) : (627000, 5983000)
Räumliche Gliederung:	Gesamte Gebietsausdehnung
Georeferenzierung:	UTM-Abbildung in Zone 32 Ellipsoid GRS80, Datum ETRS89, EPSG 25832
Lagegenauigkeit:	20 cm
Aktualität:	2022-02, Details siehe Kapitel 3.2. Aktualität
Datenformate*:	GeoPackage (GPKG)
Bereitstellung*:	Download
Änderungen gegenüber letztem Datensatz:	Daten werden erstmalig bereitgestellt
Historische Daten:	nein
Datenvolumen:	Einzelbäume: 22 GB Windkraftanlagen: 0.2 MB
Datenquellen:	Punktwolke aus LiDAR-Befliegung Digitaler Zwilling Deutschland, https://sg.geodatenzentrum.de/web_public/gdz/datenquellen/datenquellen_digiz-de.pdf

* Bitte beachten Sie, dass nicht über jede Bereitstellungsform alle Georeferenzierungen und Datenformate zur Verfügung gestellt werden können. Wenden Sie sich bei Fragen gern an das Dienstleistungszentrum.

2 Beschreibung des Datenbestandes

2.1 Feedback

Bei den bereitgestellten Datenprodukten handelt es sich um die ersten abgeleiteten Produkte aus den LiDAR-Daten des Digitalen Zwilling Deutschland. Wir möchten Sie daher bitten, uns Feedback zu den Produkten zu geben. Ihre Rückmeldungen, ob Kritik, Anregungen oder positives Feedback, helfen uns, Qualität und Nutzbarkeit der Daten kontinuierlich zu verbessern.

Für Fragen, Kritiken oder Anregungen nutzen Sie bitte unser Feedbackformular: https://social.bscw.bund.de/pub/bscw.cgi/146241883

2.2 Inhalt

Das Produkt „Digitaler Zwilling Deutschland Datenprodukte“ des Bundesamtes für Kartographie und Geodäsie (BKG) beinhaltet verschiedene Themen, die im Rahmen des Digitalen Zwilling Deutschland (DigiZ-DE) erstellt werden. Mit unterschiedlichen Methoden werden Objekte zu einzelnen Themen aus der Punktwolke abgeleitet und als 2D-Vektordaten bereitgestellt.
Die Grundlage für alle derzeitigen Themen stellt dabei die Punktwolke aus der flugzeuggestützten Befliegung mit LiDAR-Sensoren des DigiZ-DE aus dem Demonstrationsprojekt in der Metropolregion Hamburg dar.

In Tabelle 1 werden die wichtigsten Eigenschaften der Punktwolke beschrieben.
In Abbildung 1 ist die Ausdehnung des Datensatzes dargestellt.

Tabelle 1 Eigenschaften der Punktwolke DigiZ-DE

Aufnahmezeitraum	Oktober 2021 – Februar 2022
Geographische Ausdehnung (EPSG:25832)	West: 528000 Süd: 5895000 Ost:627000 Nord: 5983000
Punktdichte	Mind. 42 Punkte pro m²
Klassen (orientieren sich an den Standards der AdV und der ASPRS)	0: Nie klassifiziert 1: Unklassifiziert (0-0.2 m über Grund) 2: Boden 3: Niedrige Vegetation (0.2 - 1.8 m über Grund) 4: Mittlere Vegetation (1.8-8 m über Grund) 5: Hohe Vegetation (> 8 m über Grund) 6: Gebäude 7: Tiefpunkt 9: Wasser 18: Hohes Rauschen
Höhengenauigkeit	<10 cm
Lagegenauigkeit	<20 cm
LiDAR-Sensor	SPL100

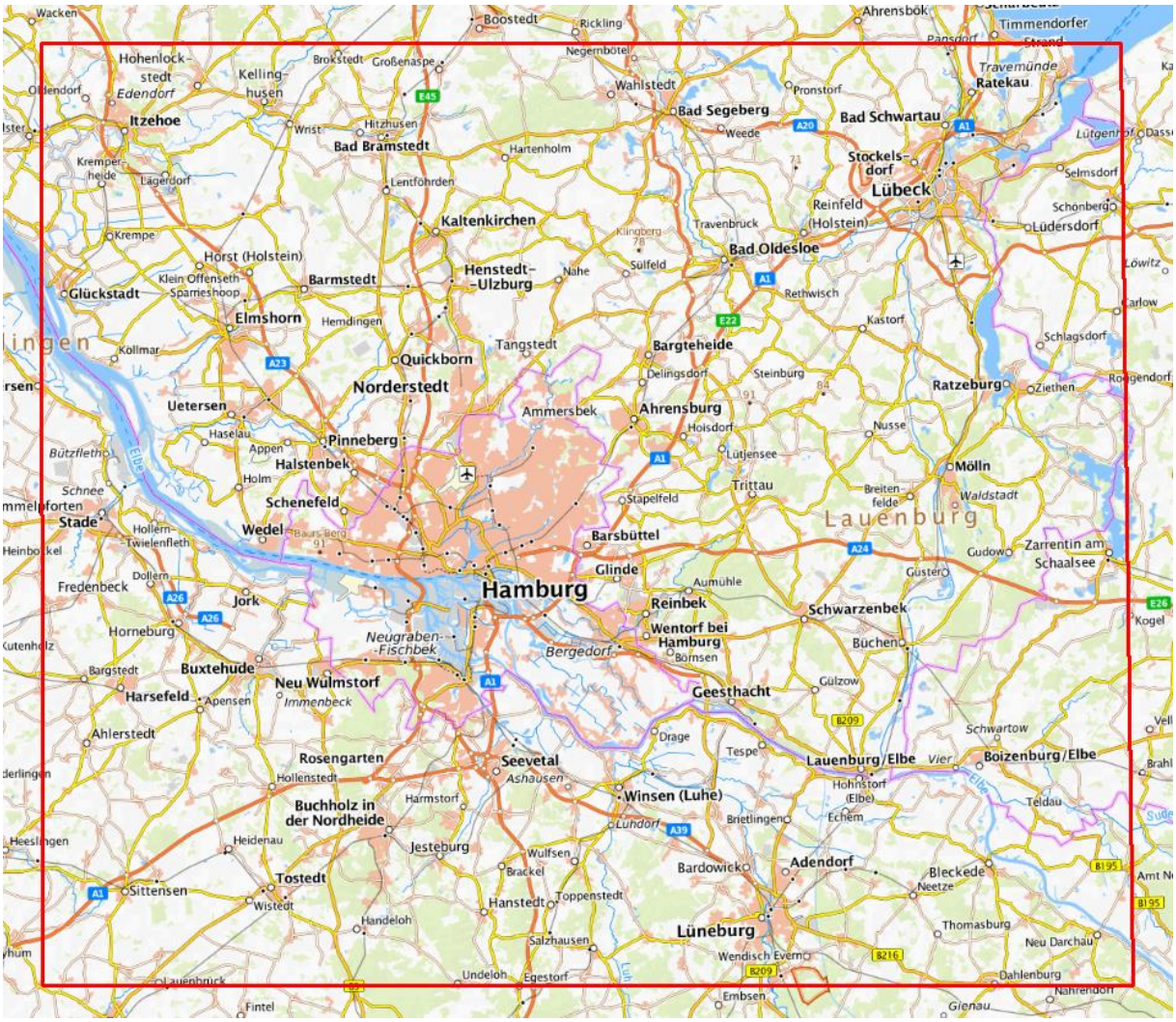


Abbildung 1: Ausdehnung des Datensatzes

2.3 Änderungen gegenüber Vorgängerdatensatz

Der Datensatz wurde erstmalig erstellt.

2.4 Beschreibung der Themen

Im Produkt DigiZ-DE Datenprodukte werden folgende Themen bereitgestellt:

- Einzelbäume
- Windkraftanlagen

Das Datenmodell zur Ableitung der Themen ist in Abbildung 2 gezeigt.

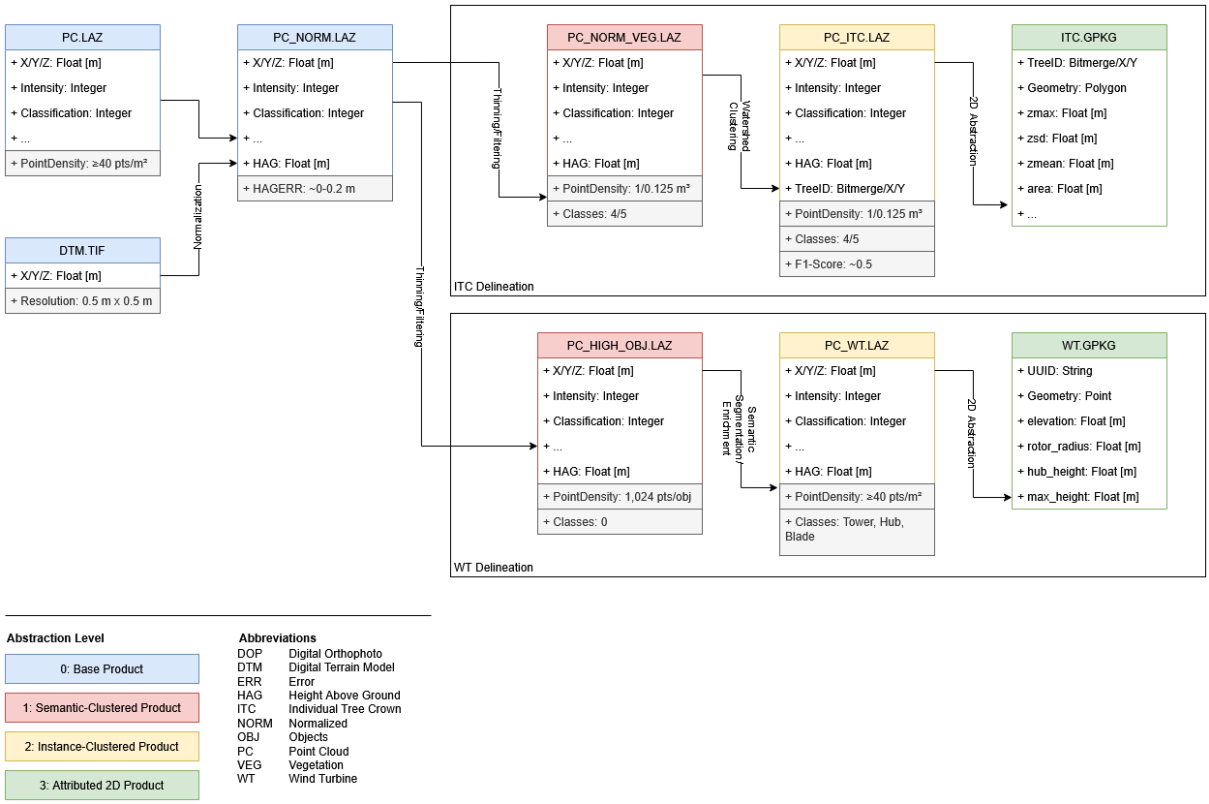


Abbildung 2: Datenmodell zur Ableitung von Einzelbaum- und Vegetationsmerkmalen (ITCs) sowie Windkraftanlagen (WT) aus der klassifizierten LiDAR-Punktwolke des Digitalen Zwillings Deutschland (DigiZ-DE).

2.4.1 Einzelbäume

2.4.1.1 Inhalt

Der Datensatz beinhaltet flächendeckende Einzelbaum- und Vegetationsstrukturinformationen in Form von attribuierten 2D-Polygonen, die aus den flugzeuggestützten LiDAR-Daten des DigiZ-DE-Projekts abgeleitet wurden. Er dient als Grundlage für forstliche, ökologische und stadtplanerische Analysen.

2.4.1.2 Attribute

Tabelle 2 Attribute der Einzelbaum- und Vegetationscluster

Attribut	Beschreibung
treeID	Eindeutige ID je Cluster
zmax	Maximale Punkthöhe des Clusters [m]
zmin	Minimale Punkthöhe [m]
zmean	Mittlere Höhe [m]

zsd	Standardabweichung der Höhenverteilung [m]
zskew	Schiefe der Höhenverteilung
zkurt	Kurtosis der Höhenverteilung
zentropy	Entropie der Höhenverteilung
pzabove2	Prozentanteil der Punkte > 2 m
pzabovezmean	Prozentanteil der Punkte > zmean
zq5, zq10, …, zq95	Höhe [m] auf dem x-ten Perzentil
zpcum1, zpcum2, …	Kumulative Prozentanteile nach Höhenklasse
itot	Gesamtsumme der Intensitätswerte aller Punkte
imax	Maximaler Intensitätswert
imean	Mittlerer Intensitätswert
isd	Standardabweichung der Intensitätsverteilung
iskew	Schiefe der Intensitätsverteilung
ikurt	Kurtosis der Intensitätsverteilung
ipground	Anteil der Intensität der Bodenpunkte relativ zur Gesamtintensität
ipcumzq10, ipcumzq30, …	Kumulative Intensität nach Höhenperzentilen
p1th, p2th, …	Anteil der Intensität nach Return Type
pground	Anteil der Bodenpunkte an der Gesamtanzahl
n	Gesamtanzahl der Punkte im Cluster
area	Fläche [m²]
perimeter	Umfang [m]
compactness	Kompaktheit
point_density	Anzahl der Punkte pro Quadratmeter innerhalb des Clusters [pts/m²]

2.4.1.3 Ableitungsmethode

Die Verarbeitungsschritte zur Ableitung von Einzelbaum- und Vegetationsstrukturinformationen aus der Punktwolke sind in Abbildung 2 gezeigt. Die Segmentierung wurde anhand der vorklassifizierten Punktwolke mit dem R-Paket lidR (Roussel et al., 2020; Roussel & Auty, 2025) zur LiDAR-Verarbeitung, dem terra-Paket (Hijmans, 2025) und dem sf-Paket (Pebesma, 2018; Pebesma & Bivand, 2023) für räumliche Analysen durchgeführt. Die Segmentierung basiert auf der Watershed-Methode aus dem EBImage-Paket (Pau et al., 2010). Die gesamte Analyse erfolgte in R (R Core Team, 2021). Alle Höhenwerte (Z-Werte) der Punktwolke wurden zunächst durch Subtraktion der jeweiligen Geländehöhe im digitalen Geländemodell (DTM) normalisiert. Anschließend wurde die normalisierte Punktwolke nach Vegetationspunkten (Klasse 4 „mittlere Vegetation“ und Klasse 5 „hohe Vegetation) und einer Mindesthöhe von zwei Metern gefiltert und zur Reduzierung des Rechenaufwands randomisiert mit einem Voxelgitter mit 0,5 Metern Kantenlänge ausgedünnt. Die Instanzsegmentierung der übrigen gefilterten und normalisierten Vegetationspunkte wurde mittels Watershed-Algorithmus umgesetzt. Das Verfahren nutzt das Prinzip der Einzugsgebiete in einem aus den Vegetationspunkten generierten Höhenmodell, bei dem lokale Maxima als potentielle Baumspitzen interpretiert und die umliegenden Punktwolkenbereiche den entsprechenden Baumspitzen zugeordnet werden. Abschließend wurden durch Ableitung der konvexen Hülle jedes Vegetationscluster aus der instanzsegmentierten Punktwolke 2D-Polygone erzeugt und mit 3D-Statistiken des jeweiligen Clusters attribuiert.

2.4.1.4 Qualität: Logische Konsistenz, Vollständigkeit und inhaltliche Genauigkeit

Alle Einzelbaum- und Vegetationspolygone sind vollständig mit allen relevanten Attributen erfasst. Eine Filterung der Einzelbaum- und Vegetationspolygone wurde entsprechend Tabelle 3 nach plausiblen Attributausprägungen durchgeführt. Die Qualität des Datensatzes unterliegt jedoch mehreren Einflussgrößen wie der Korrektheit der semantischen Klassifikation, der Normalisierung, der Ausdünnung der Punktwolke sowie der Güte der gewählten Methode zur Instanzsegmentierung. Aufgrund der hohen Komplexität der Wechselwirkungen und der großen Ausdehnung des abgeleiteten Gebiets ist eine flächendeckende Quantifizierung der Qualität schwer möglich. Bei Verwendung des Watershed-Algorithmus tritt bei Überlappung von Baumkronen oder hoher räumlicher Dichte der Vegetation Untersegmentierung auf, wobei mehrere Baumsegmente als ein zusammenhängendes Baumsegment erkannt werden können.

Die Validierung anhand eines kommunalen Baumkatasters mit mehr als 225.000 Bäumen ergab eine durchschnittliche Detektionsrate der Einzelbäume von 54 % und eine durchschnittliche Detektionsrate von 88 % für Baumkronen mit einem Mindestdurchmesser von 10 m. Das Register selbst weist Unsicherheiten bei der Position und Vollständigkeit der Einträge auf, was die exakte Bestimung der Detektionsrate erschwert. Insgesamt bietet der Datensatz eine hohe Qualität für Analysen auf Bestandesebene, während bei einzelnen Bäumen Positionsabweichungen in der Größenordnung von Dezimetern zu erwarten sind.

Tabelle 3: Filterkriterien zur Qualitätskontrolle und Klassifikation der Ergebniscluster basierend auf aus den LiDAR-Daten abgeleiteten Attributen.

Kriterium	Bedingung	Kategorie	Beschreibung
Maximale Höhe (*zmax*)	≤ 1 m	*low_zmax*	Sehr geringe maximale Höhe des Clusters
Maximale Höhe (*zmax*)	≥ 70 m	*high_zmax*	Unplausibel hohe maximale Höhe
Höhenstandardabweichung (*zsd*)	≤ 0,2 m	*low_zsd*	Geringe vertikale Variabilität, ggf. flache Objekte
Fläche (*area*)	≤ 3 m²	*small_area*	Sehr kleine Fläche, möglicherweise Rauschen oder Artefakte
Fläche (*area*)	≥ 2.500 m²	*large_area*	Sehr große Fläche, potenziell fehlerhafte Segmentierung
Fläche & *zsd* kombiniert	≥ 60 m² & *zsd* ≤ 0,6 m	*large_area_low_zsd*	Große Fläche mit geringer Höhenvarianz – ungewöhnliches Objekt
Punktanzahl (*n*)	< 10	*low_point_count*	Sehr wenige Punkte im Polygon – statistisch nicht belastbar
Kompaktheit (*compactness*)	≤ 0,4	*irregular_shape*	Unregelmäßige, stark verzerrte Form
Punktdichte (*point_density*)	< 5 pts/m²	*low_point_density*	Zu geringe Punktdichte für zuverlässige Auswertung

2.4.2 Windkraftanlagen

2.4.2.1 Inhalt

Der Datensatz beinhaltet flächendeckende Informationen zu Windenergieanlagen ab einer Höhe von 40 m über Gelände in Form von attributierten Punkten, die aus den flugzeuggestützten LiDAR-Daten des DigiZ-DE-Projekts abgeleitet wurden. Er dient als Grundlage für Analysen im Bereich Luftfahrthindernisse und erneuerbare Energien.

2.4.2.2 Attribute

Tabelle 4: Attribute der Windkraftanlagen

Attribut	**Beschreibung**
uuid	Universally Unique Identifier [128 bit]
elevation	Höhe des niedrigsten Punkts (DHHN2016) [m]
hub_height	Höhe der Nabe über Gelände [m]
rotor_radius	Radius des Rotors [m]
max_height	Höhe des höchsten Punkts über Gelände [m]

Der Datensatz enthält neben den X- und Y Koordinaten der Einzelanlagen in der Projektion utm32s / EPSG 25832 die in Tabelle 4 aufgeführten Attribute.

2.4.2.3 Ableitungsmethode

Die Ableitung der Windkraftanlagen mit entsprechenden Attributen aus der klassifizierten Punktwolke ist in Abbildung 2 dargestellt.
Zunächst wurde eine Extraktion hoher Objekte > 40 m über Gelände der Klasse „Unklassifiziert“ auf Basis regelbasierter Methoden durchgeführt (PDAL Contributors, 2025). Im Anschluss wurden die so extrahierten Objekte auf eine Punktanzahl von 1,024 Punkten pro Objekt degradiert und mit dem KI-Modell „PointNet“ in 9 Klassen klassifiziert (Qi et al., 2016). Die so erhaltenen Windkraftanlagen wurden mit einem weiteren KI-Modell „PointNet++“ in semantische Bauteile (Turm, Nabe, Rotor) segmentiert (Qi et al., 2017). Die Bauteile wurden auf ihre ursprüngliche Punktdichte mittels „k-nächste-Nachbarn-Algorithmus“ angereichert (Virtanen et al., 2020) und abschließend geometrisch vermessen. Der Mittelpunkt der Windkraftanlage wurde aus dem Mittelwert der X- und Y-Koordinaten im Bereich zwischen dem 1. und 3. Quartil der Höhe des Turms gebildet. Die Höhe, "elevation", bezeichnet den niedrigsten Punkt der Windkraftanlage in Bezug auf das Höhenreferenzsystem DHHN2026. Die Höhe der Nabe, "hub_height", bezieht sich auf die mittlere Höhe im Bereich zwischen dem 1. und dem 3. Quartil der Höhe der Nabe. Der Rotorradius, "rotor_radius", bezieht sich auf die Distanz zwischen dem am weitesten entfernten Punkt der Rotorblätter und der Position der Nabe, bestehend aus dem Mittelpunkt der Windkraftanlage und der Nabenhöhe.

2.4.2.4 Qualität

Die Vollständigkeit der Windkraftanlagen wurde gegen das Thema Windkraftanlagen im BKG-Produkt „POI Bund“ geprüft. Die Detektionsrate korrekt erfasster Anlagen beträgt 96 %. Durch den zeitlichen Versatz beider Datensätze kann die tatsächliche Genauigkeit schwanken. Wo vorhanden, wurden die Attribute „rotor_radius“, „hub_height“ und „max_height“ mit dem Referenzdatensatz verglichen (Tabelle 5).

Tabelle 5: Validierung der Attribute mit dem Thema Windkraftanlagen des BKG-Produkts "POI Bund" als Referenz

Attribut	**rotor_radius (Rotorradius)**	**hub_height** (Nabenhöhe)	**max_height** (**Maximale Höhe ü. Gelände**)
Anzahl (N)	383	383	574
Root Mean Square Error (RMSE) [m]	2,02	1,61	10,26
BIAS [m]	-0,15	0,6	-7,8
Mean Absolute Error (MAE) [m]	1,39	1,33	8,13
Bestimmtheitsmaß (R^2)	0,97	1,0	0,98

3 Qualität

3.1 Lage- und Höhengenauigkeit

Die Lagegenauigkeit der Geometrien ergibt sich aus der Lagegenauigkeit der Punktwolke, aus der die Vektor-Geometrien abgeleitet werden. Die Lagegenauigkeit der Punktwolke beträgt < 20cm.
Die Höhengenauigkeit der Geometrien ergibt sich aus der Höhengenauigkeit der Punktwolke, aus der die Vektor-Geometrien abgeleitet werden. Die Höhengenauigkeit der Punktwolke beträgt < 10cm.

3.2 Aktualität

Die Aktualität des Datensatzes hängt von der Aktualität der Punktwolke ab. Die zugrundeliegende LiDAR-Befliegung ist zwischen Oktober 2021 – Februar 2022 erfolgt.

4 Weiterführende Informationen

Weiterführende Informationen zum Produkt sind im Geodatenzentrum des Bundesamtes für Kartographie und Geodäsie unter gdz.bkg.bund.de auf der zugehörigen Produktseite bereitgestellt.Detailierte Angaben zur Ableitungsmethodik können auf Nachfrage bereitgestellt werden.

Für Fragen steht Ihnen das Dienstleistungszentrum unter dlz@bkg.bund.de zur Verfügung. Allgemeine Informationen zum Bundesamt für Kartographie und Geodäsie finden Sie auf unserer Homepage www.bkg.bund.de.

5 Quellen

Hijmans, R. J. (2025). *terra: Spatial Data Analysis*. https://rspatial.org/

Pau, G., Fuchs, F., Sklyar, O., Boutros, M., & Huber, W. (2010). EBImage—An R package for image processing with applications to cellular phenotypes. *Bioinformatics*, *26*(7), 979–981. https://doi.org/10.1093/bioinformatics/btq046

PDAL Contributors. (2025). *PDAL Point Data Abstraction Library* [Software]. https://doi.org/10.5281/zenodo.10884408

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439. https://doi.org/10.32614/RJ-2018-009

Pebesma, E., & Bivand, R. (2023). *Spatial Data Science: With Applications in R* (1. Aufl.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016

Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2016). *PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1612.00593

Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *arXiv preprint arXiv:1706.02413*.

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Roussel, J.-R., & Auty, D. (2025). *Airborne LiDAR Data Manipulation and Visualization for Forestry Applications*. https://cran.r-project.org/package=lidR

Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J.-F., De Boissieu, F., & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sensing of Environment*, *251*, 112061. https://doi.org/10.1016/j.rse.2020.112061

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … Vázquez-Baeza, Y. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2