

## Erfassung von Bäumen mittels Laserscan-Daten (LiDAR-Daten) zur Expositionsanalyse entlang des Bundesfernstraßennetzes von Nordrhein-Westfalen

FA 1.201

Forschungsstelle: Beak Consultants GmbH, Freiberg

Bearbeiter: Schipek, M. / Steffen, M. / Hamperl, A.

Auftraggeber: Bundesministerium für Verkehr und digitale Infrastruktur, Bonn

Abschluss: Juni 2020

### 1 Ausgangslage und Zielstellung

Bedingt durch den Klimawandel ist in Zukunft häufiger mit Extremwittersituationen zu rechnen. Insbesondere durch Starkwindereignisse induzierte Sturmwürfe stellen dabei sowohl für Verkehrsinfrastrukturelemente als auch für Verkehrsteilnehmer eine große Gefahr dar. Im Zuge einer Klimawirkungsanalyse soll in einem ersten Schritt unter dem Schwerpunkt "Sturmgefahren" eine Expositionsanalyse von Bäumen entlang des Bundesfernstraßennetzes durchgeführt werden, welche bei Sturmereignissen ein potenzielles Risiko für den Menschen als auch für die Verkehrsinfrastruktur darstellen können.

Das Ziel des vorliegenden Forschungsprojekts ist es deshalb, alle Bäume entlang des Bundesfernstraßennetzes zu erfassen und mit entsprechenden Parametern (Lage, Höhe, Straßenabschnitt, Entfernung zur Straße etc.) zu attributieren. Als Datengrundlage sollen die auf Länderebene verfügbaren digitalen Geländemodelle verwendet werden. Die Methodik ist dabei so zu entwickeln und aufzubereiten, dass der Auftraggeber nach Abschluss des Projekts selbstständig in der Lage ist, die Arbeiten fortzuführen und neue Gebiete hinzuzufügen sowie bereits vorhandene Bereiche zu aktualisieren. Die hierfür bereitzustellenden Werkzeuge sollen in einer ArcGIS-Desktop-Umgebung lauffähig sein.

### 2 Kenntnisstandsanalyse

Zur Vorbereitung des Projekts und als Basis der Methodikentwicklung wurde eine Analyse zum aktuellen Kenntnisstand durchgeführt. Der Fokus diesbezüglich lag bei den Themen zur

- Extraktion von Objekten aus LiDAR-Daten sowie digitalen Orthofotos,
- Erstellung und Optimierung eines Kronenhöhenmodells,
- Verwendung des Normalized Difference Vegetation Index zur Unterscheidung vegetationsloser und vegetationsbedeckter Flächen sowie der Differenzierung von Baumtypen sowie den
- Möglichkeiten zur Validierung und Beurteilung der Ergebnisse.

In der überwiegenden Anzahl der gesichteten Veröffentlichungen zur Baumerkennung werden dabei sogenannte Kronenhöhenmodelle genutzt sowie die Aspekte hinsichtlich deren Generierung, Aufbereitung und Methoden zur Extraktion der Baumkronen untersucht. Für Modelle, welche auf dieser klassischen Herangehensweise beruhen, sind in der Regel Erkennungsraten zwischen 60 – 80 % für Einzelbäume zu erwarten.

Eine große Herausforderung auf Basis dieser Herangehensweise stellt insbesondere die Minimierung der Errors of Commission (falsch erkannte Bäume) und der Errors of Omission (nicht erkannte Bäume) dar, welche von folgenden Faktoren abhängen:

- Variation der Kronengröße beziehungsweise des Kronenumfangs,
- Irregularitäten der Baumkronen im Kronenhöhenmodell,
- Lage und Dichte/Stand der Bäume im Verbund,
- Überfliegszeiträume.

Um die Fehler so gering wie möglich zu halten, wurden verschiedene Möglichkeiten und Ansätze in der Literatur diskutiert und evaluiert. Eine einheitliche Vorgehensweise zur Optimierung der Modelle existiert jedoch nicht. Die Kernaussage ist dabei jedoch, dass keine pauschal beste Methodik existiert, sondern das auszuwählende Verfahren stets vom Zweck abhängt. So eignen sich einige Ansätze besser für die Detektion von dominanten, herausragenden Bäumen, andere wiederum sind exakter bei der Erkennung von kleinen Bäumen oder Vegetation unterhalb der obersten Baumschicht.

Weiterhin existieren alternative Herangehensweisen auf Basis unterschiedlicher Datengrundlagen, wie zum Beispiel objektorientierte Ansätze, Segmentierung von 3-D-Punktwolken mit zusätzlicher Unterstützung terrestrischer Messungen oder bildbasierte Mustererkennungsverfahren.

Seit der flächendeckenden Verfügbarkeit von hochauflösenden Satellitendaten existieren zudem Möglichkeiten zur Differenzierung von Laub- und Nadelbäumen sowie zur Klassifizierung von Baumarten. Eine Anwendung für Einzelbäume ist jedoch aufgrund der noch zu geringen Auflösung nicht realistisch, sondern kann derzeit nur über größere zusammenhängende Flächen beziehungsweise Baumgruppen einer Baumart/eines Baumtyps Anwendung finden.

### 3 Datenbeschaffung und Datenaufbereitung

Zur Bearbeitung des Projekts standen die folgenden Ausgangsdaten zur Verfügung:

- das Bundesfernstraßennetz aus dem Bundesinformationssystem Straße, © NEMO/BASt (2019),

- die hochauflösenden digitalen Gelände- und Oberflächenmodelle sowie 4-Kanal-Orthofotos des Bundeslandes Nordrhein-Westfalen (© Land NRW (2019))
- Lage bekannter Bäume aus Ereigniskatastern und verschiedenen Datenbanken.

Die Daten der digitalen Höhen- und Geländemodelle wurden zunächst in Rasterdatensätze konvertiert (0,5, 1 und 2 m Zellgröße) und daraus anschließend das Kronenhöhenmodell berechnet. Die digitalen Orthofotos wurden aufgrund ihrer enorm hohen Auflösung von 10 x 10 cm<sup>2</sup> Zellgröße auf 1-m-Pixel resampled und auf dieser Grundlage der Normalized Difference Vegetation Index berechnet. Der Metadatenabgleich bezüglich der Aufnahmezeiträume ergab, dass die Höhenmodelle hauptsächlich aus Überfliegsdaten aus den Jahren 2012 bis 2018 mit den Monaten November bis April (jahresübergreifend) stammen. Dies ist sehr gut für die Erstellung des Geländemodells, jedoch kontraproduktiv mit Blick auf die zu erwartenden Unregelmäßigkeiten in den Baumkronen, da Laubbäume in diesen Jahreszeiten entweder noch kein oder nur ein schwach ausgebildetes Blattwerk besitzen. Die digitalen Orthofotos stammen aus den Jahren 2017 bis 2019. Bei dem Vergleich mit den digitalen Höhendaten wurde dabei dementsprechend deutlich, dass nur ein geringer Teil der Datensätze tatsächlich im gleichen Zeitraum aufgenommen wurde und teils mehrere Jahre zwischen den Aufnahmen liegen.

Die Datensätze hinsichtlich der Lokation bekannter Bäume basieren im Einzelnen auf dem Alleenkataster, Baumkatastern verschiedener Städte im Bundesland Nordrhein-Westfalen, der Altbaumfinder-Datenbank, dem Baumregister und ausschnittsweise dem Bestand des Landesbetriebs Straßen.NRW. Weiterhin standen Vegetationsaufnahmen und Landschaftselemente aus der Landschaftsinformationssammlung NRW zur Verfügung. Die mitgelieferten Informationsebenen sind dabei sehr heterogen und reichen von Daten ohne Informationen bis hin zu Angaben der Baumhöhe, des Kronendurchmessers, der Baumart und der Stammanzahl. Insgesamt stehen zur Validierung ca. 360 000 bekannte Individuen zur Verfügung, wobei der Großteil hiervon den Laubbäumen zuzurechnen ist (ca. 55 % Laub- und nur 2,5 % Nadelbäume, Rest unbekannt).

## 4 Konzept

### 4.1 Allgemeine Herangehensweise

Entsprechend den Ergebnissen der Kenntnisstandsanalyse existiert eine Vielzahl von Einflussfaktoren, die sich auf die Qualität der Modelle auswirken. Diese Einflüsse können sowohl natürlichen, technischen und methodischen Ursprungs sein. Für die Entwicklung der Methodik zur Einzelbaumerkennung kam daher ein iterativer Ansatz zum Einsatz, bei dem Modelle mit verschiedenen Parametern und Aufbereitungsmethoden erstellt wurden. Die grundlegende Herangehensweise basiert dabei zunächst auf der Berechnung des Kronenhöhenmodells in verschiedenen Auflösungsversionen, welches anschließend durch Anwendung

unterschiedlicher Filterarten und Filtergrößen einer Rauschfilterung unterzogen wurde (konzeptionelle Modelle). Zusätzlich wurde für jedes dieser Modelle eine Version auf Basis des vom jeweiligen Kronenhöhenmodell abgeleiteten Oberflächenparameters "Minimale Krümmung" berechnet.

Nach deren Auswertung mithilfe der vorhandenen Validierungsdatensätze wurden anschließend die vielversprechendsten Modelle ausgewählt, weiteren Änderungen unterzogen und analysiert (Feinabstimmung). Zur Optimierung der Ergebnisse wurde dieser Prozess mehrfach wiederholt. Die Anpassungen beruhen dabei auf teils grundlegenden Veränderungen, wie zum Beispiel den Settings bei der Generierung des Oberflächenmodells, einer mehrfach durchgeführten Rauschfilterung sowie insbesondere verschiedenen Ansätzen zur Identifizierung der Baumkronen. Speziell auf die Erkennungsmethoden bezogen kamen der Focal-Flow-Algorithmus, die Erfassung des lokalen Maximums sowie ein Ansatz zur automatischen Klassifizierung auf Grundlage der minimalen Krümmung (ISO Cluster Classifier) zum Einsatz. Mit dem Ziel der Reduzierung mehrfach ausgewiesener Baumkandidaten innerhalb einer Baumkrone, welche zum Beispiel aufgrund von vielen Nebenmaxima und Unregelmäßigkeiten im Kronenhöhenmodell insbesondere bei Laubbäumen auftreten können, wurden darüber hinaus verschiedene objektorientierte Ansätze getestet.

### 4.2 Validierung

Aufgrund dessen, dass die Lage der kartierten Individuen aus den Baumkatastern mit Unschärfen behaftet ist, wurden die vorhandenen Punktdaten der Baumkataster mit einem Radius von 3 m gepuffert und die Validierung der Modellergebnisse zur Einzelbaumerkennung im Wesentlichen mit diesem Datensatz realisiert. Darüber hinaus wurde ein weiterer Validierungsdatensatz erstellt, bei welchem die Puffergröße entsprechend vom attribuierten Wert des Kronendurchmessers abhängig gemacht wurde. Dies sollte zumindest theoretisch eine realistischere Abdeckung der Kronengeometrie und somit auch Einschätzung der Modellergebnisse ermöglichen. Insgesamt besteht der Validierungsdatensatz für einzelne Bäume aus ca. 225 000 Individuen (davon 49 % als Laub- und 2 % als Nadelbäume attribuiert).

Zusätzlich zur Beurteilung der Einzelbaumerkennung wurden die Modellergebnisse hinsichtlich der Anzahl von Individuen innerhalb einer ausgewählten Fläche untersucht. Hierbei kommt es nicht darauf an, dass möglichst jeder bereits bekannte Baum korrekt erkannt wird, sondern wie viele Bäume das jeweilige Modell innerhalb der zuvor definierten Fläche ausweist. Dieser Aspekt wird insbesondere für Baumgruppen und Wälder, bei denen die Übergänge der Baumkronen benachbarter Individuen lückenlos und in der Regel im Kronenhöhenmodell schwächer ausgeprägt sind, als auch für einzelne Laubbäume mit großen Kronendurchmessern – und somit potenziell mehreren Nebenmaxima, relevant. Von Interesse sind daher vor allem die Gesamtanzahl ausgewiesener Baumkandidaten sowie deren

räumliche Verteilung. Da in den vorhandenen Baumkatastern keine Angaben zu größeren Baumgruppen oder Waldflächen enthalten sind, wurde deshalb eine ca. 250 x 250 m<sup>2</sup> große Fläche händisch digitalisiert sowie die Aufnahmen der im Rahmen des Projekts durchgeführten Feldkartierung genutzt.

## 5 Ergebnisse

### 5.1 Konzeptionelle Modelle

Eine erste Sensitivitätsanalyse der Berechnungsergebnisse zeigt, dass bei den Modellen mit höherer Auflösung des Kronenhöhenmodells prinzipiell auch mehr Baumkandidaten ausgewiesen werden. Ebenfalls führt die Verwendung des Oberflächenparameters "Minimale Krümmung" als Ausgangsdatensatz zu erheblich mehr Punkten als die Versionen, die auf dem Kronenhöhenmodell basieren. Weiterhin konnte gezeigt werden, dass der Filtertyp mitunter entscheidend auf die Anzahl der Baumkandidaten einwirkt.

Darüber hinaus konnte die Gesamtpunktzahl durch die Anwendung einer Mindesthöhe teilweise um mehr als 50 % reduziert werden, was darauf schließen lässt, dass viele dieser Punkte auf Rauschen basieren und die Modelle noch zu stark darauf reagieren. Auch in diesem Punkt zeigte sich eine Abhängigkeit hinsichtlich der untersuchten Auflösungen, da die höchste Reduzierung auch bei den Modellen mit der kleinsten Zellgröße auftrat. Die Anwendung weiterer Glättungsdurchläufe mit ausgewählten Modellen zeigte weiterhin, dass die Punktzahl abhängig von der Filterart um bis zu 50 % reduziert werden kann, bei einem zusätzlichen dritten Durchlauf jedoch nur noch ca. 10 % beträgt.

Bezüglich der qualitativen Einschätzung konnten sowohl für Versionen auf Basis des Kronenhöhenmodells als auch der minimalen Krümmung hohe bis sehr hohe Wiedererkennungsraten von über 80 %, teils 90 % beobachtet werden, wobei der alternative Ansatz auf Grundlage des Oberflächenparameters modellübergreifend die höheren Werte lieferte. Auch hier zeigte sich eine auflösungsbedingte Abhängigkeit, da durch die Modelle mit geringerer Auflösung auch grundsätzlich weniger Individuen wiedererkannt wurden. Weiterhin konnte gezeigt werden, dass zwischen diesen Ergebnissen und denen der Sensitivitätsanalyse ein Zusammenhang besteht, denn in der Regel wiesen Modelle mit höheren Erkennungsraten auch insgesamt eine höhere Gesamtpunktzahl auf. Dies hat zur Folge, dass einzelne Baumkronen, insbesondere Laubbäume mit großen Kronendurchmessern, aufgrund vorhandener Nebenmaxima erheblich überklassifiziert werden und so die Anzahl der real vorhandenen Bäume um ein Vielfaches überschätzt wird. Dieser Effekt konnte zwar durch eine weitere Glättung der Modelle abgeschwächt werden, jedoch verringerte sich hierbei auch die Wiedererkennungsraten mitunter um einen niedrigen zweistelligen Prozentsatz.

### 5.2 Feinabstimmung

Für die Validierung der zur Feinabstimmung ausgewählten und weiter angepassten Modelle wurden sowohl die Datensätze der Baumkataster zur Ermittlung der Erkennungsrate von Einzelbäumen als auch die der digitalisierten und im Gelände aufgenommenen Datensätze für den Vergleich der Anzahl von Individuen in den jeweiligen Flächen genutzt. Dabei stellte sich insbesondere der Versuch, die Modelle so anzupassen, dass alle vorherrschenden Bedingungen möglichst gleichermaßen berücksichtigt werden, als nicht überwindbare Hürde heraus. Die Analyse der Berechnungsergebnisse zeigte unmissverständlich, dass Einzelbäume und Bäume in dichten Verbänden, zum Beispiel in größeren Baumgruppen oder Wäldern, nicht mithilfe eines einzigen Ansatzes in vergleichbarer Qualität identifiziert werden können. Dies trifft insbesondere für einzelne Laubbäume mit großen Kronendurchmessern im Vergleich mit dichten Waldgebieten mit Nadelhölzern und geringen Kronendurchmessern zu. Hierbei werden die Nadelbäume stets systematisch erheblich unter- oder die Anzahl der Laubbäume deutlich überschätzt. Dies ist damit begründet, dass die Modelle bezüglich der eng stehenden Nadelbäume entweder zu stark geglättet sind – oder aber im Umkehrschluss zu sensitiv auf Nebenmaxima in den Baumkronen reagieren, was wiederum zu einer Überschätzung der Laubbäume führt. Ähnliches gilt für die Beurteilung der Wiedererkennungsraten: auch hier werden insbesondere für die Laubbäume um so mehr Punkte ausgewiesen, je höher die Erkennungsrate der Einzelbäume ist.

Um dabei einen akzeptablen Wert zu erreichen, ohne eine unplausibel hohe Menge an Baumkandidaten zu erzeugen, wurden zusätzlich zur Modellberechnung mehrere objektbasierte Ansätze getestet, um die Zahl der innerhalb einer Baumkrone mehrfach ausgewiesenen Punkte zu reduzieren. Die Verringerung der Punktzahl ist dabei umso effektiver, je mehr Punkte sich innerhalb der jeweiligen Baumkrone befinden. Final konnten mit der Methodik für die Einzelbaumerkennung ca. 70 % der Laub- und 60 % der Nadelbäume identifiziert werden. Gleichzeitig liegt die Anzahl der Baumkandidaten in den Validierungsgebieten für Laubbäume zwischen 100 und 115 %. Bezüglich dichter Nadelholzwälder reduziert sich die Anzahl der Baumkandidaten mit dieser Herangehensweise jedoch teilweise auf unter 50 %.

## 6 Anwendung

### 6.1 Durchführung

Die Erfassung von Einzelbäumen auf Grundlage der ausgewählten Methodik erfolgte unter Einsatz der erstellten ArcGIS-Werkzeuge für das gesamte Bearbeitungsgebiet. Die letztendlichen Bearbeitungs- und Berechnungszeiten hängen dabei wesentlich von der Fläche des Zielgebiets und der verfügbaren Rechen-technik ab. Mit derzeit aktueller Hardware sind aufgrund der enormen Datenmengen Zeiträume zwischen 20 und 30 Tagen für die Bearbeitung realistisch.

Um das Gesamtergebnis zu verbessern, wurde zusätzlich zum finalen Modell der Einzelbaumerkennung ein weiterer Ansatz berechnet, um die Anzahl der Nadelhölzer in dichten Wäldern realistischer abschätzen zu können. Die Ergebnisse für diese Flächen befinden sich in einem Bereich um 100 %. Die Kombination beider Methoden erfolgte anschließend auf Grundlage der Punktdichte der Einzelbaumerkennung unter Verwendung eines Grenzwerts. Dieser wurde empirisch ermittelt und so gewählt, dass die Grenzen der bekannten Nadelwälder möglichst gut abgebildet werden. Da die Laubbäume im Allgemeinen eine deutlich geringe Dichte an Baumkandidaten aufweisen, ist gleichzeitig sichergestellt, dass bis auf gegebenenfalls wenige Ausnahmen auch hauptsächlich Flächen mit Nadelhölzern von dieser Methodik profitieren.

Abschließend wurden die generierten Baumkandidaten mithilfe einer minimalen Kronenhöhe von 2 m sowie einem minimalen Wert des Normalized Difference Vegetation Index (0,1) gefiltert, um rausch-induzierte Punkte und solche auf Flächen ohne Vegetation zu entfernen.

### 6.2 Attributierung

Unter Einsatz der erstellten ArcGIS-Werkzeuge wurden den Baumkandidaten verschiedene Parameter wie Lage, Höhe des Kronenmodells, Normalized Difference Vegetation Index und der Abstand zum nächstgelegenen Straßenabschnitt zugewiesen. Mithilfe dieser Attribute können anschließend die für das Straßennetz relevanten Baumkandidaten ausgewiesen werden, welche bei Baumstürzen oder Abknicken aufgrund ihrer Höhe und ihres Abstands zum Straßenrand eine potenzielle Gefahr darstellen.

Weiterhin erfolgte eine Parametrisierung des Straßennetzes mit den relevanten Baumkandidaten im Gesamten sowie in Pufferbereichen von 0 – 30 m, 30 – 60 m und mehr als 60 m Abstand zum Straßenrand. Aufbauend auf diesen Daten werden zusätzlich durch das ArcGIS-Tool zwei Rasterdatensätze mit allen und den relevanten Punkten erzeugt, die für weitere Auswertungen und insbesondere zur Visualisierung dienen können.

### 6.3 Ergebnisse

Bei dem Vergleich der Baumkandidatenanzahl beider verwendeter Ansätze wurden für das Einzelbaummodell insgesamt nur ca. 22 % der durch das Wald-Modell erhaltenen Punkte generiert. Durch die Kombination beider Herangehensweisen konnte die Gesamtzahl an Baumkandidaten (ohne Filter) um ca. 250 000 Individuen in den sehr dicht bewaldeten Gebieten erhöht werden. Nach Anwendung der Grenzwerte für die Kronenhöhenmodell-mindesthöhe von 2 m und einen NDVI von mindestens 0,1 bleiben im Einzelbaummodell noch ca. 65 % und im Wald-Modell noch ca. 52 % der Punkte erhalten. Für den kombinierten Datensatz aus beiden Modellen, welcher sich aus ca. 96 % der Einzelbaum-Modellpunkte und ca. 4 % Wald-Modell-Baumkandidaten

zusammensetzt, sind nach der Filterung noch ca. 66 % der Ausgangsmenge im Datensatz vorhanden. Final werden von den 6 098 887 ermittelten Baumkandidaten im gesamten Bearbeitungsgebiet noch ca. 16 % der Baumkandidaten als relevant eingestuft und zur Attributierung an das Straßennetz genutzt.

## 7 Schlussfolgerungen und Ausblick

Ausgehend von der gesichteten Literatur existiert eine Fülle von Ansätzen und Möglichkeiten zur Erkennung von Einzelbäumen. Bezüglich der Vorgabe zur Verwendung kostenneutraler und insbesondere flächendeckend verfügbarer Datengrundlagen sowie des hohen Rechenaufwands für die Fläche eines gesamten Bundeslandes, wurde die klassische Herangehensweise auf Basis eines Kronenhöhenmodells gewählt. Speziell aufgrund der Überfliegszeiträume für das digitale Oberflächenmodell ist jedoch damit zu rechnen, dass in keiner der durchgeführten Modellberechnungen das Optimum erreicht wurde.

Die methodische Entwicklung hat eindeutig gezeigt, dass die Leistung der getesteten Modelle vor allem aufgrund der Vegetationseigenschaften und -strukturen teils sehr deutlichen Schwankungen unterliegen kann – und auch entsprechend von diesen abhängt. Durch die Tatsache, dass die Baumerkennung bereits bei nur kleinen Veränderungen und Anpassungen in den Ausgangsdaten sehr sensitiv reagiert, ist davon auszugehen, dass es bei der Anwendung in anderen Bundesländern mit gegebenenfalls unterschiedlicher Qualität der Ausgangsdaten auch zu Abweichungen bezüglich der Modellergebnisse kommen kann, obwohl bei der Auswahl der Methodik auf eine möglichst geringe Variation diesbezüglich geachtet wurde.

Die mit dem Modell zur Einzelbaumerkennung erreichte Erkennungsrate befindet sich bei ca. 70 % für Laubbäume und ca. 65 % für Nadelbäume. Alternative Modelle mit höheren Quoten zur Lagebestimmung einzelner Bäume existieren, sind dann jedoch so sensitiv, dass die Mehrfachbelegung großer Baumkronen durch zu viele Punkte nicht mehr auf ein passables Niveau reduziert werden kann. Um dies zu korrigieren, wurde eine zweite Anwendung mit einer darauf angepassten Methodik durchgeführt und die Ergebnisse beider Modellberechnungen anschließend kombiniert.

Insgesamt ist festzuhalten, dass mithilfe der entwickelten Methode ein Großteil der Einzelbäume zuverlässig detektiert werden kann. Zur Verbesserung sollten jedoch sowohl die Erweiterung der Methodik durch Einbeziehung von Satellitendaten als auch alternative Herangehensweisen wie zum Beispiel der Mustererkennung auf Basis hochauflösender Orthofotos in Betracht gezogen werden.