

# Untersuchung zur Erkennung von Wasserflächen in Full-Waveform-Laserscanner-Daten

ALENA SCHMIDT<sup>1</sup>, FRANZ ROTTENSTEINER<sup>1</sup> & UWE SÖRTEL<sup>1</sup>

*Zusammenfassung: Die Extraktion von Wasserflächen besitzt für viele Anwendungsfelder im Bereich der Wasser- und Schifffahrtswege sowie des Küstenschutzes Relevanz. Für die flächendeckende Erfassung von Oberflächeninformation kommen oftmals flugzeuggestützte Laserscansysteme zum Einsatz. In diesem Beitrag wird untersucht, inwiefern eine Detektion von Wasserflächen in Laserscandaten möglich ist. Dazu erfolgt eine überwachte Klassifikation der Punktwolke nach einem Ansatz der Fuzzy-Logik. Die geometrischen und physikalischen Größen, die als Klassifikationsmerkmale Verwendung finden, werden ebenso wie die zusätzliche Information des Full-Waveform-Laserscannings auf ihre Eignung im Bereich der Wasserextraktion analysiert.*

## 1 Einleitung

In den letzten Jahren hat sich das luftgestützte Laserscanning als effizientes Verfahren zur flächendeckenden Gewinnung von Oberflächeninformation erwiesen. Durch Aussenden eines Laserpulses und Detektion des Empfangssignals kann über Laufzeitmessung die Höhe der reflektierenden Oberfläche abgeleitet werden. Die geometrische Information findet insbesondere für die Erstellung digitaler Geländemodelle (DGM) Verwendung. Darüber hinaus lassen sich unter Berücksichtigung der Charakteristik der Echos einzelne Objekte aus der 3D-Punktwolke extrahieren. Information über die Oberflächenbeschaffenheit kann unter anderem aus der Intensität des empfangenen Signals gewonnen werden. Eine neue Generation sogenannter Full-Waveform (FWF)-Sensoren tastet zudem den Sendepuls und das empfangene Echo mit einer sehr hohen Frequenz ab und ermöglicht auf diese Weise die Auswertung des zeitlichen Verlaufs des Signals (Wagner et al., 2006). Hieraus kann die Breite eines Einzelechos ermittelt sowie Information über die Anzahl und Nummer einzelner Echos bei Mehrfachreflexionen gewonnen werden.

Ein Anwendungsfeld, das hochaufgelöste Laserscandaten nutzt, ist die Erforschung und Überwachung von Gewässern. Das Monitoring von Flussverläufen ist Bestandteil des Gefahrenschutzes in Überflutungsbereichen, es dient unter anderem der Gewährleistung gesicherter Schifffahrtswege (Brügelmann & Bollweg, 2004). Hinsichtlich des Verständnisses hydrologischer Systeme und deren Änderungen gewinnt die kontinuierliche Überwachung von Gewässern an Bedeutung (Thoma et al., 2005).

In diesem Beitrag wird ein Verfahren zur Detektion von Wasserflächen in Laserscandaten vorgestellt und der Fragestellung nach geeigneten Merkmalen nachgegangen. Anhand eines auf Fuzzy-Logik beruhenden Ansatzes werden Full-Waveform-Laserscandaten von zwei Testgebieten klassifiziert. Zielsetzung ist dabei, das Potential von Full-Waveform-Daten zu

1) Alena Schmidt, Franz Rottensteiner, Uwe Sörgel, Institut für Photogrammetrie und GeoInformation, Leibniz Universität Hannover, Nienburger Straße 1, 30167 Hannover; E-Mail: {alena.schmidt, rottensteiner, soergel}@ipi.uni-hannover.de

untersuchen sowie geeignete Klassifikationsmerkmale für die Trennung von Wasser- und Landflächen zu analysieren. Als Merkmale kommen sowohl geometrische Größen wie Höhen und Höhenvarianzen innerhalb der näheren Umgebung als auch physikalische Eigenschaften wie die Intensität oder Breite eines Pulses zum Einsatz.

## 1.1 Verwendete Datensätze

Im Rahmen einer Befliegung von Bereichen des Nord-Ostsee-Kanals vom 8. bis 12.10.2007 wurde die östliche Ausfahrt und die Mündung in die Kieler Förde bei der Schleswig-Holsteinischen Landeshauptstadt mit dem luftgestützten Sensor RIEGL LMS-Q560 erfasst. Das Gebiet hat eine Ausdehnung von ca. 1,3 km x 3,5 km, wobei knapp die Hälfte der Fläche von Wasser eingenommen wird. Während des Messflugs wurde der gesamte Signalverlauf der rückgestreuten Pulse aufgezeichnet. Für die daraus abgeleiteten Oberflächenpunkte liegen neben den 3D-Koordinaten die Intensität, Pulsbreite sowie Anzahl und Nummer der Echos vor. Die resultierende Punktwolke ist in Abb. 1a dargestellt.

Für eine weitere Untersuchung wurden Laserscandaten eines zweiten Testgebietes über dem fränkischen Bonmland verwendet. Die Full-Waveform-Daten wurden am 13.9.2006 mit dem Sensor RIEGL LMS-Q560 aufgezeichnet. Die Punktwolke (Abb. 1b) umfasst ein Gebiet der Größe von ca. 0,8 km x 1,8 km, wobei die gleichen Merkmale der Pulse wie im ersten Testgebiet erfasst worden sind. Bei den vorkommenden Wasserflächen handelt es sich um mehrere kleine Seen mit einer geringen Flächenausdehnung. Die Testgebiete unterscheiden sich in der Beschaffenheit und der Größe der Wasserflächen.

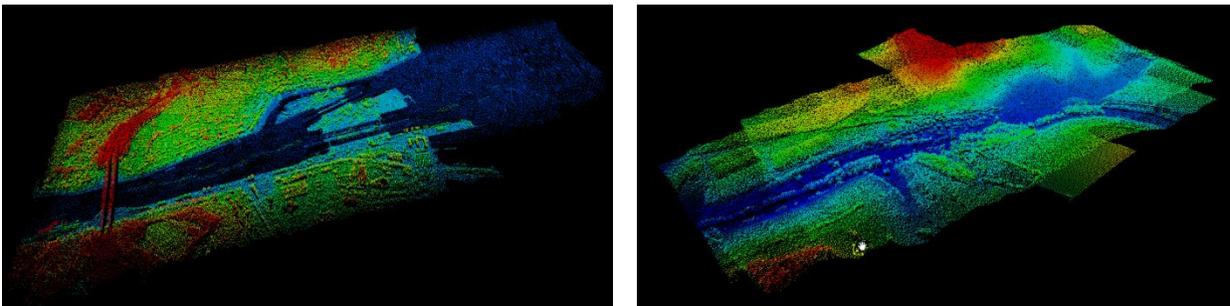


Abb. 1: Punktwolken der erfassten Gebiete Kiel (a) und Bonmland (b)

## 2 Verfahren zur Detektion von Wasserflächen

Der Klassifikationsansatz basiert ausschließlich auf der Analyse von Laserscandaten und erfolgt direkt in der Punktwolke ohne Vorverarbeitungsschritte in Form einer Filterung oder DGM-Modellierung. Grundlage bildet ein Algorithmus zur Erkennung von Restwasserflächen in Wattgebieten nach Brzank (2008), der auf andere Anwendungsgebiete und für die Nutzung von Full-Waveform-Laserscandaten erweitert ist und im Folgenden detailliert vorgestellt wird.

## 2.1 Methodik

Die Punkte der gegebenen Laserscandaten liegen zum Teil auf Wasseroberflächen und zum Teil auf dem Gelände und darauf befindlichen Objekten wie Vegetation oder Gebäuden. Um die Wasserflächen zu extrahieren, ist eine Zuordnung der Oberflächenpunkte zu den Klassen "Wasser" und "Land" erforderlich. Das verwendete Verfahren nutzt hierfür verschiedene geometrische und intensitätsbasierte Merkmale (siehe Abschnitt 2.2), die entsprechend der statistischen Verteilung mit unterschiedlicher Gewichtung in den Algorithmus einfließen. Als Ergebnis wird in einem auf Fuzzy-Logik beruhendem Verfahren für jeden Laserpunkt der Zugehörigkeitsgrad  $\mu_w$  eines Laserpunktes zu der Klasse "Wasser" bestimmt. Auf Grundlage eines Schwellwertes wird anschließend die Zuordnung zu "Wasser" oder "Land" vorgenommen. Die Bestimmung von  $\mu_w$  erfolgt anhand Gleichung 1:

$$\mu_w = \frac{\sum_{i=1}^n d_i(x_i) \cdot \mu_i}{\sum_{i=1}^n d_i(x_i)} \quad (1)$$

Dabei sind  $\mu_i$  mit Hilfe des verwendeten Merkmals  $i$  bestimmte Zugehörigkeitswerte. Die erforderlichen Parameter zur Bestimmung dieser Einzelzugehörigkeitswerte werden aus Trainingsgebieten abgeleitet. Für jedes Merkmal ergibt sich eine Zugehörigkeitsfunktion, die für den mittleren Merkmalswert  $\bar{x}_w$  im Trainingsgebiet des Wassers maximal ( $\mu_i = 1$ ) bzw. im Landgebiet minimal ( $\mu_i = 0$ ) wird. Die Eignung eines Merkmals zur Land-Wasser-Trennung wird mit Hilfe eines Gewichtes  $d_i(x_i)$  ausgedrückt. Ausschlaggebend ist hierbei die Verteilung der Merkmalswerte für Wasser und Land. Mit Hilfe eines t-Tests wird untersucht, wie gut sich die Mittelwerte beider Verteilungen unterscheiden. Die Testgröße

$$t = \frac{\bar{x}_l - \bar{x}_w}{\sqrt{\frac{s_l^2}{n_l} + \frac{s_w^2}{n_w}}} \quad (2)$$

leitet sich aus den Differenzen der Mittelwerte  $\bar{x}_w$  bzw.  $\bar{x}_l$  für Wasser und Land, sowie aus den Standardabweichungen  $s_w$  bzw.  $s_l$  und der Anzahl  $n_w$  und  $n_l$  der Wasser- und Landpunkte ab. Da  $n_w$  und  $n_l$  große Werte annehmen, nähert sich die t-Verteilung der Normalverteilung an. Für das Signifikanzniveau gilt

$$\gamma = P(-t \leq T \leq t), \quad (3)$$

wobei  $T$  der Zufallsvariablen zu  $t$  entspricht. Bei einem Signifikanzniveau nahe 1 wird ein hohes Gewicht vergeben, bei Werten  $\gamma \leq 0,5$  wird das Merkmal nicht für die Klassifikation verwendet. Der Schwellwert, der abschließend ausschlaggebend ist, ob ein Punkt Wasser oder Land zugeordnet wird, ergibt sich als Schnittpunkt der Verteilungen der Gesamtzugehörigkeitsgraden in den Trainingsgebieten von Wasser und Land. Die klassifizierte Punktwolke bildet das Ergebnis des Verfahrens.

## 2.2 Klassifikationsmerkmale

In dem dargestellten Verfahren werden acht unterschiedliche Klassifikationsmerkmale zur Trennung von Wasser und Land verwendet, die zu drei Gruppen zusammengefasst werden:

- **Merkmale nach Brzank (2008):** Höhe, Intensität, Punktdichte
- **flächenbasierte Merkmale:** Höhen- und Intensitätsvariation
- **FWF-Merkmale:** Signalbreite, normierte Echonummer, Echoanzahl (Chehata, 2009)

Bei früheren Untersuchungen haben sich die ersten drei Merkmale als geeignete Größen erwiesen, um im Wattbereich eine Unterteilung von Wasser und trocken gefallen Landflächen vorzunehmen (Brzank, 2008). Dabei wird die Annahme getroffen, dass Wasserflächen über eine geringere absolute Höhe als Landflächen verfügen, aufgrund eines geringeren Reflexionsgrades zu einer niedrigeren Intensität führen sowie bei spiegelnder Reflexion einen höheren Anteil an Datenausfällen zu verzeichnen haben (in Abhängigkeit vom Auslenkwinkel bezogen auf den Nadirpunkt), der zu einer geringen Punktdichte innerhalb einer vorgegebenen Fläche führt. Für Untersuchungsgebiete außerhalb des Watts wird statt einer geringen absoluten Höhe von einer niedrigen Streuung der Höhenwerte ausgegangen. Als weitere flächenbasierte Merkmale werden zudem die Höhen- bzw. Intensitätsvariation eingeführt. Beide Größen entsprechen der Standardabweichung des jeweiligen Merkmals innerhalb einer vorgegebenen Fläche. Im Folgenden wird hierfür eine kreisförmige Umgebung mit festem Radius  $R = 5 \text{ m}$  angenommen und ein effizienter Zugriff auf umgebende Punkte mit Hilfe eines k-d-Baums (Bentley, 1990) der Dimension 2 realisiert. Durch die Verwendung von FWF-Daten lassen sich zudem als weitere Merkmale aus der Punktwolke die Signalbreite sowie Mehrfachreflexionen in Form der Anzahl der rückgestreuten Pulse pro ausgesandtem Signal und dem Verhältnis des jeweiligen Echonummer zu dieser Größe ableiten.

## 3 Ergebnisse

Die Evaluierung der Ergebnisse erfolgt auf Grundlage manuell erstellter Referenzdaten in Orthophotos der jeweiligen Gebiete. Um die Qualität der Klassifikation zu beurteilen, werden als Genauigkeitskriterien die *Vollständigkeit* und *Korrektheit* der Ergebnisse bestimmt. Die Vollständigkeit  $V$  ergibt sich aus dem Verhältnis der richtig extrahierten Punkte zu der Anzahl der relevanten Punkte in den Referenzdaten, während unter der Korrektheit  $K$  das Verhältnis richtig klassifizierten Daten zu der Gesamtheit der Referenzdaten verstanden wird:

$$V = \frac{100 \cdot TP}{TP + FN} \%$$

$$K = \frac{100 \cdot TP}{TP + FP} \% \quad (4)$$

In Gleichung 4 entspricht  $TP$  den *True Positives* (den mit den Referenzdaten übereinstimmenden Wasserpunkten),  $FN$  der Anzahl der *False Negatives* (nicht erkannte Wasserpunkte) sowie  $FP$  den *False Positives* (als Wasser fehlklassifizierte Landpunkte).

Die Klassifikation der Laserpunkte wird für die beiden Testgebiete durchgeführt. Hierbei werden die zu analysierenden Klassifikationsmerkmale in unterschiedlicher Kombination berücksichtigt und ihr Einfluss auf die Wasserflächenextraktion anhand der Vollständigkeit und Korrektheit beurteilt.

### 3.1 Testgebiet Kiel

In dem erfassten Gebiet über Kiel wird nahezu die Hälfte der Fläche von Wasser eingenommen. Bedingt durch spiegelnde Reflexion kommt es jedoch zu einigen Datenausfällen, insbesondere im Randbereich des Flugstreifens. Hier ist der zum Empfänger gelangende Signalanteil zu gering, um als Puls erkannt zu werden. Der Vergleich der Merkmalswerte für Wasser (Tab. 1) und Land (Tab. 2) sowie der Histogramme einiger ausgewählter Größen (Abb.2) machen deutlich, dass eine eindeutige Unterscheidbarkeit beim Merkmal Höhe gegeben ist (Abb. 2a). Die Wasseroberfläche entspricht den geringsten Höhenwerten im gesamten Testgebiet. Dementsprechend wird hierfür ein hohes Gewicht im Klassifikationsalgorithmus vergeben. Die Intensitätswerte (Abb. 2b) weisen in beiden Fällen eine hohe Streuung auf. Darüber hinaus sind die Datenausfälle über Wasser signifikant höher und somit die durchschnittliche Punktdichte geringer. Bei der Signalbreite (Abb. 2c) ist zu erkennen, dass die rückgestreuten Pulse über Land eine etwas stärkere Aufweitung erfahren, was zu einer größeren Pulsbreite führt. Insbesondere bei Vegetation ist dies der Fall. Die nahe beieinander liegenden Mittelwerte erschweren jedoch die Trennung von Wasser- und Landflächen. Ebenso die Anzahl und normierte Nummer der einzelnen Echos, da bedingt durch geringe Mehrfachreflexionen außerhalb von Vegetation eine zu große Anzahl an Landpunkten der Merkmalsausprägung für Wasser ähnelt. Eine signifikante Unterscheidbarkeit ist hingegen für die flächenbasierten Merkmale der Intensitäts- und Höhenvariation gegeben.

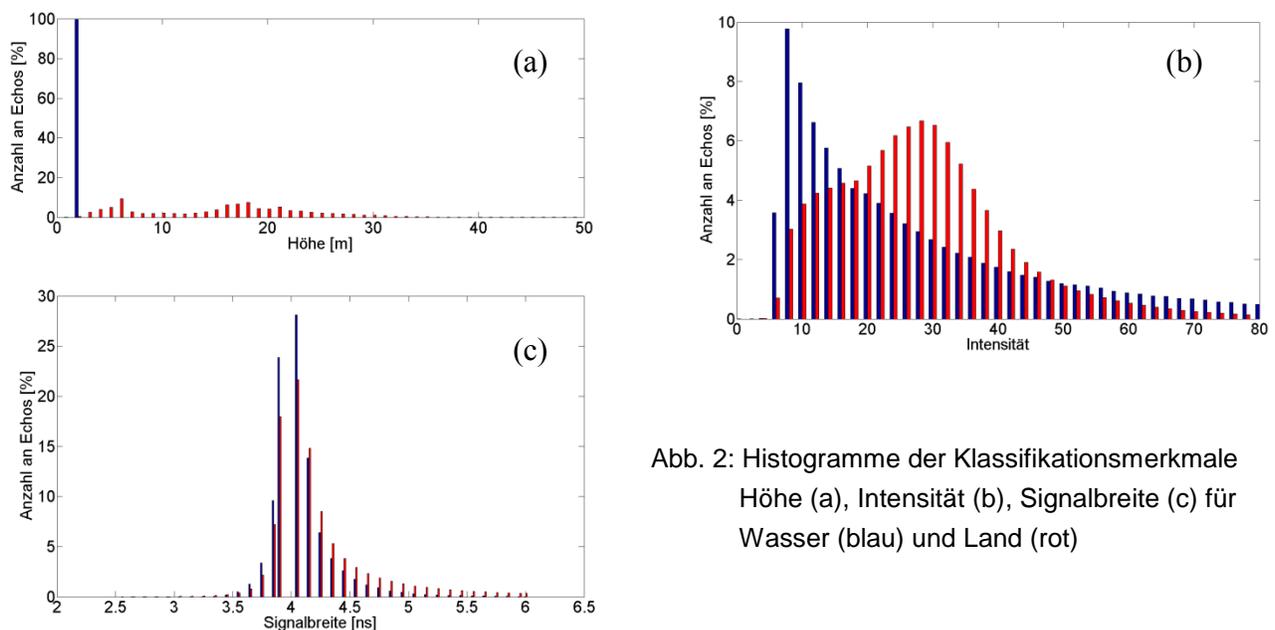


Abb. 2: Histogramme der Klassifikationsmerkmale Höhe (a), Intensität (b), Signalbreite (c) für Wasser (blau) und Land (rot)

Merkmal	Min	Max	Mean	Std
Höhe [m]	0.13	2.57	2.42	0.04
Intensität	4.02	255	41.89	49.03
Punktdichte [1/m <sup>2</sup> ]	0.01	3.39	1.47	0.60
Signalbreite [ns]	0.10	6.70	4.04	0.28
Norm. Echonr.	0.50	1.00	1.00	0.00
Echoanzahl	1.00	3.00	1.00	0.02
Höhenvar. [m]	0.01	2.57	0.04	0.14
Intensitätsvariation	0.10	95.6	41.38	15.26

Merkmal	Min	Max	Mean	Std
Höhe [m]	2.25	66.75	16.16	7.96
Intensität	1.10	255	30.20	16.14
Punktdichte [1/m <sup>2</sup> ]	0.01	4.95	2.04	0.61
Signalbreite [ns]	0.10	9.90	4.31	0.89
Norm. Echonr.	0.25	1.00	0.97	0.12
Echoanzahl	1.00	4.00	1.12	0.34
Höhenvar. [m]	0.01	15.44	0.79	0.55
Intensitätsvariation	0.10	58.88	10.91	6.57

Tab. 1 u. 2: Statistische Größen der Klassifikationsmerkmale (Minimum, Maximum, Mittelwert, Standardabweichung) für Wasser- (1) bzw. Landpunkte (2) im Testgebiet von Kiel

Die Durchführung der Klassifikation liefert bei Verwendung der Merkmale Höhe, Intensität und Punktdichte die besten Ergebnisse. In diesem Fall können die Wasserflächen mit einer Korrektheit von 95.5 % sowie einer Vollständigkeit von 99.9 % detektiert werden. Einige wenige Fehlklassifikationen treten im Bereich tief gelegener Schleusenpunkte auf. Die Klassifikation anhand der Full-Waveform-Merkmale führt aufgrund der zuvor ausgelegten Problematik zu keinen zufriedenstellenden Ergebnissen. Sobald anhand von Höhen- und Intensitätsvariation klassifiziert wird, entspricht die Korrektheit 73.6 % und die Vollständigkeit 99.6 %. Fehlklassifikationen sind insbesondere bei großen horizontalen Landflächen wie beispielsweise Sportplätzen zu beobachten.

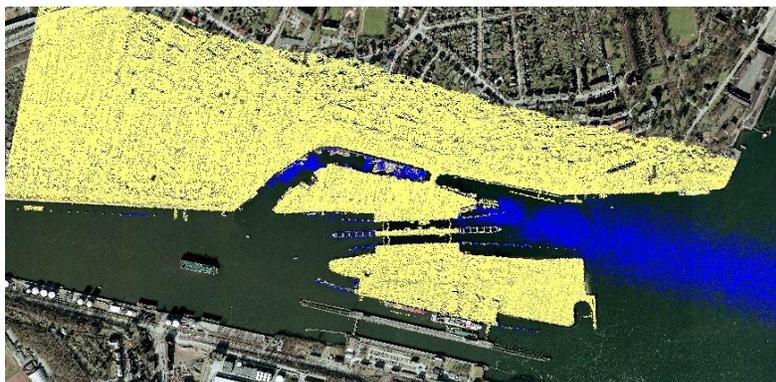


Abb. 3: Klassifikationsergebnis bei Verwendung der Merkmale Höhe, Intensität und Punktdichte (Korrektheit  $K = 95.5\%$ , Vollständigkeit  $V = 99.9\%$ )

### 3.2 Testgebiet Bonnland

Das Testgebiet um Bonnland umfasst mehrere kleine Seen. Bei der Analyse der Merkmalswerte zeigt sich, dass durch das Vorhandensein von mehr als einem Gewässer die absolute Höhe ungeeignet ist, um Wasser in der Laserpunkt wolke zu detektieren. Bedingt durch die geringe Streuung der Höhen bei Gewässer oberflächen ist jedoch anhand der Höhenvariation eine gute Unterscheidung möglich (Abb. 4a). Die Intensitätsvariation ist auf den ruhigen Seeoberflächen

gering und eignet sich daher ebenfalls für die Land-Wasser-Trennung (Abb. 4b). Sobald anhand dieser beiden Größen klassifiziert wird, können Wasserpunkte mit einer Korrektheit von 93,6 % und einer Vollständigkeit von 61,4 % detektiert werden. Die vergleichsweise geringe Vollständigkeit kann mit der Bestimmung insbesondere des Merkmals Höhenvariation erklärt werden. Hier fließen die Höhen aller umgebenden Punkte ein. Im Randbereich der Gewässer schließt dies Vegetation mit ein, was zu entsprechend großen Varianzen der Höhe führt. Objektbezogen wurde jedoch eine hohe Vollständigkeit erzielt, denn alle Gewässer sind als solche erkannt worden (Abb. 5). Eine Untersuchung bezüglich der Signalbreiten (Abb. 6) zeigt bei einer Unterteilung der Klasse "Land" in Vegetation, Gebäude, natürlichen und künstlichen Grund eine geringe Unterscheidbarkeit im Vergleich zu den Wasserpunkten. Nur im Bereich der Vegetation ist eine charakteristische Aufweitung der Pulse und starke Streuung zu beobachten. Eine Trennung der Wasserflächen insbesondere von den Bodenpunkten ist jedoch nicht möglich.

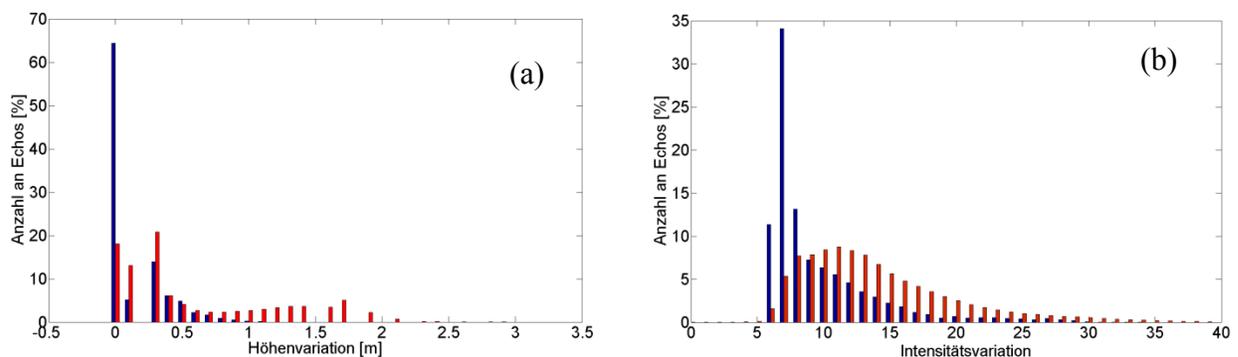
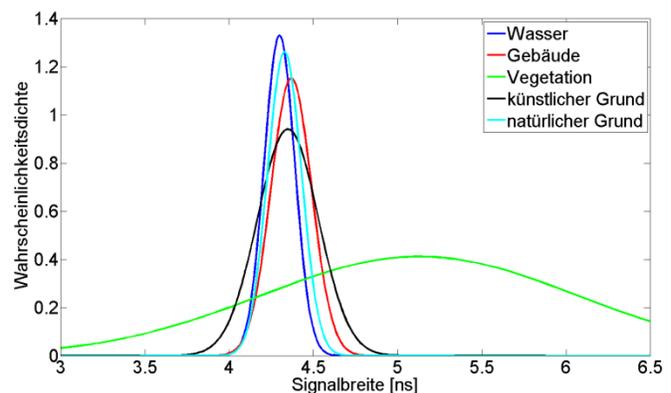


Abb. 4: Histogramme der Klassifikationsmerkmale Höhen- (a) und Intensitätsvariation (b)



Abb. 5: Klassifikationsergebnis bei Verwendung der Merkmale Höhen- und Intensitätsvariation (K = 93,6 %, V = 61,4 %). Alle drei Gewässer (blau) können detektiert werden.

Abb. 6: Verteilung der Signalbreiten für Wasser- und verschiedene Landflächen



## 4 Zusammenfassung und Ausblick

Die Zielsetzung dieses Beitrags besteht darin, Wasserflächen aus Laserscandaten zu extrahieren und die dabei verwendeten Klassifikationsmerkmale auf ihre Eignung im Bereich der Wasserdetektion zu analysieren. Die Einteilung von Oberflächenpunkten zu "Wasser" und "Land" basiert auf Fuzzy-Logik und ordnet den Laserdaten Zugehörigkeitsgrade zur Klasse "Wasser" zu. Die Auswertung zeigt, dass auf diese Weise eine Wasserflächenextraktion möglich ist. Als besonders geeignet erweisen sich Merkmale, die die Variation der Höhe und Intensität berücksichtigen. Hiermit ist es möglich, objektbezogen eine hohe Vollständigkeit der klassifizierten Wasserflächen zu erzielen. Im Randbereich der Gewässer führen umgebende Vegetationspunkte zu stärkerer Streuung der Höhenwerte, was die Detektion ufernaher Punkte erschwert. Dies macht eine Nachbearbeitung erforderlich. Bei Vorliegen eines einzelnen Gewässers stellt die absolute Höhe ein charakteristisches Merkmal dar und ermöglicht neben Größen wie der Intensität oder der Punktdichte die Detektion von Wasser. Eine Verbesserung des Klassifikationsergebnisses durch die Berücksichtigung von Full-Waveform-Information konnte in dem gewählten Ansatz hingegen nicht nachgewiesen werden. Mehrfachreflexionen und Aufweitungen der reflektierenden Pulse, die zu einer bezüglich des gesendeten Pulses höheren Signalbreite führen, können in erster Linie lediglich bei Vegetation ermittelt werden und diese von der Menge der möglichen Wasserpunkte ausschließen. Eine Trennung insbesondere von Bodenpunkten kann nicht vorgenommen werden. In diesem Zusammenhang ist der Übergang zu einem probabilistischen Verfahren vielversprechend und Gegenstand weiterführender Untersuchungen.

## 5 Literaturverzeichnis

- BENTLEY, J.L., 1990: K-d trees for semidynamic point sets. Proceedings of the 6th Annual Symposium on Computational Geometry, S. 187-197
- BRÜGELMANN, R. & BOLLWEG, A., 2004: Laser altimetry for river management. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, **35** (B2), S. 234-239
- BRZANK, A., 2008: Bestimmung Digitaler Geländemodelle in Wattgebieten aus Laserscannerdaten. Dissertation, Deutsche Geodätische Kommission Reihe C, **622**, Hannover
- CHEHATA, N., GUO, L. & MALLET, C. 2009: Airborne lidar selection for urban classification using random forests, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. **38** (Part 3/W8), S. 207-212
- THOMA, D.P., 2005: Airborne laser scanning for riverbank erosion assessment. Remote Sensing of Environment, **95** (4), S. 493-501
- WAGNER, W., ULLRICH, A., DUCIC, V., MELZER, T. & STUDNICKA, N., 2006: Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing, **60**, S. 100-112