Classifying wetland-related land cover types and habitats using fine-scale lidar metrics derived from country-wide Airborne Laser Scanning

Zsófia Koma, Arie C. Seijmonsbergen & W. Daniel Kissling
Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, P.O. Box 94240, Amsterdam 1090 GE, The Netherlands

Keywords
Habitat classification, Phragmites australis, reedbed, structural heterogeneity, vegetation complexity, wetland conservation

Abstract
Mapping 3D vegetation structure in wetlands is important for conservation and monitoring. Openly accessible country-wide Airborne Laser Scanning (ALS) data—using light detection and ranging (lidar) technology—are increasingly becoming available and allow us to quantify 3D vegetation structures at fine resolution and across broad spatial extents. Here, we develop a new, open-source workflow for classifying wetland-related land cover types and habitats using fine-scale lidar metrics derived from country-wide ALS data. We developed a case study in the Netherlands with a workflow consisting of four routines: (1) pre-processing of ALS data, (2) calculation of lidar metrics (i.e. 31 features representing cover, 3D shape, vertical variability, horizontal variability and height of vegetation as well as microtopography), (3) assessing feature importance of lidar metrics for classifying wetland habitats, and (4) applying a Random Forest algorithm for mapping and prediction. We used an expert-based vegetation map for annotation and generated 100, 500 and 1000 annotation points for each class. Using a three-level hierarchical approach, we differentiated at level 1 planar surfaces (e.g. roads and agricultural fields) from wetland vegetation with 82% mean overall accuracy, using predominantly height and horizontal variability metrics. At level 2, we classified wetland vegetation into four land cover types (forest, grassland, reedbeds, shrubs) with 71% mean overall accuracy, using lidar metrics related to vegetation height and horizontal and vertical variability. At level 3, we differentiated two types of land reed as well as water reed with 78% mean overall accuracy, using predominantly vertical variability metrics. Our results demonstrate that lidar metrics (related to vegetation height, cover, vertical and horizontal variability) derived from country-wide ALS data can differentiate land cover types and habitats within wetlands at high resolution. Given appropriate annotation data, our workflow can be up-scaled to a country-wide extent to allow the comprehensive mapping and monitoring of wetlands at national scales.

Introduction
Wetlands are key habitats for animals and plants, are among the most vital ecosystems on Earth, and provide multiple ecosystem services to humans (Mitsch et al., 2015). Wetlands consist of a large variety of different land cover types and habitats, including meadows and wet grasslands formed by herbaceous vegetation (e.g. grasses, sedges and bulrush), patches of trees and shrubs (e.g. European alder, willows), and reedbeds (i.e. mono-dominant stands of the common reed Phragmites australis) (Ellenberg, 1996; Weller, 1999). Since plants in wetlands vary in life form and physical stature, they form different structural habitat components which are important for many wetland animals, including migratory and breeding birds (Weller, 1999; Strayer and Findlay, 2010). Wetlands in general, and reedbeds in particular, have been strongly modified by changes in water chemistry, climate, land-use, urban development...
and water level regulations (Ostendorp, 1989; Davidson, 2014). Consequently, international agreements such as the Ramsar Convention on Wetlands (https://www.ramsar.org/) urge to monitor and map wetlands for effective planning and natural resource management, and to foster national action and international cooperation for the conservation and wise use of wetlands and their resources.

Remotely sensed data provide an efficient method for mapping and monitoring biodiversity and habitat change across broad spatial extents and with fine resolution (Petitorelli et al., 2016; Guo et al., 2017; Navarro et al., 2017). Manual interpretation of stereographic aerial photographs has traditionally been used to map wetland extent (Howland, 1980). More recently, satellite-based multispectral (Shaday-deh et al., 2017) and radar (Muro et al., 2016) data, as well as hyperspectral imagery (Hirano et al., 2003), has allowed the (semi-)automatic mapping and monitoring of wetlands. Light detection and ranging (lidar) technology is increasingly used because it captures the 3D physical structure of terrestrial vegetation and habitats (Lefsky et al., 2002; Vierling et al., 2008) and provides fine-scale digital elevation models for characterizing microtopography (Brubaker et al., 2013).

A common type of lidar measurement is Airborne Laser Scanning (ALS) where the sensor is mounted, for instance, on an airplane or helicopter. The technique uses the time difference between the pulse emitted from the sensor and the return signal from the ground from which $x,y,z$ coordinates of points characterizing objects can then be calculated. Besides the coordinates of the points, the intensity of the returned pulse is recorded which depends on the backscattering properties and distribution of the objects. The resulting 3D point cloud typically varies in point density depending on the type of the scanner, acquisition parameters and whether single or multiple return or full-waveform data are recorded by the instrument. The resulting high-resolution information on the vertical and horizontal distribution of vegetation enables to quantify 3D habitat structures across large spatial extents (Coops et al., 2016; Moeslund et al., 2019). Because country-wide ALS data are now increasingly becoming available, there is potential to characterize fine-scale habitat structures not only in forests (where much of lidar-based research has been conducted) but also in low-stature habitats such as wetlands.

Lidar can be used to derive ecologically relevant information on 3D vegetation structure and requires the processing of the 3D point cloud into suitable metrics, for example, by statistically aggregating the 3D point cloud information within raster cells (Davies and Asner, 2014; Baks et al., 2019). These lidar metrics can quantify vegetation height (e.g. maximum of height values within a cell), vertical vegetation variability (e.g. the standard deviation of height values within a cell) or horizontal vegetation variability (e.g. the standard deviation of height values between neighbouring cells). Moreover lidar metrics have also been operationally used in forestry to derive stand inventory parameters such as biomass or basal area (Hyyppä et al., 2008; Cao et al., 2014). Within wetlands, commonly used lidar metrics are those derived from Digital Surface Models (DSMs), Canopy Height Models (CHMs) or Digital Terrain Models (DTMs) (Lindsay et al., 2004; Goodale et al., 2007; Gilmore et al., 2008; Onojeghuo et al., 2010; Onojeghuo and Blackburn, 2011; Lang et al., 2013; Chasmer et al., 2016; Rapinel et al., 2018). Height metrics such as the mean and maximum percentiles of height (Alexander et al., 2015; Ene et al., 2018), canopy density measures (Ding et al., 2011; Kopeć et al., 2016; Corti Meneses et al., 2017; Ene et al., 2018), the standard deviation of points within cells (Alexander et al., 2015; Kopeć et al., 2016; Ene et al., 2018), or the variance of the DSM have also been used in wetland studies (Zlinszky et al., 2012; Alexander et al., 2015; Chasmer et al., 2016). Beyond these examples, ALS is not yet widely used for characterizing wetland-related land cover types and for differentiating habitats within wetlands, probably because robust and open-source workflows and ground-based annotation datasets are largely lacking.

Here, we develop a new open-source workflow that uses fine-scale lidar metrics derived from country-wide ALS data for classifying land cover types and habitats. We test this with a case study in a Dutch wetland to make a first step towards upscaling the use of country-wide ALS data. To the best of our knowledge, this is the first study that uses country-wide ALS data for classifying land cover types and habitats within wetlands in the Netherlands. We use a wall-to-wall digitized expert-based map for annotation because ground-based annotation data were not available. The openly accessible ALS dataset comes with multiple returns and a point density of 6–10 points/m². We derived 31 lidar metrics to capture the 3D structure of the vegetation by quantifying aspects related to vegetation cover, 3D shape, vertical variability, horizontal variability and height, as well as fine-scale topographic variability. We then used a Random Forest algorithm to evaluate feature importance and to classify land cover types and habitats at different hierarchical levels. We expect that from the available set of lidar metrics only a few metrics are needed to classify wetland-related land cover types and habitats, and that metrics related to vegetation height and horizontal as well as vertical vegetation variability play an important role for classification.

**Materials and Methods**

**Airborne Laser Scanning data**

The ALS data (AHN2) was captured during the second Dutch national ALS flight campaign (Actueel
Hoogtebestand Nederland, https://www.pdok.nl/) which took place from 2007 to 2012. Our study area was captured in 2008 and 2009 in the leaf-off season (Northern hemisphere winter). The AHN2 dataset covers the whole Netherlands with a point density of 6–10 points/m² from multiple return data acquisition and with an overall vertical accuracy of 20 cm (https://www.ahn.nl/). The dataset is openly accessible (https://ahn.arcgisonline.nl/ahnviewer/) and has been pre-processed by “Rijkswaterstaat” (the executive agency of the Dutch Ministry of Infrastructure and Water Management). The available dataset comes with a ground classification that separates ground points from non-ground points. The dataset consists of the x,y,z coordinates of all points which are stored in LAZ format files (https://rapidlasso.com/laszip/). Other characteristics such as the number of returns, the intensity values and other flight-related parameters (e.g. scan angle, pulse reputation frequency and flight height) are not provided by the publisher of the dataset.

Study area and annotation dataset
We selected the study area Lauwersmeer in the northern part of the Netherlands in the province of Groningen because it is a relatively large (5754 ha) wetland area with diverse land cover types and habitats (Fig. 1). This area is recognized as a Ramsar site (site no. 1247) and has the status of a Nature 2000 site (site code NL9802012) because it is an important breeding site for coastal, water and reedbed birds. Since 2008, field data on vegetation and birds have been collected for monitoring the potential consequences of soil subsidence caused by gas extraction (Kleefstra et al., 2016). Based on these surveys, a digital vegetation map has been produced by an ecological research and consultancy firm (Altenburg & Wymenga, http://www.altwym.nl/). The mapping of vegetation structure in Lauwersmeer started in 2005 on the basis of vegetation height and coverage of the shrub and forest layers as well as the coverage of Phragmites australis within reedbeds. Updates were made in 2008/2009 and 2012. The final vegetation map was made in 2015 and consists of twenty-one digitized vegetation types (Fig. 1). These capture important habitats for animals (especially wetland birds). The vegetation map was derived from an expert-based interpretation of 3D stereoscopic aerial photographs using both textural and actual height information in combination with field visits (Kleefstra et al., 2016). We used the 1:5000 expert-based vegetation map as a source for our annotation data (training and testing points). The final vegetation map was made 6–7 years after the acquisition of the lidar data. The vegetation monitoring between 2007 and 2012 indicates that no substantial changes in vegetation structure have occurred within the annually visited transects (Bijkerk et al., 2013).

We used this expert-based vegetation map (Fig. 1) to create annotation data for the Random Forest algorithm. The algorithm was applied at three different levels which were hierarchically nested (Table 1). At level 1, the wetland class was separated into planar surfaces and wetland vegetation; at level 2, four vegetated wetland-related land cover types (forest, grassland, reedbed and shrubs) were separated from each other; and at level 3, the reedbed category was separated into three different types of reedbed habitats (water reed, structurally poor land reed, and structurally rich land reed). To apply the machine learning algorithm, we aggregated and simplified the original vegetation units for the classification (Table 1). We generally excluded polygons that describe vegetation units as “open” from the training step (see Fig. 1) because these consist of mixtures of vegetation (grasslands, shrubs, trees, reedbed patches) within the user-defined polygons. Furthermore, at level 1 we included low-stature vegetation (pioneer vegetation and grassland <0.3 m; Fig. 1) into ‘planar surfaces’. At level 2, we included vegetation polygons described as high grassland (>0.3 m), dense forest (only broad-leaved trees), dense reedbeds and dense shrub (Table 1). At level 3, we used the reedbed habitats dense water reed (flooded) and dense land reed (non-flooded). Land reed was further separated into structurally poor land reed vs. structurally rich land reed (Table 1). Structurally rich land reed consists of 25–50% cover of the common reed Phragmites australis, along with herbs as co-dominant vegetation (associations Phragmitetalia and Convululo-Filipenduletalia) (Kleefstra et al., 2016). In contrast, structurally poor land reed and water reed consist of >50% of common reed. The reclassified habitat types from the expert-based map (Table 1) served as input for generating the annotation dataset.

We performed a sensitivity analysis to test how the number of training points influences the overall accuracy of the Random Forest algorithm. For this, we randomly placed either 100, 500 or 1000 annotation points in each vegetation class (as defined in Table 1). We tested between 100 and 1000 annotation points because 100 points could be feasibly collected in the field using a GPS device (e.g. for applying the workflow in another area) whereas 1000 points provided a very dense sampling of each vegetation class.

Workflow
To process, analyse and classify the ALS data, we developed an open-source workflow (Fig. 2) with four main routines: (1) pre-processing of the ALS data, (2) calculation of lidar metrics, (3) selection of lidar metrics and (4) classification using the Random Forest algorithm. We used the R package lidR (https://github.com/Jean-Roma
Figure 1. Expert-based vegetation map of the Lauwersmeer wetland area located in the north of the Netherlands (see inset). This habitat map was originally derived from an interpretation of aerial photographs in combination with field visits. The colours indicate 21 different habitat types. A true colour aerial photograph from 2012 is shown in the back (ESRI Nederland). The two cross plots (A and B, also indicated in the map) provide examples of vegetation structure as visible in the Airborne Laser Scanning (ALS) point clouds (green: vegetation, dark brown: ground points). The ALS point clouds were derived from the AHN2 dataset (see Methods). The transect width for the cross plots is 2 m.
in/lidR) for handling and processing the ALS data and for calculating the lidar metrics, and additional R libraries (raster, randomForest, caret) for analysing and classifying the raster layers. The scripts for running the developed workflow are available via GitHub (https://github.com/eEcoLiDAR/PhDPaper1_Classifying_wetland_habitats).

**Pre-processing of ALS data**

The first routine included three pre-processing steps: (1) tiling, (2) water masking and (3) normalization of the height above the ground (Fig. 2). First, the ALS data of the whole study area was tiled into 2.5 km \( \times \) 2.5 km tiles to allow parallel processing within the lidR environment. Second, a water mask was derived based on the official Digital Terrain Model product (based on AHN2 at 0.5 m resolution obtained from https://ahn.arcgisonline.nl/ahn/viewer/). In this product, not-available (NA) values indicate the presence of water and we used this information to exclude water bodies from our dataset. Third, the height (\( z \) value) of the point cloud was normalized using the officially published ground classification product. We selected the nearest 20 ground points in the neighbourhood of each non-ground point and interpolated a local terrain model using the inverse distance weighting interpolation method with default settings (i.e. power equal to 2). The height (normalized \( z \) value) of each non-ground point was then subtracted from the height of the underlying terrain surface to estimate the absolute (normalized) height of each non-ground point in the point cloud.

**Calculation of lidar metrics**

In the second routine of the workflow (Fig. 2), we calculated thirty-one lidar metrics (Table 2). These lidar metrics were calculated at four different spatial resolutions (1 m, 2.5 m, 5 m and 10 m grid cells). The 31 lidar metrics (Table 2) belong to six feature classes (vegetation cover, 3D shape, vertical variability, horizontal variability, height and topography) which reflect different aspects of the 3D vegetation structure and microtopography. All metrics were calculated using the normalized point cloud. Furthermore, buildings were removed from the datasets before the analysis by using a cadastre dataset from the Netherlands as a mask (https://zakelijk.kadaster.nl/-/top10nl). Finally, we included ground points for lidar metrics at level 1 (planar surfaces vs. vegetation) and for the topography-related metrics at all levels. The code for the implementation of each lidar metric is available from GitHub (https://github.com/eEcoLiDAR/PhDPaper1_Classifying_wetland_habitats).

**Lidar metrics selection**

In the third routine of the workflow (Fig. 2), the calculated lidar metrics were analysed with the Random Forest algorithm (Breiman, 2001). This was done in three steps. First, a sensitivity analysis was performed to identify the influence of (1) the spatial resolution of the lidar metrics (1, 2.5, 5 and 10 m), and (2) the number of annotation points on the overall accuracy of the classification. We used the randomly placed annotation points (100, 500 and 1000 points per vegetation class), split them into training (75%) and testing (25%) data, and then used the Random Forest algorithm (run 100 times and splitting the annotation data each time randomly) to classify land cover types and habitats at each of the three levels (level 1–3). The number of trees was set to 100 and the number of variables randomly sampled at each split was set to the square root of the input variables across all RF runs. We applied the RF using lidar metrics calculated either at 1 m, 2.5 m, 5 m or 10 m resolution. We then calculated the mean ± SD of Overall Accuracies.

### Table 1. Overview of the three hierarchical levels and their vegetation classes used in the Random Forest classification. The original vegetation units from the expert-based vegetation map (Fig. 1) were re-classified into vegetation classes within three hierarchical levels (wetland, land cover types and reedbed habitats).

| Hierarchical level | Vegetation classes | Original vegetation units from expert-based vegetation map |
|--------------------|--------------------|-------------------------------------------------------|
| Level 1: Wetland   | Vegetation         | Grassland (high), forest (dense), land reed structurally rich (dense), land reed structurally poor (dense), water reed, shrub low (dense), shrub mid-height (dense), shrub high (dense) |
|                    | Planar surfaces    | Open water, bare ground, pioneer vegetation, grassland (low), anthropogenic objects |
| Level 2: Land cover types | Forest | Forest (dense) |
|                    | Grassland         | Grassland (high) |
|                    | Reedbed           | Land reed structurally rich (dense), land reed structurally poor (dense), water reed |
|                    | Shrubs            | Shrub low (dense), shrub mid-height (dense), shrub high (dense) |
| Level 3: Reedbed habitats | Structurally rich land reed | Land reed structurally rich (dense) |
|                    | Structurally poor land reed | Land reed structurally poor (dense) |
|                    | Water reed        | Water reed (dense) |
In the second step, we performed a collinearity analysis of the 31 lidar metrics using the Variance Inflation Factor (VIF) (Naimi, 2015). While the Random Forest algorithm itself is not sensitive to multicollinearity among predictor variables, the subsequent ranking of the lidar metrics is, given that highly correlated metrics have a similar feature importance value which can affect a backward feature selection procedure. We set the VIF threshold to 5 (James et al., 2013) to avoid multi-collinearity, and all lidar metrics with VIF >5 were excluded from further analyses (see Appendix S1 Table S1).

In the third step, we applied the Recursive Feature Elimination (RFE) procedure (Guyon et al., 2002) to identify the most important lidar metrics. The RFE procedure uses the Random Forest algorithm with a backward selection of lidar metrics based on ranking the features (i.e., lidar metrics) by their importance. The feature importance was assessed using the mean decrease in impurity (MDI) (Breiman, 2001). The MDI is defined as the total decrease in node impurity, averaged over all decision trees within the random forest algorithm (https://cran.r-project.org/web/packages/randomForest/randomForest.pdf). We parametrized the RFE using cross-validation by dividing the annotation points into training (75%) and testing (25%) data (100 times). We then extracted the mean and confidence interval (CI) of the accuracy values across the 100 runs. The RFE then identifies the minimum set of lidar metrics, using a 5% loss in accuracy around the highest accuracy value. This analysis was done separately for each of the three hierarchical levels.
| Feature class | Name of lidar metric | Metric abbreviation | Description | Reference |
|---------------|----------------------|---------------------|-------------|-----------|
| Cover         | Pulse penetration ratio | C_puls | Ratio of number of ground points to total number of points within a cell | Cao et al. (2014) |
|               | Density of vegetation points below 2 m* | C_b2 | Ratio of number of vegetation points to the total below 2 m | |
|               | Density of vegetation points between 2 and 5 m* | C_2.5 | Ratio of number of vegetation points to the total between 2 and 5 m | |
|               | Canopy cover* | C_can | Number of returns above mean height within a cell | |
| 3D shape      | Curvature | S_curv | Ratio between the smallest eigenvalue and the sum of eigenvalues within a cell | Weinmann et al. (2017) |
|               | Linearity* | S_lin | Ratio of the difference of the largest and medium eigenvalues and the largest eigenvalues within a cell | |
|               | Planarity | S_plan | Ratio of the difference between medium and smallest eigenvalues and the largest eigenvalue within a cell | |
|               | Sphericity | S_sph | Ratio between smallest and largest eigenvalues within a cell | |
|               | Anisotropy | S_ani | Ratio of the difference between largest and smallest and the largest eigenvalue within a cell | |
| Vertical variability | Standard deviation of height | VV_sd | Standard deviation of z within a cell | Zlinszky et al. (2012) |
|               | Variance of height* | VV_var | Variance of z within a cell | Barnes, Islam and Auer (2016) |
|               | Canopy relief ratio | VV_cr | Ratio of difference between mean and minimum z to the difference between maximum and minimum z | |
|               | Vertical density ratio* | VV_vdr | Ratio between maximum and median z and maximum z within a cell | |
|               | Skewness of height* | VV_skew | Skewness of z within a cell | Bae et al. (2014) |
|               | Kurtosis of height* | VV_kurt | Kurtosis of z within a cell | |
|               | Coefficient of variation of height* | VV_coefvar | Coefficient of variation of z within a cell | |
|               | Simpson index | VV_simp | One divided by the sum of the squares of the proportion of the points within 0.5 m height layers within a cell | |
|               | Shannon index | VV_shan | The negative sum of the proportion of points within 0.5 m height layers multiplied with the logarithm of the proportion of points within 0.5 m height layers within a cell | |
| Horizontal variability | Standard deviation of DSM | HV_sd | Standard deviation of the digital surface model based on 3 × 3 neighbourhood | Zlinszky et al. (2012) |
|               | Variance of DSM* | HV_var | Variance of the digital surface model based on 3 × 3 neighbourhood | |
|               | Roughness of DSM | HV_rough | The maximum difference of the digital surface model based on 4 neighbouring pixels | Wilson et al. (2007) |
|               | Topographic position index of DSM* | HV_tpi | Difference between a central pixel and the mean of 4 neighbouring cells | |
|               | Terrain ruggedness index of DSM | HV_tri | Mean difference between a central pixel and 4 neighbouring cells | |
| Height        | Maximum height | H_max | Maximum of z | Hyppä et al. (2008) |
|               | Mean height | H_mean | Mean of z | |
|               | Median of height | H_med | Median of z | |
|               | 25th percentile of height* | H_25p | 25th percentile of z | |
|               | 75th percentile of height | H_75p | 75th percentile of z | |
|               | 90th percentile of height | H_90p | 90th percentile of z | |

(Continued)
Classification

In the fourth routine of the workflow (Fig. 2), we used the selected set of lidar metrics at each hierarchical level to classify different land cover types and habitats (Table 1). We first build the Random Forest algorithm with 100 decision trees using randomly split points for training (75%) and validation (25%) over 100 runs. We then used the confusion matrix to assess the accuracy of the classification. We note that the accuracy of the classification is quantified relative to an expert-based vegetation map which also contains unreported classification errors. For the error assessment of the classified vegetation map, we calculated the Producer’s Accuracy (precision) and User’s Accuracy (recall) for each vegetation class, as well as the Overall Accuracy and Cohen’s Kappa values (Congalton, 1991). We finally used the Random Forest algorithm to make a prediction and map of the vegetation classes across the wetland.

For the most important lidar metrics (as identified by the RFE) we used partial dependence plots (Friedman, 2001) to visualize the relationship (i.e. response curves) between the lidar metrics and the predicted probability from the Random Forest algorithm. Partial dependence plots illustrate the probabilities (ranging from 0 to 1) that a particular vegetation class is present along the gradient of the selected lidar metric. The partial dependence plots can therefore support the interpretation of the ecological relevance of each lidar metric.

Results

Lidar metrics selection

The sensitivity analysis (Table 3, Appendix S1 Fig. S1) confirmed that the highest accuracies for the classification were achieved when the spatial resolution of the lidar metrics (5 m) matched the spatial scale of the annotation data (1:5000 map). An increase in the number of annotation points for the classification at 5 m resolution resulted in a 2% increase in Overall Accuracy at level 3 (79% for 100, 80% for 500 and 81% for 1000 annotation points), and a 1% increase at level 1 (86% for 100 and 500 and 87% for 1000 annotation points). Hence, a smaller number of annotation points still provided good classification results. Moreover, the use of 1000 points resulted in a dense clustering of points, especially for vegetation classes with a small areal coverage, thus introducing a high level of pseudo-replication. We thus opted for 500 annotation points and a 5 m spatial resolution for the remaining analyses.

The results of the collinearity analysis suggested that 17 out of 31 metrics were highly correlated with other metrics (VIF > 5). The remaining fourteen metrics (Table S1) had VIF < 2.3. The maximum Pearson correlation in the remaining lidar metrics was $r = 0.6$ between linearity ($S_{lin}$) and density of vegetation points below 2 m ($C_{b2}$). All other lidar metrics had lower correlations ($r < 0.6$).

The RFE procedure identified a minimum set of lidar metrics (out of the 14) for the classification at each of the three hierarchical levels (Fig. 3a–c). At level 1 (Fig. 3a), the first 2 lidar metrics already achieved $82 \pm 2\%$ Overall Accuracy (mean ± CI, $n = 100$ runs). These two most important lidar metrics represented vegetation height measured as 25th percentile of normalized Z ($H_{25p}$) as well as horizontal variability measured as variation of the DSM ($HV_{var}$) (both with MDI values >5; Fig. 3d). At level 2, three lidar metrics were sufficient to separate different wetland-related land cover types, reaching $71 \pm 2\%$ Overall Accuracy (Fig. 3b). The three metrics represented measures of horizontal variability ($HV_{var}$), vertical variability ($VV_{var}$) and vegetation height ($H_{25p}$) (MDI values >9; Fig. 3e). At level 3, the
The Overall Accuracies (mean ± SD across 100 runs) are given for three hierarchical levels (see Table 1, also for the vegetation classes). At each resolution, the 100, 500 and 1000 annotation points were placed randomly in each vegetation class. At 1 m resolution, only 100 and 500 annotation points were tested. A visualization of the data from this table is presented in Appendix S1 Figure S1.

**Table 3.** Results of the sensitivity analysis based on a Random Forest algorithm using varying amounts of annotation points (100, 500 and 1000) at different spatial resolutions (10 m, 5 m, 2.5 m, and 1 m).

| Resolution | 10 m | 5 m | 2.5 m | 1 m |
|------------|------|-----|-------|-----|
| Number of annotation points per vegetation class | 100 | 500 | 1000 | 100 | 500 | 1000 | 100 | 500 |
| Level 1: Wetland (Overall Accuracy [%]) | 81 ± 4 | 85 ± 2 | 86 ± 1 | 86 ± 4 | 86 ± 2 | 87 ± 1 | 79 ± 5 | 87 ± 2 | 87 ± 1 | 80 ± 5 | 83 ± 2 |
| Level 2: Land cover types (Overall Accuracy [%]) | 67 ± 4 | 74 ± 2 | 76 ± 1 | 70 ± 4 | 71 ± 2 | 73 ± 1 | 70 ± 4 | 69 ± 2 | 69 ± 1 | 62 ± 4 | 66 ± 2 |
| Level 3: Reedbed habitat (Overall Accuracy [%]) | 79 ± 4 | 80 ± 2 | 80 ± 1 | 79 ± 4 | 80 ± 2 | 81 ± 1 | 69 ± 5 | 76 ± 2 | 75 ± 1 | 65 ± 5 | 65 ± 5 |

The Overall Accuracies (mean ± SD across 100 runs) are given for three hierarchical levels (see Table 1, also for the vegetation classes). At each resolution, the 100, 500 and 1000 annotation points were placed randomly in each vegetation class. At 1 m resolution, only 100 and 500 annotation points were tested. A visualization of the data from this table is presented in Appendix S1 Figure S1.

**Classification**

The fine-scale land cover and habitats predicted from the Random Forest algorithm with the minimum set of lidar metrics was mapped across the wetland (Fig. 4). A visual comparison of the lidar-based prediction (Fig. 3) and the expert-based vegetation map (Fig. 1) suggested that the general distribution of forest, shrub, and reedbed habitats was similar. One apparent misclassification was that structurally poor land reed was wrongly classified as structurally rich land reed on the lidar-based map in the southern part when compared to the expert-based map (E6°11′, N53°21′) (compare Fig. 1 and Fig. 4).

The Overall Accuracy of the classifications (Table 4) ranged from 71–82%, and Cohen’s Kappa ranged from 0.61 to 0.67. At level 1, the User’s and Producer’s Accuracy were both high (>81%; Table 4). At level 2, forest was classified with a Producer’s and User’s Accuracy of >83%, but grassland, reedbed and shrub achieved lower Producer’s and User’s Accuracy (41–75%; Table 4). At level 3, the Random Forest algorithm achieved the separation of structurally rich land reed from the other two reedbed classes with ≥92% Producer’s and User’s Accuracies (Table 4). The structurally poor land reed and the water reed were classified with lower success (53–78% Producer’s and User’s Accuracy; Table 4).

The partial dependence plots of the most important lidar metrics revealed how land cover types and habitats within the three hierarchical classification levels were separated along gradients of vegetation structure (Fig. 5). At level 1, vegetation and planar surfaces clearly differed in their 25th percentile of normalized height (H_25p), with planar surfaces having a high probability below a threshold of 0.5 m (Fig. 5a). At level 2, the variance of the DSM (HV_var) separated different wetland-related land cover types, with reedbed and grassland showing a high probability when variability in vegetation height across neighbouring cells was low (<2.5 m; Fig. 5b). Instead, shrub and forest showed a high probability when HV_var was >2.5 m (Fig. 5b). Grassland and reedbeds also showed a high probability when the variance of height (VV_var) was <2.5 m whereas shrubs and forest had a high probability >2.5 m (Fig. 5c). At level 3, water reed could be distinguished from both land reed classes when the variance of normalized height (VV_var) was >7.5 m (Fig. 5d). Water reed was also well separated from structurally rich land reed when the 25th percentile of height (H_25p) was <2 m (Fig. 5e) or when the density of vegetation below 2 m was >10% (Fig. 5f).

**Discussion**

We tested 31 lidar metrics for classifying and mapping land cover types and habitats in a Dutch wetland using country-wide ALS data. The RFE procedure together with the hierarchical classification approach revealed that five lidar metrics were of key importance to separate fine-scale habitats within this wetland. These lidar metrics reflected horizontal variability (HV_var), vertical variability (VV_var), height (H_25p) and cover (C_can, C_b2) of the vegetation. The results confirmed our expectation that the total set of lidar metrics can be substantially reduced for classifying wetland-related land cover types and habitats, and that lidar metrics related to vegetation height and horizontal as well as vertical...
Figure 3. Results of the feature selection procedure showing the classification accuracy (mean Overall Accuracy ± CI) as a function of the number of lidar metrics within each hierarchical classification level (A–C), and the feature importance of each lidar metric within these levels (D–F). Level 1 separates vegetation from planar surfaces (A and D), level 2 four different wetland-related land cover types (forest, grassland, reedbed and shrub) (B and E), and level 3 three different types of reedbed habitats (structurally poor land reed, structurally rich land reed and water reed) (C and F). The red vertical lines indicate the minimum set of lidar metrics as suggested by the Recursive Feature Elimination (RFE) algorithm. Feature importance is measured with the mean decrease in impurity (MDI) and the selected subset of lidar metric names is indicated with red font. See Table 2 for abbreviations and explanations of lidar metrics.
Figure 4. Fine-scale land cover and habitat map of the Lauwersmeer area predicted from one example tree of a Random Forest algorithm using lidar metrics derived from country-wide Airborne Laser Scanning data. Classification was performed separately at three hierarchical levels (levels 1–3) and predictions were then combined into one map, with colour coding showing the predicted land cover types and habitats at the finest hierarchical level (white boxes in the legend indicate that the respective habitat type is classified at a finer level). See Table 1 for the vegetation classes, and Figure 1 for the expert-based vegetation map that was the basis for creating training and testing data.
variability play a crucial role for ALS-based wetland habitat classifications.

**Lidar metrics selection**

The spatial resolution of lidar metrics may affect the classification accuracy. Many lidar-based wetland studies have used spatial resolutions below 5 m (Onojeghuo et al., 2010; Zlinszky et al., 2012; Luo et al., 2015; Kopeć et al., 2016; Chasmer et al., 2016; Corti Meneses et al., 2017). However, the expert-based vegetation map which we used for annotation in our study does not capture the fine-scale heterogeneity of vegetation structure within polygons. This was confirmed by our sensitivity analysis which showed that classifications at 1 m or 2.5 m resolution showed less accuracy compared to classifications at 5 m or 10 m resolution. Other annotation data (e.g. from vegetation plots) may allow better classification accuracies at resolutions below 5 m (e.g. Alexander et al., 2015; Corti Meneses et al., 2017). Our results therefore suggest that the best use of the available expert-based vegetation map for classification in our study is at a moderate (5–10 m) spatial resolution. Studies using expert-based maps for annotation could additionally explore fuzzy classification methods which may allow to handle classification uncertainties along vegetation class boundaries (Zlinszky and Kania, 2016). In our study, the availability of a wall-to-wall expert-based vegetation map allowed us to test the influence of the number of annotation points on the overall accuracy of the classification. Such expert-based maps allow to generate hundreds or thousands of annotation points (Millard and Richardson, 2013), while studies using field-based annotation data usually rely on less because they have to be collected in situ (Alexander et al., 2015; Kopeć et al., 2016). Our sensitivity analysis implies that using 1000 instead of 100 annotation points only slightly increases the overall accuracy of the classification (from 86% to 87% at level 1, from 70% to 73% at level 2 and from 79% to 81% at level 3). This suggests that a smaller amount of annotation points (e.g. 100) could be sufficient to classify vegetation classes within wetlands in the Netherlands. This would make the collection of *in situ* measurements for annotation in other wetlands feasible. It would only require to record the land cover type or habitat with a GPS at each annotation point which could be rapidly done within a few days at a given study site.

The collinearity analysis in combination with the RFE procedure allowed us to reduce the number of lidar metrics in the classification at each hierarchical level. A similar approach to metric reduction has been used in other lidar-based vegetation classifications (Millard and Richardson, 2013; Alexander et al., 2016; Koma et al., 2016). In our study, only five lidar metrics related to vegetation structure where needed to separate different land cover types and habitats within this wetland. Similar to the wetland study of Zlinszky et al. (2012), we found that vegetation height (25th percentile of height) and horizontal variability (variance of the DSM) were the most important metrics to separate planar surfaces from vegetation. However, 3D shape metrics did not play a particularly important role in separating planar surfaces from vegetation. This is in contrast to studies in urban settings (Weinmann et al., 2017) and agricultural landscapes (Lucas et al., 2019) where neighbourhood-based features based on eigenvalue characteristics of the point cloud seem to be crucial for identifying planar surfaces. For separating wetland-related land cover types such as forest, grassland, reedbed and shrubs, the variance of height

---

**Table 4.** Classification accuracies based on the confusion matrices for each of the three hierarchical levels (mean ± SD across 100 runs) using 500 annotation points and lidar metrics calculated at 5 m resolution.

| Vegetation class          | User's accuracy [%] | Producer's accuracy [%] | Overall accuracy [%] | Cohen’s Kappa |
|---------------------------|---------------------|-------------------------|----------------------|--------------|
| Level 1: Wetland          |                     |                         |                      |              |
| Planar surfaces           | 81 ± 2              | 83 ± 3                  | 82 ± 2               | 0.64 ± 0.04 |
| Vegetation                | 83 ± 3              | 81 ± 3                  |                      |              |
| Level 2: Land cover types |                     |                         |                      |              |
| Forest                    | 83 ± 3              | 90 ± 3                  | 71 ± 2               | 0.60 ± 0.03 |
| Grassland                 | 53 ± 5              | 41 ± 4                  |                      |              |
| Reedbed                   | 71 ± 3              | 75 ± 4                  |                      |              |
| Shrub                     | 67 ± 4              | 68 ± 5                  |                      |              |
| Level 3: Reedbed habitat  |                     |                         |                      |              |
| Land reed (rich)          | 92 ± 2              | 96 ± 2                  | 78 ± 2               | 0.67 ± 0.03 |
| Land reed (poor)          | 65 ± 4              | 53 ± 5                  |                      |              |
| Water reed                | 71 ± 3              | 78 ± 4                  |                      |              |

User’s Accuracy (precision) and Producer’s Accuracy (recall) are summarized for all habitats within each hierarchical level. Overall Accuracy and Cohen’s Kappa are given for each hierarchical level.
became additionally important. This mostly reflects the key structural differences between tall and low-stature vegetation (Alexander et al., 2015; Ene et al., 2018). Finally, metrics related to vegetation cover (e.g. canopy cover and density of vegetation below 2 m) together with metrics capturing vertical variability (variance of height) and vegetation height (25th percentile of height) were crucial to separate water reed from land reed. This supports evidence from other studies that have used point density, height and variability of the DSM to quantify the structure and extent of reedbeds along lakeshores (Millard and Richardson, 2013; Corti Meneses et al., 2017).

The reduced number of lidar metrics facilitated the ecological interpretation of the Random Forest algorithm by focusing on the most important lidar metrics and their specific response curves. The partial dependence plots revealed particular thresholds along gradients of vegetation structure that separate one type of vegetation class from another. For instance, the separation of water reed from land reed showed thresholds for land reed along gradients of variance of height (VV_var > 7 m), 25th percentile of height (H_25p between 0 and 2.8 m) and density of vegetation below 2 m (C_b2 < 18%). Clear thresholds also existed for separating planar surfaces from vegetation (H_25p < 0.4 m) and for using horizontal variability for separating grasslands and reedbeds (HV_var < 2.5 m) from shrubs, forests patches and individual trees (HV_var > 2.5 m). These identified thresholds together with repeated ALS surveys could guide wetland management and monitoring, for example, in terms of vegetation structural changes that are important for breeding birds. Furthermore, the most relevant information on vegetation structure derived from ALS (i.e. lidar metrics related to vegetation height and vegetation density) corresponds to the criteria used for deriving the expert-based vegetation map.

Figure 5. Partial dependence plots showing the most important lidar metrics to differentiate wetland-related land cover types and habitats within three hierarchical levels (level 1–3) using a Random Forest algorithm. (A) Probability of having planar surfaces vs. vegetation (level 1) along a gradient of vegetation height (H_25p) quantified as 25th percentile of height. (B and C) Probability of having forest, grassland, reedbed and shrub (level 2) along gradients of horizontal variability quantified with the roughness of the DSM (HV_var) and vertical variability as variance of height (VV_var). (D–F) Probability of having different types of reedbeds along gradients of variance of height (VV_var), 25th percentile of height (H_25p) and cover defined as density of vegetation below 2 m (C_b2). See Table 2 for details of the lidar metrics and Figure 3D–F for the feature importance of each metric in the classification.

Figure 5.
Classification

For wetland vegetation classifications, some studies have reported >85% Overall Accuracies (Goodale et al., 2007; Gilmore et al., 2008; Zlinszky et al., 2012; Kopce et al., 2016; Shadaydeh et al., 2017; Rapinel et al., 2018). This is slightly higher than in our study and could be explained by the inclusion of spectral information in the classification (e.g. Onojeghuo and Blackburn, 2011; Zlinszky et al., 2012; Alexander et al., 2015). In comparison to other sensors (e.g. optical), lidar captures the structural aspects of vegetation which is its main advantage. The ALS data were acquired in the leaf-off season which could result in lower accuracies because the variability of vegetation structure might not be fully captured (Onojeghuo et al., 2010; Wasser et al., 2013). To the best of our knowledge, the effect of seasonality on capturing vegetation structure from ALS data (i.e. leaf-off vs. leaf-on season) has not yet been investigated within wetlands, and needs to be addressed in the future. A further classification error could be introduced by the use of the expert-based vegetation map for annotation because it is derived from an interpretation of 3D stereoscopic aerial photographs together with field visits. This means that we validate our automatic (ALS-based) classification against a vegetation map derived from human cognitive processes. This may contain unreported classification errors. Furthermore, there was also a time difference between the mapping of the annotation data (2015 and before) and the collection of the lidar data (2008–2009) which could further lower the classification accuracies. Nevertheless, the achieved accuracies (71–78%) for separating level 2 and level 3 vegetation classes (i.e. different wetland-related land cover types and reedbed habitats) correspond to other wetland studies in which accuracies range from 57% to 88% (Goodale et al., 2007; Zlinszky et al., 2012; Millard and Richardson, 2013; Alexander et al., 2015; Corti Meneses et al., 2017). The combination of lidar data with hyperspectral information (Onojeghuo and Blackburn, 2011) or with Synthetic-Aperture Radar (SAR) (Millard and Richardson, 2013; Montgomery et al., 2019) could further improve the classification of land cover and habitat types.

Conclusion

With the increasing availability of open-access, country-wide ALS datasets, our developed workflow offers a promising step towards monitoring vegetation structure in wetlands at fine spatial resolution and over broad spatial extents. Such monitoring of habitat structures is essential for developing an observation system that informs decision makers on global biodiversity change (Navarro et al., 2017). Most of the currently existing studies using ALS data to classify wetland-related land cover types and habitats have so far focused on relatively local sites (<100 km²), with the exception of two studies covering wetland areas of >1000 km² (Zlinszky et al. 2012; Chasmer et al. 2016). A main bottleneck for upscaling local ALS-based habitat studies is the lack of consistent in-situ annotation data within and across countries and regions. The use of expert-based vegetation maps (as in our study) is limited for upscaling because mapping schemes and their vegetation units often differ among sites. Nevertheless, additional annotation data could be derived from other sources such as vegetation inventories (e.g. georeferenced vegetation plot data obtained from standardised field sampling protocols), or from sampling information on land cover and habitat types with a GPS across multiple wetland sites. Furthermore, in many cases the collection of field data is not necessarily done simultaneously with the lidar flight campaigns, which introduces some additional uncertainty in the cross-validation procedure. A stronger coordination between communities of ecologists and remote sensing scientists is therefore desirable for monitoring biodiversity change and for the conservation and management of wetlands and other habitats.

Acknowledgments

This work is part of the project “eScience infrastructure for Ecological applications of LiDAR point clouds” (eEcoLiDAR) (Kissling et al., 2017), funded by the Netherlands eScience Center (https://www.esciencecenter.nl), grant number ASD1.2016.014. We thank three anonymous reviewers for constructive comments, the ecological consultancy firm Altenburg & Wymenga for providing the expert-based vegetation map of the study area, and the eScience engineers Meiert Grootes, Christiaan Meijer and Romulo Goncalves for useful discussions.

Data Availability Statement

The AHN2 ALS data of the Actueel Hoogtebestand Nederland (https://www.pdok.nl/) is openly accessible (https://ahn.arcgisonline.nl/ahnviewer/). The expert-based vegetation map used for annotation can be requested from the ecological consultancy firm Altenburg & Wymenga (http://www.altwym.nl/). We provide a script for automatic downloading of the required ALS data on GitHub (https://github.com/eEcoLiDAR/PhDPaper1_Classifying_wetland_habitats). All calculated lidar metrics (with a 5 m × 5 m resolution) are further available from the Zenodo data repository (https://doi.org/10.5281/zenodo.3865914). All scripts for running the developed workflow are available via GitHub (https://github.com/eEcoLiDAR/PhDPaper1_Classifying_wetland_habitats).
References

Alexander, C., B. Deak, A. Kania, W. Mücke, and H. Heilmeier. 2015. Classification of vegetation in an open landscape using full-waveform airborne laser scanner data. *Int. J. Appl. Earth Obs. Geoinf.* 41, 76–87. https://doi.org/10.1016/J.IJAG.2015.04.014

Alexander, C., B. Deak, and H. Heilmeier. 2016. Micro-topography driven vegetation patterns in open mosaic landscapes. *Ecol. Ind.* 60, 906–920. https://doi.org/10.1016/J. ECOLIND.2015.08.030

Bae, S., B. Reineking, M. Ewald, and J. Mueller. 2014. Comparison of airborne lidar, aerial photography, and field surveys to model the habitat suitability of a cryptic forest species – the hazel grouse. *Int. J. Remote Sens.* 35, 6469–6489. https://doi.org/10.1080/01431161.2014.951545

Bakx, T. R. M., Z. Koma, A. C. Seijmonsbergen, and W. D. Kissling. 2019. Use and categorization of light detection and ranging vegetation metrics in avian diversity and species distribution research. *Divers. Distrib.* 25, 1045–1059. https://doi.org/10.1111/ddi.12915

Barnes, K. W., K. Islam, and S. A. Auer. 2016. Integrating LiDAR-derived canopy structure into cerulean warbler habitat models. *J. Wildl. Manag.* 80, 101–116. https://doi.org/10.1002/jwmg.995

Bijkerk, W., R. Bakker, and R. Buiks. (2013). Monitoring effecten van bodemdaling op vegetatie in de Lauwersmeer, Verslag monitorsperiodiode 2007 t/m 2012. A&W rapport 1885. Altenburg & Wymenga ecologisch onderzoek, Feanwälden.

Breiman, L. 2001. Random forests. *Mach. Learn.* 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brubaker, K. M., W. L. Myers, P. J. Drohan, D. A. Miller, and E. W. Boyer. 2013. The use of LiDAR terrain data in characterizing surface roughness and microtopography. *Appl. Environ. Soil Sc.* 2013, 1–13. https://doi.org/10.1155/2013/891534

Cao, L., N. Coops, T. Hermosilla, J. Innes, J. Dai, and G. She. 2014. Using small-footprint discrete and full-waveform airborne LiDAR metrics to estimate total biomass and biomass components in subtropical forests. *Remote Sens.* 6, 7110–7135. https://doi.org/10.3390/rs6087110

Chasmer, L., C. Hopkinson, J. Montgomery, and R. Petrone. 2016. A physically based terrain morphology and vegetation structural classification for wetlands of the boreal plains, Alberta, Canada. *Can. J. Remote Sens.* 42, 521–540. https://doi.org/10.1080/07038992.2016.1196583

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 37, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Coops, N. C., P. Tompaski, W. Nijland, G. J. M. Rickbeil, S. E. Nielsen, C. W. Bater, et al. 2016. A forest structure habitat index based on airborne laser scanning data. *Ecol. Ind.* 67, 346–357. https://doi.org/10.1016/J.ECOLIND.2016.02.057

Corri Meneses, N., S. Baier, J. Geist, and T. Schneider. 2017. Evaluation of green-LiDAR data for mapping extent, density and height of aquatic reed beds at Lake Chiemsee, Bavaria—Germany. *Remote Sens.* 9, 1308. https://doi.org/10.3390/rs9121308

Davidson, N. C. 2014. How much wetland has the world lost? Long-term and recent trends in global wetland area. *Mar. Freshw. Res.* 65(10), 934. https://doi.org/10.1071/MF14173

Davies, A. B., and G. P. Asner. 2014. Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends Ecol. Evol.* 29, 681–691. https://doi.org/10.1016/J.TREE.2014.10.005

Ding, Q., W. Chen, B. A. King, Y. Chen, Y. Liu, and H. Zhang. 2011. Coastal wetland investigations by airborne LiDAR: a case study in the Yellow River Delta, China. *Mar. Technol. Soc. J.* 45, 59–70. https://doi.org/10.4031/MTSJ.45.5.6

Ellenberg, H., and C. Leuschner. 1996. Vegetation Mitteleuropas mit den Alpen in ökologischer, dynamischer und historischer Sicht: 170 Tabellen. Ulmer.

Ene, L. T., T. Gobakken, H.-E. Andersen, E. Næsset, B. D. Cook, D. C. Morton, et al. 2018. Large-area hybrid estimation of aboveground biomass in interior Alaska using airborne laser scanning data. *Remote Sens. Environ.* 204, 741–755. https://doi.org/10.1016/J.RSE.2017.09.027

Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 29, 1189–1232. https://doi.org/10.1214/aoas1013203451

Gilmore, M. S., E. H. Wilson, N. Barrett, D. L. Civo, S. Prisloe, J. D. Hurd, et al. 2008. Integrating multi-temporal spectral and structural information to map wetland vegetation in a lower Connecticut River tidal marsh. *Remote Sens. Environ.* 112, 4048–4060. https://doi.org/10.1016/J.RSE.2008.05.020

Goodale, R., C. Hopkinson, D. Colville, and D. Amirault-Langlais. 2007. Mapping piping plover (*Charadrius melanodus melodus*) habitat in coastal areas using airborne lidar data. *Can. J. Remote Sens.* 33, 519–533. https://doi.org/10.5589/m07-058

Guo, M., J. Li, C. Sheng, J. Xu, and L. Wu. 2017. A review of wetland remote sensing. *Sensors* 17, 777. https://doi.org/10.3390/rs17040777

Guyon, I., J. Weston, S. Barmhill, and V. Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422. https://doi.org/10.1023/A:1012487302797

Hirano, A., M. Madden, and R. Welch. 2003. Hyperspectral image data for mapping wetland vegetation. *Wetlands* 23, 436–448. https://doi.org/10.1672/18-20.

Howland, W. G. 1980. Multispectral aerial photography for wetland vegetation mapping. *Photogramm. Eng. Remote Sens.* 46, 87–99. Available at: https://www.cabdirect.org/cabdirect/abstract/19800666164 (Accessed: 27 April 2019).
Kleefstra, R., P. de Boer, C. Kampichler. 2016. Land and Water Monitoring in the Lauwersmeer in 2016. Sovon Vogelonderzoek, Nederland, Nijmegen.

Koma, Z., K. Koenig, and B. Hölle. 2016. Urban tree classification using full-waveform airborne laser scanning. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. III-3, 185–192. https://doi.org/10.5194/isprs-annals-III-3-185-2016

Kopč, D., D. Michalska-Hejduk, S. Sfawiškis, T. Berezowski, M. Borowski, S. Rosadzižkis, et al. 2016. Application of multisensoral remote sensing data in the mapping of alkaline fens Natura 2000 habitat. Ecol. Ind. 70, 196–208. https://doi.org/10.1016/j.ecolind.2016.06.001

Lang, M., G. McCarty, R. Oesterling, and L.-Y. Yeo. 2013. Topographic maps for improved mapping of forested wetlands. Wetlands 33, 141–155. https://doi.org/10.1007/s11273-013-0359-8

Lefsky, M. A., W. B. Cohen, G. G. Parker, and D. J. Harding. 2002. Lidar remote sensing for ecosystem studies. Bioscience 52, 19–30. https://doi.org/10.1641/0006-3568(2002)052[0019:LSFS]2.0.CO;2

Lindsay, J. B., I. F. Creed, and F. D. Reall. 2004. Drainage basin morphometrics for depressional landscapes. Water Resour. Res. 40, W09307. https://doi.org/10.1029/2004WR003322

Lucas, C., W. Bouten, Z. Koma, W. D. Kissling, and A. C. Seijmonsbergen. 2019. Identification of linear vegetation elements in a rural landscape using LiDAR point clouds. Remote Sens. 11, 292. https://doi.org/10.3390/rs11030292

Luo, S., C. Wang, F. Pan, X. Xi, G. Li, S. Nie, et al. 2015. Estimation of wetland vegetation height and leaf area index using airborne laser scanning data. Ecol. Ind. 48, 550–559. https://doi.org/10.1016/J.ECOLIND.2014.09.024

Millard, K., and M. Richardson. 2013. Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR-SAR fusion using a random forest classifier. Can. J. Remote. Sens. 39, 290–307. https://doi.org/10.5589/m13-038

Mitsch, W. J., B. Bernal, and M. E. Hernandez. 2015. Ecosystem services of wetlands. Int. J. Biodivers. Sci. Ecosyst. Serv. Manag. 11, 1–4. https://doi.org/10.1080/21513732.2015.1006250

Moeslund, J. E., A. Złoszczzyk, R. Ejrnæs, A. K. Brunbjerg, P. K. Bocher, J. Svenning, et al. 2019. Light detection and ranging explains diversity of plants, fungi, lichens, and bryophytes across multiple habitats and large geographic extent. Ecol. Appl. 29(5), e01907. https://doi.org/10.1002/esp.1907

Montgomery, J., B. Brisco, L. Chasmer, K. Devito, D. Cobbaert, and C. Hopkinson. 2019. SAR and lidar temporal data fusion approaches to boreal wetland ecosystem monitoring. Remote Sens. 11, 161. https://doi.org/10.3390/rs11020161

Muro, J., M. Canty, K. Conradsen, C. Hütch, A. Nielsen, H. Skriver, et al. 2016. Short-term change detection in wetlands using sentinel-1 time series. Remote Sens. 8, 795. https://doi.org/10.3390/rs8100795

Naimi, B. 2015. Uncertainty Analysis for Species Distribution Models. R Software Package. Available at: https://scholar.google.com/scholar_lookup?title=UncertaintyAnalysisforSpeciesDistributionModels. R Software Package&publication_year=2015&author=B. Naimi (Accessed: 18 October 2019).

Navarro, L. M., N. Fernández, C. Guerra, R. Guralnick, W. D. Kissling, M. C. Londoño, et al. 2017. Monitoring biodiversity change through effective global coordination. Curr. Opin. Environ. Sustain. 29, 158–169. https://doi.org/10.1016/j.cosust.2018.02.005

Onoeghuo, A. O., and G. A. Blackburn. 2011. Optimising the use of hyperspectral and LiDAR data for mapping reedbed habitats. Remote Sens. Environ. 115, 2025–2034. https://doi.org/10.1016/J.RSE.2011.04.004

Onoeghuo, A. O., G. A. Blackburn, and Z. A. Latif. 2010. Characterising Reedbed habitat quality using Leaf-off LiDAR Data. Pp. 1–5 in 2010 6th international colloquium on signal processing & its applications. Mallaca City: IEEE. https://doi.org/10.1109/CPSA.2010.5545322

Ostendorp, W. 1989. “Die-back” of reeds in Europe — a critical review of literature. Aquat. Bot. 35, 5–26. https://doi.org/10.1016/0304-7708(89)90063-6

Pettorelli, N., M. Wegmann, A. Skidmore, S. Müller, T. P. Dawson, M. Fernandez, et al. 2016. Framing the concept of satellite remote sensing essential biodiversity variables: challenges and future directions. Remote Sensing in Ecology and Conservation 2, 122–131. https://doi.org/10.1002/2sre.215

Rapinel, S., B. Clément, S. Dufour, and L. Hubert-Moy. 2018. Fine-scale monitoring of long-term wetland loss using LiDAR data and historical aerial photographs: the example of the couesnon floodplain, France. Wetlands 38, 423–435. https://doi.org/10.1007/s11357-017-0985-2

Shadaydeh, M., A. Złoszczzyk, A. Manno-Kovacs, and T. Szirányi. 2017. Wetland mapping by fusion of airborne laser scanning and multi-temporal multispectral satellite imagery. Int. J. Remote Sens. 38, 7422–7440. https://doi.org/10.1080/01431161.2017.1375614
Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Results of the collinearity and sensitivity analysis.

Table S1. Variance Inflation Factor (VIF) of 14 lidar metrics used in the Recursive Feature Elimination analysis.

Figure S1. Results of the sensitivity analysis based on a Random Forest algorithm using varying amounts of annotation points (100, 500 and 1000) at different spatial resolutions (10 m, 5 m, 2.5 m, and 1 m).