Int J Appl Earth Obs Geoinformation 84 (2020) 101970

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

Improving LiDAR-based tree species mapping in Central European mixed forests using multi-temporal digital aerial colour-infrared photographs

Yifang Shi[a,*], Tiejun Wang[a], Andrew K. Skidmore[a,d], Marco Heurich[b,c]

[a] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99, P.O. Box 217, 7500 AE Enschede, the Netherlands
[b] Department of Conservation and Research, Bavarian Forest National Park, Freyunger Str. 2, 94481 Grafenau, Germany
[c] Chair of Wildlife Ecology and Management, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany
[d] Department of Environmental Science, Macquarie University, NSW, 2109, Australia

ARTICLE INFO

Keywords:
Tree species
Multi-temporal
Digital CIR orthophotos
Airborne LiDAR
Texture features

ABSTRACT

Digital colour-infrared (CIR) aerial photographs, which have been collected routinely in many parts of the world, are an invaluable data source for the monitoring and assessment of forest resources. Yet, the potential of these data for automated individual tree species mapping remains largely unexplored. One way to maximize the usefulness of digital CIR aerial photographs for individual tree species mapping is to integrate them with modern and complementary remote sensing technologies such as the light detection and ranging (LiDAR) system and 3D segmentation algorithms. In this study, we examined whether multi-temporal digital CIR orthophotos could be used to further increase the accuracy of airborne LiDAR-based individual tree species mapping for a temperate mixed forest in eastern Germany. Our results showed that the texture features captured by multi-temporal digital CIR orthophotos under different view-illumination conditions were species-specific. As a consequence, combining these texture features with LiDAR metrics significantly improved tree species mapping accuracy (overall accuracy: 77.4%, kappa: 0.68) compared to using LiDAR data alone (overall accuracy: 69.3%, kappa: 0.58). Among various texture features, the average gray level in the near-infrared band was found to contribute most to the classification. Our results suggest that the synergic use of multi-temporal digital aerial photographs and airborne LiDAR data has the potential to accurately classify individual tree species in Central European mixed forests.

1. Introduction

Tree species information are important inputs for biodiversity modelling and biomass estimation, and are indispensable for environmental, monitoring and protection activities (Waser et al., 2011). Remote sensing-assisted classification of individual tree species can aid forest management and ecosystem modelling and lead to more comprehensive and accurate forest inventories (Yin and Wang, 2016). Over the last four decades, advances in remote sensing have enabled the classification of tree species using various sensor types (e.g. LiDAR, very high resolution digital aerial photographs, multispectral and hyperspectral sensors) (Fassnacht et al., 2016).

People have acquired aerial photographs ever since the means have existed to lift cameras above the Earth's surface, beginning in the mid-19th century (Aber et al., 2010). Historically, visual interpretation of aerial photographs was the most popular form of remote sensing for mapping trees in forests (Barrett et al., 2016; Spurr, 1960). Captured at various spatial scales, aerial photographs are used for a wide range of purposes in resource management, from detailed surveys of individual trees to general land cover mapping over broad extents (Morgan et al., 2017). The ability of aerial photographs to provide tree species information is well established and has been used for decades in forest inventory (Loetsch and Haller, 1964). Traditional film-based aerial photographs are referred to as analogue images which have been routinely scanned and converted into digital images after the first decade of the 21 st century (Aber et al., 2010). With the technical advances in electronic devices and desktop computing, film-based aerial photographs are rapidly becoming obsolete, encouraging the use of digital aerial photographs with increasingly sophisticated analysis methods. Specifically, small-format digital aerial photographs (equivalent to 35- and 70-mm film cameras) have established a niche that bridges the gap in scale and resolution between ground observations and imagery acquired from airborne and satellite sensors (Aber et al., 2010).

In recent decades, very high resolution digital colour-infrared (CIR)

* Corresponding author.
E-mail addresses: y.shi-1@utwente.nl (Y. Shi), t.wang@utwente.nl (T. Wang), a.k.skidmore@utwente.nl (A.K. Skidmore), marco.heurich@npv-bw.bayern.de (M. Heurich).

https://doi.org/10.1016/j.jag.2019.101970
Received 9 January 2019; Received in revised form 27 August 2019; Accepted 6 September 2019

aerial photographs have been used to obtain information on individual tree species (Deng et al., 2016; Singh et al., 2015; Zhang and Hu, 2012). Colour-infrared (CIR) aerial photography has long been recognized for its capabilities in land use, land cover and vegetation studies, including large-area and regional analysis, species differentiation and stress detection (Colwell, 1960). Furthermore, digital CIR orthophotos obtained by georeferencing digital CIR aerial photographs using ground control points as well as GPS/INS system provide the true orthographic positions, which can be used for making direct measurements of distances, angles, directions, positions and areas without corrections for distortions. Manual (visual) interpretation of aerial photography has long been a standard procedure in developing forest inventories; however, human interpretations can be subjective, inconsistent, and labour-intensive (Koch et al., 2002). Although very high resolution digital CIR orthophotos allow specialists to recognize patterns of species, visual interpretation may not always fully reveal the spectral and texture information of individual trees. The variability among the same tree species and the similarity between different tree species present additional challenges (Dechesne et al., 2017).

Previous research has shown that there are species-specific differences in how the observed brightness changes when the viewing direction in digital aerial photographs is altered (Korpela et al., 2014). Studies also pointed out that the spectral separability of tree species is partly dependent on the view-illumination geometry and the reflectance anisotropy, and the high within-species reflectance variation in digital aerial photographs may affect the spectral signatures of trees (Dechesne et al., 2017; Korpela et al., 2011; Leckie et al., 2005). The collection of very high resolution digital aerial photographs has become a routine forest inventory procedure in many countries. This provides an opportunity to capture brightness variation patterns within tree crowns by using multi-temporal information. Korpela et al. (2011) concluded that “the brightness variation inside crowns showed band- and species- specific differences” and “the reflectance differences observed in the various bands and geometries could potentially be taken into account for the improvement of species classification”. However, for individual tree species classification, the capability of texture information derived from multi-temporal digital aerial photographs under different illumination conditions remains largely unexplored.

Recent advances in remote sensing technology and machine learning methods provide avenues to further explore the potential of digital CIR orthophotos for individual tree species mapping. Processing capabilities allow us to efficiently generate spectral and texture features from digital CIR orthophotos, while machine learning algorithms offer a means to combine these features with complementary data sources. For instance, LiDAR data provide a range of metrics related to both the geometric and radiometric characteristics of trees. While geometric features describe the architecture of tree crowns and branching patterns, radiometric features are more related to foliage material, leaf orientation and density (Kim et al., 2009). In addition, directional reflectance signatures and crown texture information derived from digital aerial photographs can be used for forest monitoring and species classification (Korpela et al. 2014; Kuzmin et al., 2016; Singh et al., 2015; Tian et al., 2017). Crown texture information often reflects the spatial variation of crown-internal shadows, foliage properties (size, density, reflectivity) and branching patterns (Sayn-Wittgenstein, 1978). On small format CIR orthophotos, leaves, branches and inner tree crown shadows all contribute to the texture information (Haara and Haarala, 2002). New perspectives for regional and national mapping approaches are emerging through digital aerial photographs, which are likely – in contrast to airborne LiDAR data – to be updated more regularly by regional or national mapping agencies (Waser et al., 2017). Although the combination of airborne LiDAR data and texture features derived from individual digital aerial photographs has been successfully employed for tree species monitoring (Kulikova et al., 2007) and land cover mapping (Silveyra Gonzalez et al., 2018), the utility of texture information derived from multi-temporal digital CIR orthophotos in aiding LiDAR-based tree species mapping has not been tested.

The current study aims to evaluate the role of multi-temporal digital CIR orthophotos in improving LiDAR-based individual tree species classification in Central European mixed forests. Specifically, we set out to (1) examine the correlation of texture features derived from multi-temporal digital CIR orthophotos; (2) integrate texture information derived from multi-temporal digital CIR orthophotos with airborne LiDAR derived metrics for tree species classification; and (3) assess the utility of combining multi-temporal digital CIR orthophotos and LiDAR data for tree species discrimination and identify the most valuable features from both datasets.

2. Materials and methods

2.1. Study sites and field data

The Bavarian Forest National Park (BFNP) was the first National Park established in Germany, in 1970, and is situated in south-eastern Germany along the border with the Czech Republic. As a mixed temperate forest reserve in Central Europe, it covers 24,218 ha with elevations between 600 and 1453 m above sea level (Cailleret et al., 2014). Within the park, three major forest types exist: above 1100 m there are sub-alpine spruce forests of Norway spruce (*Picea abies*) and some Mountain ash (*Sorbus aucuparia*) (high altitude); on the slopes, between 600 and 1100 m, are mountains of mixed forests of Norway spruce, Silver fir (*Abies alba*), European beech (*Fagus sylvatica*), and sycamore maple (*Acer pseudoplatanus*); in wet depressions that often experience cold air pockets in the valley bottoms, spruce forests of Norway spruce, Mountain ash, and white birches (*Betula pendula*) occur (valley bottoms) (Heurich, 2008).

Two study sites (approximately 25 ha for each site) were selected based on the species richness in the BFNP (Fig. 1). The field data was collected in July 2016 and July 2017, including 256 locations of individual tree species at site A and 193 locations of individual tree species at site B. A Leica Viva GS10 Plus differential GPS (Leica Geosystems AG, Heerbrugg, Switzerland) was used to record the spatial location of trees in the field (Shi et al., 2018b). After post-processing, the absolute error of the differentially corrected coordinates was within 0.25 m.

2.2. Multi-temporal digital CIR orthophotos

Three digital CIR orthophotos were used in this study, from 6 June 2015, 23 June 2016 and 11 June 2017, covering the whole BFNP. All three digital CIR images were recorded using Intergraph’s Digital Mapping Cameras (DMC) from an altitude of approximately 2900 m. Four parallel multispectral colour DMCs were utilised to form composite images, and each of them had a resolution of 3456 × 1920 pixels with a virtual pixel size of 12 µm. The nominal focal length of the cameras was 12 mm. The digital CIR images consist of three bands: NIR (near-infrared) (675–850 nm), Red (590–675 nm) and Green (500–650 nm). Finally, all three digital CIR images were radiometrically corrected and orthorectified by using optimal camera calibration observations, transformation parameters and ground control points. The procedures were conducted in the program system OrthoBox (Orthovista, Orthomaster) of the company Trimble/INPHO. The detailed information about spatial resolution, average flying altitude, flight time and sun position of three digital CIR orthophotos are shown in Table 1.

2.3. Airborne LiDAR data

An airborne LiDAR flight campaign was conducted out on 18 August 2016, using a Riegl LMS-Q680i scanner integrated in a full-waveform laser scanning system. The system operated at a wavelength of 1550 nm and was flown at approximately 300 m above terrain elevation. The

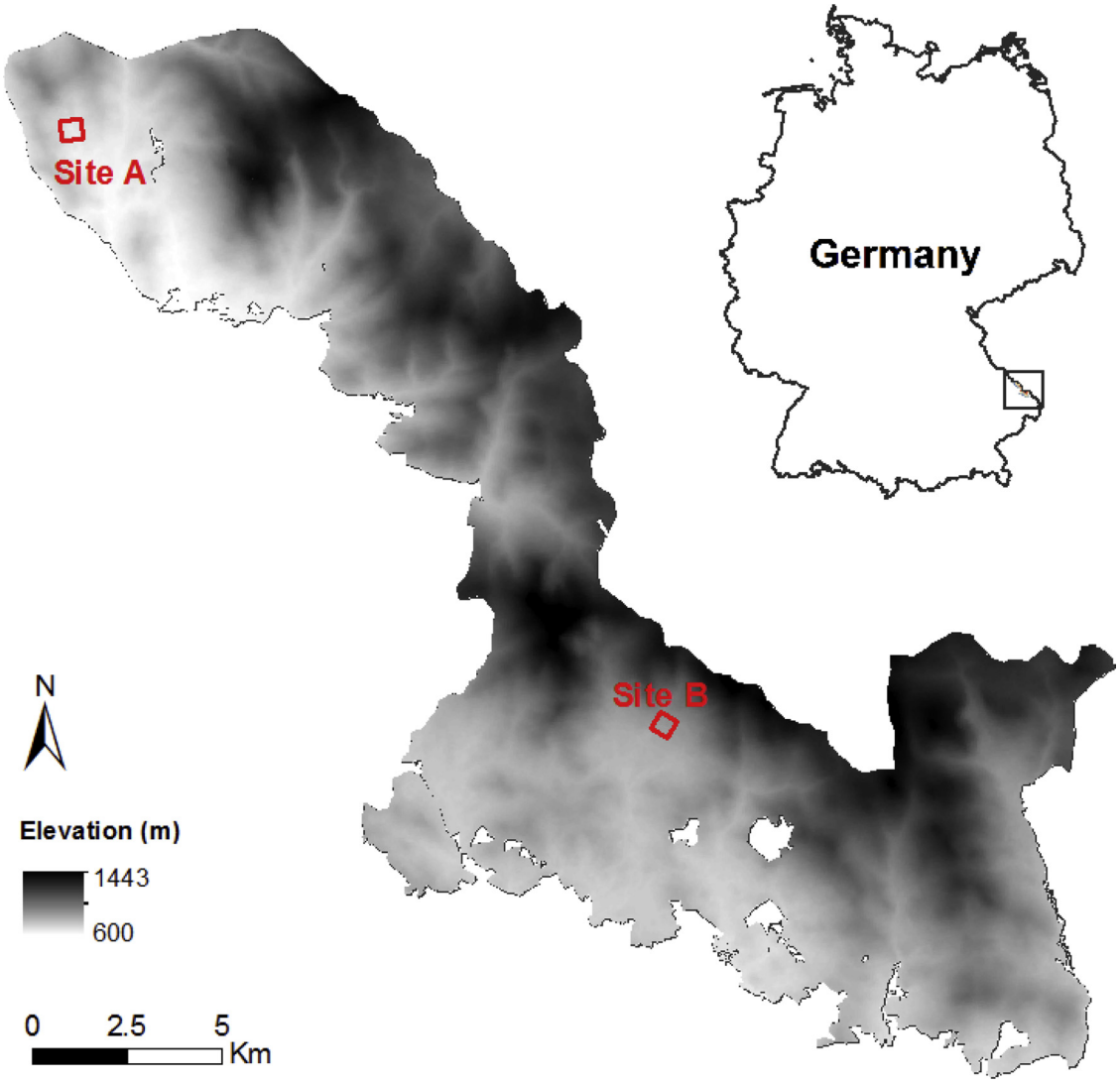


Fig. 1. Location of the two study sites in Bavaria Forest National Park, Germany.

Table 1
Detailed information of multi-temporal digital CIR orthophotos used in this study.

Digital CIR orthophotos	6 June 2015	23 June 2016	11 June 2017
Spatial resolution	20 cm	20 cm	10 cm
Average altitude	2909 m	2918 m	2879 m
Flight time	09:50–12:05	09:50–12:15	10:30–13:25
Sun position	44°–57°	43°–62°–53°	49°–64°–35°

pulse repetition frequency was 400 kHz with a maximum scanning angle of ± 15°. The average point density was about 70 points/m^2. Four transects in the park were covered by 21 flight lines with 30%–50% strip overlaps.

The airborne LiDAR data was delivered by Milan Flug GmbH, consisting of point data coordinates together with the echo width, intensity, return number, number of returns and the GPS timestamp of the returns. A Canopy Height Model (CHM) with 0.20 m resolution was generated using LAStools software package (LAStools, version 160921, rapidlasso GmbH, http://lastools.org) (Shi et al., 2018b).

2.4. Individual tree delineation

The automatic delineation of individual tree crowns was performed using an enhanced approach, which detects individual trees in multi-layered forests with an integrated 3D segmentation proposed by Yao et al. (2013). This method is based on 3D segmentation combining mean shifts with normalized cuts, which can exploit the advantages of full waveform data and detect supressed trees in the both middle and lower layers. Yao et al. (2013) achieved a detection rate of up to 70% for trees in the upper layer and found that tree detection rates using this method improved by 10–25% for all three forest layers compared to that obtained by the method presented in Reitberger et al. (2009) and Yao et al (2012).

Field measured trees that were discernible both in the LiDAR segments and in the multi-temporal digital CIR orthophotos formed the potential reference trees for this study (Fig. 2). To link the field measured trees to both LiDAR and digital CIR orthophotos, we firstly overlaid the post-processed DGPS tree locations with tree crown segments and georeferenced digital CIR orthophotos; and secondly we identified the field measured trees which located within the crown segments and visible in digital CIR orthophotos. Finally, a visual verification was carried out with the assistance of the photos of the measured trees and the species of neighbouring trees recorded during the fieldwork. To reduce linking errors, trees undetected by the segmentation or assigned to more than one segment were removed from further analysis. In the end, 58 beech trees, 40 birch trees, 60 fir trees, 55 maple trees and 57 spruce trees (270 sample trees in total) from two

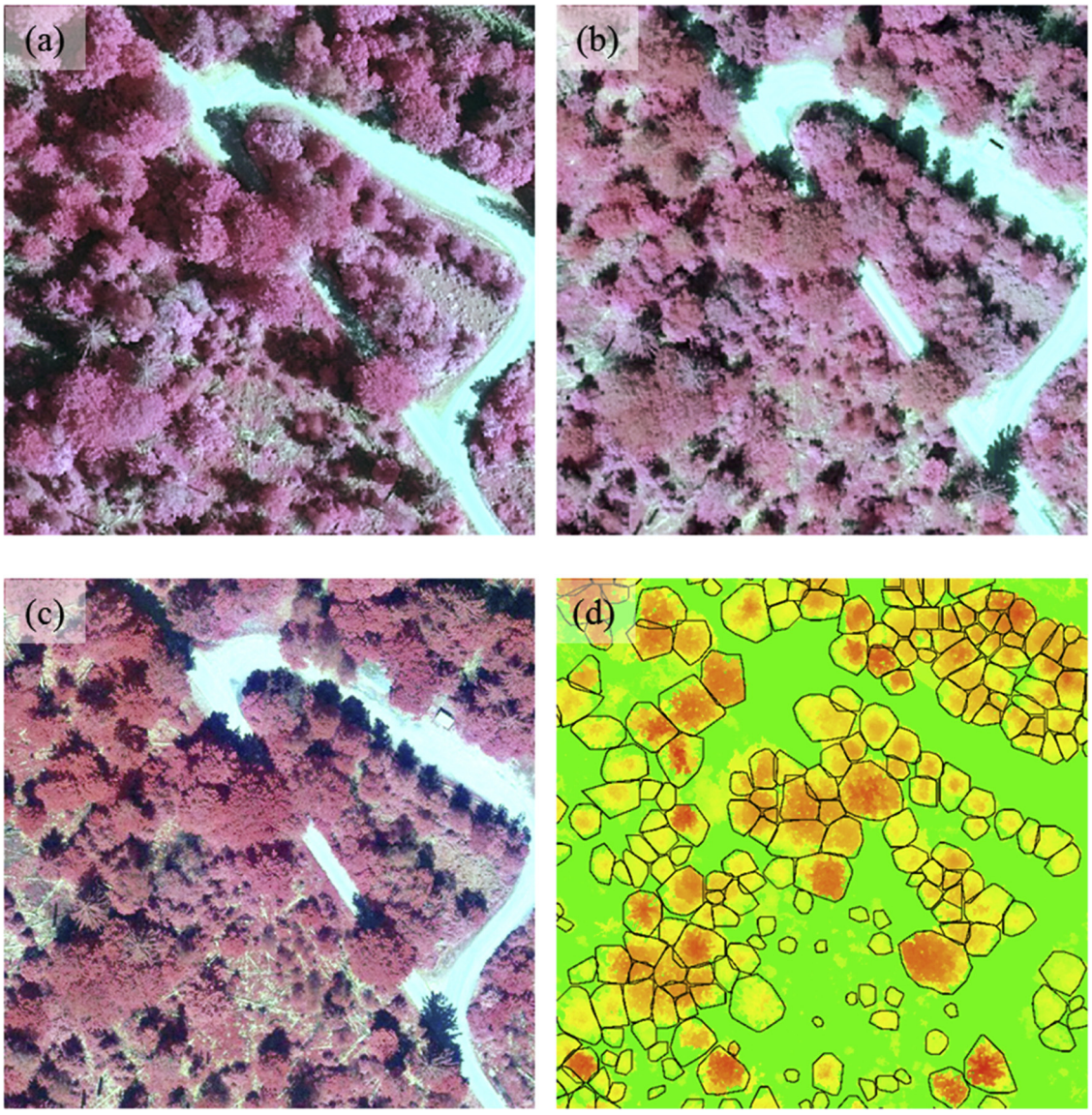


Fig. 2. The digital CIR orthophotos of a sample area from 2015 (a), 2016 (b) and 2017 (c). The 3D segmentation of tree crown is displayed with the background of the LiDAR-derived Canopy Height Model (d).

study sites were selected for further analysis (Shi et al., 2018a).

2.5. Feature derivation

2.5.1. Geometric and radiometric features from LiDAR data

Once the sample trees were correctly linked to the 3D segments, the normalized point clouds within the segments were extracted to derive LiDAR metrics for each sample tree. The derived LiDAR metrics can be grouped into four categories: (1) height distribution (including maximum, mean, standard deviation, coefficient of variation, skewness, kurtosis, variance, 25th percentile and 95th percentile of the normalized height); (2) crown shape (including the ratio of crown base height to tree height, the ratio of crown volume to crown area, and canopy relief ratio); (3) point distribution (including the percentage of first returns to all returns, the percentage of last returns to all returns, first returns above mean height, and all returns above 2 m); and (4) intensity and echo width metrics (including the abovementioned statistical variables applied for height distribution, the mean value of the first-or-single returns, and the mean value of single returns). As a result, 40 LiDAR metrics were assigned to each sample tree. Detailed description of generated LiDAR metrics can be found in our previous study, see Shi et al (2018b).

2.5.2. Texture features from multi-temporal digital CIR orthophotos

The Gray Level Co-Occurrence Matrix (GLCM) is one of the best known texture analysis methods (Haralick et al., 1973). The texture features obtained by the GLCM method have been widely employed in the remote sensing community for forest modelling (Kayitakire et al., 2006; Ozdemir et al., 2008; Pasher and King, 2010) and tree species classification (Kuzmin et al., 2016; Singh et al., 2015). Based on the crown polygons delineated by 3D segmentation, we used the GLCM to derive texture features of each sample tree. Texture features were calculated for each pixel within a tree crown, based on a window size of 3 by 3 pixels (i.e. each pixel and its 8 neighbours). We used the following eight texture parameters: mean, homogeneity, contrast, dissimilarity, entropy, variance, angular second moment and correlation. After calculating the parameters for all the pixels within an individual tree crown, we used the mean value of each parameter to represent the texture features of each sample tree. All texture features were normalized for clarity. In total, 72 texture features, generated from the NIR, red and green bands from three years of digital CIR orthophotos, were assigned to each sample tree. The description of the texture

Table 2
Description of generated texture features. In formulas, i and j are row and column numbers, respectively. N is the total number of pixels. u_i, u_j, σ_i^2, and σ_j^2 are the means and standard deviations of P_i and P_j. $P(i, j)$ is the normalized co-occurrence matrix.

GLCM texture features	Formula	Description
MEAN	$\sum_{i,j=0}^{N-1} iP_{i,j}$	The average gray level in the local window
Homogeneity (HOM)	$\sum_{i,j=0}^{N-1} i\frac{P_{ij}}{1+(i-j)^2}$	A measure of lack of variability or the amount of local of similarity in an image
Contrast (CON)	$\sum_{i,j=0}^{N-1} iP_{ij}(i-j)^2$	A measure of the amount of local variation in pixels values among neighbouring pixels
Dissimilarity (DIS)	$\sum_{i,j=0}^{N-1} iP_{i,j}\ \lvert i-j\rvert$	A measure of lack of similarity in an image
Entropy (ENT)	$\sum_{i,j=0}^{N-1} iP_{i,j}(-\ln P_{i,j})$	A measure of the degree of disorder in an image
Variance (VAR)	$\sum_{i,j=0}^{N-1} P_{ij}(i-u_i)^2$	A measure of the degree of variation in an image
Angular second moment (ASM)	$\sum_{i,j=0}^{N-1} iP_{i,j}^2$	A measure of textural uniformity or pixel pair repetitions
Correlation (COR)	$\sum_{i,j=0}^{N-1} P_{i,j}(\frac{(i-u_i)(j-u_j)}{\sigma_i\sigma_j})$	A measure of gray level linear dependencies in an image

features is shown in Table 2. The derivation of texture features was performed in the R language environment (http://www.r-project.org/).

2.6. Feature selection and tree species classification

Many of the second-order texture measures proposed by Haralick et al. (1973) have been found to be highly correlated (Clausi, 2002). A logical aim in classification using a large number of textural features should therefore be the selection of the most relevant features (Rodriguez-Galiano et al., 2012). As such, we examined the correlation coefficient, and its significance, between each pair of texture features. The test of the significance of the correlation coefficient was performed to assess whether the linear relationship between the input features was strong enough to consider elimination. We ranked the texture features by the frequency of low correlation ($|r| < 0.70$) and high insignificance ($p > 0.05$) with other features, and the top 10 texture features from each year were selected for further analysis, along with selected LiDAR metrics. The selection of LiDAR metrics was based on our previous study (Shi et al., 2018b).

The Random Forest algorithm was used for feature selection and classification. Random Forest was proposed by Breiman (2001) as a non-parametric machine learning method and uses predictions derived from an ensemble of decision trees to determine overall class assignment. We applied the Random Forest classifier for this study because: 1) it handles large datasets efficiently, 2) it quantifies the importance of each input variable, 3) it is robust to outliers and noise, and 4) it is computationally faster than other ensemble methods (e.g. Boosting) (Cutler et al., 2007; Rodriguez-Galiano et al., 2012). Although Random Forest is able to handle high dimensional data, studies have shown the performance of classification can be significantly improved by selecting the most important features (Yu et al., 2017). Thus, we ranked the input features based on the Mean Decrease Accuracy (MDA) index, the built-in measurement of feature importance in Random Forest, to select a subset of variables (texture features and LiDAR metrics combined) that contributed to the optimal classification result. To decide the number of selected variables, we calculated the accumulated contribution rate, which varies by the number of input variables, and retained the variables that resulted in the highest accuracy during classification (Shi et al., 2018a). In total, 22 variables were selected for tree species classification, including 12 LiDAR metrics and 10 texture features derived from the multi-temporal digital CIR orthophotos. Two tuning parameters are required in the Random Forest classifier: the number of trees "Ntree" and the number of predictors sampled for splitting at each node "Mtry". To obtain the optimal combination of the parameters for best classification performance, we tested Random Forest models by varying "Ntree" from 1 to 500 and "Mtry" from 1 to 30. The resulting models were evaluated using the kappa index (Cohen, 1960). The model with the highest kappa coefficient was chosen as the best model.

Accuracy of the models were assessed using overall accuracy, producer's and user's accuracy, and the kappa coefficient. We also employed McNemar's test (McNemar, 1947) to determine if statistical significant differences occur among classifications using different datasets. The feature selection and classification procedures were carried out using the R package "*randomForest*" (Liaw and Wiener, 2002).

3. Results

3.1. Correlation and feature selection

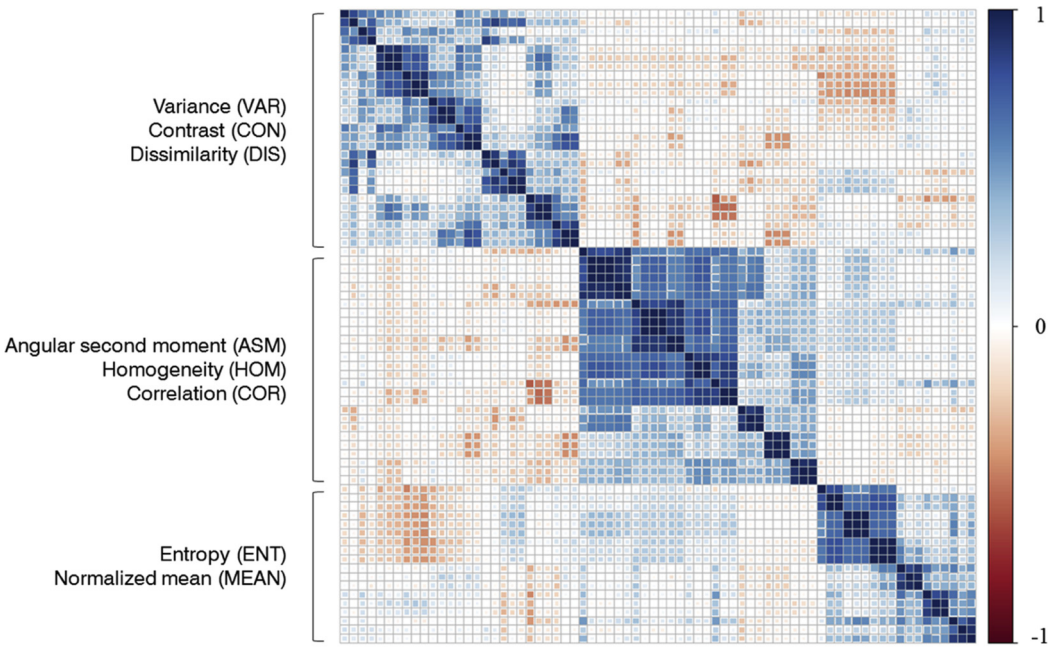
High correlations ($|r| > 0.7$) were observed between the texture features of the NIR, red and green bands generated from the digital CIR orthophotos of the same year (Fig. 3a). The average gray level (MEAN) showed the lowest correlations with other texture features between the three years. When comparing the texture features between different years, the angular second moment (ASM), homogeneity (HOM) and entropy (ENT) features all showed high correlations with each other. When comparing the texture features from the same year, high correlation occurred between the angular second moment (ASM), homogeneity (HOM), contrast (CON) and variance (VAR) features.

Fig. 3b shows the correlation and significance between texture features from 2015. The features were ranked by the frequency of low correlation ($|r| < 0.70$) and high insignificance ($p > 0.05$) with other features. As a result, the top 10 features were selected from 2015: DIS_NIR, VAR_NIR, VAR_G, CON_NIR, ASM_NIR, ASM_R, ASM_G, MEAN_NIR, MEAN_R, and MEAN_G (See Table 2 for the definitions of the texture features. R, G, and NIR represent the texture features that were derived from the red, green and near-infrared bands, respectively). Similarly, the 10 top-ranked texture features were selected from the years 2016 and 2017.

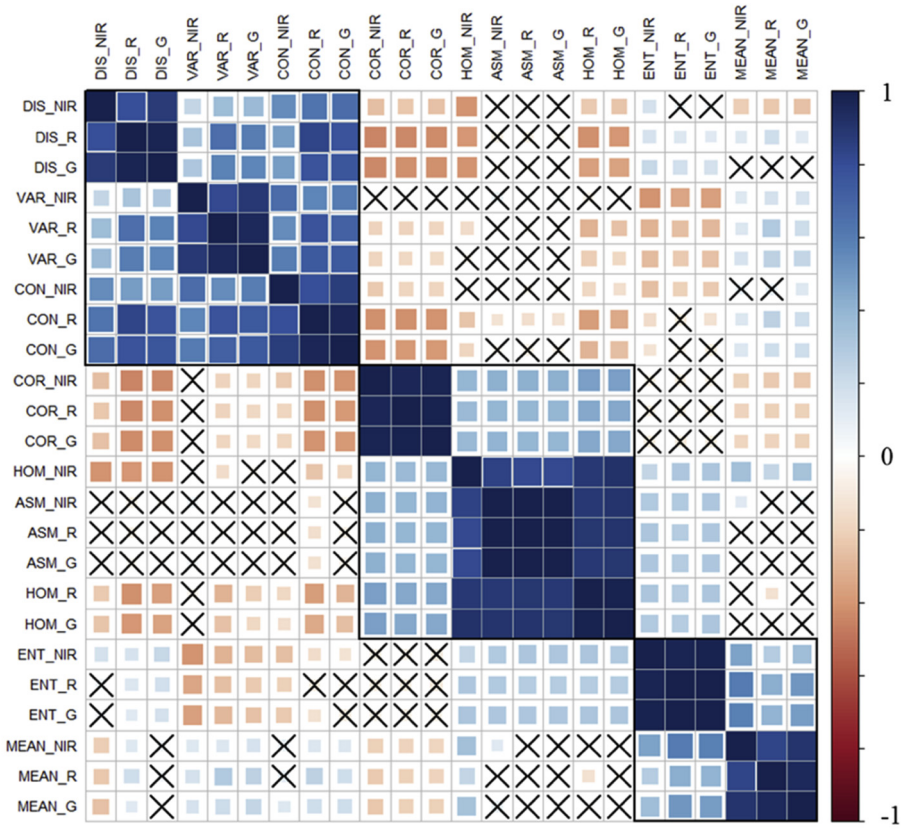
Table 3 lists the selected variables derived from airborne LiDAR and multi-temporal digital CIR orthophotos for the classification. 9 out of 12 selected LiDAR metrics were radiometric metrics (intensity and echo width related metrics). Among 10 selected texture features, 7 were generated from NIR bands. The average gray level (MEAN) and the dissimilarity (DIS) were the most frequently selected texture features.

3.2. Classification performance from different combinations of datasets

Table 4 shows the confusion matrices and corresponding producer's and user's accuracies, the overall accuracy and kappa coefficient for classification results using different feature combinations. The best classification accuracy was achieved with a model containing both LiDAR metrics and texture variables from three years of digital CIR orthophotos: overall accuracy of 77.4% with a 0.68 kappa coefficient. The classifications using only LiDAR and only CIR texture features yielded similar results, with overall accuracies of 69.3% and 66.7%,



(a)



(b)

Fig. 3. The cross-correlation matrix of the texture features derived from multi-temporal digital CIR orthophotos (a) and the combined correlogram with the significance test of 2015 (b). Blue colours indicate positive correlations and red colours indicate negative correlations. The insignificant values ($p > 0.05$) were marked with black crosses. The features were sorted by hierarchical clustering order (black rectangles) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 3
Selected variables derived from LiDAR and digital CIR orthophotos for classification.

Input variables	Index or description
LiDAR metrics (12)	
Imean_first	The mean intensity of first-or-only returns
Ip75	The 75th percentile of intensity
Height	The height of tree
Ewmean_first	The mean echo width of first-or-single returns
EWmean	The mean value of echo width
Icv	The coefficient variation of intensity
Ewmean_single	The mean echo width of single returns
Imean_single	The mean intensity of single returns
Imean	The mean intensity
Ivar	The variation of intensity
Hvar	The variation of height
Hsd	The standard deviation of height
Texture features (10)	
2015_MEAN_NIR	The average gray level of NIR band from 2015
2016_MEAN_NIR	The average gray level of NIR band from 2016
2017_DIS_NIR	The dissimilarity of NIR band from 2017
2015_MEAN_G	The average gray level of green band from 2015
2015_DIS_NIR	The dissimilarity of NIR band from 2015
2017_MEAN_NIR	The average gray level of NIR band from 2017
2016_DIS_R	The dissimilarity of red band from 2016
2015_ASM_NIR	The angular second moment of NIR band from 2015
2016_DIS_G	The dissimilarity of green band from 2016
2017_VAR_NIR	The variance of NIR band from 2017

respectively. The birch trees had the highest user's accuracy in each of the models. The lowest user's and producer's accuracies occurred for the fir and maple trees using only multi-temporal texture features. The highest misclassifications were found between the pairs of beech and maple, and fir and spruce.

Table 5 shows the significance levels between classification results generated from different combinations, based on the p values in McNemar's test. Combining LiDAR and texture features from three years of digital CIR orthophotos significantly improved the accuracy compared to using LiDAR and digital CIR orthophotos alone. However, the combination of one or two years of digital CIR orthophotos with LiDAR data did not yield statistically significant improvements compared to using LiDAR data alone. Moreover, there was no statistically significant difference between the classification performance of using only LiDAR or digital CIR orthophotos.

3.3. The contribution of selected features for tree species discrimination

Fig. 4 shows the relative importance of the selected variables for the classification based on the MDA index. In the combined model (LiDAR metrics and texture features), the three most important variables were LiDAR metrics (Imean_first, Ip75 and Height), followed by the average gray level of the NIR bands from 2015 and 2016. Specifically, the top-ranked LiDAR metric was the mean intensity of first-or-single returns (Imean_first) and the top-ranked texture feature was the average gray level of the NIR band from 2015 (2015_MEAN_NIR). The average gray level and the dissimilarity of NIR bands were found to be more important than other texture features (e.g. contract, variance and the angular second moment).

Table 6 shows five top-ranked features which contributed most to the discrimination of one tree species from the others in the classification. The texture features derived from the NIR bands of the multi-temporal digital CIR orthophotos contributed most to the discrimination of beech and spruce from the other species. On the contrary, birch and maple were more distinguishable by using LiDAR metrics – echo width and height related metrics contributed most for the classification of birch, while intensity related metrics were more important for the identification of maple.

Fig. 5 shows the inter-species comparison of different variables derived from multi-temporal digital CIR orthophotos and LiDAR data. The average gray level and the dissimilarity of NIR bands (MEAN_NIR and DIS_NIR) from different years have similar patterns, however, the differences between species vary between years. For example, the difference of the average gray level of the NIR band (MEAN_NIR) between beech and birch was more pronounced in 2015 (first row in Fig. 5), while the difference in the dissimilarity of the NIR band (DIS_NIR) became more distinguishable between maple and spruce in 2016 (second row in Fig. 5). Out of the most important LiDAR metrics (e.g. Imean_first, Ewmean and Height), the mean intensity of first-or-single returns (Imean_first) distinguished beech and maple from other species. In contrast, the mean value of echo width (Ewmean) was the most important feature for the discrimination of birch (Table 6).

4. Discussion

In this study, we explored the potential of using texture features derived from multi-temporal digital CIR orthophotos to improve LiDAR-based individual tree species classification in a temperate mixed forest in Germany. Our study found that adding texture features from multi-temporal digital CIR orthophotos to LiDAR-based tree species classification significantly increased classification accuracy. In particular, the average gray level and the dissimilarity of the NIR band within a tree crown, derived from multi-temporal digital CIR orthophotos, were found to provide valuable information for discriminating tree species. However, we did not find significant improvements between the results of using only one or two years of digital CIR orthophotos with LiDAR data and that of using LiDAR data alone. This finding aligns with Korpela et al. (2014), who used a single year of aerial photography and concluded that interspecies differences in directional reflectance anisotropy do not constitute a significant improvement to tree species classification. However, our results showed that the texture features derived from three years of digital CIR orthophotos resulted in a significant improvement to the classification performance compared with only using LiDAR data. The texture features derived from the same year are more likely to be highly correlated than those from different years, which may partly explain why adding one year or two years of digital CIR orthophotos did not yield significant improvements. The spectral

Table 4
Confusion matrices for five tree species using different selections of variables. PA is producer's accuracy, UA is user's accuracy, OA is overall accuracy.

	LiDAR + CIRs (22)						LiDAR (12)						CIRs (10)					
Species	Beech	Birch	Fir	Maple	Spruce	PA(%)	Beech	Birch	Fir	Maple	Spruce	PA(%)	Beech	Birch	Fir	Maple	Spruce	PA(%)
Beech	48	0	0	10	0	82.8	40	2	2	11	2	69.0	46	1	2	7	2	79.3
Birch	3	33	4	0	0	82.5	5	33	1	1	0	82.5	3	30	5	1	1	75.0
Fir	3	2	42	2	11	70.0	4	2	37	2	15	61.7	4	2	39	4	11	65.0
Maple	12	2	0	40	1	72.7	15	2	1	36	1	65.5	13	6	5	27	4	49.1
Spruce	0	1	10	0	46	80.7	0	2	14	0	41	71.9	2	0	14	3	38	66.7
UA(%)	72.7	86.8	75.0	76.9	79.3		62.5	80.5	67.3	72.0	69.5		67.6	76.9	60.0	64.3	67.9	
OA(%)	77.4						69.3						66.7					
Kappa	0.68						0.58						0.56					

Table 5
McNemar's test for pairwise comparison between classification results using different combinations.

	LiDAR + CIR (×3)	LiDAR + CIR (×2)	LiDAR + CIR (×1)	LiDAR	CIR (×3)
LiDAR + CIR (×3)	–	**	***	***	***
LiDAR + CIR (×2)		–	NS	NS	*
LiDAR + CIR (×1)			–	NS	NS
LiDAR				–	NS
CIR (×3)					–

CIR (×3) means all three years of digital CIR orthophotos; CIR (×2) means any two years of digital CIR orthophotos; CIR (×1) means any one year of digital CIR orthophoto. ***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; NS: $p > 0.05$.

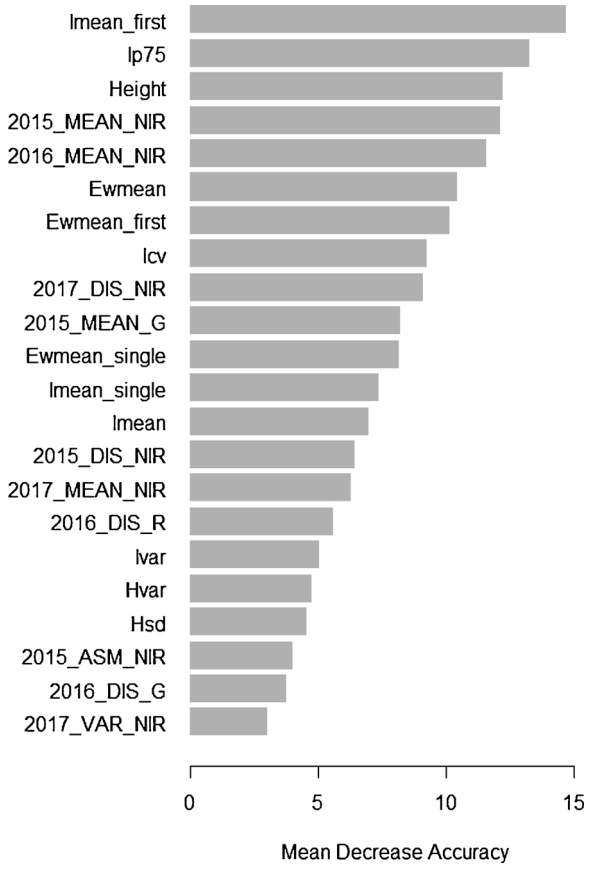


Fig. 4. The relative importance of the selected variables for tree species classification.

variance within species hampered the ability of using only digital CIR orthophotos for species identification, which highlights the benefits of fusing multiple sources of data.

When digital aerial photographs have been used as complementary data sources for LiDAR-based individual tree species mapping, they have been mainly used to visually identify tree tops, tree crowns and tree types (conifers or deciduous trees) (e.g. Korpela, 2006; Koukoulas and Blackburn, 2005; Persson et al., 2004). In our study, the tree crowns were delineated from high density airborne LiDAR data using a 3D segmentation algorithm. In comparison to manual delineation of tree crown from digital aerial photographs, which usually eliminates the shaded parts (e.g. Heinzel et al., 2008; Persson et al., 2004), LiDAR-based tree crown delineation provides an opportunity to automatically generate more complete tree crowns, without the influence of shadow effects. This allows species-specific information to be captured by the texture patterns within a tree crown, especially the texture features derived from multi-temporal digital aerial photographs with different illumination geometry. Although multi-temporal digital aerial photographs are not always able to capture seasonal changes, the variation of brightness patterns under different view-illumination conditions still provides valuable information for individual tree species classification. Our study demonstrates the potential of improving tree species mapping with the assistance of archived digital aerial photographs, which are widely available in many countries. It should be noted that our study was conducted in a temperate mixed forest in Central Europe with relatively low variability regarding forest type and tree species. Therefore, the potential capability of multi-temporal digital aerial photographs for LiDAR-based individual tree species classification in other forest ecosystems or species-rich habitats needs further investigation.

Our results highlighted the importance of texture features derived from the NIR band of digital aerial photographs in discriminating between tree species. The average gray level of the NIR band was found less correlated with other texture features and contributed more in the classification. This is most likely due to its sensitivity to illumination variation and foliage. Similar results were reported by Coops et al. (2004), who found a greater discrimination between canopy and understory vegetation in the NIR band than the green band. In our study, beech and spruce showed distinct differences in the NIR band, and in fact the texture features of the NIR band were considered more important than LiDAR metrics for discriminating beech and spruce from other species (Table 6). The dissimilarity of the NIR band was another texture feature that was identified as relatively important in tree species classification. In agreement with our previous study (Shi et al., 2018a), radiometric LiDAR metrics (intensity and echo width related metrics) played an important role in the tree species classification – by well representing the structural and morphological characteristics of the tree crown. Our results showed that LiDAR metrics were more capable than texture features in distinguishing tree species in certain cases. For example, the mean value of echo width (Ewmean) was able to distinguish birch from other four species.

Digital aerial photographs have several advantages over hyperspectral and LiDAR data, including lower deployment and data

Table 6
Top five most important variables for discriminating each tree species in the classification. See Table 3 for the definitions of the metrics.

	Beech	Birch	Fir	Maple	Spruce
1	2015_MEAN_NIR	Ewmean	lmean_first	Ip75	2016_MEAN_NIR
2	2016_MEAN_NIR	Ewmean_first	Ip75	lmean_first	2015_DIS_NIR
3	2015_MEAN_G	Height	lmean_single	Icv	2017_DIS_NIR
4	2017_DIS_NIR	Hvar	Height	lmean	Ewmean_single
5	lmean_first	Hsd	2015_MEAN_NIR	Ivar	Height

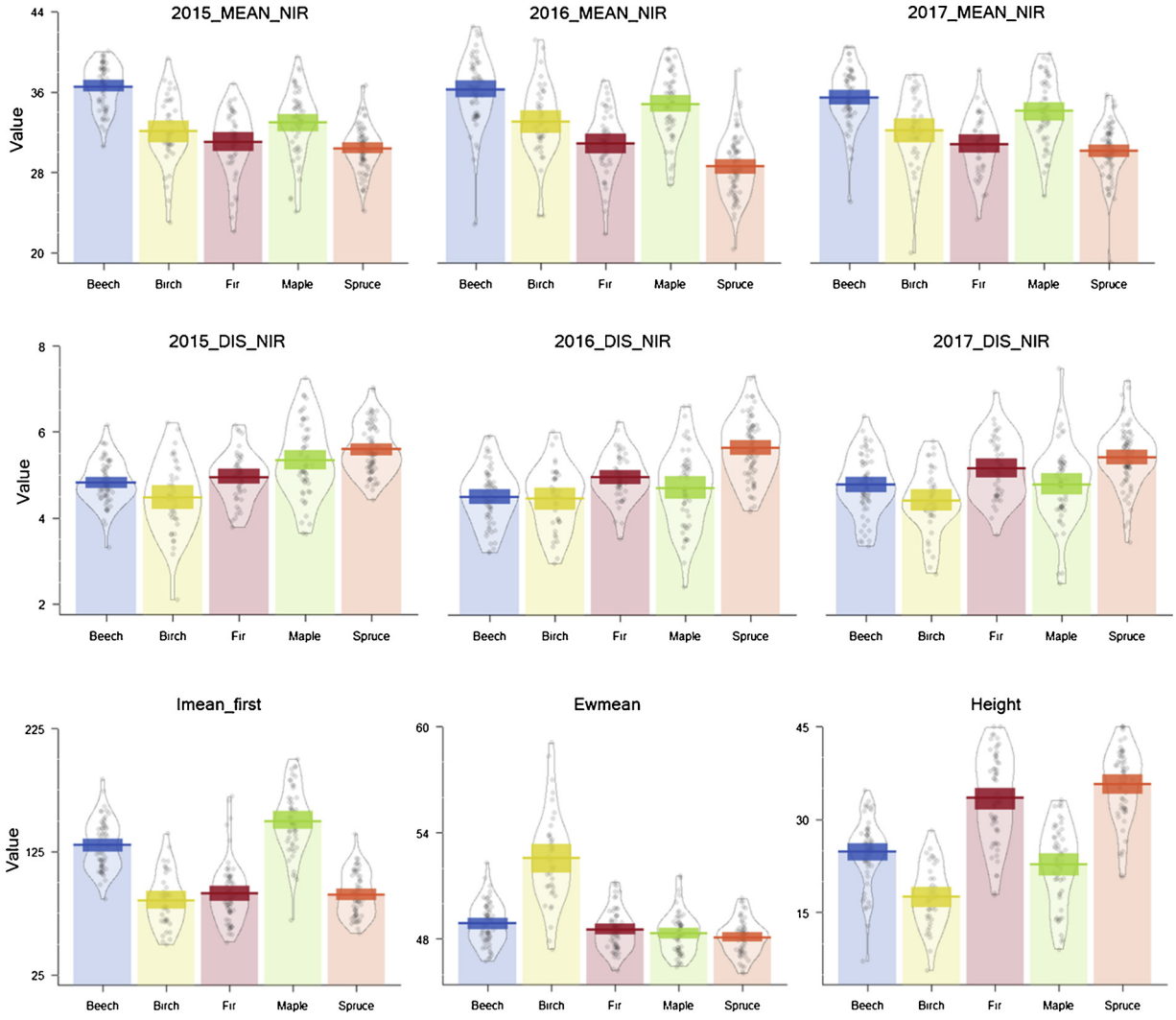


Fig. 5. Interspecies comparison of different variables derived from multi-temporal digital CIR orthophotos and LiDAR data. The first row is the average gray level of the NIR band (MEAN_NIR) from 2015, 2016 and 2017 digital CIR orthophotos; the second row is the dissimilarity of the NIR band (DIS_NIR) from 2015, 2016 and 2017 digital CIR orthophotos; and the last row is the mean intensity of first-or-only returns (Imean_first), the mean value of echo width (Ewmean), and the tree height (Height) derived from airborne LiDAR data.

collection costs, easier access to a variety of platforms, and the opportunity for forest managers to capture and process data over a long time span. As reported by Key et al. (2001), phenological information such as spring leaf-out, leaf maturity, and autumn senescence captured by multi-temporal digital aerial photographs are important indicators in the spectral separation of individual tree species. Digital aerial photographs taken in the spring, shortly after the flushing of leaves, or in autumn, after the trees have turned color (Holmgren et al., 2008), would be able to provide more species-specific information for species identification. We note that system differences and misregistration between images may hamper the utility of multi-temporal datasets, especially those from different platforms and sensors. Furthermore, the spectral bands available in most commercial sensors are located in the blue, green, red and near-infrared parts of the spectrum, which only describe the differences between the tree species across limited wavelengths. As discussed by Heikkinen et al. (2010), tree species classification based on multispectral images could be improved by including the red edge region of the spectrum, which is more unique to tree species, age and health (Cho and Skidmore, 2006). Future decisions regarding the use of multispectral sensors for mapping tree species at the individual tree level should be made with consideration of the trade-offs between spatial resolution, spectral resolution, flight season and revisit cycle.

5. Conclusions

To our knowledge, this is the first study that investigates how archived multi-temporal digital aerial colour-infrared photographs can be used to improve LiDAR-based individual tree species mapping. Our results show that the texture features generated from multi-temporal digital CIR orthophotos under different view-illumination conditions are species-specific. Combining these texture features with LiDAR metrics was shown to significantly improve the individual tree species mapping accuracy in a temperate mixed forest in eastern Germany. Specifically, the average gray level and the dissimilarity of the NIR band contributed most (among texture features) to the classification. Our results demonstrate that the fusion of multi-temporal digital aerial photographs with airborne LiDAR data can accurately classify individual tree species in Central European mixed forests.

Acknowledgements

This work was supported by the China Scholarship Council under Grant 201506650001 and co-funded by the ITC Research Fund of the

University of Twente. We acknowledge the support of the “Data Pool Initiative for the Bohemian Forest Ecosystem” data-sharing initiative of the Bavarian Forest National Park. We thank Prof. Dr. Peter Krzystek from the University of Applied Science Munich for conducting individual tree 3D segmentation.

References

Aber, J.S., Marzolff, I., Ries, J., 2010. Small-Format Aerial Photography. Elsevier.

Barrett, F., McRoberts, R.E., Tomppo, E., Cienciala, E., Waser, L.T., 2016. A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. Remote Sens. Environ. 174, 279–289.

Breiman, L., 2001. Random forests. Mach. Learn. 45.

Cailleret, M., Heurich, M., Bugmann, H., 2014. Reduction in browsing intensity may not compensate climate change effects on tree species composition in the Bavarian Forest National Park. For. Ecol. Manage. 328, 179–192.

Cho, M.A., Skidmore, A.K., 2006. A new technique for extracting the red edge position from hyperspectral data: the linear extrapolation method. Remote Sens. Environ. 101, 181–193.

Clausi, D.A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote. Sens. 28, 45–62.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20, 37–46.

Colwell, R.N., 1960. Manual of Photographic Interpretation. American Society of Photogrammetry.

Coops, N.C., Wulder, M.A., Culvenor, D.S., St-Onge, B., 2004. Comparison of forest attributes extracted from fine spatial resolution multispectral and lidar data. Can. J. Remote Sens. 30, 855–866.

Cutler, D.R., Edwards Jr, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.

Dechesne, C., Mallet, C., Le Bris, A., Gouet-Brunet, V., 2017. Semantic segmentation of forest stands of pure species combining airborne lidar data and very high resolution multispectral imagery. ISPRS J. Photogramm. Remote. Sens. 126, 129–145.

Deng, S., Katoh, M., Yu, X., Hyyppä, J., Gao, T., 2016. Comparison of tree species classifications at the individual tree level by combining ALS data and RGB images using different algorithms. Remote Sens. (Basel) 8, 1034.

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., 2016. Review of studies on tree species classification from remotely sensed data. Remote Sens. Environ. 186, 64–87.

Haara, A., Haarala, M., 2002. Tree species classification using semi-automatic delineation of trees on aerial images. Scand. J. For. Res. 17, 556–565.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, SMC-3 610–621.

Heikkinen, V., Tokola, T., Parkkinen, J., Korpela, I., Jääskeläinen, T., 2010. Simulated multispectral imagery for tree species classification using support vector machines. IEEE Trans. Geosci. Remote. Sens. 48, 1355.

Heinzel, J.N., Weinacker, H., Koch, B., 2008. Full automatic detection of tree species based on delineated single tree crowns–a data fusion approach for airborne laser scanning data and aerial photographs. Proceedings of SILVILASER, 2008, 8th.

Heurich, M., 2008. Automatic recognition and measurement of single trees based on data from airborne laser scanning over the richly structured natural forests of the Bavarian Forest National Park. For. Ecol. Manage. 255, 2416–2433.

Holmgren, J., Persson, Å., Söderman, U., 2008. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. Int. J. Remote Sens. 29, 1537–1552.

Kayitakire, F., Hamel, C., Defourny, P., 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. Remote Sens. Environ. 102, 390–401.

Key, T., Warner, T.A., McGraw, J.B., Fajvan, M.A., 2001. A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. Remote Sens. Environ. 75, 100–112.

Kim, S., McGaughey, R.J., Andersen, H.E., Schreuder, G., 2009. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. Remote Sens. Environ. 113, 1575–1586.

Koch, B., Svoboda, J., Adler, P., Dees, M., 2002. Automatic tree species detection based on digitised CIR aerial photos. Allgemeine Forst Und Jagdzeitung 173, 131–140.

Korpela, I., 2006. Incorporation of Allometry into Single-Tree Remote Sensing With Lidar and Multiple Aerial Images. Department of Forest Resource Management. University of Helsinki, Finland, pp. 6.

Korpela, I., Heikkinen, V., Honkavaara, E., Rohrbach, F., Tokola, T., 2011. Variation and directional anisotropy of reflectance at the crown scale—implications for tree species classification in digital aerial images. Remote Sens. Environ. 115, 2062–2074.

Korpela, I., Mehtätalo, L., Markelin, L., Seppänen, A., Kangas, A., 2014. Tree Species Identification in Aerial Image Data Using Directional Reflectance Signatures.

Koukoulas, S., Blackburn, G.A., 2005. Mapping individual tree location, height and species in broadleaved deciduous forest using airborne LIDAR and multi-spectral remotely sensed data. Int. J. Remote Sens. 26, 431–455.

Kulikova, M.S., Mani, M., Srivastava, A., Descombes, X., Zerubia, J., 2007. Tree species classification using radiometry, texture and shape based features. Signal Processing Conference, 2007 15th European. IEEE, pp. 1595–1599.

Kuzmin, A., Korhonen, L., Manninen, T., Maltamo, M., 2016. Automatic segment-level tree species recognition using high resolution aerial winter imagery. Eur. J. Remote. Sens. 49, 239–259.

Leckie, D.G., Tinis, S., Nelson, T., Burnett, C., Gougeon, F.A., Cloney, E., Paradine, D., 2005. Issues in species classification of trees in old growth conifer stands. Can. J. Remote. Sens. 31, 175–190.

Liaw, A., Wiener, M., 2002. Classification and Regression by Random Forest. R news *2/3*.

Loetsch, F., Haller, K., 1964. Forest Inventory. Volume I. Statistics of Forest Inventory and Information From Aerial Photographs.

McNemar, Q., 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12, 153–157.

Morgan, J.L., Gergel, S.E., Ankerson, C., Tomscha, S.A., Sutherland, I.J., 2017. Historical aerial photography for landscape analysis. Learning Landscape Ecology. Springer, pp. 21–40.

Ozdemir, I., Norton, D.A., Ozkan, U.Y., Mert, A., Senturk, O., 2008. Estimation of tree size diversity using object oriented texture analysis and aster imagery. Sensors 8, 4709–4724.

Pasher, J., King, D.J., 2010. Multivariate forest structure modelling and mapping using high resolution airborne imagery and topographic information. Remote Sens. Environ. 114, 1718–1732.

Persson, Å., Holmgren, J., Söderman, U., Olsson, H., 2004. Tree species classification of individual trees in Sweden by combining high resolution laser data with high resolution near-infrared digital images. Int. Arch. Photogr. Remote Sens. Spatial Inform. Sci. 36, 204–207.

Reitberger, J., Schnorr, C., Krzystek, P., Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS J. Photogramm. Remote Sens. 64, 561–574.

Rodriguez-Galiano, V.F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P.M., Jeganathan, C., 2012. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. Remote Sens. Environ. 121, 93–107.

Sayn-Wittgenstein, L., 1978. Recognition of Tree Species on Aerial Photographs.

Shi, Y., Skidmore, A.K., Wang, T., Holzwarth, S., Heiden, U., Pinnel, N., Zhu, X., Heurich, M., 2018a. Tree species classification using plant functional traits from LiDAR and hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 73, 207–219.

Shi, Y., Wang, T., Skidmore, A.K., Heurich, M., 2018b. Important LiDAR metrics for discriminating forest tree species in Central Europe. ISPRS J. Photogramm. Remote. Sens. 137, 163–174.

Silveyra Gonzalez, R., Latifi, H., Weinacker, H., Dees, M., Koch, B., Heurich, M., 2018. Integrating LiDAR and high-resolution imagery for object-based mapping of forest habitats in a heterogeneous temperate forest landscape. Int. J. Remote Sens. 1–26.

Singh, M., Evans, D., Tan, B.S., Nin, C.S., 2015. Mapping and characterizing selected canopy tree species at the Angkor World Heritage site in Cambodia using aerial data. PLoS One 10, e0121558.

Spurr, S.H., 1960. Photogrammetry and photo-interpretation. Soil Sci. 90, 380.

Tian, J., Schneider, T., Straub, C., Kugler, F., Reinartz, P., 2017. Exploring digital surface models from nine different sensors for forest monitoring and change detection. Remote Sens. (Basel) 9, 287.

Waser, L., Ginzler, C., Rehush, N., 2017. Wall-to-Wall tree type mapping from country-wide airborne remote sensing surveys. Remote Sens. (Basel) 9, 766.

Waser, L.T., Ginzler, C., Kuechler, M., Baltsavias, E., Hurni, L., 2011. Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. Remote Sens. Environ. 115, 76–85.

Yao, W., Krzystek, P., Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sens. Environ. 123, 368–380.

Yao, W., Krzystek, P., Heurich, M., 2013. Enhanced detection of 3D individual trees in forested areas using airborne full-waveform LiDAR data by combining normalized cuts with spatial density clustering. ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. 1, 349–354.

Yin, D., Wang, L., 2016. How to assess the accuracy of the individual tree-based forest inventory derived from remotely sensed data: a review. Int. J. Remote Sens. 37, 4521–4553.

Yu, X., Hyyppä, J., Litkey, P., Kaartinen, H., Vastaranta, M., Holopainen, M., 2017. Single-sensor solution to tree species classification using multispectral airborne laser scanning. Remote Sens. (Basel) 9, 108.

Zhang, K., Hu, B., 2012. Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles. Remote Sens. (Basel) 4, 1741–1757.