Remote Sensing of Environment 186 (2016) 64–87

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Review of studies on tree species classification from remotely sensed data

Fabian Ewald Fassnacht [a,*], Hooman Latifi [b], Krzysztof Stereńczak [c], Aneta Modzelewska [c], Michael Lefsky [d], Lars T. Waser [e], Christoph Straub [f], Aniruddha Ghosh [g]

[a] Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany
[b] Department of Remote Sensing in Cooperation with German Aerospace Center, University of Wuerzburg, Oswald-Kuelpe-Weg 86, D-97074 Wuerzburg, Germany
[c] Department of Forest Resources Management, Forest Research Institute, Sekocin Stary, 3 Braci Lesnej Street, 05-090 Raszyn, Poland
[d] Center for Ecological Applications of LiDAR, College of Natural Resources, Colorado State University, 400 University Ave, Fort Collins, USA
[e] Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zuercherstrasse 111, 8903 Birmensdorf, Switzerland
[f] Department of Information Technology, Bavarian State Institute of Forestry (LWF), Hans-Carl-von-Carlowitz-Platz 1, D-85354 Freising, Germany
[g] Department of Environmental Science and Policy, University of California, 1023 Wickson Hall, Davis, USA

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form 19 July 2016
Accepted 8 August 2016
Available online xxxx

Keywords:
Forestry
Remote sensing
Scale
Tree species
Classification
Mapping
Validation

ABSTRACT

Spatially explicit information on tree species composition of managed and natural forests, plantations and urban vegetation provides valuable information for nature conservationists as well as for forest and urban managers and is frequently required over large spatial extents. Over the last four decades, advances in remote sensing technology have enabled the classification of tree species from several sensor types.

While studies using remote sensing data to classify and map tree species reach back several decades, a recent review on the status, potentials, challenges and outlooks in this realm is missing. Here, we search for major trends in remote sensing techniques for tree species classification and discuss the effectiveness of different sensors and algorithms based on a literature review.

This review demonstrates that the number of studies focusing on tree species classification has increased constantly over the last four decades and promising local scale approaches have been presented for several sensor types. However, there are few examples for tree species classifications over large geographic extents, and bridging the gap between current approaches and tree species inventories over large geographic extents is still one of the biggest challenges of this research field. Furthermore, we found only few studies which systematically described and examined the traits that drive the observed variance in the remote sensing signal and thereby enable or hamper species classifications. Most studies followed data-driven approaches and pursued an optimization of classification accuracy, while a concrete hypothesis or a targeted application was missing in all but a few exceptional studies.

We recommend that future research efforts focus stronger on the causal understanding of why tree species classification approaches work under certain conditions or – maybe even more important - why they do not work in other cases. This might require more complex field acquisitions than those typically used in the reviewed studies. At the same time, we recommend reducing the number of purely data-driven studies and algorithm-benchmarking studies as these studies are of limited value, especially if the experimental design is limited, e.g. the tree population is not representative and only a few sensors or acquisition settings are simultaneously investigated.

Contents

* Corresponding author.
E-mail addresses: fabian.fassnacht@kit.edu (F.E. Fassnacht), hooman.latifi@uni-wuerzburg.de (H. Latifi), k.sterenczak@ibles.waw.pl (K. Stereńczak), a.modzelewska@ibles.waw.pl (A. Modzelewska), lefsky@gmail.com (M. Lefsky), lars.waser@wsl.ch (L.T. Waser), Christoph.Straub@lwf.bayern.de (C. Straub), anighosh@ucdavis.edu (A. Ghosh).

http://dx.doi.org/10.1016/j.rse.2016.08.013
0034-4257/

1. Introduction and review approach

1.1. Importance of tree species information

Remote sensing-assisted classification of tree species is motivated by a wide variety of applications confronting the forest management and conservation sectors. These applications include questions linked to resource inventories (van Aardt and Wynne 2007), biodiversity assessment and monitoring (Shang and Chisholm, 2014), hazard and stress management (Cho et al., 2010; Fassnacht et al., 2014), monitoring of invasive species (Boschetti et al., 2007), wildlife habitat mapping (Jansson and Angelstam, 1999) as well as the overarching aim of a sustainable forest management (European Environmental Agency, 2007). Many studies highlighted the importance of tree species maps either as standalone products for forest management (e.g. Dalponte et al., 2012; Heinzel and Koch, 2012) or as an input for species-specific growth and yield models (e.g. Vauhkonen et al., 2014) or any species-specific allometric model (Ørka et al., 2013). In this realm, Korpela and Tokola (2006) demonstrated the importance of tree species information in remote sensing-based single tree inventory approaches to avoid unwanted averaging effects (e.g. when calculating growing stock volume). Knowledge on tree species distribution may also affect forest harvesting and management policies (Dalponte et al., 2012; Jones et al., 2010; Plourde et al., 2007). In urban areas, the sustainable management of urban trees requires species information as well and remote sensing approaches have been discussed as an efficient alternative to field inventories (Jensen et al., 2012).

Spatially explicit information on tree species composition over large areas are also relevant for an improved understanding of the ecology of tree species for example concerning community dynamics and the contribution of species to ecosystem functions and services (Chambers et al., 2013; Van Ewijk et al., 2014). Other environmental studies for instance focusing on wildlife habitat mapping (Pausas et al., 1997) or estimation of insect abundances in forests (Kennedy and Southwood, 1984) also benefited from tree species information.

1.2. Objectives

The objectives of this review on tree species classification are to:

1. Quantify general trends in remote sensing studies focusing on tree species classification
2. Provide a detailed overview of the current approaches for classifying tree species from typical sensor types
3. Identify research gaps and future trends for tree species classification using remote sensing data

1.3. Review approach

We present a literature review supported by a set of descriptive statistics derived from 101 selected studies. We focused the search for research articles published in peer-reviewed journals in English language within the time span between January 1980 and December 2014. We did not restrict scales or sensor types. Thereby we considered a wide variety of spatial scales from leaf to stand level, as well as all sensor types including active (Light Detection and Ranging (LiDAR), Synthetic Aperture Radar (SAR)) and passive (multispectral, hyperspectral, thermal) systems. We browsed ISI Web of Science and Google Scholar databases using the following keywords: Remote Sensing AND Vegetation OR Tree OR Plant OR forest AND Classif* OR Map* OR Identi* OR Detect* OR Discriminat*. We performed alteration in the keywords to refine the search for specific sensors, for example, replacing Remote Sensing with specific sensors (multispectral, hyperspectral, LiDAR, SAR, thermal). >1200 studies were found satisfying the above conditional search. These studies were further pruned based on the following criteria:

1. The research must discriminate at least two tree species.
2. The research must not focus on broader forest types (for example coniferous, broadleaved).
3. The research must mainly consider the spatial presence/absence of tree species.
4. The research must report detailed accuracy assessment results (at least a measure of overall accuracy).

In exceptional cases we considered studies in the review that did not fulfill all criteria. These studies were mostly considered because they provided very relevant information on the topic or due to a generally sparse number of papers for a certain sensor type (e.g., SAR and thermal). The number of SAR studies was further limited as we excluded older studies that had already been covered in the review of Boerner et al. (1998).

To derive the supporting descriptive statistics we occasionally dismantled multi-sensor approaches or comparative studies in sub-studies and treated them as individual cases based on the number of variant data and methods they used. For example, a number of studies integrated multiple investigations on variable number of tree species (e.g. Heinzel and Koch, 2011) or scales (e.g., Ghosh et al., 2014; Ørka et al., 2013; Suratno et al., 2009). The final number of reviewed studies was 101 (references of all studies are listed in the Supplementary data) and the final number of cases was 129.

The remainder of the study is structured as follows: Section 2 will give a short overview over potential users of remote-sensing derived tree species products and their demands; Section 3 will depict current trends in tree species classification studies based on descriptive statistics obtained from the 101 reviewed studies; Section 4 focuses on tree species traits that cause species-related differences in the signal observed from various sensor types; Section 5 highlights the most important methods which are used for remote sensing-assisted tree species classifications; Section 6 discusses current constraints; Section 7 gives a short overview over new upcoming sensors and platforms and their potential for tree species classification; in Section 8 conclusions are drawn.

2. User demands and cost-efficiency

The motivation to derive tree species information from remote sensing data arises from the assumption that there are potential users for this information (either in form of a tree species map or tree species layers as auxiliary information in other models).

The demands concerning spatial detail and desired accuracy vary with the targeted application and it is therefore hard to define generally valid minimum accuracy percentages (compare also Section 5.2.1) or an optimal target unit. For example, single-tree species information might be desirable in precision forestry while for the identification of forest stands endangered under climate change scenarios, tree species composition information at one hectare resolution can be sufficient (Immitzer et al., 2015). Forest enterprises and administrations in central Europe would often already benefit from accurate stand-wise estimates of coniferous to broadleaved ratios. The generally large interest of practitioners in remote sensing derived tree species information is mirrored in the survey of Felbermeier et al. (2010) who analyzed 347 questionnaires sent to professionals working in the forestry sector. Two thirds of the interviewees reported forest information deficiencies and 90% of them expected improvements by the application of remote sensing. Out of 63 parameters, tree species were ranked first when the interviewees were asked which parameters have to be addressed by remote sensing applications. This figure might be biased, as for some of the interviewees - not being remote sensing experts - species information just might have been the most logical parameter to obtain from remote sensing data.

One crucial point when thinking about the operational provision of remote sensing products is their cost-efficiency. In this realm, most competitive sensor systems are satellite based systems including free of charge medium spatial resolution systems (e.g., Landsat, Sentinel-2) but also high resolution commercial systems (Rapid Eye, Ikonos, Geoeye) with an approximate cost of 1–14 € per km^2 (Ørka and Hauglin, 2016). Amongst airborne systems, aerial imagery (approximately 35–62 € per km^2) is still more affordable than hyperspectral (120–180 € per km^2) and LiDAR data (62–240 € per km^2) (Ørka and Hauglin, 2016) but the information content might be higher in the last-mentioned. From the perspective of a private or state forest owner, the costs associated with the remote sensing data have to be compensated by a reduction of field measurement expenses or by an increase in management efficiency that leads to increased income based on improved decision making. As an example, in Poland the number of field plots for total growing stock estimation would have to be reduced by approximately 50% to compensate for the costs of the airborne laser scanning data acquisition. To the best of our knowledge detailed assessments of financial benefits generated by species information have not been presented yet. However, approaches to evaluate differing forest inventory approaches via cost-plus-loss analysis (e.g., Eid et al., 2004; Bergseng et al., 2015) or the net present value (Holopainen and Talvitie, 2006) exist. In these studies, the suitability and cost-efficiency of differing forest inventory methods are compared by confronting the total cost of the inventory with the expected economic losses due to incorrect future decisions based on wrong inventory information (Eid et al., 2004). As one of the few examples concerning tree species, Jones et al. (2010) found automated methods from combined hyperspectral and LiDAR data (approximately 6 USD per ha) to be competitive against traditional aerial photograph interpretation (approximately 12 USD per ha) in terms of accuracy and cost for a study area in South-western Canada.

It can be summarized that an interest in remotely sensed tree species information exists on practitioners' side. However, more cost efficiency analyses are needed in order to justify its use. In the reviewed studies only few papers mainly conducted in Scandinavia (e.g., Dalponte et al., 2013; Pant et al., 2013) and North America (e.g., Bauer et al., 1994; Jones et al., 2010) considered cost as a relevant factor in the discussion.

3. Trends in tree species classification

3.1. Literature trends

The number of studies focusing on tree species classification has constantly increased over the last 35 years (Fig. 1) which is well-supported by the general trend of increased publication activity. However, the almost exponential growth between the time periods 2005–2010 and 2010–2015 is also a function of the increased availability of airborne hyperspectral and LiDAR data which is illustrated in the sensor-specific frequencies (dashed curves in Fig. 1). Both data sources have been frequently applied in a forest inventory context with tree species being

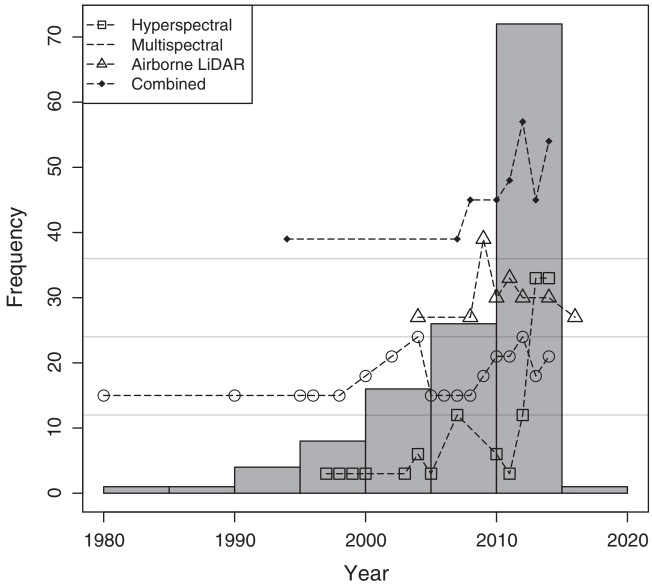

Fig. 1. Frequency of studies per five-year time interval. The histogram is overlaid with sensor-specific frequencies per year. The plotted values are scaled by a factor of three on y - axis to emphasize the trends and offsets were added to improve the readability of the figure. Light horizontal lines represent the y-axes (at y = 0) for each sensor type.

one of the more popular target variables besides total growing stock volume and biomass.

Most studies were conducted within temperate forest ecosystems (67 cases), followed by studies in boreal ecosystems (27 cases) and tropical forests (18 cases). Savannah systems (10 cases) were examined mostly in South Africa. There were 7 cases which we could not allocate to any of the defined classes. The patterns of biome frequency was relatively stable over all sensor types, although studies using ALS and combined approaches show a trend to be more frequently situated in boreal biomes than the studies using other sensor types.

The trends concerning the biomes are also reflected in the frequency of studies per nation (Fig. 2). Most studies were conducted in North America (USA + Canada) and Europe (especially Germany and Scandinavia). The studies conducted in the Krueger National Park lead to another small hot-spot in South Africa.

Focusing on sensor types, most cases were observed for hyperspectral or imaging spectroscopy studies (42 cases). The second most frequently applied sensor type (36 cases) were multispectral systems which range from moderate (Landsat, Hyperion) to high (SPOT, RapidEye) and very-high (IKONOS, WorldView satellites, airborne sensors) spatial resolution systems. There were also a large number of studies making use of more than one sensor type (28 cases). Except for one case, all studies combined an active system (in all but one cases LiDAR) with a passive optical system (very often airborne multi- or hyperspectral systems). This explains the slightly reduced number of LiDAR-only studies (17 cases). Only a few studies made use of SAR or thermal sensors to separate tree species (in total 6 cases) although there have been considerable efforts to map broader forest types (e.g. coniferous and broadleaved forests) with SAR sensors. The number of species that were considered in the studies of each sensor type are summarized along with the obtained accuracies in Fig. 3. Accuracy values are in almost all cases overall accuracy values while in few cases % kappa was used as no overall accuracy was reported. LiDAR and SAR studies typically considered fewer species than multispectral, hyperspectral, and combined sensors. Nevertheless, the reported accuracies were comparable.

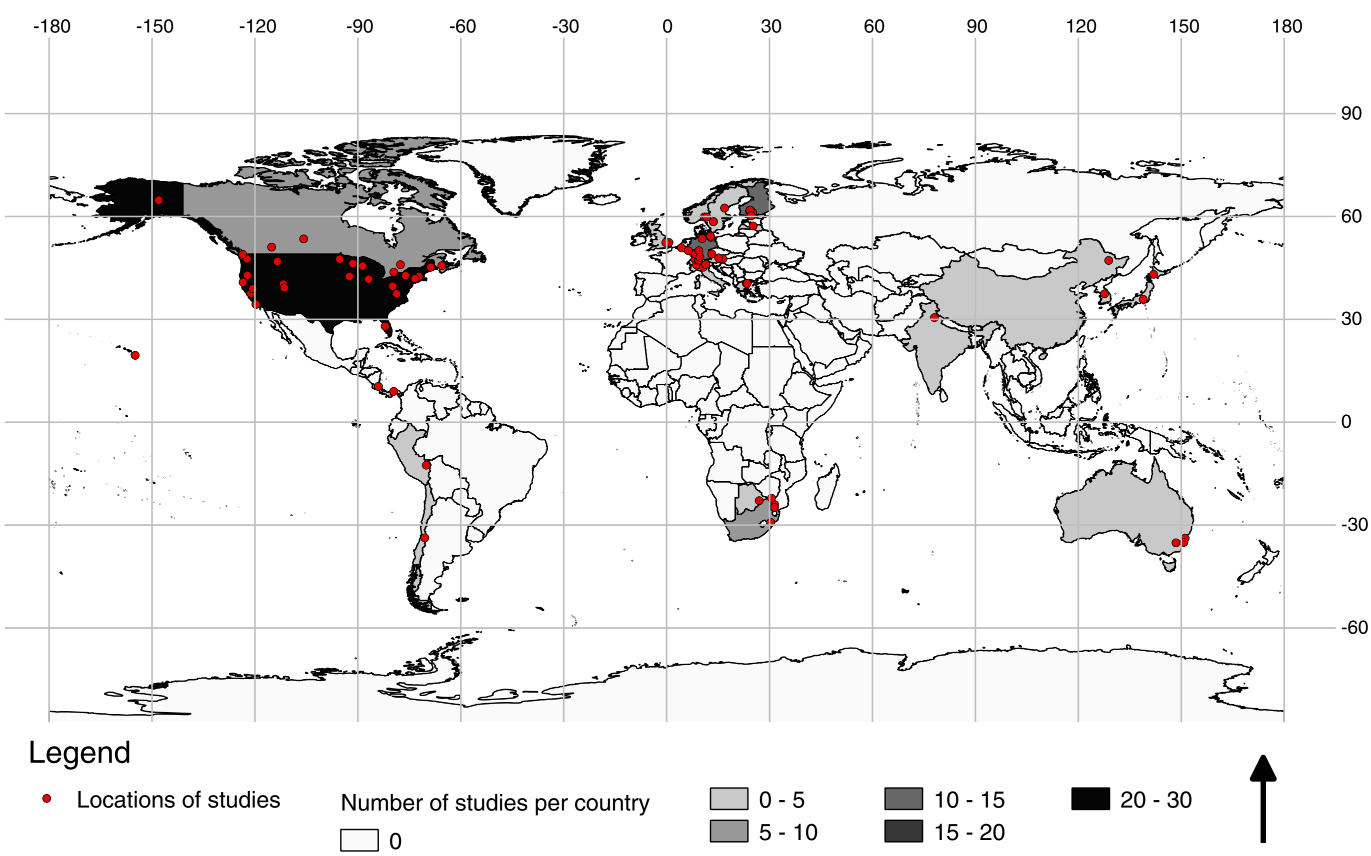

Fig. 2. World map displaying where the 101 selected studies have been conducted throughout the world (red dots). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

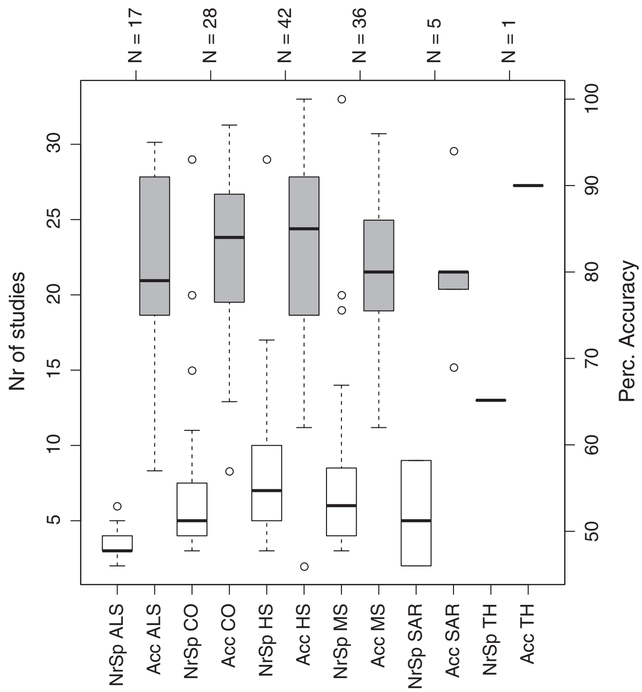

Fig. 3. Number of species and obtained accuracies per sensor type. Abbreviations are: NrSp = number of species; Acc = obtained accuracies; ALS = airborne laser scanning; CO = combined sensor systems; HS = hyperspectral; MS = multispectral; SAR = synthetic aperture radar; TH = thermal; N = number of studies.

It must be noted that the presented numbers are affected by the study selection process followed in the review approach and might lack completeness.

Concerning scale, we observed a clear dominance of studies working on a single tree scale (77 cases). Studies making use of all different sensor types contributed to this group. Second ranked was forest/stand scale (26). Multispectral systems contributed the most to this category. Except for LiDAR, studies using all sensor types contributed to the tree group scale (16). Finally, we considered eight studies that exclusively focused on leaf-level (laboratory experiments with individual leaf samples) of which seven used hyperspectral and one thermal data.

3.2. Optimal ground sampling density and spatial unit

Two questions related to scale arise when classifying tree species from remote sensing: (1) what is the spatial unit on which species information should be obtained? And (2) what is the optimal ground sampling density of a given sensor to derive tree species information?

3.2.1. Terminology

Forestry experts characterize species composition typically through single-tree distributions or proportions of stem-volume per spatial unit (e.g., a stand). The probability that a given pixel has exactly the size corresponding to a single tree or a stand and also matches its extent is low. The presence of multiple age classes and species makes this case virtually impossible as sensor-visible (dominant) crowns vary in size. In this realm, scale is also related to a terminological problem. Any species-level classification should by definition be for single trees or stands consisting only of a single species. Therefore, approaches focusing on larger spatial units (either stand objects or pixel sizes that do not allow for the separation of individual trees) are rather focusing on forest stand composition classification or on the assessment of species mixtures than on tree species classification (except for the case when stands consist of only a single species). In forestry, "forest stand composition" and also "species mixture" are often defined based on species-specific basal area or standing timber volume. These definitions do not match the remote sensing perspective which typically describes stand composition or species mixture as percentages of species-specific crown cover present in the upper canopy layer that is visible in the remote sensing data.

3.2.2. Spatial units and background signal

From a technical perspective, classification of either individual trees or management units (e.g., forest stands) requires object-based classification approaches. From the 129 reviewed cases, 63 studies applied object-based classifications. Of these, 56 were conducted on the single tree level and seven on other object-types (mostly management units or tree groups). The remaining cases operated at pixel level or used field spectra. Studies operating at stand level have to consider that the influence of the background signal may differ in stands with identical species composition but differing canopy closure. Furthermore, the type of background signal (e.g., bare soil versus understory vegetation) will also vary with location. This will notably increase the intra-species variability of the remote sensing signal and may even complicate the discrimination of stands consisting only of a single species. This is particularly important, when attempting to map larger geographic extents where the co-occurrence of stands of the same species with different canopy closure and background signal types is more likely. Some of the reviewed studies following pixel-based approaches circumvented problems related to background signals by only considering reference data derived from dense mono-species stands (e.g., Adelabu et al., 2013; Carleer and Wolff, 2004; Ghosh et al., 2014).

To characterize stands with mixtures of several species, classification approaches are of limited values and spectral unmixing or similar approaches should be preferred if the spatial resolution of the data does not allow for an individual tree approach within the stand (see e.g., Somers and Asner, 2014). Replacing discrete species classes by continuous fields related to spectral properties of the species could be an alternative approach. Schmidtlein et al. (2007) proposed such an approach for grassland areas.

While these coherences concerning background signal and object-size appear to be intuitive, they were actively considered in the methodology in only 11 cases while in 23 cases problems with background signals were mentioned in the discussion. For example, Shang and Chisholm (2014) reported reduced classification accuracies for a community (stand)-level classification as compared to a single tree classification and explained it by presumably increased influence of the background signal. Jensen et al. (2012) attributed most of their misclassifications to understory show-through and atypical individual canopy signatures. Mallinis et al. (2008) integrated an additional step in their object-based decision tree approach to discriminate between dense and open stands before classifying each object to species level.

3.2.3. Relations between ground sampling density and separating features

The second scale-related question, the identification of an optimal ground sampling distance, has been addressed by few hyperspectral studies. Focusing on coarser scales than the leaf level, the reflectance observed by optical sensors results from a complex interplay of radiation with crown tissues (foliage, stems, branches, fruits, lianas and flowers) (Clark et al., 2005) and background signal (stemming from soil, herbaceous vegetation, etc.). It is further dependent on the structural arrangement of foliage (number of layers, clumping and leaf angles) and their corresponding shadow fractions. View-illumination geometry is another important factor that adds to the complexity of the observed signal (see also Sub-section 5.7).

Due to this complex interplay, switching from individual leaf to twig/branch scale will notably increase the spectral variability of the observed reflectance (Roberts et al., 2004). Going to even coarser spatial scales (e.g., pixel sizes approximating a single crown) will presumably reduce the spectral variability, as the corresponding larger pixel sizes will lead to reflectance signals that average larger ranges of surface conditions (compare results and discussion in Ghosh et al., 2014). However, problems due to mixed species in a single pixel will increase with increasing pixel size. It yet remains unclear whether these processes affect species separation in a positive or negative way, and what the best spatial resolution for the successful separation of tree species is. An argument for small pixel sizes is that the increased spectral variability can be methodically addressed by applying object-based approaches. Furthermore, the distribution of spectral signatures of the pixels within a crown object could vary amongst species and therefore contain relevant information. The possibility of using distribution measures of pixel values in tree objects instead of averaged spectra for classification purposes is still underrepresented in the literature. Current studies rather average pixel values within an object to reduce intra-species variability (e.g., Pant et al., 2013; van Aardt and Wynne, 2007). For example, Peña et al. (2013) also explicitly state that very high spatial resolution leads to undesirable spectral complexity that complicates tree species classifications.

It is very likely that an optimal spatial resolution will also depend on the applied methods and the forest types under investigation. Marceau et al. (1994) stated that the discrimination of tree species is maximized when the pixel size of the applied data allows for depicting the intrinsic spatial properties of the examined trees. Only few studies in the passive optical domain have focused on the search for an optimal pixel size. Furthermore, the results of these studies are not in all cases directly comparable as the examined spatial scales and spatial properties of the forests differ: while Ghosh et al. (2014) observed their best classification results with a pixel size of 8 m (approximately 120% of an average crown diameter in the study area) when comparing three hyperspectral scenes with pixel sizes of 4 m, 8 m and 30 m resolution, Peña et al. (2013) observed their best classification accuracies for the highest spatial resolution (pixel size 0.3 m, approximately 5% of an average crown diameter) when spatially degrading their data to a pixel size of 2.4 m. Dalponte et al. (2013) also observed reduced accuracies when spatially degrading their data from 0.4 m (approximately 10% of an average crown diameter) to 1.5 m pixel size. One could interpret these results as a confirmation of the above-mentioned hypothesis concerning the evolvement of spectral variability with pixel size: Assuming a relatively dense canopy, accuracies decrease when switching from close-to-leaf level scales to twig/branch scales (Dalponte et al., 2013; Peña et al., 2013) while accuracies increase again when the pixel size reaches more or less a single-crown level (Ghosh et al., 2014). This hypothesis only holds true for classifications at pixel level, while object-based approaches are likely to generally benefit from higher spatial resolutions.

The hypothesis receives further support by studies that classified hyperspectral data at the leaf level which often reached high (e.g., 80.4–83.5% OA for 17 species (Hesketh and Sanchez-Azofeifa, 2012)) to very high accuracies (e.g., 96% OA under cross-validation for 13 species (Pu and Liu, 2011)) and typically reached higher accuracies in comparison to airborne approaches in the few studies that conducted a direct comparison of the two approaches (Clark and Roberts, 2012; Shang and Chisholm, 2014). However, the leaf-level datasets in these studies might not only benefit from the decreased variance due to the homogeneous structure of the leaves, but also from comparably good measurement conditions that can be considered more or less free from observation errors, shadow casting and background signals. At this point the number of studies focusing on scale-related issues in the optical domain is sparse and more systematic research is needed. Good examples for systematic investigations related to scale are given by the studies of Korpela et al. (2014) and Roberts et al. (2004).

3.2.4. Ground sampling density of LiDAR data

Focusing on LiDAR data, scale relates to the degree of illumination of a certain area by laser pulses. In few studies, two related variables - pulse density and footprint size - have been investigated in a tree species classification context. There is a general agreement that higher pulse density (up to a threshold of approximately 10 points per m^2) and especially the application of waveform LiDAR data (Yu et al., 2014) improves the detection of single trees and thereby has additional positive effects on the stability of structural metrics used as input to species classification at single tree level (e.g., Reitberger et al., 2008; Hovi et al., 2016). Low density discrete return LiDAR was reported to be of limited value for species discrimination (Suratno et al., 2009; Yu et al., 2014). For simulated waveform data, Hovi and Korpela (2014) reported reduced classification accuracies for footprint sizes smaller than 0.36 m when classifying juvenile forest vegetation (compare Section 5.9).

4. Species-related traits measured by remote sensing sensors

4.1. Passive optical (multispectral/hyperspectral) data

Passive optical sensors can be divided into multispectral and hyperspectral (also called imaging spectroscopy) systems. Whereas most of the multispectral sensor systems typically have 4–8 bands, hyperspectral imagery is acquired in narrow, contiguous bands that can cover the visible (VIS), near-infrared (NIR) and shortwave-infrared (SWIR) portions of the electromagnetic spectrum (0.4–2.5 μm). Both sensor systems provide useful information to separate tree species by measuring the spectral response of directional electromagnetic radiation emitted by the sun and reflected by the canopy (and other surfaces) in sensor-specific wavelengths regions. These measurements are then typically transformed to spectral radiance and later to at-surface reflectance (see Schaepman-Strub et al., 2006 for a discussion of the definition of at-surface reflectance) (compare Section 5.6). In tree canopies the amount of radiation that is reflected in the different wavelengths regions is related to 1) plant chemical properties of the tissue which include water, photosynthetic pigments and structural carbohydrates, 2) leaf morphology (thickness of cell-walls, air spaces and cuticle wax) (Asner, 1998; Clark et al., 2005; Grant, 1987) as well as 3) canopy structure (leaf and branch density, angular distribution, clumping) and tree size compared to neighboring trees (Leckie et al., 2005) which also relates to view-illumination geometry. These properties vary not only with species but also with vertical leaf area density (Treuhaft et al., 2002), leaf age (Roberts et al., 1997; Einzmann et al., 2014) and health status (Waser et al., 2014). Further influences on the amount of reflected radiation are raised by background signals relating to bare soil, litter, lichens, mosses, herbaceous vegetation, lianas or other epiphytes, bark (Clark and Roberts, 2012) and from neighboring trees (Korpela et al., 2011).

The spectrum of green leaves in the VIS part of the spectrum (400–700 nm) is dominated by the absorption features of the photosynthetic pigments chlorophyll *a* and *b* and to a lower extent by absorptions of carotenoids and anthocyanins (Clark and Roberts, 2012; Ustin et al., 2009). At the leaf level, water is the major controller for the chemically driven absorption of the reflectance in the NIR (700–1300 nm) with two major absorption features at 970 and 1200 nm (Asner, 1998; Gao and Hoetz, 1990). The shapes of all of these absorption features are typically very similar across species. Focusing on leaf morphology, vegetation typically shows strong reflectance in the NIR which is caused by photon scattering within air-cell wall interfaces (Clark et al., 2005). In the SWIR region (1300–2500 nm) water absorbs a large amount of the incoming radiation and thereby masks also existing absorption features of lignin, cellulose and nitrogen which therefore can often only be observed in dry tissue (Clark and Roberts, 2012). Differences in the mentioned plant chemical and physical properties between species result in differing levels of reflectance (amplitude) but hardly affect the spectrum (shape). These differences in levels of reflectance are the main drivers to discriminate species in the VIS-SWIR region.

4.1.1. Important wavelength regions

The importance of chemical plant components for differentiating tree species is underlined by findings of studies that combined fine spectral resolution hyperspectral sensors with feature selection approaches (e.g., Dalponte et al., 2012; Fassnacht et al., 2014). With such an approach the importance of the different wavelengths regions and the associated absorption features can be examined. Fig. 4 illustrates important wavelength regions synthesized from 13 studies based on hyperspectral data. The important wavelength regions mainly overlap with absorption features of plant pigments and water, thereby supporting the hypothesis that these factors are the main drivers for species separation in the VIS-SWIR region. Fig. 4 also reveals that there were hardly any spectral regions that were not considered important in at least one of the 13 studies which might be interpreted as an argument for hyperspectral sensors with continuous coverage of the VIS to SWIR region. Concerning the three non-selected wavelength regions, two of them (1400–1450 nm; 1850–1900 nm) correspond to the centers of atmospheric water absorption bands (which typically got removed in a pre-processing step) while the third is positioned at the edge of the spectral range of the applied sensors (2450–2500 nm) and is therefore likely to be affected by noise.

The important wavelengths regions reported in Fig. 4 also relate reasonably well to findings of studies applying multispectral data. The importance of the visual bands have been reported by various studies using spaceborne multispectral data (Immitzer et al., 2012a; Moore and Bauer, 1990; Peerbhay et al., 2014; Waser et al., 2014) but is also underscored by studies using aerial imagery (Rohde and Olson, 1972; Erikson, 2004; Katoh et al., 2009; Waser et al., 2010). Some studies have pointed out the special importance of the blue wavelength region (Key et al., 2001; Peerbhay et al., 2014; Waser et al., 2011). Pu and Liu (2011) discuss the special relevance of the blue region to classify coniferous tree species due to their relatively lower photosynthetic activity in the blue light. Despite the importance of the red edge region reported in hyperspectral datasets and as indicated in Fig. 4, the red edge region was found to be less important in two multispectral studies including the one of Adelabu et al. (2013) who attempted to classify tree species in a semi-arid woodland in Botswana and Peerbhay et al. (2014) who examined WorldView-2 data to classify Eucalyptus plantations in South Africa.

4.1.2. Texture information

Crown texture information has also been exploited to improve tree species classification. Crown texture information is mainly related to crown-internal shadows, foliage properties (size, density, reflectivity) and branching (Sayn-Wittgenstein, 1978). On coarser scales, crown size, crown closure, crown shape, stand density and forest type (broadleaved, coniferous) are the main driver for texture in passive optical imagery.

Combining spectral and texture features often improves the accuracy of tree species classifications (Johansen and Phinn, 2006; Mallinis et al., 2008). According to Franklin et al. (2000) texture layers are known to improve the classification accuracy by up to 10–15%. The study also advises selective application of texture for specific classes in hierarchical classification systems as often not all classes are better separable based on texture information. The multitude of available texture variables can be computed at different directions and window sizes (Ghosh and Joshi, 2014; Kim et al., 2009; Pu and Landry, 2012). Owing to the multiscale perspective of texture, the optimum window size varies for example with the crown diameter of a specific tree.

At stand level, texture information is mainly related to variability in the stand structure (Franklin et al., 2000) which is amongst others driven by species composition. For example Mallinis et al. (2008) found texture information useful to differentiate between oak and deciduous/coniferous mixed stands.

On single-tree level, some textural measures have been explicitly developed for tree species classification tasks. For example, applying very high resolution aerial imagery, Zhang and Hu (2012) reported overall accuracy improvements from ≤75% to 84.5% and 86.1% when including longitudinal crown profiles to a decision tree classifier in an urban environment. Zhang and Hu (2012) argue that due to the low competition in urban environments, trees are more likely to develop

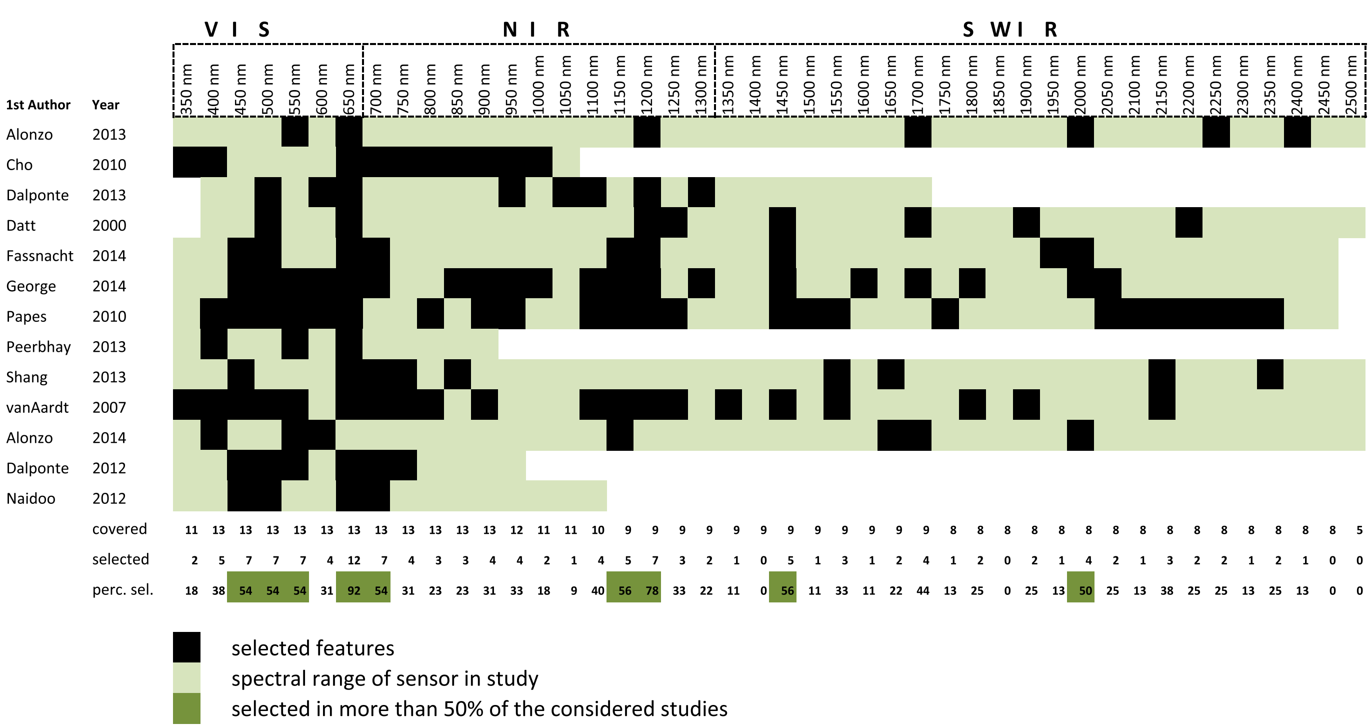

Fig. 4. Summary of important wavelength regions as identified by 13 studies making use of hyperspectral data and feature selection approaches. Covered = number of studies that covered the wavelength region; perc.sel. = number (percentage) of studies that selected the wavelength region as being relevant for tree species classification.

their genetically determined size and shape. Therefore, the shape of trees is likely to be species-specific in urban areas.

4.1.3. Phenology

A further useful trait for species discrimination is phenology. Phenology embraces very obvious processes such as the coloring of leaves in deciduous temperate forests in autumn due to leaf senescence (related mainly to the faster decomposition of chlorophyll pigments in comparison to Anthocyanins and Carotenoids), the intense green colors of fresh leaves and needles in spring time as well as flowering events. Since phenology varies with species, species-specific knowledge of phenology is preferable over broad knowledge of forest phenology (Chuine and Beaubien, 2001). It is therefore desirable to align the time of image acquisition with the phenological cycle of the species under investigation (Gärtner et al., 2016).

Concerning the frequency and the ideal time-point of image acquisitions to capture phenology-related information, two studies using airborne and spaceborne imagery suggested that a combination of two images, one during spring time (green-up) and one during autumn (senescence) supports high classification accuracies (Hill et al., 2010; Mickelson et al., 1998). Adding more images was not found to be of high value. Furthermore, it was recommended to avoid late fall imagery after most of the leaves have fallen (Wolter et al., 1995). In this stage influences from the background will increase and the separation of broadleaved species will become more problematic (Cho et al., 2012). Voss and Sugumaran (2008) also reported no improvements in overall accuracy when applying hyperspectral data from fall as compared to a summer dataset. Hesketh and Sanchez-Azofeifa (2012) found notable differences in leaf spectra of tree species collected in two seasons. While with both single datasets high classification accuracies were achieved, classifying tree species across seasons was not possible. Main driver for the differences across seasons were varying water and pigment contents. For broadleaved forests, variations in spectral properties due to seasonal changes have been observed both within (Blackburn and Milton, 1995) and between the tree species (Key et al., 2001). For example, the rate and nature of change in brightness and greenness of sugar maple (from May to October) differ from that of aspens. This acts as key information to discriminate the two species (Dymond et al., 2002). Combination of multiple phenological periods has produced very detailed classification results (Mickelson et al., 1998; Wolter et al., 1995).

Leveraging phenological information for tree species classification can become challenging if areas with larger ecological gradients are targeted. For example, Wolter et al. (1995) reported that the temperature buffering effect caused by large water bodies (in this case Lake superior in the USA) can cause senescence gradients in the study area which can hamper the classification. Similar effects can be expected for altitudinal gradients. Furthermore, depending on the applied data type, costs for multiple acquisitions might become a restriction of phenological approaches.

4.1.4. Ecotypes, site conditions and leaf age

Concerning the stability of spectral signatures of trees over larger geographic areas, subtle differences in canopy structure or pigment composition have been observed for different ecotypes (defined as populations of the same species that have genetically adapted to differing ecological conditions) and plasticity types (individuals with differences due to differing ecological conditions but with identical genetics) of the same species. These differences are mostly related to variable site conditions (Dinuls et al., 2012). Hesketh and Sanchez-Azofeifa (2012) provide a detailed summary of earlier research exploring the influence of spatial and temporal dynamics on reflectance properties of trees. They mention studies focusing on the relation between leaf reflectance and leaf age (Roberts et al., 1997) as well as on reflectance differences between the same species at different locations (Castro-Esau et al., 2006; Martin et al., 2007; Sanchez-Azofeifa et al., 2009).

When focusing on large geographic extents, some of this additional variability could be accounted for by integrating topo-climatic variables to the classification approach. So far opposing results were obtained when integrating this kind of data. While Engler et al. (2013) reported limited benefit of topo-climatic information to improve remote sensing-based models from a study site of 200 km^2, Zimmermann et al. (2007) reported that topo-climatic variables are the most important predictors for species distribution modeling, while remote sensing only adds limited information (study site with 60,000 km^2). As also discussed with more detail in Engler et al. (2013), these contrasting results can be explained by the differing scales of the two studies.

4.2. Mid-infrared and thermal-infrared sensors

As outlined in Section 4.1, it is hardly feasible to identify species-specific absorption features in the VIS-SWIR region (Asner, 1998). In the mid-Infrared (MIR) and thermal infrared (TIR) part of the spectrum contrasting observations have been made. Salisbury (1986) presented leaf-level TIR spectra of beech (*Fagus grandifolia*), red oak (*Quercus rubra*) and two cherry species (*Prunus* sp.) and identified well-defined spectral features that differed notably across the four species. Salisbury (1986, p. 1884) hypothesized that the observed features are related to "strong hydrocarbon bands of the waxy cuticle superimposed on strong continuum absorption of water". As cuticles are unique for every plant species (Holloway, 1982) this observation seems to be very promising for species discrimination. Furthermore, Salisbury and Milton (1988) obtained close-range thermal reflectance measurements of several other species and also reported differences in the spectra of most of the investigated species. In a more recent study Ribeiro da Luz and Crowley (2007) found that the TIR signal associates with several plant chemical and structural compounds such as cellulose, silica, xylan and oleanolic acid. These compounds are part of the plant tissue that "form the external surfaces, namely, the cuticular membrane and the outer cell wall" (Ribeiro da Luz and Crowley, 2010, p. 404). Ribeiro da Luz and Crowley (2007) again point out that the signal in the TIR domain is much more species-specific than the reflectance signal observed in the VIS-SWIR region. They additionally mention the potential influence of secondary plant metabolites on the TIR. Many plants have unique strategies to develop chemical compounds to (1) fight being fed by herbivores and (2) attract potential pollinators. These chemical compounds might help to define species-specific TIR signatures. First indications strengthening this assumption have been provided by Ullah et al. (2012) who searched for important wavelength regions in the MIR and TIR to separate plant species on leaf-scale with the help of a genetic algorithm. One of the identified wavelengths region at 11.5 μm has earlier been suggested to relate to different aromatic compounds (Ribeiro da Luz, 2006) which form part of the secondary plant metabolites.

4.3. LiDAR

LiDAR data can provide a range of features related mainly to the structure of trees. While the geometric part of LiDAR information relates to the architecture of crowns, branching, and foliage (Riaño et al., 2004, Coops et al., 2007), the intensity of the backscattered signal is additionally connected to foliage type (Suratno et al., 2009), leaf size, leaf orientation, leaf clumping and foliage density (Korpela et al., 2010a; Kim et al., 2009). These properties can all vary within and between tree species and are at least partly complementary to the data gathered by passive optical remote sensing sensors (Alonzo et al., 2014).

4.3.1. Height information

LiDAR derived height information alone is of limited value for tree species separation as discussed in Section 5.1. However, the separation of broadleaved and coniferous trees can be accomplished in dense stands with height information from a multi-temporal (i.e. leaf-on and leaf-off) LiDAR dataset (Reitberger et al., 2008). This multi-temporal approach entails recording two LiDAR datasets at the same site under leaf-off and leaf-on conditions (e.g. Liang et al., 2007). A straightforward comparison between the canopy height models (CHMs) of the two acquisitions allows for the identification of the broadleaved stands as the CHM will show notably lower values for the leaf-off acquisition as compared to the leaf-on acquisition. This is due to the missing foliage and a thereby notably higher number of LiDAR returns on the ground and the stems which decreases the average height of the canopy surface model (Kim et al., 2009; Wasser et al., 2013). However, deciduous conifers (i.e. conifers which defoliate in autumn) cannot be identified as conifers with this approach. Improvements of tree species classification using combined leaf-on and leaf-off data were also reported by Kim et al. (2009). Furthermore, they found leaf-off data more valuable than leaf-on data when both datasets were examined independently, although these differences might have been a result of the differing sensor systems.

4.3.2. Other geometric features

At the single tree level, the application of more complex geometric features was found to be of value for species separation. For example, Holmgren et al. (2008) stated that the relative crown base height is an effective measure when separating scots pine from Norway spruce and a group of broadleaved species (mainly birch), with the rationale that scots pines normally feature a higher crown base height than other species. This is connected to the typical habitus of light demanding pine species which often show a comparably short crown on the top of a high stem. However, Holmgren et al. (2008) also stated that the relative crown base height is highly varied within a forest stand by factors such as stem density, site quality, and forest management practices. Further motivation for using structural features arises from the assumption that features like crown shape (for example conical shape of spruce trees compared to rounded shapes for many broadleaved trees) and the allocation of biomass in the crown (broadleaved trees tend to have more biomass in the upper crown parts) are species-dependent (Ørka et al., 2009). However, Ørka et al. (2009) pointed out that any LiDAR feature describing crown properties may depend on the tree height and the presence of different height distributions across species may cause problems. The authors discuss previously proposed normalization procedures (e.g., Brandtberg, 2007) as one option to reach height-invariant LiDAR features but also state that the suitability of such approaches still needs to be proven.

4.3.3. Return or echo type

Another family of LiDAR features that was found to be beneficial for species separation is related to return or echo types. For example, Holmgren et al. (2008) stated that the proportion of single laser returns were generally higher for coniferous trees which often feature a denser canopy than broadleaved trees in boreal forests. Similarly, Suratno et al. (2009) stated that the proportion of first returns is helpful for species identification. Ørka et al. (2009) also discuss the relationships between crown shapes and return types. They argue that in birch trees (*Betula pendula*; *Betula pubescens*) more first and single returns are observed due to denser foliage in the upper part of the crown, while spruce trees show a higher proportion of last returns due to their lower crown base. This contradicts the statement of Holmgren et al. (2008) but is based on the examination of only two species. One reason for the contradictory observations may be that the described interactions also depend on the footprint diameter and the signal-to-noise ratio of the sensor system.

4.3.4. LiDAR intensity

The probably most promising group of LiDAR features for tree species classification are those related to the recorded intensity (e.g. Kim, 2007; Korpela et al., 2010a; Hovi et al., 2016). The definitions of intensity of the LiDAR signal in the literature are ambiguous and often the exact algorithm leading to the applied intensity values is also unknown (Hovi et al., 2016). Of the 17 reviewed studies using intensity data only 9 provided a definition. These definitions were all related to the amount of reflected energy at the peak amplitude of the returned signal (e.g., Heinzel and Koch, 2011; Hovi et al., 2016; Korpela et al., 2010a, 2010b) or the general return signal strength (Kim et al., 2009) further defined as "total backscattered radiation from scatterers along the pulse path" by Hovi et al., (2016, p. 235). Intensity-related features were found to be amongst the most relevant predictors in numerous studies (e.g., Heinzel and Koch, 2011; Holmgren and Persson, 2004; Reitberger et al., 2008). LiDAR sensors typically record intensity information in the NIR-SWIR regions. From the reviewed studies, 19 out of 27 relevant cases applied a sensor working at 1064 nm, 6 used a sensor working at 1550 nm and in three cases the wavelength was not evident from the data descriptions.

According to Vauhkonen et al. (2014) coniferous trees have lower average intensity values compared to most broadleaved trees. This is in accordance with Korpela et al. (2010a, 2010b) and Ørka et al. (2009) who found that birch had generally higher intensity values compared to spruce and pine. Yet, Korpela et al. (2010a, 2010b) stated that the intensities from birches also greatly depend on tree size (crown and height). These intensity differences are mainly caused by the differing structures of broadleaved trees featuring larger single leaves with continuous surfaces and coniferous needle leaves with a continuous interruption of the leaf surface which is caused by the linear structure of the needles (Kim et al., 2009). For coniferous species, Kim et al. (2009) observed higher intensity values for species with single needles as compared to species with clumped needles with the assumption being that species with clustered needles have a higher proportion of exposed branches (and branches having a lower reflection in the NIR compared to foliage). However, it must be considered that intensity is not only a function of the reflectivity of the surface, but it also depends on issues like scattering effects, which are related to the architecture of a given tree (including also leaf size, leaf orientation, foliage density and branches) (Korpela et al., 2010a). The architecture of a tree, as measured by intensity, in turn, can be partly explained by its age and the fertility of the site (Korpela et al., 2010a). Furthermore, photons emitted by each beam and returned to the sensor from the canopy are often affected by multiple scattering (interaction of photons of a laser beam with more than one object within the canopy). Multiple scattering can falsify the return laser signals in peak-amplitude range detection systems as they appear to stem from lower parts of the canopy. In addition and according to Kotchenova et al. (2003) multiple-scattering might enhance the signal magnitude, especially at lower parts of the canopy and in large footprint scanners. One way to reduce such scattering effects in discrete return systems could be to focus on intensity values of echoes recorded as first or single return/echo where the scattering effects are typically less pronounced. Ørka et al. (2009) observed improved accuracies for such an approach at least for large trees, while they hardly found any benefit from using intensity information for the species classification of smaller trees. Intensity measures calculated from first or single return echoes were also recommended by Hovi et al. (2016) and Korpela et al. (2010a), though the intensity values in discrete return data were shown to be affected by the applied range detection algorithm (Constant fraction discriminator). For sparse density LiDAR the advantage of using first or only echoes might not apply as contrasting results were reported by Suratno et al. (2009) who obtained better results when using all returns. All in all, factors such as bidirectional reflectance and the geometry of the volumetric target surfaces are amongst the most influential factors on the LiDAR intensity values. The forest canopy is an extreme case for volumetric and heterogeneously-shaped target objects that shapes numerous variations creating (and exacerbating) the bidirectional reflectance, and has thus been stated as a main source of uncertainty when dealing with intensity data (Höfle et al., 2008).

Furthermore, 'raw' intensity values are affected by the acquisition geometry. A range normalization of the 'raw' intensity values based on the distance between the sensor and the reflecting object was proposed to improve species classification (Korpela et al., 2009; Korpela et al., 2010a, 2010b). However, one critical point during the normalization is that the choice of best optimization parameters is species dependent (Korpela et al., 2010a, 2010b). This has been theoretically proven by Wagner et al. (2008) and practically shown by Gatziolis (2011). Furthermore, it has been reported that the optimal normalization parameters also depend on the return/echo type (Korpela et al., 2010a, 2010b). In addition, the range variation (given fixed transmitted power) ultimately causes the signal-to-noise ratio to fluctuate, thus influencing the echo numbers and types (Korpela et al., 2013, Wagner et al., 2006). This makes an optimal normalization for tree species classification tasks quasi impossible, especially in mountainous terrain (Korpela et al., 2010a, 2010b). Nevertheless, Korpela et al. (2010a, 2010b) estimated an expected 2–3% gain in accuracy when applying a normalization of the intensity values under boreal conditions with three species. Intensity measures (~peak WF amplitude) can also be affected by the applied sensor system and Korpela et al. (2010a, p. 335) reported "slightly improved" classification results when weighting was applied to the intensity data to normalize the scales of the intensity observations of two separate LiDAR instruments. It is worth noting that the intensity normalization requires a radiometric calibration of the intensity values that links instantaneous at-sensor power with the recorded intensity. In the study of Korpela et al. (2010a, 2010b) this was the case only for one of the two tested datasets.

4.3.5. Waveform LiDAR

Most points discussed so far are relevant for both discrete echo recording and pulsed waveform LiDAR systems. For tree species classification, a number of studies have reported improvements in accuracy when applying waveform datasets (e.g., Hovi et al., 2016, Reitberger et al., 2008, Vaughn et al., 2012, Yu et al., 2014). One reported reason for the improvement are the increased number of echoes that can be extracted from the waveform recordings in comparison to typical discrete return data acquisitions (Reitberger et al., 2008). These denser point clouds were reported to for example enhance the discrimination amongst different coniferous species (Vaughn et al., 2012). However, converting waveform signals to discrete points is likely to generalize the detailed information contained in the waveforms which is at least partly connected to species information (see results of Vaughn et al., 2012). For example waveform features such as the mean and standard deviation of the echo width within single laser beams have been stated to be amongst essential variables when classifying broadleaved and coniferous tree groups from waveform recording (Hollaus et al., 2009; Reitberger et al., 2008). Yu et al., (2014, p. 1026) stated that the "vertical distribution of FWF (full-waveform) information" contains species-specific characteristics that do not exist in discrete return data.

A recent study of Hovi et al. (2016) focused on a systematic analysis of the discrimination potential of waveform features by analyzing the sources behind the within-species variation. They considered phenology, age, site type and scan angle as well as the tree effect (a variable that describes the variance that is created by the individual structure of the tree itself; compare Section 5.7) as potential sources. Besides smaller influences of phenology (comparison of several leaf-on acquisitions) they mainly found that the tree effect explained up to 65% of intra-species variance. Furthermore, they reported that all waveform features depended on tree size. Therefore, stratification into size classes was recommended.

4.4. SAR

Most SAR studies focused on the discrimination of broad forest types in the framework of land-cover classification omitting the species level. Forest information by SAR relates mainly to canopy structure and water content. These traits depend on the frequency or wavelength of the SAR system, the look angle and direction, the spatial resolution as well as on the polarization modes. Furthermore, environmental variables such as terrain and weather conditions have an influence on the Radar information. However, methods to account for the grazing angle and topography can be incorporated in the pre-processing of the SAR data (e.g., Ranson et al., 2001). In this context, Ortiz et al. (2012) reported a notable influence of the applied digital elevation model on the obtained classification accuracies.

4.4.1. SAR bands (wavelength)

SAR systems are specified by the applied wavelength (λ). The penetration depth and the photon pathway through the canopy depend on λ. Shorter wavelengths (e.g. X-band, $\lambda = 4$–2.5 cm) interact mainly with the top of the canopy and are influenced by the leaf orientation (Leckie and Ranson, 1998), while C-band systems ($\lambda = 4$–7.5 cm) penetrate deeper into the canopy (Rignot et al., 1994). Longer wavelengths (e.g. L-band; $\lambda = 15$–30 cm) penetrate through crowns and yield more backscattering from large branches, trunks or even the ground. In-depth information on these interactions is summarized in Leckie and Ranson (1998).

Concerning the selection of an appropriate SAR-band, there is some empirical evidence that shorter wavelengths have a higher potential than longer wavelengths for classifying tree species. Rignot et al. (1994) stated that C-band data is more useful than L- or P-bands (P-band: $\lambda =$ ca. 70–110 cm) when applying AirSAR data over a boreal forest. The branching geometry and foliage of certain tree species (e.g., white spruce and leafless balsam poplar) varies notably affecting the C-band SAR signal and facilitating the separation of the species. Similar statements concerning the potential of shorter wavelength systems were made for X-band systems (Ranson et al., 1995). In addition, Wollersheim et al. (2011) mention that C-band systems have lower sensitivity to effects caused by slanted terrain as compared to L- and P-band systems. The higher abilities of shorter wavelengths have also been confirmed by the studies of Hoekman (1985) and Pierce et al. (1998) who found good and notably improved results, respectively, when applying X-band data. However, Ortiz et al. (2012) reported only moderate results when trying to separate broadleaved and coniferous species using X-band TerraSAR-X data. Furthermore, Saatchi and Rignot (1997) reported P and L bands to be of high relevance for discriminating jack pine and black spruce as well as coniferous from broadleaved species which contradicts the findings of Rignot et al. (1994). The choice of an appropriate band presumably also depends on the examined forest as the structure of a forest varies widely with age, management and biome (e.g., multi-layered tropical rainforest with huge individual trees as compared to more homogeneous boreal forest with comparably small trees). Furthermore, phenology might influence the suitability of a band. Ortiz et al. (2012) reported slightly better accuracies for X-band SAR images acquired under leaf-off conditions compared to leaf-on data. Notable improvements were reported for the use of the combined dataset.

4.4.2. SAR polarization information

Canopy structure is not only important for the differences in backscatter signal obtained by different Radar bands but also for the additional information obtained from differing polarizations. The relationship between polarized signal and the observed target is often complex, but the dependency of the polarized signal on the scattering behavior of the observed target (either specular, diffuse or volume scatterer) is established. Vegetation areas and especially forests are mainly volumetric scatterers that tend to depolarize the signal due to multiple scattering in the canopy (Morain and Simonett, 1967). That is, some of the incident polarized power is reflected as un-polarized power. This depolarization is a complex process that can only be retraced if reliable information on characteristics such as the size of the leaves, diameter of boles and branches, foliage density, the height of the vegetation as well as presence or absence of understory vegetation is available (Trevett, 1986). Scattering behavior also varies with incident angle and wavelength which makes the whole system even more complex. The potential usefulness of polarization information to separate broad forest types has been already mentioned in the early study of Knowlton and Hoffer (1981) who reported brightness differences between coniferous and broadleaved trees in the HH image of an X-band system which were not apparent in the HV image.

The application of advanced polarimetric measures to separate tree species has been investigated in a number of studies. A review on the application of polarimetry for species identification in forests can be found in Boerner et al. (1998). In our review we mainly focused on studies that were not covered in this earlier work. Pope et al. (1994) applied four polarimetry-based biophysical indices including a volume scattering index, a canopy structure index (CSI), a biomass index (BMI) and an interaction type index (ITI) for P, L and C band data. They found most of the indices relevant for separating certain land-cover classes (including forested and non-forested areas and some forest species-related classes) in a mixed landscape in Belize, Central America. The CSI and ITI of P-band were found to be especially valuable to separate different swamp forest types and the BMI of the P-band was able to discriminate swamp forests from upland forests. For the separation between different upland forest types, they found ITI of C-band to be most suitable. Wollersheim et al. (2011) applied a total number of 61 polarimetric SAR variables in a two-stage classification approach where they first grouped stands into pine species, spruce species and broadleaved classes and then attempted to classify each broad class into individual species. The most informative polarimetric SAR variables included two circular polarization ratios, Touzi's degree of polarization variables and the HHVV coherence and pedestal characteristics. Furthermore, the authors found a large importance of variance and texture variables that had not been frequently applied in former studies. Further potentially important polarimetric SAR variables are listed in the study of Maghsoudi et al. (2012) who combined a feature selection approach with a Wishart classifier and compared it to a Support Vector Machines (SVM) classifier to classify seven tree species and two additional classes.

4.4.3. Water and SAR information

Besides the size and arrangement of forest stand and canopy components, the water content and surface moisture of trees and the surface background contribute to the Radar signal by affecting the dielectric constant. The dielectric constant of objects varies with their material and temperature. Most natural materials have a dielectric constant between 3 and 8 while water reaches values of approximately 80 (Lewis et al., 1998) which, however, drops to 3 if water is present in ice-form. Thus, it has been observed that under frozen conditions with a very low dielectric constant the interaction of the Radar signal with the canopy are notably reduced (Kwok et al., 1994). Ranson and Sun (1994, p. 151) even reported that the needles were "nearly transparent" for the Radar signal.

Due to this extraordinary property of water, the moisture content in vegetation can help to separate species with differing amounts of water in the crown layer (for example dense Beech crowns versus sparse Pine crowns). However, this signal may become contaminated by other short term effects such as rainfall or drought events. Dobson et al. (1991) reported notable changes in the Radar signal if a Radar dataset was acquired shortly after a rain event compared to scenes acquired three days later. Ortiz et al. (2012) hypothesized that imagery obtained shortly after a rainfall event may have higher potential to separate broad forest type classes, since the amount of interception depends on the leaf-type and the effect of rainwater tends to be larger for broadleaved compared to coniferous forests.

5. Methods for tree species classification

5.1. Reference data

The selection of reference or training and validation data in an experimental research has to fulfill certain criteria to ensure that they can solve the question under investigation. That is, (1) the considered classes have to match the research question; (2) the data should be representative for the site under investigation; (3) the spatial scale of the data should match the problem under investigation; (4) the data should acknowledge the underlying assumptions of the applied methodology (e.g., minimum number of samples per class, balanced number of samples across classes, etc.); (5) observation errors should be known and their impact on the results should be discussed; (6) its samples should be spatially independent.

Criterion (1) was often simple to achieve in the reviewed studies as the objective was in almost all cases to develop an approach to classify major tree species without any further focus on certain key species or an ecological (e.g., trees that serve as habitat for endangered bird species) or economical (high-value species) context which might have complicated the experiment.

Criterion (2) was presumably not achieved in all studies. Mentioned criteria for selecting certain items as reference data were: visual detectability in the remote sensing data (canopy-emergent trees, dominating trees, large and isolated trees etc., e.g. Clark and Roberts, 2012); sufficient number of samples for a species (e.g., Clark et al., 2005); accessibility in the field (close to roads, e.g. Jensen et al., 2012); homogeneous surrounding (pure stands, e.g. van Aardt and Wynne, 2007); and minimization of background signal (dense stands or crowns, e.g. Youngentob et al., 2011). In some studies criteria concerning terrain (avoidance of steep slopes, Franklin et al., 2000); ecological gradients (soil water balance, soil fertility, growing degree days, e.g. Engler et al., 2013); and different tree ages (Yu et al., 2014) were also considered for selecting adequate reference data.

Many of the criteria applied in the reviewed studies lead to an improved classification result as the within-species variance is reduced (e.g. by minimizing background or terrain effects or by only considering dense crowns) and in some cases only a subset of the full population of tree species is considered in the experiment. Only few studies applied criteria to increase the representativeness of the dataset by considering e.g., ecological gradients or tree age (e.g., Engler et al., 2013). Therefore, it is likely that in many studies the selected reference data are not completely representative for the full population of trees in the study area and reported accuracies are hence optimistic. This is particularly critical as in many studies the missing representativeness of the reference data is not further discussed when reflecting the obtained results (e.g., Ghosh et al., 2014, Fassnacht et al., 2014).

Depending on the overall number of samples available in the dataset, one straightforward way to increase the representativeness of the dataset while avoiding increased intra-class variability could be the definition of several reference classes per species (e.g., spruce on shaded and sunlit slopes). The use of such internal sub-classes was found to be useful for example by Leckie et al. (2005). The classes can then again be merged in a post-classification step. Another issue was that reference samples were often biased towards dominant trees. That is, any reported accuracies might be representative for a larger part of the wood volume of a forest, but not necessarily for the larger fraction of individual trees, as most trees in the understory are not even considered in remote sensing approaches. The few approaches that attempted to classify understory species confirmed that this a very challenging task (e.g., Korpela et al., 2012; Korpela et al., 2008). Ørka et al. (2009) also observed that the classification of smaller trees with LiDAR data is especially challenging. They suggested that area-based approaches might be more appropriate than single-tree approaches for classifying smaller trees.

Concerning criterion (3), the reviewed studies showed that single tree objects were the most commonly applied observation unit, followed by individual pixels (corresponding to various scale levels depending on the spatial resolution of the sensor) and fewer studies in which stand objects or spectra collected on leaf or sub-crown scale were applied. The scale and the selected observation units typically agreed well with the spatial resolution of the applied data. One reason for this was that in most studies the training and validation data for a supervised classification were sampled from the same pool of data being used as classification input. One exception were combined classification approaches, in which often LiDAR data were used to identify the observation unit (tree crowns) and then passive optical data were added for the classification. However, the LiDAR and passive optical datasets used in the combined studies typically matched well in scale. Another exception was a study applying thermal data from a spectral library to identify tree species in an independent image (Ribeiro da Luz and Crowley, 2010). While field spectrometer measurements are often incomparable to measurements of airborne sensors in hyperspectral datasets (e.g., Chance et al., 2016; Ribeiro da Luz and Crowley, 2010) mainly due to the variability of the canopy structure affecting the VIS-SWIR (Roberts et al., 2004), it seems to be feasible in the thermal Infrared domain. The successful detection of >20 species from an airborne TIR image with the help of a spectral library containing spectra measured under laboratory conditions seems to confirm that a notable number of plant species indeed have a unique spectral signature in the MIR and TIR (compare Section 4.2).

Criterion (4) was considered in most of the studies and species with too few samples were normally excluded from the classification. Nevertheless, the number of samples per species was in some cases very low which can be challenging, particularly in presence of a large number of predictor variables (see Sections 5.2.2 and 5.4).

Although criterion (5) seems to be trivial, it can be hard to achieve in remote sensing of forests. In the reviewed studies, a re-occurring procedure was to combine samples gathered from field surveys or forest inventory information with additional samples obtained from photo-interpretation of aerial photographs or satellite imagery (e.g., Immitzer et al., 2012a). Although the success-rate of a photo-interpretation expert to identify major tree species is high (e.g., Heller et al., 1964), a certain level of error still has to be expected. The scale of the imagery also affects the success rate. Further sources of error include wrong species identification in the field, the misalignment of field-visited trees with remote sensing data, overlapping tree crowns and possible time lags between field surveys and image acquisitions. The identification of homogeneous areas stocked with a single species and covering an area of at least 3×3 pixels (to ensure at least one pure center-pixel) can be troublesome or even impossible depending on the spatial resolution of the applied sensor and depending on how high the degree of mixture is in the examined forest. As a consequence, some studies only required that at least a certain percentage of the area of a reference sample is stocked with the target species (Boschetti et al., 2007; Adelabu et al., 2013). A further possibility is to account for the often non-discrete nature of forests by applying fuzzy validation strategies (e.g., Mickelson et al., 1998) as also further discussed in Section 5.2.

Criterion (6) refers to problems arising from spatial auto-correlation in the reference data. Following the general assumption that adjacent pixels or objects of a class are more similar, it is likely that the classification performance measures will be positively biased if many samples of a class (e.g., trees) are clustered in a small area (e.g., a single stand). Therefore, problems with spatial auto-correlation can be reduced if the samples of each target class are well distributed over the whole study area and keep a certain minimum distance from each other. In managed forests, maximizing the number of sampled stands in the study area increases the representativeness of the reference samples. If this is not possible for some reason, the issue should be clearly stated and validation strategies that consider the location of the reference data during the data-splitting are recommended (compare Section 5.2.2).

5.2. Calibration and validation of supervised classification algorithms

5.2.1. Definition of a successful result

The definition of optimum classification accuracy varies with the viewpoint of the user or the application (Foody, 2002). As an example, it might be interesting from a nature conservation perspective to detect individual trees of a rare species serving as habitat for an endangered bird species. In this case, mainly the accurate detection of the target species (presumably making up only a small proportion of the considered samples) is relevant. On the other hand, from the viewpoint of a forest enterprise, the detection of rare species might be less relevant, as they only play a minor economic role.

From the studies reviewed here, only three defined a targeted accuracy level (Goodenough et al., 2003; Heikkinen et al., 2010; Korpela et al., 2010a, 2010b). These targeted minimum overall accuracies varied between 80% (forest type mapping in the Pacific Northwest of the USA) and 95% (Finland for single tree remote sensing of three species). This is approximately in line with what Thomlinson et al. (1999) suggested for land-cover maps. Their minimum accuracy requirements were 85% overall accuracy with no class-specific accuracy below 70%. Target accuracies are obviously also connected to the biomes and their species diversity, where in biomes with more species lower accuracies might be satisfactory than in biomes containing only few species.

In summary, finding a general definition for minimum classification accuracies for tree species classifications is impossible and always depends on where the study takes place and what the objectives of the study are. Except for some studies in North America and Scandinavia that referred to concrete forest inventory tasks (e.g., Jones et al., 2010; Ørka et al., 2013), hardly any study mentioned a concrete application for the created product but were referring to general motivation points as summarized in Section 1.1.

5.2.2. Independent validation strategies

Many recent classification algorithms can reach high classification accuracies during the training phase by strongly adapting their discrimination rules to a given dataset. As the feature space in a supervised classification is enlarged, for example by adding numerous LiDAR and hyperspectral variables, the risk of overfitting the classification model increases. The situation is particularly exacerbated in case of small training sets (Dalponte et al., 2014). Another problem associated with the use of small training data sets and numerous predictor variables is the Hughes effect that will be discussed in Section 5.4.

To avoid overfitting of the classification model, a commonly applied method is to split the initial reference dataset into training and validation samples. This data-splitting can be accomplished in various manners. Amongst the most popular approaches are (1) simple data splitting (often 70% of the samples for training and 30% of the samples as validation); (2) x-fold cross-validation (samples are randomly split into x parts (folds) of equal sample size, then x classifications are run in which each of the x parts serves as validation sample once; leave-one-out (LOO) cross-validation is a special case of x-fold cross-validation where x = number of samples); (3) bootstrap-resampling (the n reference samples get sampled n times with replacement; some samples get selected several times, other samples do not get selected at all; the non-selected samples are then used as validation set; such a strategy is also applied in the internal out-of-bag validation of random forest (RF)). All those data-splitting options have certain statistical properties with associated prediction tendencies that are discussed in the corresponding literature (for example Kuhn and Johnson, 2013). Approaches (1) and (2) can furthermore be enhanced by splitting the datasets based on their location. Such “leave-one-side-out” or “leave-one-stand-out” approaches are typically more conservative as they are sensitive to non-stationarity effects in the dataset (e.g., Engler et al., 2013) which also relates to criterion (6) of Section 5.1.

From the 129 reviewed cases, 101 applied some type of simple data-splitting. Of these 38 used a cross-validation (including LOO), 14 used bootstrap-resampling while the remaining cases mostly used simple splits into training and validation data. The other cases did not specify their validation strategy or reported the training accuracies without independent validation.

All reported data-splitting approaches are likely to obtain meaningful classification accuracies if the reference dataset is sufficiently large and the requirements concerning reference samples described in Section 5.1 are fulfilled. If the reference dataset is comparably small, accuracies obtained from a single split into training and validation samples often do not fully represent the classification accuracy of the obtained map but vary notably depending on the composition of the training and validation samples. By applying iterative data-splitting, several studies using passive optical data have shown that a variation in overall accuracy of 5–10% (depending on which samples are used as training and which as validation samples) is rather standard than the exception in the presence of a comparably small number of samples (Fassnacht et al., 2014; Gougeon, 1995; Ghosh et al., 2014). Such iterative procedures were applied in only 18 of the reviewed cases (excluding cross-validation and internal out-of-bag validation of RF).

In conclusion, we recommend the application of an iterative data-splitting approach and an additional completely independent test set as a gold standard for tree species classification studies. While the iterative data-splitting enables reporting of prediction variance of all accuracy measures, the completely independent test set could, in an ideal case, be composed from a few fully sampled plots that enable not only a sample-based but also a direct wall-to-wall comparison between reference and classification maps. To avoid biases from spatial auto-correlation induced for example by similar stand-histories these fully sampled plots should be selected in a stratified-random sampling to ensure that the existing variation within the study area is fully represented. While the acquisition of fully sampled plots might not be feasible in all studies due to restricted resources, a sufficient amount of reference samples to enable iterative data-splitting and a completely independent sample-based test set should always be pursued. Furthermore, it should be considered that some machine learning classifiers such as SVM or RF might even require an additional set of reference samples to accomplish a sound parameter tuning.

Finally, it is worthwhile mentioning that in some cases the application of fuzzy accuracy assessment (which allows each pixel to have multiple and partial class memberships) can deliver more meaningful accuracy estimates than the traditional accuracy assessment based on discrete classes (Foody, 2002). As an example: a confusion between different pine-species might be considered less critical in a classification than the confusion between a pine and a beech.

5.3. Single tree detection

The detection of single tree crowns with automated algorithms is an own research field that has been reviewed before (e.g., Koch et al., 2013; Brandtberg and Warner, 2006). In the reviewed studies, single tree delineation was either performed manually (25 out of 65 relevant cases) or with automated algorithms that were often based on earlier studies (40 out of 65 relevant cases). In few cases, semi-automatic approaches have been presented. For example Korpela et al. (2011) manually identified tree tops and then applied an automated algorithm to define the corresponding crown regions. It is important to stress that the delineation quality has a direct influence on subsequent tree species classifications. Accurate single-tree delineations should lead to improved spectral signatures (reduced number of pixels outside of the actual crown) and thereby help to reduce the intra-species variability of the remote sensing signal when calculating averaged spectra for crown objects (see Immitzer et al., 2012a). Accurate delineations are also a requirement for majority voting approaches which have been used in a number of studies (e.g., Alonzo et al., 2014; Kim et al., 2011; Meyer et al., 1996). One suggestion to increase the applicability is to use manually delineated crowns in the training stage followed by an application of the trained algorithm to crown objects automatically delineated from LiDAR data (Dalponte et al., 2014).

LiDAR data was found to be advantageous for single tree delineation as it is less affected by occlusions and shading than passive optical data (Culvenor et al., 1999; Voss and Sugumaran, 2008). Therefore, LiDAR-based delineation algorithms often result in more detected crowns and higher delineation quality compared to approaches based on passive optical data. This was for example confirmed by Dalponte et al. (2014) who compared classification performances based on tree crowns delineated manually, as well as automatically from LiDAR and from hyperspectral data. They achieved best results based on the manually delineated crowns (kappa = 0.89) followed by LiDAR (kappa = 0.79) and hyperspectral (kappa = 0.76). These kappa values were obtained with different numbers of crowns as the crown detection rate differed amongst the approaches (manual = reference; LiDAR 48.5%; hyperspectral only 28.4%).

For waveform LiDAR, Hovi et al. (2016) reported an influence of the applied tree delineation algorithm on the values of the extracted waveform-features. On the contrary the values of the waveform features were found to be comparably stable when extracted from two differing LiDAR datasets, given the same delineation algorithm was applied.

In addition to the studies that examined the influence of the delineation procedure on subsequent processing steps, Heinzel and Koch (2012) presented an approach to improve crown delineation quality by integrating pixel-based species classification maps to post-process the delineated segments. The approach followed the assumption that under-segmentation during the delineation process can be improved by splitting single segments that contain two pixel-clusters of different classes (species) and thereby indicate that the segment contains two trees instead of one.

Finally, new high spatial resolution SAR systems might also enable the classification of SAR data on a single tree level. A prototypical processing chain for the reconstruction of individual broadleaved trees from SAR has recently been presented (Schmitt et al., 2015).

5.4. Feature reduction

The derivation of numerous predictors from passive optical, LiDAR or Radar information has become standard in the remote sensing literature. Typical examples are texture measures, vegetation indices, derivations of the spectral signature, linear transformations such as principal component and Tasseled Cap analysis (Moisen et al., 2006) as well as LiDAR based height quantiles or density related predictors. These numerous predictors raise the methodical problem that a large hyper-dimensional predictor feature space faces a frequently sparse number of reference samples. The gravity of this problem depends partly on the algorithm. In particular, parametric classification algorithms were found to suffer in solving classification problems with a low number of training samples compared to the number of predictors, as the estimation of the within-class covariance matrix is often hampered (Clark et al., 2005) or even impossible. This problem is referred to as the curse of dimensionality (e.g., Dalponte et al., 2013) or the Hughes effect (Hughes, 1968). While today hyper-dimensional feature spaces can be handled from a computational perspective, issues like the bias-variance trade-off (describing the issue of finding a sound balance between fitting the training data well but avoiding overfitting to allow for a general application of the model) and related effects should be considered.

In the reviewed studies, the predictors of hyper-dimensional feature spaces are often reduced by using feature reduction algorithms. Feature reduction algorithms can be subdivided into feature extraction and feature selection methods. The main difference between these two is the former methods select a subset of the original predictor variables, whereas the latter ones calculate new predictor variables that typically summarize the content of several original predictors (Kuhn and Johnson, 2013). In many cases, feature selection algorithms are less efficient than feature extraction algorithms for improving classification performance (see e.g., results of Fassnacht et al., 2014). However, feature selection algorithms have the advantage of enabling a meaningful interpretation of the selected predictors (compare e.g., Fig. 4) and thereby increase the understanding of what exactly drives the discrimination of the species. This is often hard to achieve in the newly composed predictors calculated in a feature extraction approach.

In the reviewed studies, the most common feature extraction methods were principal component analysis (PCA, 5 cases) and minimum noise fraction transformation (MNF, 11 cases). Most frequent feature selection approaches included stepwise procedures (15 cases), manual band selection (10 cases) and RF (4 cases). An overview of additional feature selection approaches with a focus on their implementation for tree species classification is provided by Fassnacht et al. (2014). While the application of feature reduction methods improved classification accuracies in a number of studies (Clark and Roberts, 2012; Fassnacht et al., 2014), other studies also reported declines in accuracies when reducing the number of bands (Dalponte et al., 2013). Fassnacht et al. (2014) performed a systematic comparison of four feature reduction approaches and demonstrated that the combination of a subset of the first components of a MNF transformation and a SVM classifier significantly outperformed classifications based on all available bands. These results were found to be consistent over three datasets gathered with two different sensors from three test sites. The application of MNF components has also led to best accuracies for classifying tree species and vegetation classes in studies of Ghosh et al. (2014) and Zhang and Xie (2012).

5.5. Classification algorithms

Remote sensing-based tree species discrimination and classification evolved with methodological developments in the domain of statistical learning. In early studies the most widely used classification techniques included supervised maximum likelihood classifiers (MLC) and unsupervised clustering (K-means, ISODATA) (Moore and Bauer, 1990; Walsh, 1980). These methods were easily applicable and available in the prevalent image processing software packages at that time (Lu and Weng, 2007).

After 1995, non-parametric decision tree based classifiers and neural networks emerged as an alternative to the other classifiers. These classifiers do not require the input data to be normally distributed (Mallinis et al., 2008; Wolter et al., 1995; Zhang and Qiu, 2012). Some recent studies using mixed sets of input variables (spectral, texture, geometric, indices) have preferred the use of non-parametric machine learning methods like RF or SVM (Immitzer et al., 2012b; Pant et al., 2013). This trend was facilitated by the rapid improvements in computational capacities in hardware and software (e.g. freely available algorithms in open source environments such as R, Python or the Orfeo Toolbox).

Table 1 summarizes the most frequently applied classification algorithms and some of their advantages and disadvantages mentioned in the reviewed studies. It has to be stressed that the mentioned advantages and disadvantages are mainly extracted from the reviewed studies. A sound comparison of all algorithms would require standardized settings and experimental testing that is beyond the scope of this work. From the experience of the authors, the choice of the classifier itself is often of low importance if the data is adequately pre-processed to match the requirements of the classifier. A more detailed review on classification approaches in a remote sensing context is for example provided by Lu and Weng (2007).

5.6. Atmospheric correction

When classifying tree species from remote sensing data, atmospheric correction of passive optical data in the VIS-SWIR range theoretically always becomes important if a study bases on imagery that was collected during a larger time interval (e.g., an airborne campaign that lasted for several hours) or over a large geographic area. In both cases, it cannot be assumed that the illumination conditions and atmospheric conditions were stable for all collected datasets. Therefore, any classification approach will be affected by the increased variance in the imagery.

Table 1
Overview over advantages and disadvantages of most commonly applied classification algorithms.

Classification model # number of cases that used the approach	Advantages	Disadvantages
Parametric approaches		
Discriminant Analysis (linear, quadratical, canonical, stepwise, regularized, penalized) # 34	LDA does not require the tuning of parameters Accepts multiple input variables Easier interpretation of Between-class differences	Assumes Gaussian distribution of training data Classical discriminant analyses are less sensitive to ill-posed problems and outliers Noisy results in complex landscapes Limited ability to deal with multi-colinearity
Maximum Likelihood # 20	Consistent approach for a variety of estimation problems Approx. unbiased in presence of larger sample sizes Many software implementations	Assumes Gaussian distribution of training data Biased for small samples Sensitive to the number of input variables
Spectral Angle Mapper # 5	Suitable if sufficient information on a known spectra is available Insensitive to illumination	Does not consider the sub-pixel value (problematic when working spatial)
Bayesian # 4	Allows incorporating prior knowledge on data Interpretable outputs Inferences which are exact and conditional to the underlying data	Complications on how to select a prior Posterior results are heavily affected by the choice of prior High computational costs, especially when using higher number of input variables
Generalized Linear Model # 4	Suitable for modeling binary response variables Enables using binomial distributions to conduct classifications	Response variables must be independent Variable selection is necessary to avoid the curse of dimensionality
Spectral Mixture Analysis # 3	Ideal for sub-pixel analysis Less sensitive to noise or error in atmospheric correction	Terrain and illumination effects in selected endmembers
Logistic # 3	Probabilistic outputs Relaxed assumptions on the distribution of data Able to use both categorical and continuous input variables Hypothesis testing Handles non-linear effects	Needs high amount of data to achieve stable results Sensitive to outliers Might be subject to overfitting Problems with unbalanced input data
Fuzzy logic # 3	Useful in mixed-class areas More flexible than discrete classifications Possibility to integrate many other decision rules such as decision trees	Problem in how to present the output Subjectivity due to lack of learning mechanism in the algorithm Manual optimization of fuzzy parameters in complex systems is impossible
Thresholding # - (unclear)	Simple to use Also used as basis in other methods such as segmentation	No incorporation of the relationships between pixels Exacerbated when data is noisy, or when shadow or different illumination effects are present
Non parametric approaches		
SVM # 31	No distributional assumption required	Optimal design of a multi-class SVM is

Table 1 (*continued*)

Classification model # number of cases that used the approach	Advantages	Disadvantages
	Suitable when incorporating non-remote sensing variables into classification robust to noise and high-dimensional data fast predictions (sparse model due to support vectors) Comparably few training data needed Possibility to easily access probability values instead of only discrete classes	demanding Comparatively high computational cost (algorithmic complexity) for training Selection of kernel function parameters
Random Forest (RF) # 26	No distributional assumption required Less sensitive to the number of input variables Less sensitive to overfitting Intuitive derivation of accuracy and variable importance	Might overfit in presence of noisy data Might be biased in case response classes have different number of levels
Neural Networks # 3	Able to extract patterns and identify trends from pool of data Implemented by many software packages Suitable to deal with classification problems which are hardly mathematically definable	Difficult to train, since the results are ultimately dependent on the initial parameters Black box-like setup

Atmospheric correction of airborne and satellite-based measurements in the passive optical domain transforms the measured values into physically-based quantities in which the influence of the atmosphere is reduced or in ideal case eliminated. Generally, the process consists of at least three steps: The chip on the remote sensing sensor first measures unit-less Digital Numbers (DN), which base on the radiometric resolution of the sensor (the number of grey levels the chip can separate) and the number of photons of a wavelength-range reaching a given pixel of the chip. In a second step, typically at-sensor radiance (ASR) is obtained with an absolute calibration. That is the measured DN values are corrected with sensor-specific functions that eliminate or at least reduce noise originating from the sensor itself (e.g. related to shutter speed, f-number) (Korpela et al., 2014). According to Schaepman-Strub et al., (2006, p. 29) spectral radiance is defined as "the radiant flux in a beam per unit wavelength and per unit area and solid angle of that beam, and is expressed in the SI units [$W\ m^{-2}\ sr^{-1}\ nm^{-1}$]". In a last step, the ASR values are used as input to atmospheric correction algorithms to obtain reflectance values that are defined as "the ratio of the radiant exitance (M [$W\ m^{-2}$]) with the irradiance (E [$W\ m^{-2}$]) … following the law of energy conservation, the value of the reflectance is in the inclusive interval 0 to 1." (Schaepman-Strub et al., 2006, p. 29). As outlined in the review of Schaepman-Strub et al. (2006) there is a frequent confusion of reflectance terminology in the remote sensing community.

Although many tree species classification studies (58 out of 101 relevant cases) have applied atmospheric correction, only few studies have focused on the effect of atmospheric correction on the obtained results (e.g., Korpela et al., 2011, 2014; Pu et al., 2015). Pu et al. (2015) compared three atmospheric correction methods while classifying seven urban tree species from a WorldView-2 sensor. They found that neither of the atmospheric correction algorithms improved the classification accuracy. Similarly, Korpela et al., (2014, p. 17) conclude that their results "imply that the gain in species identification performance will be low, even if the reflectance calibration of images is accurate". At the same time, both studies point out that these results are also likely to be a consequence of comparably small datasets that were consisting either of a single image acquisition (Pu et al., 2015) or were collected during a relatively short airborne campaign (a 2 h time window in Korpela et al., 2011, 2014). As a consequence, Korpela et al. (2014) state that ASR values can be sufficient if the scene illumination and the atmosphere does not notably change during data acquisition. However, they also stress that for larger flight campaigns calibration to reflectance values might become important as overlapping flight lines might be collected with a several hours' time-gap and the corresponding changes in illumination and atmospheric conditions will affect the images. Then the application of training data collected in one image to classify an image collected under other illumination conditions will be problematic for obvious reasons (e.g., Haara and Haarala, 2002). In summary, these findings agree with Song et al. (2001) who suggested that atmospheric correction of spaceborne Landsat data is in most cases unnecessary when focusing on classification problems where training and image data are obtained from the same single image.

Compared to studies working in the VIS-SWIR region, studies focusing on thermal sensors more strongly emphasized the importance of atmospheric correction. The emissivity features of leaves in the TIR are generally subtle and therefore, besides a possibly high signal-to-noise ratio, sophisticated methods for atmospheric compensation and temperature-emissivity separation are crucial (Ribeiro da Luz and Crowley, 2010). As a consequence, measurements at leaf-level in the laboratory are already critical but situation gets even more challenging with airborne acquisitions on canopy level. Irregular canopies with multiple leaf-angles and shading effects lead to heterogeneous temperature distributions and canopy voids can cause further signal attenuation (Ribeiro da Luz and Crowley, 2007).

5.7. Anisotropy effects

Anisotropy is defined as the property of a natural object or phenomenon being directionally dependent. In case of reflectance, anisotropy means that the brightness of an observed object depends on the view-illumination geometry (Schaepman-Strub et al., 2006). Directional reflectance anisotropy causes variation in the spectral signal observed by sensors in the passive optical domain. Notably changing view-illumination geometries can occur during airborne campaigns with differing flight directions or by repeated satellite acquisitions with pointable sensor systems. Comparably large view-angles (close to 30°) were found to be problematic when classifying tree species from Worldview-2 data due to a notably increased amount of occlusion and in the given case shading (Pu and Landry, 2012) while slight deviation from nadir were hypothesized to be of value due to a decreased influence of soil (Gong et al., 1997). The study of Gong et al. (1997) based on in-situ data but its findings are likely to apply also for airborne and spaceborne sensors.

Variance of the spectral signal caused by anisotropy has been considered both noise and information in forestry related applications (Fassnacht and Koch, 2012). In remote sensing, the anisotropy of reflectance is described with the Bidirectional Reflectance Distribution Function (BRDF) (Nicodemus et al., 1977, Schaepman-Strub et al., 2006). Therefore, many remote sensing studies refer to BRDF effects when they address anisotropic properties of canopies observed in remote sensing data. On canopy level, anisotropy has been examined as an additional source of information for differing target parameters such as forest type classification or the estimation of biochemical and biophysical parameters. Fassnacht and Koch (2012) give a review on studies in this field and found that at forest stand level the application of images from several view angles improved the obtained classification accuracies in all of the examined studies (e.g., Kayitakire and Defourny, 2004; Dyk et al., 2006). While the studies reviewed in Fassnacht and Koch (2012) mostly used satellite data, Koukal et al. (2014) presented an approach to combine data from overlapping airborne imagery with two semi-empirical BRDF models. Imagery was obtained with an UltraCam-D frame camera and imagery was converted to reflectance data via empirical line correction. Parameters obtained by RTM inversion and added as explanatory variables to the RF classification were found to be useful for improving forest type classification on the stand level as compared to applying single-angle multi-spectral data. Overall accuracy improvements from 72% (kappa: 0.64) to 85% (kappa: 0.81) were observed when adding three directional parameters of the inverted Pinty-Verstraete model.

At single tree level, the influence of anisotropy has only been examined in detail in two well-designed studies conducted in a boreal forest (Korpela et al., 2011, 2014). Korpela et al. (2011) examined several factors affecting the reflectance variation of single trees as well as directional reflectance anisotropy by applying airborne multispectral imagery collected with an ADS40-SH52 three-line scanner. The three-line scanning system allows for simultaneously collecting data in nadir, backward and forward looking mode which in combination with multiple overflights ensured that for all the investigated single-tree objects reflectance information was gathered from differing view-illumination geometries. Concerning reflectance variation they identified low influences of age and site type and stronger influence of general between-stand variation. Accordingly, Peerbhay et al. (2014) reported no improvements in classification accuracy when adding forest age as variable to a classification based on multispectral Worldview-2 data. Korpela et al. (2011) also observed influences on reflectance variation connected to tree height but stated that these might have also been due to the experimental design. The study clearly observed directional reflectance anisotropy in sunlit crowns. Reflectance varied between $\pm 30\%$ in the examined view-angle range of $\pm 30°$ and the anisotropy was found to differ amongst species. The authors concluded that species-specific anisotropy information could be valuable for species separation at a single-tree level.

In Korpela et al. (2014) this assumption was tested. Korpela et al. (2014) found that differences in directional reflectance anisotropy cannot significantly improve tree species classifications (based on a simulated dataset an improvement of 1–3% was estimated). One reason for this was that the within-species reflectance variance, especially in the NIR domain, was explained by up to 58–70% by the tree effect, whereas directional reflectance anisotropy only explained 4–14%. The tree effect bases on the idea that the tree-specific crown structure and illumination incident influence the observed reflectance in all view geometries. Therefore, the tree effect basically describes the variance that is created by the individual structure of the tree itself in contrast to effects caused by atmospheric and species-related anisotropic effects. Korpela et al. (2014) acknowledge and list potential deficits in the experiment, e.g. the soundness of the crown models (specific to a tree), as potentially contributing to the observations. Based on the low importance of a shade fraction metric in a RF classification, Clark and Roberts (2012) stated that illumination variation is not an important factor in discriminating species at a pixel size of 1.6 m. They furthermore support the tree effect hypothesis as they also state that one of the main drivers of reflectance variability is the observed proportions of leaf and bark tissues in a crown which is directly related to the crown structure of an individual tree.

Another approach to account for view-illumination geometry is the separation of the crown into sunlit and shaded pixels. With this step, the intra-species variability of the spectral signatures can be reduced. Korpela (2004) used a photogrammetric approach to map tree tops in 3D and used this as basis to separate tree crowns into shaded and sunlit parts. Differences in self-shading properties of Scots pine and Norway spruce supported species identification in this study. A similar approach to separate sunlit and shaded parts of tree crowns using a combination of photogrammetric and LiDAR information was applied in Korpela et al. (2011, 2014). Other approaches to extract sunlit pixels include NDVI-masks (e.g., Engler et al., 2013, Alonzo et al., 2013) and more sophisticated statistical approaches applied to the pixel values in a crown object (e.g., Clark et al., 2005, Dalponte et al., 2014). A number of studies reported that the use of only sunlit pixels or crowns helped to improve the classification accuracies notably (e.g., Leckie et al., 2005; Waser et al., 2011) while in other cases accuracies did not improve (e.g., Alonzo et al., 2013) or even dropped (Clark and Roberts, 2012).

5.8. Spectral derivatives of passive optical data

The inherent high-dimensionality of multi- and hyperspectral data is often further increased by adding derivatives from the original data. Typical derivatives embrace the first and second derivative of the spectral signatures which are either calculated directly (e.g., Datt, 2000; Ghiyamat et al., 2013) or combined with a smoothing procedure (e.g. Savitzky-Golay derivation applied in Fassnacht et al., 2014), continuum removal (Fassnacht et al., 2014; Youngentob et al., 2011) as well as vegetation or spectral indices (Clark and Roberts, 2012; Ghosh et al., 2014; Jensen et al., 2012). The motivation of applying spectral derivatives lies in their potential to reduce influences of the soil background and other noise (Kokaly and Clark, 1999; King et al., 1999) and to enhance pigment absorption features (Blackburn, 2007). However, the benefit of applying derivatives in classification procedures is unclear as mirrored in the results of the reviewed studies: Ghiyamat et al., (2013, p. 190) examined several classification approaches with reflectance as well as the first and second derivatives (calculated as the slope of the reflectance/ first derivative in each band) as input data and found that "derivative spectra perform better than reflectance when the discrimination (problem) is simpler". For complex classification problems the derivative spectra were found to be less stable than the reflectance spectra. Contrasting findings were reported by Datt (2000) who found best results when applying the first derivative spectra. This might be explained by the increased noise level in the study of Ghiyamat et al. (2013) who applied airborne data as compared to the leaf-level data applied by Datt (2000).

Focusing on the application of vegetation or spectral indices, none of the reviewed studies demonstrated a clear advantage of their application. Generally, the application of the original reflectance or features derived from a feature extraction approach always led to at least similar but in most cases notably better results.

Concerning the application of continuum-removal (CM), Youngentob et al. (2011) reported improved classification performance when combining CM with multiple endmember spectral mixture analysis. Fassnacht et al. (2014) also found CM to improve the results compared to the application of original reflectance in several of the tested cases. However, as also discussed for the derivative spectra, these results may also depend on the remaining methodical work-flow and in particular on the approach applied to deal with the high-dimensionality of the data.

In summary, the benefit of using spectral derivatives for species classification remains unclear and in our opinion the probability that future studies identify a clear advantage of applying them in classification scenarios with single images are rather low. As also discussed for atmospheric correction procedures, their use might become more relevant in cases where several images with differing view-illumination geometries and acquisition conditions are analyzed simultaneously.

5.9. Physically-based models

The application of physically-based models for tree species classification has been limited so far. In the passive optical domain, radiative transfer models (RTM) have been successfully applied to obtain biochemical and biophysical leaf parameters such as chlorophyll content or leaf area index (Jacquemoud et al., 2009) in comparably homogeneous, often agricultural, vegetation. Only few studies have worked with RTM in forests (e.g., Essery et al., 2008; Kötz et al., 2004) and we found only a single study using a RTM in the context of tree species classification (Féret and Asner, 2011).

One obvious reason for the few studies using RTM for species classification is that species are typically not considered as variable in RTM. This makes a direct inversion of the model to obtain tree species impossible. On the other hand, species are likely to have typical values for certain biochemical and biophysical parameters that are considered in RTM. Therefore, an inversion of the model to obtain these parameters could be followed by a second step to identify species based on these derived biochemical and biophysical traits (Féret and Asner, 2011). The second step can either again be a supervised classification as presented in Féret and Asner (2011) or alternatively a decision tree based on expert knowledge which might have a certain potential for biomes with fewer species. One advantage of such an indirect RTM approach would be that well-calibrated canopy RTM are able to consider and thereby reduce effects related to view-illumination geometry and background (soil) signal (Jacquemoud et al., 2009). However, this is contrasted by high requirements concerning the atmospheric correction quality and by a larger number of typically unknown variables of canopy RTM which complicate an accurate model inversion.

While we found only one study in the passive optical domain that used a physically-based model to obtain species information, there have been a few SAR studies applying physically-based models to separate between major architectural growth forms of tree canopies based on the principle of "self-consistent structural regions" (Dickinson et al., 2013). Most of the studies working on this topic were connected to the Wishart classifier (Lee et al., 1999) or the SIR-C classifier of the University of Michigan (Dobson et al., 1996; Ulaby et al., 1990). The challenge that is addressed by these classifiers is the ill-posed problem to deconvolve the influences of vegetation structure and the dielectric properties of the vegetation. Dobson et al. (1995) describe the principle of such simulation models which combine simple architectural growth-forms with plant- and community level growth models to simulate the combined influence of vegetation structure and biomass. Such models enable the user to vary sensor characteristics and the structural characteristics of vegetation. Different growth forms or structural models were proposed for grasses and sedges, woody and herbaceous shrubs as well as trees. Trees are typically defined based on the growth form of their stem and their leaf shape and type (broadleaved or coniferous). The typical growths forms as applied in various studies are excurrent (large, vertical stem, typically coniferous), decurrent (complex branching structure, typically broadleaved) and columnar (main stem with a "top-knot", for example palm trees). When coupling these growth forms with radiative transfer models, the simulated backscatter patterns often agree well with the patterns observed by the SAR sensor. Further studies making use of physically-based scattering models have been presented by Dickinson et al. (2013) and van Zyl (1989).

Focusing on LiDAR, a recent study by Hovi and Korpela (2014) presented the first validated waveform-LiDAR simulator based on Monte Carlo ray tracing. The simulations in the study were based on geometric-optical models of three species typical for juvenile boreal forests (fireweed, raspberry and birch). The simulator was able to produce waveform signals that matched the real data well and the authors stated that the observed species-related differences were preserved in the modelled signal of the simulator. One reason for the development of the simulator was the possibility to study how different acquisition settings influence the observed LiDAR signal and how the setting can be adapted to improve species classification. Hovi and Korpela (2014) observed improved classification accuracies when increasing the footprint up to a threshold of approximately 0.30–0.36 m while the pulse width and sampling rate were found to be less important when classifying the three examined species. While the current version of the simulator is not yet able to consider multiple tree species, the study marks a first step towards physically-based LiDAR models that might be relevant for optimizing flight campaigns or may even enable model inversion approaches to classify tree species from waveform LiDAR data with no or little use of reference data.

5.10. Data fusion

Data fusion in remote sensing typically refers to the combination of (2D and 3D) data from several sensors (Alonzo et al., 2014). This combination can occur at pixel, feature or decision level (Pohl and Van Genderen, 1998). Many of the reviewed studies (18 out of 27 relevant cases) combined data at feature level with the feature typically being a single tree object.

A combination of LiDAR and hyperspectral data was identified to be the most common constellation amongst the combined approaches (17 out of 29 relevant cases). As mentioned in Section 4.3.1, canopy height information often failed to notably improve results based on hyperspectral data alone (e.g., Ghosh et al., 2014, Jones et al., 2010). Canopy height per se is an illogical predictor to classify tree species as the absolute height of a tree varies mainly with age, site conditions and competition and only to a minor degree with species. In the presence of several age classes of a single species, height information rather leads to confusion than to improving the discrimination power. However, height information was found to improve classification accuracies of smaller trees (Naidoo et al., 2012) which might be related to the previously mentioned dependency of spectral values from the size of the tree and its neighbors. Furthermore, LiDAR-derived (Jakubowski et al., 2013; Jones et al., 2010) or stereo image-based (e.g., Waser et al., 2010, 2011) vegetation height can be used to separate non-canopy from canopy elements using a height threshold, followed by a species classification based on passive optical information. The CHM can also be used - either independently or in combination with passive optical information – as input to segmentation algorithms for subsequent object-based classification. Area-based and single tree-based approaches (compare also Section 5.3) have been demonstrated (Jakubowski et al., 2013; Ke et al., 2010).

Recent studies were focusing on the derivation of more complex predictor variables from LiDAR data. In an urban environment, Alonzo et al. (2014) observed improved results when combining passive optical information not only with LiDAR CHM information but also with variables describing the crown morphology and variables presumably related to the structural arrangement of crowns and leaves. In the study, the crown objects were derived from the CHM model and the LiDAR metrics were extracted from the 3D point cloud within each crown object. Such 3D metrics might be harder to obtain in closed forests where borders of tree crowns are less clear. In a temperate forest, Dalponte et al. (2012) compared various combinations of geometric LiDAR data (high and low density) with hyperspectral as well as multispectral data. The tested LiDAR features were height statistics of the point clouds within tree objects. Best accuracies were obtained when fusing the hyperspectral data with the LiDAR predictors. The high density LiDAR data delivered better accuracies than the low-density data.

Studies that directly compared the use of single band LiDAR-only and multi-band passive optical-only data often reported better results for the passive optical datasets. In some cases this might have also been related to the suboptimal processing of the LiDAR data. In an urban area, Alonzo et al. (2014) reported an overall classification accuracy of 79.2% for hyperspectral-only and 32.9% for LiDAR-only datasets when classifying 29 urban tree species in California. Fusion of the datasets resulted in 83.4% overall accuracy (here, correctly classified canopy area). Heinzel and Koch (2012) reported contrary results and found better overall accuracies (79.2%) when classifying 4 central-European tree species with waveform LiDAR data as compared to results with hyperspectral data (64.7%). However, as shown in the studies of Ghosh et al. (2014) and Fassnacht et al. (2014) using the same hyperspectral imagery, the low results for the hyperspectral data reported in Heinzel and Koch (2012) are not fully representative. Holmgren et al. (2008) classified three species in a boreal forest in

Sweden and reported kappa of 0.87 for multispectral DMC data collected in autumn as compared to kappa of 0.82 for LiDAR data (TopEye Mk II, $\lambda = 1064$ nm). In summer, the LiDAR data performed better than the multispectral data (kappa = 0.76). Sarrazin et al. (2011) reported overall accuracies of 73% for hyperspectral data as compared to 53% obtained with LiDAR data when classifying 4 broadleaved savanna tree species. Unfortunately, many other studies skipped the results for using LiDAR-only data and only reported the gain in accuracy when adding LiDAR to the passive optical data. However, we interpret this as an indirect proof that the obtained accuracies with LiDAR-only data were notably worse. In conclusion, we think that the benefit of using spectral and especially hyperspectral data instead of LiDAR-only data will increase with increasing species numbers.

In addition to typical feature and pixel level fusion approaches, few studies have also combined passive optical and LiDAR information in a more complex way. For example Zhang and Qiu (2012) applied LiDAR data to select only the single highest hyperspectral pixel in a crown-object to ensure that the corresponding spectrum is collected from the sunlit portion of the crown. Similar approaches have also been presented by Puttonen et al. (2010a) and Yao et al. (2012). More advanced approaches have been presented by Korpela et al. (2010b, 2014) who used LiDAR information to model crown shapes as well as the surrounding canopy to identify self-shaded, shaded and sunlit pixels within single-tree crowns. A drawback of such methods is that data with high spatial resolution is required for both passive optical and LiDAR data and the co-registration of the two datasets has to be highly accurate. In this realm, Packalén et al. (2009) presented a method to link aerial imagery with known orientation parameters to airborne LiDAR data. In this process each LiDAR point is identified in various images and the average of the DN values extracted from these images are used for the classification of tree species. This approach increased the overall accuracies from 85.8% (using LiDAR data only) to 92.1%. The suitability of this approach has been confirmed by Korpela et al. (2014) and the authors additionally pointed out that – given the view-illumination geometry is sampled in a balanced way - unwanted anisotropic effects can be canceled out with this approach. Based on the reviewed studies, this is an under-examined but powerful approach which should be further investigated.

Finally, in some cases decision-level fusion was found to be an efficient approach to improve tree species classification results. Stavrakoudis et al. (2014) presented an interesting approach to fuse fuzzy classification maps obtained from very high spatial resolution multispectral and medium spatial resolution hyperspectral satellite data. In the heterogeneous Mediterranean landscape, the suggested comparably simple fusion approach improved overall accuracies by 8%.

6. Discussion of current constraints and future work

The current literature on tree species classification suggests that there is a large group of studies focusing on data- or algorithm-driven questions. The most common objective was the assessment of the potential of a certain sensor type or algorithm to classify tree species in a comparably small test site. This is connected to one of the most obvious constraints of the current state-of-knowledge. Most of the tested approaches (especially at the single tree level) were not developed to account for a large variety of ecological conditions. Although not stated in many of the reviewed studies, the selection of the test sites or at least the collected reference samples might have been biased towards favorable conditions (e.g., gentle terrain, dense stands, low degrees of species mixture, sparse number of age classes, homogeneity of silvicultural practices, good accessibility). As the optimization of classifications algorithms and processing chains is important, future work should generally pursue to integrate more than a single test site in any comparative study on tree species classification. The literature demonstrates that a large number of comparable datasets already exists in the community. An exchange of these datasets should be encouraged to enable meaningful comparisons that are also of interest for forest practitioners outside of the remote sensing community who are interested in robust approaches.

Furthermore, it would be desirable to increase the number of studies that try to actually understand and quantify the drivers for classification errors. Accordingly, studies attempting to formulate a hypothesis or a research question comparable to a hypothesis were rare (only 21 out 129 cases). Except for some studies conducted in Scandinavia (e.g., Korpela et al. (2014); Korpela et al. (2010a); Hovi and Korpela (2014)) hardly any study associated the remote sensing predictor variables used in the classification process to tree and stand parameters to understand the factors supporting or hampering the classification of tree species. For example, measures that describe the density of the foliage in the crowns, the influence of neighboring trees, the size of the crowns, the stand density as well as the composition of the understory are likely to contribute to the success of tree species classification from any sensor type (compare Korpela et al., 2014 for a case study based on aerial photographs). Except for the mentioned studies, we are not aware of any attempts trying to explain the misclassification of samples based on such or other attributes. One obstacle for such examinations is the need for more detailed reference information which could either be obtained in the field or from independent remote sensing datasets. For instance, the application of very high spatial resolution imagery collected by Unmanned Aerial Vehicles (UAVs) is one efficient way to get a more detailed characterization of canopies and background signals in single tree-based classification approaches.

A further general point to be considered in future studies is the definition of concrete applications with pre-defined quality or accuracy requirements for the final data products. The direct collaboration with practitioners from wildlife management, forest management and conservation could be a viable option to ensure that the research efforts of the remote sensing community match the need of potential users.

Besides these general future prospects, there are a number of sensor-specific technical points that still require more investigation. For example, the application of multi-temporal hyperspectral data to improve classification results has rarely been examined. Particularly in temperate forests, seasonal effects differ amongst species and were found to support the successful separation of tree species in multispectral datasets (e.g., Mickelson et al., 1998, Dymond et al., 2002). Currently, multi-temporal hyperspectral datasets might be mainly of scientific interest as the costs and effort connected to multiple hyperspectral data acquisitions (big data problem) are not very likely to justify the achievable gain in accuracy from a practitioners' perspective.

Studies using hyperspectral MIR and TIR information to discriminate vegetation and tree species are currently still in an early experimental stage. Yet, the observed unique spectral signature of various plant species in hyperspectral MIR and TIR signals are motivating and research efforts towards species identification from MIR and TIR sensors should be increased. Future research should improve both sensor technology and the understanding of the interactions between biophysical and biochemical plant compartments and the MIR and TIR signals. Furthermore, improvements of atmospheric correction algorithms are needed as they may still increase the uncertainty for species classification. Finally, the development of algorithms to enable precise temperature-emissivity separation (that is separating reflected and emitted energy) was found to be a further subject for future studies as this will be the key to distinguish subtle leaf emissivity features, potentially related to vegetation species (Ribeiro da Luz and Crowley, 2007, 2010).

For waveform LiDAR data, Hovi et al. (2016) pointed out that the effects of sensor parameters such as footprint size, the length of the emitted pulse and multiple wavelengths could be further investigated. This recommendation also bases on the findings of an earlier paper where simulations showed at least a notable effect of the footprint size on the obtainable classification accuracy (Hovi and Korpela, 2014). Another current research field in LiDAR methodology is the multi-temporal use of intensity measures by for example using discrete return intensity data from multi-temporal acquisitions (Kim et al., 2009). While just adding intensity variables from multiple acquisitions to a single classification is straightforward, the comparison of intensity measures across time requires calibration of the intensity variables which are affected by sensor and acquisition parameters.

The main problem in applying SAR data to separate tree species is its high variability with a large number of potentially complicating factors. Hence, the most promising approach to enable a reliable tree species classification from SAR data might be the application of physically-based models that are able to account for all or at least the most important mentioned varying factors. Several studies (e.g., Dickinson et al., 2013; Hoekman and Quiñones, 2002) have indicated that such approaches have a comparably high potential, especially when they are combined with forest classes defined based on biophysical (structural) characteristics rather than species-driven classes. However, the application of such physically-based model might be hampered by unknown parameters which either have to be fixed based on pre-knowledge or estimated from other data sources (e.g., terrestrial laser scanning (Côté et al., 2009)). Furthermore, the application of textural measures seems to be an under-examined approach which was investigated in only few studies (Wollersheim et al., 2011). However, these recommendations have to be seen in the context of other available sensor system and it might be questionable whether research efforts to derive forest information on species level from SAR data are really worthwhile if alternative data sources are available.

For sensor fusion approaches, the sole combination of predictors from LiDAR and passive optical (mostly airborne) sensors has been well examined. However, the potential to apply the structural information of the LiDAR data to normalize the passive optical data (e.g., in terms of leaf area density) on a single tree level is still under-explored. An important technical issue in this field (particularly when working on single-tree level) is to ensure an accurate co-registration of the two datasets. Further research could also focus on the integration of LiDAR data on the mapping of species information at the stand level. That is, in areas where airborne LiDAR but no airborne passive optical data is available, density information obtained from LiDAR data could support medium resolution passive optical data (Landsat, Sentinel-2) to map stand composition over larger geographic extents by delivering estimates of the intensity of the background signal via proxies describing the canopy density.

A second potential research field is the combination of data from several multispectral missions (Landsat-8, Sentinel-2, SPOT, commercial very high resolution systems) to derive tree species information at the stand level from a coarser pixel size but over larger geographic extents (compare also Sections 3.2 and 4.1). As such approaches cannot resolve single trees the target classes either have to be broader forest types or have to cope with non-discrete classes and corresponding methods (e.g., spectral unmixing approaches). Studies focusing on larger geographic extents have generally been underrepresented and the problems associated with such tasks (as discussed in Section 3.2) have rarely been addressed. Nevertheless, such tree species maps over larger geographic extents are relevant for national forest administrations to for example assess risks of forest damages caused by tree species that have limited abilities to adapt to climate-change processes (Immitzer et al., 2015).

Finally, scale-related effects are a further topic that is still under-examined for all sensor types. The identification of optimal scales to simultaneously maximize species-specific differences in remote sensing variables and minimize the workload to acquire and process the data needs to be further explored.

Concerning classification and validation strategies, one point that is of special interest when classifying tree species is the question of how to classify rare species. Engler et al. (2013) give one example where pre-knowledge is integrated into an ensemble classifier to avoid strong over-classification of a rare species. We think it is worthwhile to continue research on classification approaches that include pre-knowledge on the expected frequencies of species (for example from existing inventory data).

7. Outlook on new sensors and platforms

Recent developments in remote sensing sensors include two major trends. A fleet of new satellite sensors including Landsat-8 and the Sentinel satellites deliver free, quasi global data at short time intervals. In the case of passive optical sensors, this also increases the chance of obtaining cloud-free imagery. On the other hand, very detailed data gathered from remotely piloted aircrafts (RPA) and micro-satellites (Sandau et al., 2008) become more and more available.

Due to increased information on phenology, dense series of multi-temporal, multispectral Landsat-8 and Sentinel-2 data can serve as a good basis for the mapping of forest composition at national scales, considering major tree species typically occurring in temperate and boreal ecosystems. In such settings, an interesting approach to avoid problems with inter-scene differences of view-angle as well as terrain and atmospheric effects is to make use of a sufficiently large database of reference data in combination with a spatially adaptive automatized training procedure. A good example for such a spatially-adaptive approach where independent classifiers are trained for different ecoregions in the image has been recently presented for RapidEye and SPOT data (Stoffels et al., 2015). Such efforts can be further supported by upcoming hyperspectral missions such as EnMAP and HySPIRI which deliver more detailed spectral information at a lower spatial resolution. Further data may be added from recently launched SAR missions such as Sentinel-1 or ALOS-2 and enduring missions such as TanDEM-X. Although strong increases in accuracy cannot be expected from adding Radar data, the free availability of most of the data could be a motivation to investigate towards such approaches. While these new data sources could be powerful for biomes with lower tree species diversity (e.g. 5–10 dominating species) they might still be insufficient for a detailed assessment of tropical areas, especially when targeting forest composition assessment at the species level. In our opinion, it is unlikely that comprehensive classification and mapping at the species level will soon become feasible with remote sensing methods in the tropics. Mapping functional diversity as proposed in earlier papers (Asner et al., 2015) is more likely to be accomplished with remote sensing and may at the same time be a more meaningful approach in tropical areas.

We are confident that at the single tree level, research on the application of airborne hyper- and multispectral but also LiDAR data will continue. The acquisition of these airborne data have reached an operational level and can be realized over large continuous areas as demonstrated in the national ecological observatory network (NEON, http://www.neoninc.org) (Keller et al., 2008). Airborne surveys may become even more powerful with the current development of multispectral LiDAR-systems which are about to reach an operational level (see e.g., the "Titan" LiDAR system of Teledyne Optech with three wavelengths of 532 nm, 1064 nm and 1550 nm).

Based on own experiences we think that UAV systems have established as a data source for very high spatial resolution data at a highly flexible temporal resolution. Recently, LiDAR sensors and off-the-shelf hyperspectral cameras for UAVs have been introduced. UAV data may be applied either directly as database to classify tree species recognition on local scales or serve as source for additional reference information for spaceborne and airborne datasets. The latter option is more viable as the cost per area of UAV data is high (see Section 2). The reference data could be either species information or more detailed properties of reference crowns such as fractions of foliage, branches and background which could then serve as explanatory variables to explain the success or failure of classification algorithms.

Local scale studies may further benefit from data of recent commercial very high spatial resolution, super-spectral sensors such as WorldView-3 (superspectral defined here as a sensor with more and finer spectral channels as the typical multispectral sensor but lack of spectral continuity). The combination of multiple narrow bands in the VIS, NIR and SWIR regions with a very high spatial resolution that allows for species classification at a single tree level is a promising alternative to the still relatively laborious and expensive acquisition of airborne hyperspectral data. Early studies applying data from the preceding WorldView-2 sensor to discriminate tree species (e.g., Immitzer et al., 2012a, 2012b, Peerbhay et al., 2014, Waser et al., 2014) have demonstrated the potential of these kind of data to derive single-tree based information on species and health status of trees. This potential of the WorldView data is expected to be further increased with the new SWIR channels of WorldView-3.

Finally, there are also still ongoing developments in airborne sensor technology. Recent works have tested a combined use of LiDAR and hyperspectral devices in an integrated terrestrial-based system for in-situ applications, which is, however, still in an early experimental phase (Puttonen et al., 2010a; Vauhkonen et al., 2013, Junttila et al., 2015). Other experiments with multi- or even hyperspectral laser scanner systems are currently reaching operationality. The first experimental results for vegetation and land-cover classification are promising, though the multi return effects were stated to bias the spectral signatures from rough and volumetric targets such as trees (Wichmann et al., 2015).

Concerning the development of multispectral passive sensors, Heikkinen et al., (2010) showed that the placement, number and the width of VIS-NIR bands influence the classification performance of single trees. The simulated performance of the four-band, three-line Leica ADS40 sensor could be improved by applying a fifth band near the red-edge. The optimal width of this band was only 15 nm. The addition of the red-edge band to the existing VIS-NIR bands improved classification accuracies by 5–13%. The overall sub-optimality of the band placement and width in the ADS40 became obvious. Under boreal conditions, such a 5-band system performed similar to a hyperspectral system which was tested simultaneously.

8. Conclusions

The main findings of this review can be summarized with the following points:

(1) The number of studies focusing on tree species classification increased constantly over the last 40 years which is partly due to generally increased publication rates but still mirrors an ongoing interest in the topic as species information is a bottleneck in many applications. Most studies were conducted in temperate and boreal forest ecosystems of North America and Europe while the number of studies conducted in Asia, Africa, South America and Australia was sparse.

(2) Data-driven or sensor-driven investigations focusing on the maximization of classification accuracy at a single often comparably small test site by far dominated the reviewed studies. The value of these studies is limited as the generalization of the reported results is often not possible. We recommend fostering the exchange of comparable datasets amongst researchers to support research efforts targeting on the development of operational classification approaches.

(3) Studies working with a concrete hypothesis, trying to systematically understand and quantify the factors affecting tree species classification or having a well-defined application in mind are still sparse and research efforts towards such objectives should be increased.

(4) Passive optical systems and especially hyperspectral systems generally showed higher potential for classification of tree species (similar accuracies but with higher number of tree species) than active SAR or LiDAR sensor systems. However, LiDAR data have proven to be suitable for regions with low numbers of species and might further increase their potential with new multispectral sensor systems. Studies on hyperspectral thermal sensors are still sparse but first results indicate that species-specific spectra can be observed in the thermal domain.

Acknowledgements

The first author acknowledges the Fulbright scholar program that partly supported this work with a visiting scholar scholarship. We acknowledge Dr. Somidh Saha who contributed to early versions of this study. Furthermore, this work was partly supported by the project REMBIOFOR „Remote sensing based assessment of woody biomass and carbon storage in forests", supported by The National Centre for Research and Development in Poland under BIOSTRATEG program, agreement no. BIOSTRATEG1/267755/4/NCBR/2015. Finally, we would like to thank five anonymous reviewers who helped to improve an earlier version of the manuscript. We would like to especially acknowledge the tremendous amount of feedback of Ilkka Korpela which was of great help to increase the quality of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2016.08.013.

References

Adelabu, S., Mutanga, O., Adam, E., Cho, M.A., 2013. Exploiting machine learning algorithms for tree species classification in a semiarid woodland using RapidEye image. J. Appl. Remote. Sens. 7 (1) (073480–1–073480–13).

Alonzo, M., Bookhagen, B., Roberts, D.A., 2014. Urban tree species mapping using hyperspectral and lidar data fusion. Remote Sens. Environ. 148, 70–83.

Alonzo, M., Roth, K., Roberts, D., 2013. Identifying Santa Barbara's urban tree species from AVIRIS imagery using canonical discriminant analysis. Remote Sens. Lett. 4 (5), 513–521.

Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sens. Environ. 64 (3), 234–253.

Asner, G.P., Martin, R.E., Anderson, C.B., Knapp, D.E., 2015. Quantifying forest canopy traits: imaging spectroscopy versus field survey. Remote Sens. Environ. 158, 15–27.

Bauer, M.E., Burk, T.E., Ek, A.R., Coppin, P.R., Lime, S.D., Waish, T.A., Walters, D.K., Befort, W., Heinzen, D.F., 1994. Satellite inventory of Minnesota forest resources. Photogramm. Eng. Remote Sens. 60 (3), 287–298.

Bergseng, E., Ørka, H., Næsset, E., Gobakken, T., 2015. Assessing forest inventory information obtained from different inventory approaches and remote sensing data sources. Ann. For. Sci. 72, 33–45.

Blackburn, G.A., 2007. Hyperspectral remote sensing of plant pigments. J. Exp. Bot. 58 (4), 855–996.

Blackburn, G.A., Milton, E.J., 1995. Seasonal variations in the spectral reflectance of deciduous tree canopies. Int. J. Remote Sens. 16 (4), 709–720.

Boerner, W.M., Mott, H., Lueneburg, E., Livingstone, E., Brisco, B., Brown, R.J., Paterson, S., 1998. Polarimetry in radar remote sensing: basic and applied concepts. In: Henderson, F.M., Lewis, A.J. (Eds.), Principles and Applications of Imaging Radar. John Wiley & Sons, Ltd.

Boschetti, M., Boschetti, L., Oliveri, S., Casati, L., Canova, I., 2007. Tree species mapping with airborne hyper-spectral MIVIS data: The Ticino Park study case. Int. J. Remote Sens. 28 (6), 1251–1261.

Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne LIDAR. ISPRS J. Photogramm. Remote Sens. 61, 325–340.

Brandtberg, T., Warner, T., 2006. High spatial-resolution remote sensing. In: Shao, G., Reynolds, K.M. (Eds.), Computer Applications in Sustainable Forest Management: Including Perspectives on Collaboration and Integration, pp. 19–41.

Carleer, A., Wolff, E., 2004. Exploitation of very high resolution satellite data for tree species identification. Photogramm. Eng. Remote Sens. 70 (1), 135–140.

Castro-Esau, K., Sanchez-Azofeifa, G., Rivard, B., Wright, S., Quesada, M., 2006. Variability in leaf optical properties of Mesoamerican trees and the potential for species classification. Am. J. Bot. 93, 517–530.

Chambers, D., Périé, C., Casajus, N., de Blois, S., 2013. Challenges in modelling the abundance of 105 tree species in eastern North America using climate, edaphic, and topographic variables. For. Ecol. Manag. 291, 20–29.

Chance, C.M., Coops, N.C., Crosby, K., Aven, N., 2016. Spectral wavelength selection and detection of two invasive plant species in an urban area. Can. J. Remote. Sens. 41, 1–14.

Cho, M.A., Debba, P., Mathieu, R., Naidoo, L., van Aardt, J., Asner, G.P., 2010. Improving discrimination of savanna tree species through a multiple-endmember spectral angle mapper approach: canopy-level analysis. IEEE Trans. Geosci. Remote Sens. 48 (11), 4133–4142.

Cho, M.A., Mathieu, G.P.R.A., Naidoo, L., van Aardt, J., Ramoelo, P.A., Debba, W.K., Main, R., et al., 2012. Mapping tree species composition in South African

savannas using an integrated airborne spectral and LiDAR system. Remote Sens. Environ. 125, 214–226.
Chuine, I., Beaubien, E.G., 2001. Phenology is a major determinant of tree species range. Ecol. Lett. 4, 500–510.
Clark, M.L., Roberts, D.A., 2012. Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier. Remote Sens. 4 (6), 1820–1855.
Clark, M.L., Roberts, D.A., Clark, D.B., 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sens. Environ. 96 (3–4), 375–398.
Coops, N.C., Hilker, T., Wulder, M.A., St-Onge, B., Newnham, G., Siggins, A., Trofymow, J.A., 2007. Estimating canopy structure of Doulgas -fir forest stands from discrete-return LiDAR. Trees-Struct. Funct. 21, 295–310.
Côté, J.-F., Widlowski, J.-L., Fournier, R.A., Verstraete, M.M., 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. Remote Sens. Environ. 113 (5), 1067–1081.
Culvenor, D.S., Coops, N., Preston, R., Tolhurst, K.G., 1999. A spatial clustering approach to automated tree crown delineation. In: Hill, D.A., Leckie, D.G. (Eds.), International Forum On Automated High Resolution Digital Imagery For Forestry, pp. 67–80 Victoria, B.C., Feburary 10–12, 1998.
Dalponte, M., Bruzzone, L., Gianelle, D., 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. Remote Sens. Environ. 123 (0), 258–270.
Dalponte, M., Ørka, H.O., Ene, L.T., Gobakken, T., Næsset, E., 2014. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. Remote Sens. Environ. 140, 306–317.
Dalponte, M., Ørka, H.O., Gobakken, T., Gianelle, D., Næsset, E., 2013. Tree species classification in boreal forests with hyperspectral data. IEEE Trans. Geosci. Remote Sens. 51 (5), 2632–2645.
Datt, B., 2000. Recognition of Eucalyptus forest species using hyperspectral reflectance data. Proceedings of IEEE Geoscience and Remote Sensing Symposium, 2000. 4, pp. 1405–1407.
Dickinson, C., Siqueira, P., Clewley, D., Lucas, R., 2013. Classification of forest composition using polarimetric decomposition in multiple landscapes. Remote Sens. Environ. 131, 206–214.
Dinuls, R., Erins, G., Lorencs, A., Mednieks, I., Sinica-Sinavskis, J., 2012. Tree species identification in mixed Baltic Forest using LiDAR and multispectral data. IEEE J. Sel. Top. in Appl. Earth Obs. Remote Sens. 5 (2), 594–603.
Dobson, M.C., Pierce, L.E., Mcdonald, K., Sharik, T., 1991. Seasonal change in radar backscatter from mixed conifer and hardwood frorests in northern Michigan. Proceedings of: International Geoscience and Remote Sensing Symposium, 1991. IGARSS '91. Remote Sensing: Global Monitoring for Earth Management. 3, pp. 1121–1124.
Dobson, M.C., Pierce, Ulaby, F., 1996. Knowledge-based land-cover classification using ERS-l/JERS-1 SAR composites. IEEE Trans. Geosci. Remote Sens. 34 (1), 83–99.
Dobson, M.C., Ulaby, F.T., Pierce, L.E., 1995. Land-cover classification and estimation of terrain attributes using synthetic aperture radar. Remote Sens. Environ. 51, 199–214.
Dyk, A., Goodenough, D.G., Li, J.Y., Niemann, K.O., Guan, A., Chen, H., Duong, J., 2006. Multi-temporal, multi-angle evaluation with CHRIS of coastal forests. IGARSS'06, Proceedings of the IEEE International Conference on Geoscience and Remote Sensing Symposium, 2006. Denver, Colorado, USA.
Dymond, C.C., Mladenoff, D.J., Radeloff, V.C., 2002. Phenological differences in Tasseled Cap indices improve deciduous forest classification. Remote Sens. Environ. 80 (3), 460–472.
Eid, T., Gobakken, T., Næsset, E., 2004. Comparing stand inventories for large areas based on photo interpretation and laser scanning by means of cost-plus-loss analyses. Scand. J. For. Res. 19, 512–523.
Einzmann, K., Ng, W.-T., Immitzer, M., Bachmann, M., Pinnel, N., Atzberger, C., 2014. Method analysis for collecting and processing in-situ hyperspectral needle reflectance data for monitoring Norway spruce. Photogrammetrie, Fernerkundung, Geoinformation 5 (2014), 0423–0434.
Engler, R., Waser, L.T., Zimmermann, N.E., Schaub, M., Berdos, S., Ginzler, C., Psomas, A., 2013. Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. For. Ecol. Manag. 310, 64–73.
Erikson, M., 2004. Species classification of individually segmented tree crowns in high-resolution aerial images using radiometric and morphologic image measures. Remote Sens. Environ. 91 (3–4), 469–477.
Essery, R., Bunting, P., Hardy, J., Link, T., Marks, D., Melloh, R., Pomeroy, J., Rowlands, A., Rutter, N., 2008. Radiative transfer modelling of a coniferous canopy characterized by airborne remote sensing. J. Hydrometeorol. 9, 228–241.
European Environmental Agency, 2007. European Forest Types: Categories and Types for Sustainable Forest Management Reporting and Policy. (No. EEA Terchnical Report No 9/2006). EEA, Copenhagen Retrieved from http://www.env-edu.gr/Documents/European%20forest%20types.pdf.
Fassnacht, F.E., Koch, B., 2012. Review of forestry oriented multi-angular remote sensing techniques. Int. For. Rev. 14 (3), 285–298.
Fassnacht, F.E., Neumann, C., Förster, M., Buddenbaum, H., Ghosh, A., Clasen, A., Joshi, P.K., Koch, B., 2014. Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. IEEE J. Sel. Top. in Appl. Earth Obs. Remote Sens. 7 (6), 2547–2561.
Felbermeier, B., Hahn, A., Schneider, T., 2010. Study on user requirements sensing applications in forestry. In: Wagner, W., Székely, B. (Eds.), ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7, 2010, IAPRS. Vol. XXXVIII, Part 7B.
Féret, J.-B., Asner, G.P., 2011. Spectroscopic classification of tropical forest species using radiative transfer modeling. Remote Sens. Environ. 115 (9), 2415–2422.
Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201.
Franklin, S.E., Hall, R.J., Moskal, L.M., Maudie, A.J., Lavigne, M.B., 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. Int. J. Remote Sens. 21 (1), 61–79.
Gao, B.C., Hoetz, A.F.H., 1990. Column atmospheric water-vapor and vegetation liquid water retrievals from airborne imaging spectrometer data. J. Geophys. Res.-Atmos. 95, 3549–3564.
Gärtner, P., Förster, M., Kleinschmit, B., 2016. The benefit of synthetically generated RapidEye and Landsat 8 data fusion time series for riparian forest disturbance monitoring. Remote Sens. Environ. 177, 237–247.
Gatziolis, D., 2011. Dynamic range-based intensity normalization for airborne, discrete return lidar data of forest canopies. Photogramm. Eng. Remote. Sens. 77 (3), 251–259.
Ghiyamat, A., Shafri, H.Z.M., Amouzad, M.G., Shariff, A.R., Mansor, S., 2013. Hyperspectral discrimination of tree species with different classifications using single- and multiple-endmember. Int. J. Appl. Earth Obs. Geoinf. 23, 177–191.
Ghosh, A., Joshi, P.K., 2014. A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. Int. J. Appl. Earth Obs. Geoinf. 26, 298–311.
Ghosh, A., Fassnacht, F.E., Joshi, P.K., Koch, B., 2014. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. Int. J. Appl. Earth Obs. Geoinf. 26, 49–63.
Gong, P., Pu, R., Yu, B., 1997. Conifer species recognition: an exploratory analysis of in situ hyperspectral data. Remote Sens. Environ. 62, 189–200.
Goodenough, D.G., Dyk, A., Niemann, K.O., Pearlman, J.S., Chen, H., Han, T., Murdoch, M., West, C., 2003. Processing Hyperion and ALI for forest classification. IEEE Trans. Geosci. Remote Sens. 41 (6), 1321–1331.
Gougeon, F.A., 1995. Comparison of possible multispectral classification schemes for tree crowns individually delineated on high spatial resolution MEIS images. Can. J. Remote. Sens. 21, 1–9.
Grant, L., 1987. Diffuse and specular characteristics of leaf reflectance. Remote Sens. Environ. 22, 309–322.
Haara, A., Haarala, M., 2002. Tree species classification using semi-automatic delineation of trees on aerial images. Scand. J. For. Res. 17 (6), 556–565.
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 (3), 1355–2010.
Heinzel, J., Koch, B., 2011. Exploring full-waveform LiDAR parameters for tree species classification. Int. J. Appl. Earth Obs. Geoinf. 13 (1), 152–160.
Heinzel, J., Koch, B., 2012. Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. Int. J. Appl. Earth Obs. Geoinf. 18, 101–110.
Heller, R.C., Doverspike, G.E., Aldrich, R.C., 1964. Identification of tree species on large-scale panchromatic and color aerial photographs. Agriculture Handbook, No. 261. U.S. Department of agriculture, Forest Service 16 pages.
Hesketh, M., Sanchez-Azofeifa, G.A., 2012. The effect of seasonal spectral variation on species classification in the Panamanian tropical forest. Remote Sens. Environ. 118, 73–82.
Hill, R.A., Wilson, A.K., George, M., Hinsley, S.A., 2010. Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. Appl. Veg. Sci. 13 (1), 86–99.
Hoekman, D.H., 1985. Radar backscatter of forest stands. Int. J. Remote Sens. 6 (2), 325–343.
Hoekman, D.H., Quiñones, M.J., 2002. Biophysical Forest type characterization in the Colombian Amazon by airborne polarimetric SAR. IEEE Trans. Geosci. Remote Sens. 40 (6), 1288–1300.
Höfle, B., Hollaus, M., Lehner, H., Pfeifer, N., Wagner, W., 2008. Area-based parameterization of forest structure using full-waveform airborne laser scanning data. In: Hill, R.A., Rosette, J., Suárez, J. (Eds.), Proceedings of SilviLaser 2008: 8th International Conference on LiDAR Applications in Forest Assessment and Inventory. Edinburgh, UK, pp. 227–236.
Hollaus, M., Mücke, W., Höfle, B., Dorigo, W., Pfeifer, N., Wagner, W., Bauerhansl, C., et al., 2009. Tree species classification based on full-waveform airborne laser scanning data. Proceedings of Silvilaser 2009. College Station, Texas, USA, pp. 1–9 (wl27www658.webland.ch).
Holloway, P.J., 1982. Structure and histochemistry of plant cuticular membranes: an overview. In: Cutler, D.F., Alvin, K.L., Price, C.E. (Eds.), The plant cuticle. Academic Press, New York.
Holmgren, J., Persson, Å., 2004. Identifying species of individual trees using airborne laser scanner. Remote Sens. Environ. 90, 415–423.
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 (5), 1537–1552.
Holopainen, M., Talvitie, M., 2006. Effect of data acqusistion accuracy on timing of stand harvests and expected net present value. Silva Fennica 40 (3), 531–543.
Hovi, A., Korpela, I., 2014. Real and simulated waveform-recording LiDAR data in juvenile boreal forest vegetation. Remote Sens. Environ. 140, 665–678.
Hovi, A., Korhonen, L., Vauhkonen, J., Korpela, I., 2016. LiDAR waveform features for tree species classification and their sensitivity to tree- and acquisition related parameters. Remote Sens. Environ. 173, 224–237.
Hughes, G.F., 1968. On the mean accuracy of statistical pattern recognizers. IEEE Trans. Inf. Theory 14 (1), 55–63.
Immitzer, M., Atzberger, C., Einzmann, K., Böck, S., Mattiuzzi, M., Wallner, A., Seitz, R., Pinnel, N., Müller, A., Frost, M., 2015. Fichten- und Kiefernkarte für Bayern. LWF

aktuell (Bayerische Landesanstalt für Wald und Forstwirtschaft) 106, 30–37 in German.

Immitzer, M., Atzberger, C., Koukal, T., 2012a. Tree species classification with random Forest using very high spatial resolution 8-Band WorldView-2 satellite data. Remote Sens. 4 (9), 2661–2693.

Immitzer, M., Atzberger, T.C., Koukal, 2012b. Suitability of WorldView-2 data for tree species classification with special emphasis on the four new spectral bands. Photogrammetrie, Fernerkundung, Geoinformation 2012 (5), 573–588.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. PROSPECT + SAIL models: a review of use for vegetation characterization. Remote Sens. Environ. 113, 56–66.

Jakubowski, M.K., Li, W., Guo, Q., Kelly, M., 2013. Delineating individual trees from Lidar data: a comparison of vector- and raster-based segmentation approaches. Remote Sens. 5, 4163–4186.

Jansson, G., Angelstam, P., 1999. Threshold levels of habitat composition for the presence of long-tailed tit (*Aegithalos audatus*) in a boreal landscape. Landsc. Ecol. 14, 283–290.

Jensen, R.R., Hardin, P.J., Hardin, A.J., 2012. Classification of urban tree species using hyperspectral imagery. Geocarto International 27 (5), 443–458.

Johansen, K., Phinn, S., 2006. Mapping structural parameters and species composition of riparian vegetation using IKONOS and Landsat ETM+ data in Australian tropical savannahs. Photogramm. Eng. Remote. Sens. 72 (1), 71–80.

Jones, T.G., Coops, N.C., Sharma, T., 2010. Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. Remote Sens. Environ. 114 (12), 2841–2852.

Junttila, S., Kaasalainen, S., Vastaranta, M., Hakala, T., Nevalainen, O., Holopainen, M., 2015. Investigating bi-temporal hyperspectral LiDAR measurements from declined trees- experiences from laboratory test. Remote Sens. 2015 (7), 13863–13877.

Katoh, A., Moskal, L.M., Schiess, P., Swanson, M.E., Calhoun, D., Stuetzle, W., 2009. Capturing tree crown formation through implicit surface reconstruction using airborne lidar data. Remote Sens. Environ. 113 (6), 1148–1162.

Kayitakire, F., Defourny, P., 2004. Forest type discrimination using multi-angle hyperspecral data. Proc. of the 2nd CHRIS/Proba Workshop. ESA/ESRIn, Frascatiy, Italy.

Ke, Y., Quackenbush, L.J., Im, J., 2010. Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. Remote Sens. Environ. 114 (6), 1141–1154.

Keller, M., Schimel, D.S., Hargrove, W., Hoffman, F.M., 2008. A continental strategy for the National Ecological Observatory Network. Front. Ecol. Environ. 6, 282–284.

Kennedy, C.E.J., Southwood, T.R.E., 1984. The number of species of insects associated with British trees: A re-analysis. J. Anim. Ecol. 53 (2), 455–478.

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 (1), 100–112.

Kim, S., 2007. Individual Tree Species Identification Using LIDAR-Derived Crown Structures and Intensity Data Dissertation University of Washingtonhttp://depts.washington.edu/rsgal/pubs/SooyoungKimPhDDissertation.pdf.

Kim, S.-R., Lee, W.-K., Kwak, D.-A., Biging, G.S., Gong, P., Lee, J.H., Cho, H.-K., 2011. Forest cover classification by optimal segmentation of high resolution satellite imagery. Sensors 11, 1943–1958.

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 (8), 1575–1586.

King, R.L., Ruffin, C., LaMastus, F.E., Shaw, D.R., 1999. The analysis of hyperspectral data using Savitzky-Golay filtering – practical issues 2. Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Hamburg, Germany, June 28–July 2. 2, pp. 398–400.

Knowlton, D.J., Hoffer, R.M., 1981. Radar imagery for forest cover mapping. Proceedings of: 7th International Symposium on Machine Processing of Remotely Sensed Data with Special Emphasis on Range, Forest and Wetlands Assessment Retrieved from http://docs.lib.purdue.edu/lars_symp/481.

Koch, B., Kattenborn, T., Straub, C., Vauhkonen, J., 2013. Segmentation of forests to tree objects. In: Maltamo, M., Naesset, E., Vauhkonen, J. (Eds.), Forestry Applications of Airborne Laser Scanning. Springer.

Kokaly, R.F., Clark, R.N., 1999. Spectroscopic determination of leaf biochemistry using band-depths analysis of absorption features and stepwise linear regression. Remote Sens. Environ. 67 (3), 267–287.

Korpela, I., 2004. Individual tree measurements by means of digital aerial photogrammetry. Silva Fennica, Monographs 3, 2004 (93 pp).

Korpela, I.S., Tokola, T.E., 2006. Potential of aerial image-based monoscopic and multiview single-tree forest inventory: a simulation approach. For. Sci. 52 (2), 136–147.

Korpela, I., Ørka, H.O., Hyyppä, J., Heikkinen, V., Tokola, T., 2010a. Range and AGC normalization in airborne discrete-return LiDAR intensity data for forest canopies. ISPRS J. Photogramm. Remote Sens. 65, 369–379.

Korpela, I., Ørka, H.O., Maltamo, M., Tokola, T., Hyyppä, J., 2010b. Tree species classification using airborne LiDAR—effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. Silva Fennica 44 (2), 319–339.

Korpela, I., Heikkinen, V., Hokavaara, 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., Hovi, A., Korhonen, L., 2013. Backscattering of individual lidar pulses from forest canopies explained by photogrammetrically derived vegetation structure. Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci. XL-1/W1, 171–176.

Korpela, I., Hovi, A., Morsdorf, F., 2012. Understory trees in airborne LiDAR data – selective mapping due to transmission losses and echo-triggering mechanisms. Remote Sens. Environ. 119, 92–104.

Korpela, I., Koskinen, M., Vasander, H., Holopainen, M., Minkkinen, K., 2009. Airborne small-footprint discrete-return LiDAR data in the assessment of boreal mire surface patterns, vegetation, and habitats. For. Ecol. Manag. 258, 1549–1566.

Korpela, I., Mehtätalo, L., Seppänen, A., Kangas, A., 2014. Tree species identification in aerial image data using directional reflectance signatures. Silva Fennica 48 (3), article id 1087.

Korpela, I., Tuomola, T., Tokola, T., Dahlin, B., 2008. Appraisal of seedling stand vegetation with airborne imagery and discrete-return LiDAR – an exploratory analysis. Silva Fennica 42 (5), 753–772.

Kotchenova, S.Y., Shabanov, N.V., Knyazikhin, Y., Davis, A.B., Dubayah, R., Myneni, R.B., 2003. Modeling lidar waveforms with time-dependent stochastic radiative transfer theory for remote estimations of forest structure. J. Geophys. Res. 108 (D15), 4484.

Kötz, B., Schaepman, M., Morsdorf, F., Bowyer, P., Itten, K., Allgöwer, B., 2004. Radiative transfer modeling within a heterogeneous canopy for estimation of forest fire fuel properties. Remote Sens. Environ. 92, 332–344.

Koukal, T., Atzberger, C., Schneider, W., 2014. Evaluation of semi-empirical BRDF models inverted against multi-angle data from a digital airborne frame camera for enhancing forest type classification. Remote Sens. Environ. 151, 27–43.

Kuhn, M., Johnson, K., 2013. Resampling Techniques. Chapter in: Applied Predictive Modeling. Springer, New York, NY, pp. 69–73.

Kwok, R., Rignot, E., Way, J.B., Freeman, A., Holt, J., 1994. Polarization signatures of thawed and frozen forests of varying biomass. IEEE Trans. Geosci. Remote Sens. 32, 371–381.

Leckie, D.G., Ranson, K.J., 1998. Forestry Applications Using Imaging Radar. In: Henderson, F.M., Lewis, A.J. (Eds.), Chapter in: Principles and Applications of Imaging Radar Vol. 3. John Wiley & Sons, Ltd.

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 (2), 175–190.

Lee, J.S., Grunes, M.R., Ainsworth, T.L., Du, L., Schuler, D.L., Cloude, S.R., 1999. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier. IEEE Trans. Geosci. Remote Sens. 37 (5), 2249–2258.

Lewis, A.J., Henderson, F.M., Holcomb, D.W., 1998. Radar fundamentals: the geoscience perspective. In: Henderson, F.M., Lewis, A.J. (Eds.), Principles and Applications of Imaging Radar. John Wiley and Sons, Inc.

Liang, X., Hyppä, J., Matikainen, L., 2007. Deciduous-Coniferous tree classification using difference between first and last pulse laser signatures. IAPRS. Volume XXXVI, Part 3/W52, pp. 253–257 Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.222.3625&rep=rep1&type=pdf.

Lu, D., Weng, Q., 2007. Survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. 28, 823–870.

Maghsoudi, Y., Collins, M., Leckie, D.G., 2012. Polarimetric classification of Boreal forest using nonparametric feature selection and multiple classifiers. Int. J. Appl. Earth Obs. Geoinf. 19, 139–150.

Mallinis, G., Koutsias, N., Tsakiri-Strati, M., Karteris, M., 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. ISPRS J. Photogramm. Remote Sens. 63 (2), 237–250.

Marceau, D.J., Gratton, D.J., Fournier, R.A., Fortin, J.P., 1994. Remote sensing and the measurement of geographical entities in a forested environment. 2. The optimal spatial resolution. Remote Sens. Environ. 49, 105–117.

Martin, R., Asner, G., Sack, L., 2007. Genetic variation in leaf pigment, optical and photosynthetic function among diverse phenotypes of *Metrosideros polymorpha* grown in a common garden. Oecologia 151, 387–400.

Meyer, P., Staenz, K., Itten, K.I., 1996. Semi-automated procedures for tree species classification in high spatial resolution data from digitized colour infrared-aerial photography. ISPRS J. Photogramm. Remote Sens. 51, 5–16.

Mickelson, J.G., Civco, D.L., Silander, J.A., 1998. Delineating forest canopy species in the northeastern United States using multi-temporal TM imagery. Photogramm. Eng. Remote. Sens. 64 (9), 891–904.

Moisen, G.G., Freeman, E.A., Blackard, J.A., Frescino, T.S., Zimmermann, N.E., Edwards Jr., T.C., 2006. Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. Ecol. Model. 199 (2), 176–187.

Moore, M.M., Bauer, M.E., 1990. Classification of forest vegetation in North-Central Minnesota using Landsat multispectral scanner and thematic mapper data. For. Sci. 36 (2), 330–342.

Morain, S.A., Simonett, D.S., 1967. K-band radar in vegetation mapping. Photogramm. Eng. 33 (7), 730–740.

Naidoo, L., Cho, M.A., Mathieu, R., Asner, G., 2012. Classification of savanna tree species, in the Greater Kruger National Park by integrating hyperspectral and LiDAR data in a random forest data mining environment. ISPRS J. Photogramm. Remote Sens. 69, 167–179.

Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W., Limperis, T., 1977. Geometrical considerations and nomenclature for reflectance. NBS Monogr., No. 160. National Bureau of Standards, U.S. Department of Commerce 52 p.

Ørka, H.O., Hauglin, M., 2016. Use of remote sensing for mapping of non-native conifer species. INA Fagrapport 33 (76 pp).

Ørka, H.O., Dalponte, M., Gobakken, T., Næsset, E., Ene, L.T., 2013. Characterizing forest species composition using multiple remote sensing data sources and inventory approaches. Scand. J. For. Res. 28 (7), 677–688.

Ørka, H.O., Næsset, E., Bollandsås, M., 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. Remote Sens. Environ. 113 (6), 1163–1174.

Ortiz, S.M., Breidenbach, J., Knuth, R., Kändler, G., 2012. The influence of DEM quality on mapping accuracy of coniferous- and deciduous-dominated forest using TerraSAR-X images. Remote Sens. 4, 661–681.
Packalén, P., Suvanto, A., Maltamo, M., 2009. A two stage method to estimate species-specific growing stock. Photogramm. Eng. Remote Sens. 75 (12), 1451–1460.
Pant, P., Heikkinen, V., Hovi, I.A., Korpela, Hauta-Kasari M., Tokola, T., 2013. Evaluation of simulated bands in airborne optical sensors for tree species identification. Remote Sens. Environ. 138 (0), 27–37.
Pausas, J.G., Austin, M.P., Noble, I.R., 1997. A forest simulation model for predicting eucalypt dynamics and habitat quality for arboreal marsupials. Ecol. Appl. 7 (3), 921–933.
Peerbhay, K.Y., Mutanga, O., Ismail, R., 2014. Investigating the capability of few strategically placed worldview-2 multispectral bands to discriminate Forest species in KwaZulu-Natal, South Africa. IEEE J. Sel. Top. in Appl. Earth Obs. Remote Sens. 7 (1), 307–316.
Peña, M.A., Cruz, P., Roig, M., 2013. The effect of spectral and spatial degradation of hyperspectral imagery for the Sclerophyll tree species classification. Int. J. Remote Sens. 34 (20), 7113–7130.
Pierce, L.E., Bergen, K.M., Dobson, M.C., Ulaby, F.T., 1998. Multitemporal land-cover classification using SIR-C/X-SAR imagery. Remote Sens. Environ. 64, 20–33.
Plourde, L.C., Ollinger, S.V., Smith, M.-L., Martin, M.E., 2007. Estimating species abundance in a northern temperate forest using spectral mixture analysis. Photogramm. Eng. Remote. Sens. 73 (7), 829–840.
Pohl, C., Van Genderen, J., 1998. Review article multisensor image fusion in remote sensing: concepts, methods and applications. Int. J. Remote Sens. 19 (5), 37–41.
Pope, K.O., Rey-Benayas, J.M., Paris, J.F., 1994. Radar remote sensing of forest and wetland ecosystems in the Central American tropics. Remote Sens. Environ. 48, 205–219.
Pu, R., Landry, S., 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. Remote Sens. Environ. 124 (0), 516–533.
Pu, R., Liu, D., 2011. Segmented canonical discriminant analysis of in situ hyperspectral data for identifying 13 urban tree species. Int. J. Remote Sens. 32 (8), 2207–2226.
Pu, R., Landry, S., Zhang, J., 2015. Evaluation of atmospheric correction methods in identifying urban tree species with WorldView-2 imagery. IEEE Journal of selected topics in applied earth observation and remote sensing 8 (5), 1886–1897.
Puttonen, E., Litkey, P., Hyyppä, J., 2010a. Individual tree species classification by illuminated-shaded area separation. Remote Sens. 2, 19–35.
Ranson, K.J., Sun, G., 1994. Northern Forest classification using temporal Multifrequency and Multipolarimetric SAR images. Remote Sens. Environ. 47, 142–153.
Ranson, K.J., Saatchi, S., Sun, G., 1995. Boreal Forest ecosystem characterization with SIR-C/XSAR. IEEE Trans. Geosci. Remote Sens. 33 (4), 867–876.
Ranson, K.J., Sun, G., Kharuk, V.I., Kovacs, K., 2001. Characterization of forests in western Sayani Mountains, Siberia from SIR-C SAR data. Remote Sens. Environ. 75, 188–200.
Reitberger, J., Krzystek, P., Stilla, U., 2008. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. Int. J. Remote Sens. 29 (5), 1407–1431.
Riaño, D., Chuvieco, E., Condés, S., González-Matesanz, J., Ustin, S.L., 2004. Generation of crown bulk density for *Pinus sylvestris* L. From lidar. Remote Sens. Environ. 92, 345–352.
Ribeiro da Luz, B., 2006. Attenuated total reflectance spectroscopy of plant leaves: A tool for ecological and botanical studies. New Phytol. 172, 305–318.
Ribeiro da Luz, B., Crowley, J.K., 2007. Spectral reflectance and emissivity features of broad leaf plants: Prospects for remote sensing in the thermal infrared (8.0–14.0 μm). Remote Sens. Environ. 109, 393–405.
Ribeiro da Luz, B., Crowley, J.K., 2010. Identification of plant species by using high spatial and spectral resolution thermal infrared (8.0–13.5 μm) imagery. Remote Sens. Environ. 114, 404–413.
Rignot, E.J.M., Williams, C.L., Way, J., Viereck, L.A., 1994. Mapping of Forest types in Alaskan boreal forests using SAR imagery. IEEE Trans. Geosci. Remote Sens. 32 (5), 1051–1059.
Roberts, D.A., Green, R.O., Adams, J.B., 1997. Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS. Remote Sens. Environ. 62, 223–240.
Roberts, D.A., Ustin, S.L., Ogunjemiyo, S., Greenberg, J., Dobrowski, S.Z., Chen, J., et al., 2004. Spectral and structural measures of northwest forest vegetation at leaf to landscape scale. Ecosystems 7, 545–562.
Rohde, W.G., Olson Jr., C.E., 1972. Multispectral sensing of forest tree species. Photogramm. Eng. Remote. Sens. 38, 1209–1215.
Saatchi, S., Rignot, E., 1997. Classification of boreal Forest cover types using SAR images. Remote Sens. Environ. 60, 270–281.
Salisbury, J.W., 1986. Preliminary measurements of leaf spectral reflectance in the 8-14 μm region. Int. J. Remote Sens. 7 (12), 1879–1886.
Salisbury, J.W., Milton, N.M., 1988. Thermal infrared (2.5- to 13.5-μm) directional hemispherical reflectance of leaves. Photogramm. Eng. Remote. Sens. 54 (9), 1301–1304.
Sanchez-Azofeifa, G.A., Castro-Esau, K.L., Wright, S.J., Gamon, J., Kalacska, M., Rivard, B., Schnitzer, S.A., et al., 2009. Differences in leaf traits, leaf internal structure, and spectral reflectance between two communities of lianas and trees: Implications for remote sensing in tropical environments. Remote Sens. Environ. 113, 2076–2088.
Sandau, R., Roeser, H.-P., Valenzuela, A., 2008. In: Sandau, R., Roeser, H.-P., Valenzuela, A. (Eds.), Small Satellite for Earth Observation. Springer.
Sarrazin, M.J.D., van Aardt, J.A.N., Asner, G.P., McGlinchy, J., Messinger, D.W., Wu, J., 2011. Fusing small-footprint waveform LiDAR and hyperspectral data for canopy-level species classification and herbaceous biomass modeling in savanna ecosystems. Can. J. Remote. Sens. 37 (6), 653–665.
Sayn-Wittgenstein, L., 1978. Recognition of tree species on aerial photographs. Information Report FMR-X-118. Forest management Institute, Ottawa, Ontario.
Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S., Martonchik, J.V., 2006. Reflectance quantities in optical remote sensing – Definitions and case studies. Remote Sens. Environ. 103, 27–42.
Schmidtlein, S., Zimmermann, P., Schüpferling, R., Weiß, C., 2007. Mapping the floristic continuum: Ordination space position estimated from imaging spectroscopy. J. Veg. Sci. 18, 131–140.
Schmitt, M., Shahzad, M., Zhu, X.X., 2015. Reconstruction of individual trees from multi-aspect TomoSAR data. Remote Sens. Environ. 165, 175–185.
Shang, X., Chisholm, L.A., 2014. Classification of Australian native Forest species using hyperspectral remote sensing and machine-learning classification algorithms. IEEE J. Sel. Top. in Appl. Earth Obs. Remote Sens. 7 (6), 2481–2489.
Somers, B., Asner, G.P., 2014. Tree species mapping in tropical forests using multi-temporal imaging spectroscopy: wavelength adaptive spectral mixture analysis. Int. J. Appl. Earth Obs. Geoinf. 31, 57–66.
Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., Macomber, S.A., 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote Sens. Environ. 75, 230–244.
Stavrakoudis, D.G., Dragozi, E., Gitas, I.Z., Karydas, C.G., 2014. Decision fusion based on hyperspectral and multispectral satellite imagery for accurate forest species mapping. Remote Sens. 6, 6897–6928.
Stoffels, J., Sachtleber, T., Mader, S., Buddenbaum, H., Stern, O., Langshausen, J., Dietz, J., et al., 2015. Satellite-based derivation of high-resolution forest information layers for operational forest management. Forests 6, 1982–2013.
Suratno, A., Seielstad, C., Queen, L., 2009. Tree species identification in mixed coniferous forest using airborne laser scanning. ISPRS J. Photogramm. Remote Sens. 64 (6), 683–693.
Thomlinson, J.R., Bolstad, P.V., Cohen, W.B., 1999. Coordinating methodologies for scaling landcover classifications from site-specific to global: steps toward validating global map products. Remote Sens. Environ. 70, 16–28.
Treuhaft, R.N., Asner, G.P., Law, B.E., Tuyl Van, S., 2002. Forest leaf area density profiles from the quantitative fusion of radar and hyperspectral data. J. Geophys. Res. 107 (D21), 4568.
Trevett, J.W., 1986. Imaging Radar for Resources Surveys. Chapman and Hall, NY.
Ulaby, F.T., Sarabandi, K., McDonald, K., Whitt, M., Dobson, M.C., 1990. Michigan microwave canopy scattering model (MIMICS). Int. J. Remote Sens. 11 (7), 1223–1253.
Ullah, S., Schlerf, M., Skidmore, A.K., Hecker, C., 2012. Identifying plant species using mid-wave infrared (2.5–6 μm) and thermal infrared (8–14 μm) emissivity spectra. Remote Sens. Environ. 118, 95–102.
Ustin, S., Gitelson, A.A., Jacquemoud, S.M., Asner, G.P., Gamon, J., Zarco-Tejada, 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sens. Environ. 113, 67–77.
Van Ewijk, K.Y., Randin, C.F., Treitz, P.M., Scott, N.A., 2014. Predicting fine-scale tree species abundance patterns using biotic variables derived from LiDAR and high spatial resolution imagery. Remote Sens. Environ. 150, 120–131.
Van Zyl, J., 1989. Unsupervised classification of scattering behavior using radar polarimetry data. IEEE Trans. Geosci. Remote Sens. 22 (1), 36–45.
Vaughn, N.R., Moskal, L.M., Turnblom, E.C., 2012. Tree species detection accuracies using discrete point lidar and airborne waveform lidar. Remote Sens. 4 (2), 377–403.
Vauhkonen, J., Hakala, T., Suomalainen, J., Kaasalainen, S., Nevalainen, O., Vastaranta, M., Holopainen, M., et al., 2013. Classification of spruce and pine trees using active hyperspectral LiDAR. IEEE Geosci. Remote Sens. Lett. 10 (5), 1138–1141.
Vauhkonen, J., Ørka, H.O., Holmgren, J., Dalponte, M., Heinzel, J., Koch, B., 2014. Tree species recognition based on airborne laser scanning and complementary data sources. Forestry Applications of Airborne Laser Scanning, Managing Forest Ecosystems. Springer, Netherlands, pp. 135–156.
Voss, M., Sugumaran, R., 2008. Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach. Sensors 8 (5), 3020–3036.
Wagner, W., Hyyppä, J., Ullrich, A., Lehner, H., Briese, C., Kaasalainen, S., 2008. Radiometric calibration of full-waveform small-footprint airborne laser scanners. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 37, 163–168.
Wagner, W., Ullrich, A., Ducic, V., Melzer, T., Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS J. Photogramm. Remote Sens. 60 (2), 100–112.
Walsh, S.J., 1980. Coniferous tree species mapping using LANDSAT data. Remote Sens. Environ. 9 (1), 11–26.
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 (1), 76–85.
Waser, L.T., Klonus, S., Ehlers, M., Küchler, M., Jung, A., 2010. Potential of Digital Sensors for Land Cover and Tree Species Classifications - A Case Study in the Framework of the DGPF-Project. J. Photogramm. Remote Sens. Geoinf. Process. 2, 141–156.
Waser, L.T., Küchler, M., Jütte, K., Stampfer, T., 2014. Evaluating the potential of WorldView-2 data to classify tree species and different levels of ash mortality. Remote Sens. 6 (5), 4515–4545.
Wasser, L., Day, R., Chasmer, L., Taylor, A., 2013. Influence of vegetation structure on lidar-derived canopy height and fractional cover in forested riparian buffers during leaf-off and leaf-on conditions. PLoS One 8 (1). http://dx.doi.org/10.1371/journal.pone.0054776.
Wichmann, V., Bremer, M., Lindenberger, J., Rutzinger, M., Georges, C., Petrini-Monteferri, F., 2015. Evaluating the potential of multispectral airborne lidar for topographic mapping and land cover classification. ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. II-3/W5, 113–119.

Wollersheim, M., Collins, M.J., Leckie, D., 2011. Estimating boreal forest species type with airborne polarimetric synthetic aperture radar. Int. J. Remote Sens. 32 (9), 2481–2505.
Wolter, P.T., Mladenoff, D.J., Host, G.E., Crow, T.R., 1995. Improved forest classification in the Northern Lake States using multi-temporal Landsat imagery. Photogramm. Eng. Remote. Sens. 61 (9), 1129–1143.
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.
Youngentob, K.N., Roberts, D.A., Held, A.A., Dennison, P.E., Jia, X.P., Lindenmayer, D.B., 2011. Mapping two Eucalyptus subgenera using multiple endmember spectral mixture analysis and continuum-removed imaging spectrometry data. Remote Sens. Environ. 115 (5), 1115–1128.
Yu, X., Litkey, P., Hyyppä, J., Holopainen, M., Vastaranta, M., 2014. Assessment of low density full-waveform airborne laser scanning for individual tree detection and tree species classification. For. Trees Livelihoods 5 (5), 1011–1031.
Zhang, K., Hu, B., 2012. Individual urban tree species classification using very high spatial resolution airborne multi-spectral imagery using longitudinal profiles. Remote Sens. 4, 1741–1757.
Zhang, C., Qiu, F., 2012. Mapping individual tree species in an urban forest using airborne lidar data and hyperspectral imagery. Photogramm. Eng. Remote Sens. 78 (10), 1079–1087.
Zhang, C., Xie, Z., 2012. Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. Remote Sens. Environ. 124, 310–320.
Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S., Blackard, J.A., 2007. Remote sensing-based predictors improve distribution models or rare, early successional and broadleaf tree species in Utah. J. Appl. Ecol. 44, 1057–1067.
van Aardt, J.A.N., Wynne, R.H., 2007. Examining pine spectral separability using hyperspectral data from an airborne sensor: an extension of field-based results. Int. J. Remote Sens. 28 (2), 431–436.

Further reading

Puttonen, E., Soumalainen, J., Hakala, T., Räikkonen, E., Kaartinen, H., Kaasalainen, S., Litkey, P., 2010b. Tree species classification from fused active hyperspectral reflectance and LIDAR measurements. For. Ecol. Manag. 260 (10), 1843–1852.