3.08 Vegetation Structure (LiDAR)
M van Leeuwen and M Disney,
University College London, London, United Kingdom
© 2018 Elsevier Inc. All rights reserved.

3.08.1 Introduction

Advances in remote sensing technology since the mid-2000s have drastically increased the potential to acquire traditional forest inventory variables as well as information about the effects of forest canopy structure on radiative transfer, and its implications for tree and ecosystem physiology. Light detection and ranging (LiDAR) has provided means to record in high-resolution (cm-scale), the three-dimensional (3D) structure of the canopy, including height and leaf area density and in certain cases even information related to the branching architecture of trees (Wulder et al., 2012b; Dubayah and Drake, 2000). LiDAR data are currently being acquired from airborne and ground-based platforms with pulse repetition frequencies over 500,000 pulses per second. The many billions of returned laser pulses that reflect off the ground surface and other objects, such as trees, shrubs, and rocks, captured by the detector and for which the exact 3D positions are determined based on accurate measurements of the instrument’s position, are accessible in easy-to-use tabular data files, and can be easily displayed on modern graphics hardware through a handful of command line expressions or even user-friendly software packages, often available free of charge to academic and noncommercial users (e.g. SPDlib.org, see Bunting et al., 2013a,b; cloudcompare.org; rapidlasso.com). Over the past decades, the combined effort of a broad community of academic researchers and corporations has delivered a range of algorithms for the processing of these data, such as the extraction of terrain models and surface shapes, including stems and tree crowns, or wall-to-wall forest inventory attributes, including but not limited to basal area, merchantable timber volume, biomass, tree stocking density, and tree height. An overview of these algorithms and some of the accuracies by which forest attributes can be extracted from the data can be found in review articles and in the introductions of many research papers. A query on the Scopus search engine for academic research papers shows over 2000 hits on the topic of “LiDAR” and “forestry” alone (Disney, 2016), indicating both the great success of the technology in acquiring forest inventory information and the perceived value of the technology to the field.

Airborne LiDAR is generally viewed as a cost-effective means to collect forest inventory information over broad scales (e.g., regional to national scale). The data are often summarized into simple statistical descriptors such as percentile distributions of the return heights above the digital terrain model (DTM) and then stored in two-dimensional raster files, similar in representation to the raster files, which are used for passive-optical satellite or airborne data. The percentile distributions can then be used in linear regression models to estimate a range of inventory attributes and the model coefficients can be determined using training data sets of field measurements. This empirical method has been backed by many research articles from which a good sense of estimation accuracies can be derived for a range of forest types and attributes of interest. The empirical nature of this remote sensing method, however, does not warrant an accurate estimation of attribute values when the correlation models are used outside the calibration areas, or when other constraints (e.g., time, resources) mean that the homogeneity cannot be assumed or quantified across the study areas. Hence, these purely empirical approaches lack generality, predictive ability, and, most importantly, physically based parameters or process understanding.

The exact effect of parameter retrieval uncertainty on the cartographic information being produced is not always known in detail and it is not unusual that maps are short-lived and replaced by the latest insights as new remote sensing technologies advance. Nevertheless, the wealth of information provided by remote sensing enables decision makers and policy makers, more than ever before, to base their arguments on the latest information available. In many ways, forestry and forest conservation have been combined commercial and political efforts, where the collection of information over broad spatial scales has been a very challenging one until these recent advances in remote sensing emerged. In the early days of forestry practice, information was gathered through observers who would go out to the field and estimate attributes as forest stocking density and tree height by means of visual assessment (Kangas and Maltamo, 2006). The results were neither necessarily accurate nor necessarily worse than tape measurements or
angle gauges and finger counts. The emergence of remote sensing technology has greatly advanced the consistency and transparency by which attribute estimates can be acquired and used to guide decision making (Herold and Johns, 2007). A prediction accuracy of 60% may sound weak to some, but on the other hand, it can provide valuable insight for a decision maker who is confronted with the difficult, but necessary task of marshalling finite, scarce resources—for example deciding on areas to keep for recreational purposes while selecting others to exploit for resource extraction. Similarly, when timber properties can be predicted only with weak accuracy based on extrinsic observations of tree vigor and stem diameter or lean angle, significant value might arise from the allocation of downed trees to specialized sawmills.

The technical accuracy of LiDAR instruments may far exceed our ability to link point cloud data to forest attributes. In this case, a further reduction of measurement error would not greatly improve the estimation of biophysical or biochemical attributes. This leaves some questions around the need for repetition in studies on the potential of LiDAR remote sensing in forestry. While being the cornerstone of science, the repetition with which the suitability of LiDAR for forest attribute retrieval has been established seems to some extent excessive, and one might even expect that the explosive drift to conduct and report LiDAR studies that we have seen over the years, has had something to do with the emphasis that has been put on publication metrics. That said, significant research potential remains in exploring the use of LiDAR combined with other remote sensing technologies in studies that focus on the interaction between forest structure and physiological functions. For example, branching architecture can take on a range of angular form, planophile or erectophile or anywhere in between; canopies can be dense or sparse, and consist of mixed species or mono-cultures, and can be multilayered vertically. All of these traits have an impact on the radiation budget and hence, the photosynthetic and respiratory functioning of the canopy. While standard passive optical, multispectral, and hyperspectral remote sensing can provide coarse information about canopy structure, it can only generally do so for the upper layer(s) of dense canopies and requires empirical or process-based models to do so. LiDAR (and to some extent radar) is the only method able to provide near-direct measures of structural information that are easy to interpret and work with.

A significant source of high-resolution (cm to mm-scale) point cloud data is terrestrial laser scanning (TLS), including mobile or hand-held laser scanning (Caldern et al., 2015a; Disney, 2016). TLS allows trees to be scanned from multiple directions to increase the fidelity of the acquired data while reducing the effects of occlusion where multiple objects shadow one another. TLS is characterized by a much higher point sampling density and smaller laser footprint that is often of the order of millimeters. This allows a near-continuous sampling of surface geometry from which solid geometry models can be derived through subsequent processing. The increase in the utility of TLS facilitated the production of lower-cost instruments in the range of $10 K to $100 K, where cheaper models are often limited to recording pulses up to 30 m in range (Kelbe et al., 2015) and the more expensive scanners allow ranges of (up to) 100s m (Calders et al., 2016; Jupp et al., 2009). This distinction can be important when the emphasis is placed on an accurate acquisition of canopy structure. However, for plot-level characterization of stem attributes or for the autonomous, unattended acquisition of data that reveals structural change over time (Eitel et al., 2013) faster and more lightweight scanning devices prove highly efficient. Great advances in TLS processing have been achieved since the early 2010s, and include the extraction of branching architecture based on an application of graph network theory (Bucksch et al., 2010; Raumonen et al., 2013; Verroust and Lazarus, 2000; Runions et al., 2007; Côté et al., 2009, 2011, 2012; Hackenberg et al., 2015; Newnham et al., 2015). Also, the use of some high-resolution point cloud data for the assessment of leaf orientation angles (Zheng and Moskal, 2012) and marker-free coregistration of scans that exploits patterns in the positions of tree stems (Kelbe et al., 2016, 2017) are among the many examples that have advanced the state of art of TLS in forestry. Ongoing efforts, originating from the 1990s, include the use of returned laser intensity information to approximate the apparent surface reflectance (Strahler et al., 2008; Jupp et al., 2009), which in the light of dual-wavelength laser scanners (Douglas et al., 2012; Danson et al., 2015) introduces exciting new potential to derive 3D, geometrically explicit maps of vegetation indices that may reveal vital information about attributes such as leaf moisture (Gaulton et al., 2013).

The remainder of this chapter will explore the various uses of LiDAR remote sensing and reviews some of the potentials that it holds for forestry and tree physiology. First, we discuss temporal and spatial domains in which LiDAR remote sensing offers new opportunities, ranging from the global scale to the stand and plot levels and further down to the scale of individual trees and tree crowns. Following this, we discuss matters of structural assessment as well as how canopy structural information can be used to study the radiation budget and impacts on physiological functions. Finally, we present a brief review of potential synergies arising from a combination of LiDAR technology and other remote sensing technologies.

### 3.08.2 LiDAR: The Technology

The development of LiDAR scanning has resulted from the confluence of a number of technologies, including lasers, inertial navigation, global positioning, as well as timing electronics and signal digitization at very high sampling speeds. The effective use of LiDAR in forestry may date back to the beginnings of the 1980s (Lim et al., 2003). Some initial efforts have focused on the acquisition of LiDAR data from satellite instrumentation for monitoring of the Earth surface from space. The NASA IceSat/Geoscience Laser Altimeter System (GLAS) project collected surface topographic data in addition to cloud height information from January 2003 until its failure in 2010. However, data collection was always intermittent due to issues around the powering of the laser
source (Breon et al., 2005). The GLAS laser footprint was approximately 60 m on the ground, with each footprint separated along
the track by around 1 km, and with no side scanning. At this resolution, extraction of attributes such as tree height can be chal-
lenging when the underlying terrain is sloped (Sun et al., 2008; Harding and Carabajal, 2005). GLAS was not designed for forest
applications, but proved extremely useful for mapping canopy height, in flat, tropical, and boreal regions (Saatchi et al., 2011;
Baccini et al., 2012). Perhaps more convincing evidence of the power of LiDAR for canopy applications has come from airborne
data collection campaigns. Flown at around 1000 m above ground, the laser footprint may be somewhere around 30 cm or less,
with points collected at scan rates of 10s of kHz, resulting in dense point cloud data that are visually pleasing and immediately
speaks to the imagination, even of those with an untrained eye: Terrain and build-up structures such as roofs, roads, and trees
are immediately recognized.

Some of the more recent contributions include the collection of LiDAR data from unmanned airborne systems (UAS) and hand-
held (and carried) devices where an operator simply sweeps or walks the instrument around to collect data in areas that would
otherwise be hard to cover with conventional remote sensing platforms (e.g. Lin et al., 2011; Bosse et al., 2012; Parker et al.,
2004). Much is also to be expected from the automotive industry where self-driving vehicles require real-time acquisition of the
surrounding 3D world. These demands are likely to drive down both the cost as well as the size of laser scanning equipment.

Some evolution in LiDAR scanners can also be discerned in the technique by which individual returns are registered, although
different techniques may continue to coexist due to trade-offs in cost and utility value. Initially, commercial scanners were equipped
to digitize the incoming waveform of laser intensity as a few discrete pulses. Different techniques for registering discrete returns exist,
including time-of-flight and phase-based scanning, where the latter is faster and often cheaper but suffers the problem of ‘ghosting’
when targets overlap along the photon path. Many “research-grade” LiDAR instruments are equipped to record the entire waveform.
The full-waveform scanners are more expensive due to data storage requirements. Discrete scanners only record a range distance
when the returned laser intensity exceeds a set threshold. Upon the registration of a return, the instrument electronics require
time to process the returned signal and, as a result, the discrete-return LiDAR scanners suffer from an instrument-dead time
when two targets are within short range of one another (Armston et al., 2013). The full-waveform data are more challenging to
process and analyze, typically requiring a deconvolution step where the entire waveform is described as a superposition of Gaussian
distributions. The full-waveform scanners have the added benefit that by integrating waveform-intensity values over a range, an esti-
mate of target reflectance can be made given assumptions about the backscattering cross-section and normal angle of the intercepted
targets (e.g. Jupp et al., 2009; Cawse-Nicholson et al., 2014). It is worthwhile mentioning that most commercial airborne scanners
now have full-waveform recording capabilities.

A newer type of LiDAR scanner comprises the single photon LiDAR that registers incoming light more efficiently and, hence,
can be used to acquire denser point clouds than the more traditional discrete-return or full-waveform scanners offer. Single
photon LiDARs can also operate at a higher pulse repetition frequencies and are said to bear some potential to penetrate
ground fog and thin cloud layers. The precise significance of the single photon LiDAR technology to forestry applications
remains a topic of ongoing research and efforts are required to filter these data from noise induced by solar radiation (Swatantran
et al., 2016).

### 3.08.3 Global Scale

Mapping forest canopy characteristics at a global scale is beneficial to sustainable resource management as well as climate science.
Notably, one instrument, the GLAS instrument aboard the ICESat satellite, has been used to infer tree height information world-
wide. Due to its discontinuous sampling pattern along tracks (i.e., spaced 170 m along the track and several kilometers across track),
studies have generally focused on the fusion of GLAS data with other Earth observation data. For example, Lefsky (2010) used the
Medium Resolution Imaging Spectroradiometer (MODIS) to extrapolate tree height estimates that were derived from GLAS foot-
print data. The MODIS data were segmented based on their spectral and textural attributes and a regression model was calibrated
to relate spectral and textural indices derived from the image segments to the waveform-derived height estimates. The latter esti-
mates were obtained from coincident GLAS footprints and tree height measurements from select field sites. Simard et al. (2011)
produced a similar map based on extrapolated GLAS tree height estimates using 1-km MODIS data and ancillary climatology
data (Baldocchi et al., 2001). Field mensuration data was gathered from 66 FLUXNET sites that are distributed globally for the
collection of micro-meteorological data. Simard et al. found that the former map by Lefsky generally produced taller estimates
compared to theirs for regions that are characterized by a sloping terrain, whereas on flat terrain the latter map produced generally
taller estimates. A cross-validation where Lefsky’s tree height estimates were compared against the 66 validation sites used by Simard
et al., and the authors found no correlation ($r^2 = 0.01$) at all, shedding doubt on the map quality. Authors have emphasized that
estimation errors may be large over sloping terrain and for complex or dense forest canopies and processing methodologies may
require sophisticated techniques to decompose the waveform (Wang et al., 2016).
Despite the large uncertainties around global height estimates, some authors have advocated the use of such maps with inference techniques to estimate global biomass (Hese et al., 2005; Hu et al., 2016). For example, Hu et al. (2016) used random forest classifiers and ancillary EO data to extrapolate GLAS footprints and produce a global map of waveform parameters. Using a Monte Carlo simulation (n = 100), regression models were used to link biomass reports from literature studies to the spatially continuous waveform-parameter and other EO data. This produced 100 aboveground biomass maps from which mean and standard deviations were derived. Similar ensemble methods have often been demonstrated to be successful in remote sensing literature; However, cross-validation results often indicate regression coefficients around 50% and it is not always clear whether such global products are rightfully characterized as "globally consistent" (e.g. Healy et al., 2012) or whether further research into the overall map quality is required before such predication can be assigned. For example, the study by Hu et al. (2016), reported an $R^2$ and RMSE from selected validation sites of 56% and 87.53 Mg/ha (with a mean global above-ground biomass estimate of 210.09 Mg/ha), while at the ecozone level the $R^2$ was even reduced to 0.38%.

Much more effort into the development and accessibility of LiDAR instruments and data seems to be needed to deliver global tree height and biomass information with an accuracy that is relevant in the light of monitoring and reporting commitments such as within the framework of the United Nations’ Reduction of Emissions from Degradation and Deforestation (UN REDD+). Uncertainty also exists around the extrapolation of allometric relationships that is required to produce global biomass estimates. Due to practical reasons, allometric relationships have predominantly been derived among smaller trees and only from partial tree measurements, adding great uncertainty around the applicability of these relationships for taller trees (Disney et al., 2016). Proposed missions include the ICESat-2 and the Global Ecosystem Dynamics Investigation (GEDI) that both provide promising venues for global forest mapping efforts (Qi and Dubayah, 2016; Moussavi et al., 2014).

### 3.08.4 Regional to Stand Level

LiDAR acquisitions at regional levels currently fall entirely within the domain of airborne remote sensing; historical acquisitions of LiDAR data from orbit are, of course, still available from the IceSat/GLAS instrument. Regional, wall-to-wall acquisitions are often conducted with an interannual regularity as part of a national effort to monitor the change of the visible landscape, including land-use change, tectonic movements, and the integrity of natural or man-made structures. Monitoring efforts may also include the regular safeguarding of the integrity of power lines or ensuring that tax is paid at household level for structures such as sheds and garages that were built as add-ons to existing real estate. It is within these overarching monitoring programs that information about the land surface can be collected in a cost-effective manner and be made available to the various domains of investigation. It is through the same collaborative effort that goes into the collection of the data, that researchers have been able to learn about the means to process LiDAR data to useful end products. One of the most elementary of approaches is to aggregate LiDAR returns in percentile distributions for use in linear regression models to predict the forest inventory attributes of interest (see e.g. Naesset, 2014, and references therein). Of these percentiles, the median, 75%, 95%, and 99% percentiles have been proven most useful. While the median is strongly correlated with the transparency of the canopy medium to light propagation and relates to inventory attributes such as biomass or stocking density, the upper percentiles are used to estimate tree height or Lorey’s mean height, defined as the average height of all trees in the stand weighted by their basal area. The upper percentiles remove false returns from the airspace above the trees that may result from the interception of laser light by avifauna or an otherwise undesired behavior of the digitizing electronics inside the LiDAR scanner.

Accurate retrieval of tree height, naturally, also requires that terrain elevation can be modeled with a suitable fidelity. Radar-derived terrain maps such as the Shuttle Radar Topography Mission (SRTM; Farr et al., 2007) provides terrain topology at a 30 m or 90 m resolution, which may not be adequate for many forestry applications. Besides, detailed information about the terrain underneath the forest canopy can often be extracted from LiDAR data. Perhaps the most challenging step in producing DEMs from LiDAR data is the separation of returns into ground and nonground hits (Liu, 2008); however, in order to speed up processing or to facilitate processing on inexpensive computer hardware, reduction of data dimensionality is of critical importance (Isenburg et al., 2006). Effects of forest structure and acquisition parameters on DEM accuracy have been reported in a number of studies. For example, Reutebuch et al. (2014) reported DEM errors of around 20–30 cm in heavily forested areas. Khotrevipour et al. (2015) investigated effects of terrain slope on treetop detection and found that treetops were best identified on the raw (i.e., not normalized with respect to terrain elevation) LiDAR data. Bater and Coops (2009) investigated a number of different DEM extraction techniques and found that natural neighbor interpolation produced superior results against linear, quintic, spline-based and finite difference methods, while Su and Bork (2006) found inverse distance weighing to produce superior results over kriging and spline-based interpolation for a number of different vegetation types in Alberta, Canada.

This statistical approach to LiDAR analysis has proved widely useful and much of the literature to this end focuses on the sampling designs needed to balance estimation accuracy with the costs related to field-data collection. These issues are discussed in Maltamo et al. (2014) and Van Leeuwen and Nieuwenhuis (2010) among others. Significant early collections from research instruments include the NASA campaigns conducted with the Land Vegetation and Ice Sensor (LVIS; Blair et al., 1999), and the Scanning LiDAR Imager of Canopies by Echo Recovery (SLICER; Harding et al., 2001). A low-cost profiling instrument, the Portable Airborne Laser System (PALS), was designed and employed by Nelson et al. (2003). Commercial scanners evolved rapidly as a result of the growing demand for high-resolution data. Europe proved an important consumer of such high-resolution data and some of its smaller countries were the first to adopt nation-wide, interannual, wall-to-wall LiDAR surveys, for a range of applications, not just
forestry. For example, between 1997 and 2003, the Netherlands started a program to acquire national LiDAR data with point sampling densities around 1 point per 16 square meter, which around 2009 was being replaced by a second survey with over 10 points per square meter (http://www.ahn.nl/). Initiated by the national office for dike safety and water affairs (Rijkswaterstaat), these data were made available to a wide group of users including landscape designers and nature conservationists. The acquisition of interannual LiDAR data facilitates the modeling of land surface change. However, initial attempts with early commercial instrumentaton were not without problems. The low point sampling density in combination with the often unknown or proprietary characteristics of the individual scanners used (Disney et al., 2010; Armstrong et al., 2013) meant that temporal changes in the data could not be attributed exclusively to land surface changes, but often to differences between scanners and even survey characteristics. These latter effects can be mitigated by using higher ground-sampling densities, such that details in terrain elevation can be more clearly distilled from the point cloud data. A more insightful account of the effects of aging on the structure of forests might be derived when the same instrument is flown across a natural gradient of forests of different ages and developmental stages or when forest growth is evaluated against a gradient of environmental conditions. An example of this type of campaign has been provided by Wulder et al. (2012a) where a 25,000 km trajectory was flown across the Canadian boreal shield, providing an excellent source of data to study gradients in forest structure across the landscape and how these gradients may relate to other variables of interest such as primary productivity and stand recovery after a disturbance event (Bolton et al., 2013).

Airborne LiDAR remote sensing in forest research is not limited to the characterization of traditional inventory attributes or growth rates. Studies have investigated the potential to also predict species or species groups on the basis of laser intensity data (Korpela et al., 2010; Ørka et al., 2009). Perhaps not surprisingly does availability of leaf-on and leaf-off LiDAR intensity data (i.e., relating to different seasons) improve the capacity to discriminate between evergreen and deciduous species, even though effects of scans angle and flying altitude may have to be minimized through calibration procedures (Kim et al., 2009). Perhaps greater use is derived from merging spectral information from imaging spectroscopy with additional information on the tree crown shape, either to delineate individual tree crowns (Baldeck et al., 2015) or to provide information about branching heights and crown radii (Brandtberg et al., 2003; Holmgren and Persson, 2004). Fusion of LiDAR with radar data has also successfully been demonstrated as a means to estimate vegetation biomass. For example, Mitchard et al. (2012) used radar-derived products to classify a 5000 km² area in Gabon into vegetation classes for which average biomass estimates were assigned using LiDAR and field data, while Tsui et al. (2013) used kriging, cokriging, and regression kriging to predict aboveground biomass across a 25 km² area in British Columbia with combined LiDAR transects and radar datasets, and found synergistic value arising from the frequency of observations and all-weather capacity of the radar data, with potential to reduce cost of flight campaigns or the need for a complete wall-to-wall coverage of LiDAR data.

LiDAR has also been used to predict wood quality attributes (e.g. Hilker et al., 2013). The rationale behind this endeavor is the linkage between stem attributes, visible on the outside of the bark or from the social status of trees within the stand, and intrinsic wood attributes such as ring width, microfibril angle or lignin content, all of which affect the mechanical properties of the wood. While many of these relations have long been known in silviculture, many studies have also reported contradicting results (Van Leeuwen et al., 2011). Some of these disagreements potentially arise from the limitations in the sampling design or the fact that it is simply not feasible to compile a comprehensive database of wood fiber properties large enough to accurately cover the variation within trees as well as throughout the landscape. The wide-scale acquisition of LiDAR data offers a unique potential to characterize stands and even individual tree structural properties. By linking these data to quality measurements taken at the receiving mills, it may become feasible to greatly enhance the study of relations between wood fiber traits and stand location.

### 3.08.5 Plot to Tree Level

An important aspect of the success of LiDAR for vegetation applications is the range of spatial scales it covers. The technology provides for wide-area coverage, but at the same time, each laser return represents a footprint area that can be as small as a few centimeters in diameter. As a result, contours of individual tree crowns in airborne LiDAR data sets can often be matched by field observations, including from TLS. Fig. 1 illustrates the match than can be seen between two independent acquisitions of point cloud data: an airborne LiDAR point cloud and its derived canopy height model (CHM; in blue) and a terrestrial laser scan and its derived CHM (in green). The match between the collocations is clearly visible and any discrepancy between the crown outlines is, besides mis-registration, at least partially due to the movement of the crowns in the wind, which can amount to several meters depending on weather conditions.

Traditionally, ground-truth data are required to calibrate and validate models that relate remotely sensed reflectance (LiDAR, optical) or back-scattered signals (radar) to the inventory attributes of interest. TLS has drastically increased the possibilities for calibration and validation of remote sensing data products by facilitating the mapping of 3D forest structure. A good overview of TLS applications is provided by Liang et al. (2016) which includes works related to forest resource management as well as ecology. Some of the earliest uses of TLS in forestry were demonstrated by Lichti et al. (2002) and Hopkinson et al. (2004). Popular applications included the retrieval of stem diameter and taper and this has since remained an area of interest for its significance to the industry as well as biomass modeling (Bienert et al., 2006; Gorte and Winterhalder, 2004; Pueschel et al., 2013; Kelbe et al., 2015; Calders et al., 2015b). Recent extensions of this work include an intercomparison of stem retrieval algorithms, led by the Finnish Geodetic Research Institute (http://www.eurosdr.net/). A major challenge in developing these algorithms is the need for a robust treatment of outlier detection, due to the potentially large amounts of occlusion and scattering by leaves, needles, and branches encroaching.
into the region of interest. For this, hypothesis-generating-and-testing approaches are often utilized, such as the Hough transform (Hough, 1962) and Random Sample Consensus (RANSAC; Fischler and Bolles, 1981). These approaches rely on a certain implementation of a voting procedure where the individual data samples, or laser returns, cast a vote over the plausibility of a previously generated hypothesis or set of hypotheses. Least squares approaches may be used to fit stem diameters after outliers have been removed (Liang et al., 2012). A reliable retrieval of the complete branching structure proved difficult with early scanners as a result of the low point sampling densities. Côté et al. (2009) demonstrate the reconstruction of branching topology based on Dijkstra’s algorithm. Given a root node, the algorithm searches for the shortest path between points in the point cloud and the root node. By binning the paths together, the authors were able to derive a branching topology. Especially for fine twigs, the algorithm does not necessarily find the correct branching topology, but one that is demonstrated to be close to the real topology. Others, including Livny et al. (2010) and Raumonen et al. (2013), demonstrated similar techniques based on graph network algorithms. Other than for simple inventory studies, the recovery of branching topology requires high-resolution scans and, hence, a denser spacing of scanner locations to acquire the data. Since the point cloud data is often too sparse to reliably detect fine branches, the reconstruction of tree branching structures is typically characterized by two stages: one where the major branches are detected and a second step in which small twigs are added based on a space colonization algorithm (Runions et al., 2007) or simply by imputing crown structures that were derived with the use of tree modeling software (Morton et al., 2016; Van Leeuwen et al., 2013; Weber and Penn, 1995; Pradal et al., 2009). Space colonization follows and extends the work of Prusinkiewicz and Lindenmayer (2004) who investigated the occurrence of fractals, or self-similar mathematical patterns, in plant architectures. The reconstruction of individual trees has grown out to the reconstruction of plots of increasingly larger size. Perhaps the largest to date has been achieved in Wytham Woods, Oxford, UK (Calders et al., 2016) where an area of 300 m by 200 m has been scanned and reconstructed. Scan locations were laid out in a 20 m by 20 m regular grid and were taken with a RIEGL VZ-400 at an angular interval of 0.04° in both azimuth and zenith directions (www.riegl.com). Sites of this kind are crucial for the validation of theories and models around canopy radiation transfer and provide a way for scaling between leaf-level physiology and stand-level, top-of-canopy reflectance signals as they are observed by remote sensing platforms. This is achieved using rendering applications such as ray tracing and path tracing (Disney et al., 2000; Govaerts and Verstraete, 1998; Goodenough and Brown, 2012). Sites scanned in detailed 3D using TLS like this facilitate an end-to-end traceability chain of radiative and physiological quantities. The reconstructed 3D canopy architecture can be used directly in radiative transfer schemes and growth models. However, these digitized representations of the real stand are not only significant as surrogate “truth” against which radiative transfer models can be validated; perhaps more important is the potential to modify the 3D model scene and, hence, the ability to predict changes in top-of-canopy reflectance resulting from the particular structural or spectral attributes (Yao et al., 2016). After all, and despite the high level of realism that can be achieved using modern scanning and computer graphics techniques, a 3D model derived from laser scanning may never be interpreted as being a copy of the real forest. Many structural details must be simplified or even ignored, such as sunfleck dynamics caused by the swaying of branches and leaves in the wind. While the significance of these effects on modeling results is unclear, they can potentially be tested and quantified, if required. As a result, uncertainties in the radiative transfer model predictions can be traced explicitly and there is great potential to study the validity of various modeling assumptions across a range of different plant architectural types or across spatial scales ranging from leaf to stand level.

Fig. 1 Illustration of coregistered point clouds that were acquired from the ECHIDNA® terrestrial laser scanner (green) and an airborne LiDAR system (blue). Corresponding canopy height models are shown in matching colors. Data was acquired from site DF49, Vancouver Island, BC, 2008.
Early contributions to TLS include also the development of the ECHIDNA® laser scanner, by the Commonwealth Scientific and Industrial Research Organization (CSIRO), Melbourne, Australia. The ECHIDNA® scanner features a relatively broad-beam laser and full-waveform recording capabilities. The wide laser footprint facilitates capturing multiple returns, a feature principally used to derive a gap probability distribution as a function of range, from which a foliage profile with canopy depth can be computed. Given assumptions around the leaf angle orientation distribution and reflective (transmissive) properties, the returned light intensity at the specified range that is recorded by the instrument is integrated into a postprocessing stage and is used to derive the fraction of uncollided beam as a function of range, or the gap probability. This measure is calibrated using laser footprints in open sky and those fully intercepted by tree stems and the hemispherical sampling pattern of the instrument facilitates that gap probability as a function of range can be specified for different zenith angles. Since the gap probabilities for different leaf angle distributions are similar around the 57-degree zenith angle (Welles and Norman, 1991), this portion of the scan is often used to derive a vertical profile of the gap probability and, from this, a vertical foliage profile (Jupp et al., 2009). The hemispherical data also provide for the quantification, in angular terms, of the gap size distribution that can be related to the foliage clumping index (Zhao et al., 2012) that is unity when foliage is dispersed randomly throughout the canopy space and is within the range zero to one for clumped canopies. The gap size distribution has earlier been used to quantify clumping (Chen and Cihlar, 1995) using transect measurements and hemispherical photography (Leblanc et al., 2005), however, the availability of range information provides, in principle, for analysis of any vertical stratification in foliage clumping.

The DWEL (Dual-Wavelength Echidna LiDAR) is a follow-up of the ECHIDNA instrument and features two laser sources operating at different wavelengths and is principally being developed by Boston University in collaboration with CSIRO. While the first band is located at 1064 nm the second band is located at 1548 nm and is used to differentiate water absorptions such that woody and foliage material can more effectively be discriminated (Douglas et al., 2012, 2015). Another dual-wavelength TLS system, the Salford Advanced Laser Canopy Analyser (SALCA) has been under development by Salford University (Danso et al., 2014) and features similarly positioned wavebands (centered at 1063 and 1545 nm, respectively) and studies have demonstrated both the calibration of reflectance as well as the detection of water stress from data acquired with the SALCA scanner (Gaulton et al., 2013).

### 3.08.6 Simulation

The ability to describe individual crown outlines facilitates modeling of intercrown shading and its effects on remote sensing observations. Hilker et al. (2008, 2012) used a hill-shade model derived from a CHM and estimated the shadow fraction within the field of view of a tower-mounted revolving spectrometer. This shadow fraction was used to correct for the effect of a changing illumination environment on the state of the photochemical reflectance index that relates to the light-use efficiency of photosynthesis (Gamon et al., 1992). While a hill-shade model does not provide insight into the effects of intracrown shading, these effects can be included if a suitable tree architectural model were to be used at the respective treetop locations. Instead of using an opaque surface, the shading would then be derived from a porous model of canopy structure that matches, at a minimum, the topology of the top of the canopy and the constellation of dominant trees in the stand. This might also include the morphology of individual branches where such information is available. To illustrate, while a typical workstation might have registered a maximum of 32GB of RAM in 2008, at the time of writing (2017) we find desktop workstations with up to 1 TB RAM. As a result, increasingly complex vector models can be handled with ease to allow a separation of foliage into sunlit and shaded leaves as well as how each of these fractions are visible to a remote sensor. Examples of such efforts can be found in ray tracing (path tracing) literature, where paths of individual rays of light are followed inside virtual scenes that comprise a geometrically explicit representation of the forest canopy, including the position, orientation, and shape of leaves, needles and branches. Upon interaction with objects within a 3D scene, rays are either scattered (reflected, transmitted) or absorbed based on the spectral composition of the illumination and material properties. Ray tracing simulations can be used to compute angular dependence in top-of-canopy reflectance (Widlowski et al., 2013, 2015; Lewis, 1999; Gastellu-Etchegorry et al., 2015) as well as canopy transmission (Yao et al., 2016) and interception (Van der Zande et al., 2011) and can be used to calibrate or validate other models that link remote sensing observations to surface characteristics. While the virtual, 3D scenes may not represent the true structure and spectral composition of any real forest, they have become increasingly accurate and sophisticated. These models provide excellent means to study issues of scale in radiative transfer theory and have therefore primarily been used for the validation and calibration of other models using a series of well-documented benchmark scenes (Widlowski et al., 2013, 2015). Fig. 2 shows an example of the simulation of surface reflectance from a vector model of a conifer and illustrates how shadow effects appear. Path tracing differs from ray tracing in the way splitting of the scattering photon packets is achieved. While ray tracing implements an actual branching of the photon packet in multiple directions, a path tracer samples an individual direction, which minimizes the number of secondary rays (i.e., those that propagate after a first intersection with the scene) that need to be queued for evaluation and this proves beneficial in simulations in the near-infrared where scattering is significant.

The ability to model shadow fractions across multiple scales simultaneously opens new opportunities for the design and parameterization of vegetation models. Multiresolution models reduce the need for clumped parameters, hence the effect of different morphological traits on physiological functioning can be studied in greater detail. This provides for the study of effects of shoot morphology on the radiation budget or for a decomposition of stand-level productivity into the contributions made by different canopy height strata (e.g. Van Leeuwen et al., 2013) but this remains an area of ongoing research. The virtual scenes as such act as laboratories where parameter sensitivities can be traced across different spatial scales (Lewis, 1999). However, parameter
uncertainties derived from these virtual models are to be interpreted with care and the high level of realism of these virtual scenes can perhaps encourage overconfidence in the verisimilitude of these scenes; validation is ever more important. Since these virtual scenes can be developed from TLS point clouds that were already acquired with state-of-the-art laser scanning equipment, a rigorous means to validate these scenes is often not available. Moreover, many scattering characteristics may be excluded, such as specular scattering effects, fine surface texture and scattering within leaves or needles. While ray tracing (or path tracing) does not necessarily neglect such effects, computational tractability often requires that certain details about scattering characteristics must be omitted. This raises questions as to how confidence intervals around retrieved model parameters ought to be interpreted.

When models are used for inference, a model state serves as an estimate of reality (i.e., system state), but of course is in itself not an attribute of reality. Similarly, a confidence interval should be regarded as an attribute of the model; provided that the modeling assumptions are valid, this confidence interval would coincide with the true distribution of (real-world) system states. However, it is inconsistent to assume that both a 1D layered representation as well as a detailed multiscale 3D model bears the same relation to reality as may be implicit in the underlying modeling assumptions. Thus, while laser scanning has drastically increased the level of realism by which radiative transfer models can be parameterized, more work remains in assessing the accuracy of scene representation relative to the real forest stands they resemble. Validation efforts may often have to resort to quantifying morphological and geometric discrepancies between the algorithmic reconstructions from point cloud data versus manual reconstructions conducted by experienced researchers (Boudon et al., 2014).

3.08.7 Data Fusion

Considerable effort has also been paid to the fusion of LiDAR remote sensing data with other remotely sensed data, notably hyperspectral. For example, the Carnegie Airborne Observatory (CAO) team lead by Gregory Asner has demonstrated significant potential to simultaneously acquire biochemical and structural information about the forest landscape (Baldeck et al., 2015; Asner et al., 2015). Other recent efforts in the development of multisensor airborne mapping systems that include LiDAR have been made by the US-based National Ecological Observation Network (NEON) (Kampe et al., 2010); the National Aeronautics and Space Administration (NASA) with the G-LiHT instrument that combines a hyperspectral, a LiDAR, and a thermal imager (Cook et al., 2013); and the UK-based Natural Environment Research Council (NERC) with the Airborne Research Facility (ARF) (www.bas.ac.uk). The simultaneous access to structural and spectral information allows modelers to place further constraints on radiative transfer parameterizations. This information can be used, for example, to discriminate between ground-level, understory, and overstory reflectance which may be impossible based on either hyperspectral or LiDAR data alone (Koetz et al., 2006, 2007). Other
applications may be to improve classification using information about stand height and canopy volume profiles (Jones et al., 2010), or improve timber volume estimates (Breidenbach et al., 2010). Much is yet to be learned, however, about techniques to invert radiative transfer models using combined hyperspectral and LiDAR data. The use of multispectral LiDAR scanning systems could greatly facilitate biophysical characterization as the spectral information is then stored as attribute of the point cloud (or waveform) data, such that effects of spectral mixing at the pixel level are better understood (Woodhouse et al., 2011; Morsdorf et al., 2009; Hakala et al., 2012). The broad-scale coverage of fine-resolution data facilitates the estimation of surface characteristics with a marked multiscale dependence. This may be beneficial for highly heterogeneous landscapes or regions with very high biodiversity (Ter Steege et al., 2013; Anderson et al., 2004). Climate-related research into vegetation response requires a characterization of biogenic emissions from the land surface. However, nitrous oxide (NOx) and methane fluxes remain challenging to estimate with a desirable accuracy since both are strongly dependent on gradients in terrain elevation and soil ecology and hydrology (Bridg-ham et al., 2013). The use of spectral and topographic information, in tandem, could possibly be used to infer spatial and seasonal patterns in soil hydrology, with relevance to the modeling of fine-scale biogenic fluxes (Mohammed, 2015). Similarly, when remote sensing observations over forest canopies are aggregated to a coarse resolution (e.g., > 100 m) and then used for the estimation of primary productivity, the underlying physiological processes of photosynthesis and respiration may not be understood in detail even if the computed estimates match validation data closely. The reason is that any model, whether process-based or empirical, has at least some potential to emulate a set of observations for reasons other than a correct understanding of the underlying physical processes—the right answer but for the wrong reasons. Indeed, many detailed physical and physiological underpinnings are necessarily avoided in models that are to be evaluated at regional to global scales. A proper estimation framework may therefore require not only a coupling of models of various scale, such as demonstrated in many process-based models of stand productivity (e.g. Ibrom et al., 2006; Medlyn et al., 2002; Wang and Jarvis, 1990) but also a translation of these process-based models to empirical “drop-in” functions that are able to mimic the overall model behavior at a lower computational cost (Gomez-Dans et al., 2016). This allows the complexity of the process-based models to be distilled, to reduce the number of model parameters needed, and often speeds up computations by several orders of magnitude, while the emulated model is approximated with a known tolerance. Efforts to this aim seem increasingly beneficial for the assimilation of knowledge and observations from a diverse set of disciplines and it is by combining knowledge from these various disciplines that progress in understanding the Earth system can be achieved.

3.08.8 Conclusion

This chapter summarizes research in LiDAR remote sensing of vegetation structure. The breadth of work conducted in this area has grown significantly since the early 2010s, indicating both the development and expansion of remote sensing technology as well as the challenge in applying it. As a direct measurement technique, LiDAR data are more accessible and easier to handle than many other remote sensing products. LiDAR allows for new development and application of both empirical and process-based models. The latter have the benefit that site-specific calibrations may not be needed, possibly reducing costs associated with the data acquisition campaigns. In addition to providing near-direct measurements of vegetation structure, LiDAR has also been used to provide indirect measures related to biodiversity, energy balance, and biogenic fluxes. The coming years will most certainly see continuing progress in the use of LiDAR data as newer systems become available, including multiwavelength and photon-counting LiDAR systems. Future datasets are likely to provide ever-increasing point densities, allowing new research questions to be addressed. Many of the future opportunities and applications are likely to involve data fusion and integration, as well as the improved calibration and validation of other remote sensing data products. Much of the knowledge we gained over the last decades includes processes at microscopic and macroscopic scale, whereas the intermediate range has been paid comparatively little attention (Passioura, 1979). With the technologies that have become available to the field of remote sensing, an ever-increasing potential is anticipated to further integrate the knowledge that has been gained from both microscopic and macroscopic domains, hopefully throwing light (literally and metaphorically) on this intermediate region.

See also: 1.15. Lidar Sensors From Space.

References


