Advances in forest characterisation, mapping and monitoring through integration of LiDAR and other remote sensing datasets

Richard Lucas¹, Alex Lee², John Armston³, Johanna Breyer¹, Peter Bunting¹ & João Carreiras⁴

¹Institute of Geography and Earth Sciences, Aberystwyth University, Aberystwyth, Ceredigion, SY23 2EJ rml@aber.ac.uk
²Defence Imagery and Geospatial Organisation, Russell, Canberra, ACT 2600, Australia. Alexander.Lee@defence.gov.au
³Joint Remote Sensing Research Program, Remote Sensing Centre, Department of Natural Resources and Water, Climate Building, 80 Meiers Road, Indooroopilly, Queensland, 4068, Australia john.armston@nrw.qld.gov.au
⁴Tropical Research Institute, Department of Natural Sciences, Rua João de Barros, 27, Lisbon, Portugal, jmbcarreiras@gmail.com

Abstract

The diversity of scales and modes in which ground, airborne and spaceborne LiDAR operate has increased opportunities for quantitatively assessing forest structure, biomass and species composition and obtaining more general information on dynamics and ecological/commercial value. However, the level of information extracted can be increased even further by integrating data from other sensor types, including hyperspectral and Synthetic Aperture Radar (SAR). Examples include the generation of species-specific tree and stand level maps of biomass through inclusion of fine spatial resolution hyperspectral data and the use of LiDAR data and derived products for better interpreting the information content of SAR and optical data and parameterising models that simulate and assist understanding of the interaction of electromagnetic energy with forest components. Applications where synergistic use of LiDAR and other remote sensing data are advantageous include commercial forest inventory, quantifying carbon dynamics and biodiversity, and detecting change at scales from individual trees to landscapes. Recognition of the value of integrating other forms of remote sensing data with LiDAR is leading to the development of techniques for data fusion and also new synergistic sensors on platforms ranging from Unmanned Airborne Vehicles (UAVs) to satellites (e.g., DESDynI).

Keywords: LiDAR, hyperspectral, forests, biomass, structure, biodiversity, carbon

1. Introduction

For forest studies, ground-based, airborne and spaceborne LiDAR have been used primarily to retrieve basic structural attributes, including height, canopy cover and vertical profiles from which indirect measures (e.g., basal area, timber volume and biomass) have been derived (Lefsky et al., 2005; Tickle et al., 2006; Goodwin et al., 2006; Brandtberg, 2007; Popescu and Zhao, 2008). Increasingly, however, studies are recognising or demonstrating that by integrating data from other sensors, including optical (e.g., hyperspectral) and Synthetic Aperture Radar (SAR), forests can be better characterised in terms of their structure, biomass and species composition (Hyde et al., 2005; 2006; Chen et al., 2007; Nelson et al., 2007). Opportunities for detecting changes in these attributes over time and at various scales are also enhanced (Wulder et al., 2007). Approaches to integration have varied but have typically involved combining data and derived products from other sensors to better quantify forest attributes (e.g., Hyde et al., 2006; Nelson et al., 2007; Lucas et al., 2008) or using LiDAR-derived information to better
interpret data acquired by other sensors (e.g., Lucas et al., 2006a; Simard et al., 2008). Using our own case studies and based on a review of current literature, this paper provides an overview of such approaches and gives application examples relating to the inventory and conservation of forest resources. Future synergies of LiDAR and other forms of remote sensing data are noted, focusing particularly on integration techniques and the deployment of new platforms and sensors.

2. Measures derived using LiDAR data alone

Most early studies using LiDAR focused on retrieval of simple descriptors of forest structure (Table 1a), with the majority utilising height information in the form of canopy height surfaces/models interpolated from outer canopy point data. More recent studies have derived additional attributes (Table 1b), including diameter at breast height (DBH), basal area and density (Hudak et al., 2008), timber volume and biomass (Naesset and Gobakken, 2008). In most cases, these attributes have been determined by establishing relationships with those directly measured (e.g., height or crown dimensions; Hyypä et al., 2001), summaries of the LiDAR data themselves (e.g., canopy geometric volume or profile area; Chen et al., 2007; Wulder et al., 2007), or LiDAR-based indices (e.g., the Height Scaled Crown Openness Index (HSCOI); Lee and Lucas, 2007).

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Derived from:</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>Diameter</td>
</tr>
<tr>
<td>Crown/canopy cover</td>
<td>H or crown dimensions, HSCOI²</td>
</tr>
<tr>
<td>Crown canopy depth</td>
<td>Basal area</td>
</tr>
<tr>
<td>Crown/canopy profile</td>
<td>Volume CGV</td>
</tr>
<tr>
<td>Outer canopy ruggedness</td>
<td>Biomass</td>
</tr>
<tr>
<td>Gap fraction</td>
<td>Density</td>
</tr>
<tr>
<td>Forest type/species</td>
<td>LAI⁵, PAI⁶</td>
</tr>
<tr>
<td>FPC⁷</td>
<td>Crown dimensions</td>
</tr>
</tbody>
</table>

Table 1: Examples of structural measures derived a) directly and b) indirectly from LiDAR data.

3. Terrestrial Laser Scanners and links with airborne LiDAR

Terrestrial Laser Scanners (TLS) provide detailed reconstructions of trunk, branch and leaf distributions from which tree locations, diameter and height (Maas et al., 2008; Watt and Donoghue, 2005), timber volume by size class (Jupp et al., 2005), and canopy gap fraction (Danson et al., 2007; Henning and Radtke, 2006) can be quantified. Potential exists also for retrieving the woody biomass of individual trees, either by considering the sizes of the stems scanned or multiplying the volume of scanned branches and trunks by wood density. Although limited by survey times and occlusion as a function of stand density, TLS provide a permanent record of forest structure. A close correspondence between forest height (Breyer, 2008) and, to a
lesser extent, foliage profiles (Jupp et al., 2005) retrieved separately using co-registered TLS and airborne LiDAR has also been reported. Linking TLS data with other remote sensing datasets (e.g., airborne LiDAR) does however require a high level of geolocational accuracy (Figure 1). Hence, establishment of a comprehensive and precise network of ground survey points and the use of high quality Inertial Navigation System (INS) for airborne systems is essential if all scan points are to be correctly located in three-dimensional space.

Figure 1: Airborne (full waveform) LiDAR point cloud (grey) with a sub-plot acquired by TLS (white) included, Lake Vyrnwy, mid-Wales.

4. Linking fine spatial resolution multi/hyperspectral data

Studies are increasingly incorporating data acquired by finer (typically < 1 m) spatial resolution multi/hyperspectral airborne (e.g., Compact Airborne Spectrographic Imager; CASI) and/or spaceborne sensors (e.g., Quickbird) to enhance descriptions of forests. The desire to simultaneously acquire complementary LiDAR and multi/hyperspectral datasets has also led to sensors being flown on the same platform (e.g., the Carnegie Airborne Observatory (CAO); Asner et al., 2008). More commonly, however, data are acquired using different platforms and on a similar or proximal date and algorithms for automatic rather than manual co-registration of data are then desirable.

Accurate co-registration of datasets significantly increases the diversity of information that can be extracted. St-Onge et al. (2008), for example, used a LiDAR-derived digital terrain model (DTM) as a base for increasing the accuracy of tree height estimates generated from historical stereo aerial photography. Within co-registered datasets, stand density can be estimated by counting a) extracted high points in LiDAR or ‘bright points’ in multi/hyperspectral data (Wulder et al., 2000) and/or b) tree crowns/clusters delineated using algorithms ranging from valley following to template matching (Bunting and Lucas, 2006). For open forests and orchard sites, retrieval accuracies have exceeded 70 % (Lee and Lucas, 2007) and 99 % (Jang et al., 2008) respectively. The advantage of having co-registered datasets is that trees identified within one can be attributed with measures (e.g., height or species; Chen et al., 2007) from the other, thereby leading to better descriptions of the forest. As an example, Bunting and Lucas (2006) applied an algorithm developed within Definiens Developer software and CASI data to delineate tree crowns of varying dimension. Once delineated, crowns were associated with a species type using spectra extracted from the sunlit portions as input to a linear discriminant function. A subsequent step then applied species-specific allometric equations relating LiDAR-derived height to the above ground and component (leaf, branch and trunk) biomass (Figure 2). Whilst performing well for isolated trees, the biomass was found to be over-estimated where trees with large expansive crowns occurred but was under-estimated where stem density was high (more than several per m²). Whilst hyperspectral data provide superior classifications of tree species, several studies have discriminated species or broad forest types using LiDAR intensity data (Antonarakis et al., 2008), relative height differences between
the first and last vegetation returns (Moffiet et al., 2005) and directed graphs (Brandtberg, 2007). Holmgren et al. (2008) reported, however, best discrimination when using a combination of LiDAR and multi-spectral data.

5. LiDAR for interpreting SAR data

5.1 Empirical relationships established between SAR and LiDAR-derived data

As with LiDAR, SAR is an active sensing technique and as emitted wavelengths at different frequencies and polarisations interact with components of the forest volume, the backscattered intensity relates partly to the overall structure and biomass of the forest. LiDAR-derived estimates of biomass and structural attributes can provide a basis for supporting the development of SAR-based retrieval algorithms, particularly as field-based measurements are often limited in amount and spatial distribution. As an example, and focusing on wooded savannas in Australia, Lucas et al. (2006a) established a relationship between LiDAR metrics and biomass ($r^2 = 0.92$). Relationships established subsequently between the LiDAR derived biomass and airborne SAR backscatter at different frequencies and polarisations (and for the equivalent of 4500 0.25 ha plots) then revealed differences in the saturation of backscatter above certain thresholds of biomass between SAR channels and suggested that L-band (~25 cm wavelength) cross polarised data acquired at incidence angles > ~ 40° provided the best option for biomass retrieval. The LiDAR-derived estimates of biomass also provided opportunities to evaluate existing biomass retrieval algorithms. For example, Le Toan (2008) proposed a Bayesian approach that utilised a priori knowledge of forest biomass to increase the accuracy of biomass retrieval and quantify uncertainties such that:

$$P(B | \gamma^0) \propto P(\gamma^0 | B)P(B)$$

Equation 1.

where $P(B | \gamma^0)$ is the probability of biomass given a value of backscatter (gamma0; $\gamma^0$) and $P(\gamma^0 | B)$ is the probability of $\gamma^0$ given a value of biomass. In the case of woody savannas, the a priori information was obtained from the Gaussian probability distribution function for biomass ($P(B)$) derived from the LiDAR data. The algorithm of Saatchi et al. (2007) uses L and/or P-band (~68 cm wavelength) data to separately estimate the biomass of the trunk and crown, which are then summed to give total above ground biomass. This model was newly parameterised for wooded savannas by applying crown:trunk ratios to the LiDAR-derived biomass as a function of species type (e.g., conifer, eucalypt, acacia), as described using co-registered stereo aerial photography (Tickle et al.,

Figure 2: Estimates of A) branch and B) trunk biomass generated using a combination of LiDAR and CASI data. Dots/red circles indicate trunk locations (Lucas et al., 2008a).
1996). The resulting model coefficients were then used to map biomass across the landscape, with best retrieval obtained using a combination of L-band (trunk biomass) and L- and P-band (crown biomass). By using these same biomass data, but adjusting for clearing since 2000 using time-series of Landsat sensor data, a modified version of the Saatchi et al. (2007) model was developed that used L-band dual polarimetric data acquired in 2007 from the Advanced Land Observing Satellite (ALOS) Phased Arrayed L-band SAR (PALSAR) as input. For regional application, further calibration and validation and consideration of a greater range of forest structural types is necessary. For this purpose, the Queensland Department of Natural Resources and Water (QDNRW) previously acquired discrete return LiDAR data for sites across Queensland ranging from sparse woodlands to dense tropical rainforests. Whilst biomass has yet to be estimated from these data, a close correspondence between LiDAR-derived attributes including height, foliage projected cover (FPC) and crown density and ALOS PALSAR data has been observed (Figure 3; Armston et al., 2008) suggesting that characterisation and mapping across the wider landscape is achievable from regional coverages.

**Figure 3: Correlations of PALSAR L-band HH $\sigma^0$ against LiDAR-derived overstorey structural attributes for 33 sites randomly sampled from the area of 19 LiDAR surveys covering open and closed forests.**

5.2 Retrieval through integration of LiDAR and SAR

Whilst some success has been achieved in retrieving biomass and structural attributes from both SAR intensity data and LiDAR, the mechanism of retrieval differs because of the modes of observation. Within SAR data, the increase in backscatter with biomass, for example, is largely a consequence of the greater number, diversity and size of scatterers (leaves, branches, trunks) within the forest volume. Within LiDAR, biomass is retrieved because of an inherent relationship with height and also the canopy volume. As these sensors are responding to different elements of the forest volume, additional information might be retrieved through their combination. A useful example is that of high (e.g., > 15 m) mangroves dominated by *Rhizophora stylosa*. These mangroves exhibit a SAR backscatter (L- and P-band) that approaches that of non-forest because of microwave attenuation by the extensive root system (Lucas et al., 2007). However, by integrating height information (e.g., from LiDAR), these mangroves can be identified as being of high biomass. Nelson et al. (2007) compared the retrieval of biomass using LiDAR and SAR (multi-angle BioSAR), concluding that whilst better retrieval was obtained when in combination, the small improvement over the use of LiDAR alone may not justify the increased resources required. Nevertheless, their combined use was considered to have greatest potential in retrieving the biomass of high biomass forests with excurrent growth forms.

5.3 Parameterisation of radar simulation models.

Radar simulation models are used primarily to understand the interaction of microwaves with different components of the forest volume and also the ground surface. These models are
typically two-dimensional and assume a random distribution of elements (e.g., disks or cylinders representing leaves and branches respectively). Comparisons between actual and simulated backscatter provided an indication of how well the SAR data are simulated and also permit then the signal to be decomposed such that the contributions from the different scattering mechanisms (e.g., trunk-ground) to the overall backscatter can be better quantified and understood. More recently, models that consider the distribution of elements in three-dimensional space have been established, but often the scattering elements are assumed to be distributed randomly within the volume space. However, LiDAR provides the opportunity to establish more precisely the location of scattering elements, which is particularly useful where the distribution is clumped rather than random. As an example, Lucas et al. (2006b) parameterised a coherent SAR image simulation model with structural attributes derived from discrete return LiDAR data. Key elements of the parameterisation involving LiDAR were a) the identification of stem locations based on low values within a Height Scaled Crown Openness Index (HSCOI) surface and estimation of biomass (based on allometric equations applied to diameter, as estimated from a relationship with the HSCOI, and/or height), b) the generation of voxels based on the three-dimensional distribution and frequency of LiDAR returns within 1 m$^3$ integer intervals from ground level to the maximum height of the stand and assignment of an estimate of leaf and tertiary branch biomass (derived from allometrics and based on the number of voxels associated with each stem), and c) the approximation of primary and secondary branch locations based on position, distance and angle from the main stem and primary branches respectively and an association with volume based on logical rules. Ground surface parameters were also determined from the LiDAR digital terrain model (DTM). A close correspondence was observed between simulated and actual (AIRSAR) data, suggesting effective modelling of the SAR backscatter. The model is now available within the European Space Agency (ESA) software package POLSARPRO (http://earth.esa.int/polsarpro/). Whilst parameterisation is complex, the approach provides considerable insight into the interaction of microwaves with different components of the forest volume.

5.4 Evaluation of InSAR and polInSAR data based on LiDAR

As with LiDAR data, there is increasing demonstration of the potential of retrieving information on the distribution of scattering elements within the vertical profile of forests using SAR (repeat or single pass) interferometry (InSAR) and polarimetric SAR interferometry (PolInSAR). A particular advantage of such approaches is that the coverage of SAR is much greater compared to LiDAR (Hyde et al., 2006; Baltzer et al., 2007). Nevertheless, LiDAR data can play a key role in the verification of retrieved profiles (Slatton et al., 2001), thereby leading to fine-tuning of algorithms. LiDAR can also provide a yardstick for assessing the retrieval of structural attributes from other sensors. For example, several studies have demonstrated differences of only a few metres in the errors associated with height retrieval from InSAR and LiDAR (e.g., Baltzer et al., 2007, Breidenback et al., 2008). Hyde et al. (2006) also suggested that InSAR was best suited for structurally homogeneous forests and that LiDAR provided better estimates of the height of larger trees.

6. Application examples

Relative to optical and SAR data sources, LiDAR technology is a new advance in the remote sensing of forests. As costs were high and the logistics of acquisition were complex, initial activities were concentrated largely in the research and government sectors, which were also in a better position to advance the development and evaluation of new sensor types (e.g., LVIS, SLICER). However, with the increased availability of commercial LiDAR (terrestrial and airborne) and freely accessible spaceborne LiDAR combined with an increasing capacity to integrate data from other sensors, significant expansions in both research and operational applications have occurred in recent years. The following sections give a brief overview of several applications.
6.1 Commercial forestry

The uptake of LiDAR for operational forestry applications has, until recently, been relatively low in many countries because of the perceived inability to retrieve the same level of information obtained through traditional forest survey techniques, the high costs involved, and also the lack of expertise within the intended user community (Suarez et al., 2005). The uptake of LiDAR is, however, variable with Naesset et al. (2004) and Nelson et al. (2007) noting an increasing number of organisations using these data routinely for forest structural measurement and management planning, particularly in Scandinavia and North America. Furthermore, recognition of the wide range of information that can now be obtained from LiDAR (e.g., DTMs, forest structure and the identification of areas that cannot be logged such as habitat trees and riparian zones) and the potential of integrating with other remote sensing datasets has led to an increase in utilisation in many countries.

6.2 Carbon stocks and dynamics

The retrieval of biomass (carbon) from LiDAR metrics through empirical relationships with field measurements has been demonstrated in many studies (Culvenor et al., 2005; Naesset and Gobakken, 2008; Lim and Treitz, 2004) and such estimates can potentially support carbon trading and national accounting (Patenaude et al., 2004). As examples, Naesset and Gobakken (2008) used LiDAR-retrieved canopy height and density to estimate the biomass of boreal forests in Norway. Lefsky et al. (2005) integrated time-series of Landsat sensor data to age stands and, in conjunction with LiDAR-derived stem height and biomass, generated estimates of wood Net Primary Productivity (NPP). Using LiDAR combined with Landsat and SPOT sensor data in New Zealand (Ministry of the Environment, 2008), estimates of carbon stocks for extensive areas of forest have been generated as part of a national sampling program of greenhouse gas emissions monitoring. By contrast, the Australian National Carbon Accounting System (Brack et al., 2006) has not integrated LiDAR data to the same extent because of the difficulty of calibration associated with the complexity of forests structures. Nevertheless, the potential benefits of using LiDAR for calibrating other forms of remote sensing data and supporting carbon accounting and reporting schemes in many countries have been recognised.

6.3 Biodiversity assessment

The high diversity of fauna and non-tree flora associated with forests is attributable to the diversity of habitats, which, in part, is reflected in the spatial distribution and arrangement of structural elements within the volume space that trees create and occupy. Several studies have noted that the distribution and richness of bird species in particular are closely linked to forest canopy structure (Hyde et al., 2005) and heterogeneity (Goetz et al., 2007), both of which can be quantified using airborne LiDAR. Hill et al. (2005) and Hinsley et al. (2006) also reported a link between habitat quality (defined by forest canopy structure and height) and the breeding success of Great Tits (Parus major). Such assessments might be improved by integrating information on tree species and the age and condition of stands, as obtained using, for example, multi/hyperspectral data (Hill and Thomson, 2005). Most studies focusing on biodiversity are confined to relatively small areas because of the limited coverage of airborne acquisitions. Extrapolation to regional areas requires the establishment of forest height and structural maps over larger areas, which can potentially be generated using SAR interferometry and/or IceSAT data. Such information would complement habitat maps generated at a commensurate scale using airborne/spaceborne optical datasets.

6.4 Environmental change

The detection of changes in forest cover (deforestation, degradation and regeneration/afforestation) using LiDAR is limited primarily by spatial coverage and the cost of
data acquisition. Nevertheless, the requirement for change mapping based on such data is compelling. Within Australia, trees of certain species identified in 2000 from a combination of LiDAR and hyperspectral data were noted to have died back in 2006, partly because of the intense drought (Lucas et al., 2008b). This dieback has been observed previously over large areas and so repeat acquisition by both sensors and subsequent assessment of changes in structure, biomass and species can inform on the impacts of adverse change but also better understand how these might be detected using sensors with wider spatial coverage (e.g., ALOS PALSAR). Wulder et al. (2007) compared two profiling LiDAR transects 600 km in length across boreal forests in Canada in 1997 and 2002, indicating that global comparisons of structural attributes were less informative than spatially explicit comparisons undertaken for local areas (in this case, defined by segmenting Landsat ETM+ data). The local approach allowed LiDAR profiles to be treated as samples of a population, with the latter defined as a Landsat segment, thereby avoiding the issue of geolocation error. This study also raised the issue that long time-periods between comparisons are often required for monitoring certain processes such as vegetation growth and dieback in response to climate change. This has further implications for comparison of different LiDAR datasets as recent years have seen a rapid advancement of LiDAR technology from single discrete return profiling instruments to full waveform small-footprint scanning systems (e.g. Wagner et al., 2004). Therefore, care needs to be taken when using different airborne LiDAR systems for monitoring because geolocation errors and different acquisitions specifications of the LiDAR surveys may cause differences in estimates of structural attributes that are not the result of real change (Goodwin et al., 2006; Wulder et al., 2007).

7. Overview and future opportunities

Through a series of case studies and with reference to the published literature, this review has highlighted the benefits of integrating LiDAR with other remote sensing datasets for furthering the characterisation, mapping and monitoring of forests at a range of scales. In particular, the integration of datasets can lead to: a) an increase in the diversity and accuracy of information on forest structure, biomass and species composition and change, particularly at the individual tree and stand level, b) a greater capacity to establish empirical models with moderate to coarse spatial resolution (e.g., spaceborne optical and SAR) data, thereby facilitating retrieval across wider areas, c) unique opportunities for providing detailed parameterisation of simulation models that can be used to better understand the interaction of electromagnetic energy with forest components and/or be inverted to allow retrieval of a greater range of biophysical properties from remote sensing data (Koetz et al., 2006), and d) greater provision of data or derived products for inclusion within multi-sensor biophysical retrieval algorithms.

Recognition of the benefits of data integration has led to the design of sensors that combine LiDAR with other optical or multi-angular (e.g., Carbon3D) or radar sensors (e.g., DESDynl) and also the development of systems (e.g., The Carnegie Airborne Observatory) and new platforms (e.g., Unmanned Airborne Vehicles) with capacity to support both LiDAR and multi/hyperspectral sensors. Such developments are anticipated to lead to a greater uptake of LiDAR for a range of applications. Advances in approaches to the integration of such data (e.g., automated registration) and algorithms for retrieving biophysical attributes of forests have also been ongoing and are anticipated to lead to their greater use within forest-related applications.

References

Sensing and Photogrammetry Conference, Darwin (in press)


Le Toan. 2007. Technical Notes to ESA on Intensity Inversion (unpublished)


Moffett, T., Mengersen, K., Witte, C., King, R., and Denham, R., 2005. Airborne laser scanning:
Exploratory data analysis indicates potential variables for classification of individual trees or forest stands according to species. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(5), 289-309.


