

Integration of airborne LiDAR, satellite imagery, and field measurements using a two-phase sampling method for forest biomass estimation in tropical forests

B. R. Gautam¹, T. Tokola², J. Hamalainen¹, M. Gunia¹, J. Peuhkurinen¹, H. Parviainen¹, V. Leppanen¹, T. Kauranne¹, J. Havia³, I. Norjamaki⁴ and B. P. Sah⁵

¹Arbonaut Ltd. Koskikatu 5 B, 80100 Joensuu/Latokartanon tie 7 a, 00700 Helsinki, Finland
Tel: +358 13 2591911, Fax: +358 13 2591943
E-mail: basanta.gautam@arbonaut.com Web: www.arbonaut.com

²University of Eastern Finland, School of Forest Sciences, P. O. Box 111, 80101 Joensuu, Finland
Tel: +358 9 19158180, Fax: +358 9 919158159
Email: timo.tokola@uef.fi Web: www.uef.fi

³FM-International Oy FINNMAP, Malminkaari 5, FI-00700 Helsinki, Finland
Tel: +358 9 25221700, Fax: +358 9 251717
E-mail: jussi.havia@finnmap.com Web: www.finnmap.com

⁴Indufor Oy, Toolonkatu 11 A, FI-00100 Helsinki, Finland
Tel: +358 9 6840110, Fax: +358 9 1352552
E-mail: ilkka.norjamaki@indufor.fi Web: www.indufor.fi

⁵International Division, PASCO Corporation, 1-1-2 Higashiyama, Meguro-ku, Tokyo 153-0043, Japan
Tel: +81 3 64122888, Fax: +81 3 37158640
E-mail: bhpa_s5512@pasco.co.jp Web: www.pasco.co.jp

Abstract

Accurate and reliable monitoring of biomass in tropical forest has been a challenging task because a large proportion of forest is inaccessible. For effective implementation of REDD-plus and fair benefit sharing, monitoring methodology should be based on scientifically robust estimation of sources and sinks to meet MRV requirements. Though there have been major advances in satellite remote sensing technologies in recent years, none of them have been able to overcome the saturation problem that makes it hard to detect forests with high above-ground biomass volume and assess degradation. The saturation problem in biomass estimation can be overcome by adopting airborne LiDAR, because laser pulses penetrate even through a dense multi-layered canopy and there is a strong correlation between LiDAR data and biomass. Integrating different remote sensing and field reference data provides an accurate, precise, and affordable monitoring solution for tropical forests. In this regard, a two-phase sampling scheme optimizes field data collection efforts for model calibration and assists with objective and efficient positioning of sample plots. In the second sampling phase, estimates for a set of LiDAR transects are available, and the radiometric properties of satellite imagery are applied to identify the best estimators for target variables. Such a method is proposed here. It integrates sample plots with LiDAR transects and satellite images, and it attains a relative RMSE of 25 to 35 percent in above-ground biomass already on an area of 0.5 ha. Alternative methods, such as National Forest Inventories based on permanent sample plots, optical satellite imagery, and k-NN estimation or visual inspection, attain this error level only on areas of 100 ha or even more. Such high spatial resolution is crucial for awarding REDD credits to local forest owners. Arbonaut Ltd. has developed a forest inventory process and user-friendly tools (ArboLiDAR) to estimate above- and below-ground carbon stocks. The estimation process relies on a unified Bayesian statistical methodology, making it possible to incorporate various information sources such as direct measurements, quantities interpreted from remote sensing, and the results of modelling of carbon sinks such as below-ground carbon. These estimates can also be simply updated whenever new data becomes available.

Keywords: LiDAR, satellite image, two-phase sampling, field plot measurement, Bayesian regression, biomass, REDD-plus MRV

Introduction

Tropical forests play a very significant role in mitigating global climate change. While forests are estimated to sequester about 15% of global carbon emissions, global deforestation and forest degradation accounts for up to 20% of annual greenhouse gas emissions (Achard et al. 2007; Angelsen 2008). There has been significant progress on climate change mitigation options, such as the REDD-plus scheme (Reducing Emissions from Deforestation and Forest Degradation in Developing Countries), after the Fifteenth Conference of the Parties (COP15) of the United Nations Convention on Climate Change (UNFCCC) in December 2009. REDD-plus was able to make significant progress politically and financially despite the failure to reach a legally binding global pact on cutting greenhouse gases during the climate negotiations at the COP15. Discussion on climate change will continue, but it is already clear that tropical forests have great potential to mitigate global climate change. REDD-plus is a cost-effective, efficient and equitable approach to climate change mitigation worldwide and it is expected to be included in a post-2012 international emissions reduction treaty (Blom et al. 2010; Phelps et al. 2010).

Accurate and reliable estimation of biomass in tropical forest has been a challenging task because a large proportion of forest area is inaccessible. Lack of reliable up-to-date data has been a fundamental obstacle to understanding the scale of deforestation and forest degradation, and to monitoring the extent of forest reduction (Springate-Baginski and Wollenberg 2010). The importance of remote sensing technology in the mapping of tropical forests to support the monitoring needs of REDD-plus has been growing. For effective implementation of REDD-plus and fair benefit sharing, the monitoring methodology should be based on scientifically robust estimation of sources and sinks to meet the monitoring, reporting and verification (MRV) requirements (Bottcher et al. 2009). Topographical complexity and the lack of availability of suitable remote sensing material are two main reasons for incorrect forest inventory data in tropical countries. Although there have been major advances in satellite remote sensing technology in recent years, none have been able to overcome the saturation problem that makes it difficult to detect forests with high above-ground biomass volume and to assess degradation (Næsset 2009). This has led researchers to develop robust methods by

integrating different remote sensing materials, an approach that facilitates the direct calculation of quantitative estimates of forest attributes (Hollaus et al. 2006; Gonzalez et al. 2010; Asner et al. 2010) that support REDD-plus monitoring. The present study is a demonstration of integrating data from various sources, such as ALS technology, satellite remote sensing, and field measurements for mapping tropical forest biomass.

Airborne Laser Scanning (ALS)

Airborne laser scanning, or LiDAR (Light Detection and Ranging), is an active remote sensing technique that permits observation of the vertical structure of forests. This ability distinguishes ALS from conventional remote sensing approaches. A sensor mounted on a fixed-wing plane or helicopter emits laser pulses towards the ground and records the elapsed time between beam launch and return signal registration (Gautam and Kandel 2010). Each laser hit is geo-referenced with 3-dimensional coordinate values. LiDAR data has a great advantage because laser pulses penetrate even through a dense multi-layered canopy and hence do not suffer from the saturation problem (Næsset 2009).

Satellite technology

Satellite remote sensing can collect large amounts of image data over a wide geographical area with a high temporal frequency. Some forest attributes that are difficult to estimate from LiDAR data, such as species composition, are easy to interpret from satellite imagery.

Advanced Land Observation Satellite (ALOS)

ALOS was launched on January 24th, 2006 by JAXA (Japan Aerospace Exploration Agency). ALOS is able to cover the Earth's surface over three times per year at 10-, 20-, and 100-meter spatial resolutions (Kellndorfer et al. 2007). ALOS is a satellite well-suited to monitor tropical forests, as it scans the Earth's surface using both optical (multi-spectral and panchromatic bands) and radar sensors, which produces cloud-free imagery (LAPAN and JAXA 2009). The ALOS satellite is mounted with two optical sensors, AVNIR and PRISM, and carries a radar sensor, PALSAR, onboard.

Landsat 7

Landsat 7, the latest NASA satellite in the Landsat series, can produce an uninterrupted multispectral

record of the Earth's land surface since 1972. Landsat 7 was launched in 1999 with a 16-day repeat cycle, and is equipped with the Enhanced Thematic Mapper Plus (ETM+) sensor. Landsat 7 data cover a swath width of 185 km in eight different wavelength bands. The data are available in two spatial resolutions: 15m (pan) and 30m (ms), with improved radiometric calibration (Trigg et al. 2006).

Two-Phase Sampling with LiDAR

A two-phase sampling scheme optimizes field data collection for model calibration through efficient field sample plot positioning. The first sampling phase is based on full coverage of satellite imagery and other ancillary data, and the subsequent second phase is based on ALS data and field measurements. In the first phase, a wall-to-wall map of broad categories, such as forest, non-forests and uncertain areas, is produced using satellite imagery. In the second phase, a sample of LiDAR transects is collected over the project area. The point cloud properties such as pulse height and density characterize variation in the forest structure and successional stage (Falkowski et al. 2009). Optimally, the whole range of structural variation is covered with field plot samples in the second phase for each stratum. Field sample data is used to calibrate statistical models based on LiDAR metrics. The estimates for LiDAR transects, which are highly accurate, are utilized as a reference when estimation is carried out for a complete wall-to-wall area covered with satellite data.

ArboLiDAR Forest Inventory

Arbonaut Ltd. has developed a forest inventory process and user-friendly tools. ArboLiDAR is an inventory process that integrates airborne laser scanning, field measurements, and satellite imagery with full geographical coverage for highly accurate estimates of bio-physical forest attributes. Estimates with high spatial resolution guide operational forest management for effective maintenance and enhancement of above- and below-ground carbon stocks.

The ArboLiDAR estimation process relies on a unified Sparse Bayesian regression (Tipping 2001) that makes it possible to incorporate various information sources, such as direct measurements, quantities interpreted from remote sensing, and results of modelling of carbon sinks such as below-ground carbon.

The Sparse Bayesian regression is a general, non-parametric, locally linear Bayesian regression method (Junttila et al. 2008) that automatically determines the rank of an appropriate regression model based on the variance in measurements of the variable to be predicted (such as biomass), and the variable's correlation with a set of LiDAR metrics and satellite image spectral features. With ArboLiDAR, Sparse Bayesian regression is conducted either on an estimation grid with cells of sample plot size, or on homogeneous micro-stands, which are automatically generated by the stand delineation tool incorporated in ArboLiDAR that uses LiDAR height and density, as well as spectral and textural features, for accurate delineation.

Study area, sensors and data

The study area is situated in Vientiane, Lao PDR. The data was collected under the SUFORD Forest Monitoring Component. The SUFORD project was a bilateral co-operation between governments of Finland and Lao PDR. The leading company of the project was Indufor Ltd. Finland. The LiDAR survey area covered 25 000 hectares. LiDAR data and color-infrared (CIR) aerial images were collected with a Leica ALS 40 sensor and a Leica MP 39 digital camera. The flight campaign took place between February 6th and 8th in 2009. A detailed list of parameters for LiDAR acquisition is presented in Table 1. The ALOS AVNIR-2 imagery was taken in September 2006 and the Landsat 7 imagery was taken in November 2000. Using old satellite data provides a way to define accurate forest carbon baselines in the past with the method proposed here.

A total of 328 sample plots were established across the study area with the help of the local officials from the Department of Forests and the District Agriculture and Forest Extension Offices. Local villagers were also involved in field data collection in Songkhone and Thapangthong districts. Sample plots were rectangular in shape with a dimension of 20 x 20 metres each. The location of each plot was recorded using Global Positioning System (GPS) handheld devices.

Table 1: Parameters for LiDAR survey.

Parameter	Value
Field of View (FOV)	30 degrees
Sidelap	20%
Point density	1/m ²
Flying altitude	2000 meters
Speed	120 knots
Sun position	>20 degrees
Optimal distance	19 + 1 trajectories
Range of flying	410 km
GPS base station	40 km apart

Estimating biomass with datasets from various sources

LiDAR sampling has been suggested as a feasible way to combine the accuracy of LiDAR with the coverage and affordability of satellite imagery (Næsset et al. 2009; Asner 2009). In our tests, the target area was covered by wall-to-wall LiDAR for verification. For testing the integrated method proposed here, only every tenth LiDAR flight line was retained on which a sufficient number of sample plots was present, covering 10 percent of the area.

In the first phase, the ArboLiDAR method was used to obtain accurate estimate of biomass for the LiDAR transects and an additional plot sample of surrogate plots was created to complement the field survey plots. In the second phase, the remaining 90 percent of the area was interpreted using satellite data alone, using both the real and the complementary sample as the teaching set.

Mapping biomass for the LiDAR transects

An accurate biomass map was created for the LiDAR transects using LiDAR, ALOS AVNIR and Landsat imagery (Figure 1). Features that correlated with the modelled data were extracted from LiDAR according to the method described by Næsset (2002). ALOS and LANDSAT bands were included directly, and a linearization transformation was performed on selected bands to make their response linear. Furthermore, red edge (Seager et al. 2005) was extracted from both ALOS and Landsat bands as

$$redge = NIR - R$$

and linearized. In total, 28 LiDAR and 19 satellite features were included. The area of LiDAR transects was covered by a regular grid of 20 x 20 meter cells and estimated using the available field plots. The following targets were estimated: dominant height (m), dominant diameter (cm), basal area (m^2/ha), total volume (m^3/ha), stem biomass (tons/ha), above- and below-ground biomass (tons/ha), and above- and below-ground carbon (tons/ha).

The estimation accuracy was assessed using leave-one-out cross-validation, in which the sample plot is estimated using all remaining sample plots and the estimation result is compared to the measured value.

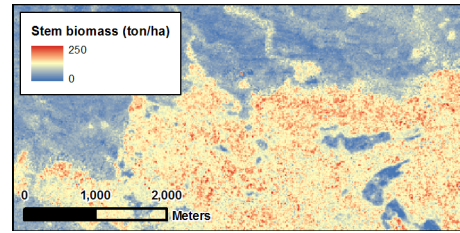


Figure 1: Stem biomass estimated from LiDAR, ALOS and Landsat data and field sample plots.

Generating surrogate plot sample

The original field sample was extended by generating additional samples, which we referred to as surrogate samples. In order to acquire plot sample that well represent the variation in biomass in the area, the new sample plots were placed randomly with probability proportional to the estimated biomass. The accurate estimates obtained from LiDAR and satellite data were taken as new ground truth and used in training the satellite-based models.

Estimating above- and below-ground biomass

Above- and below-ground biomass was calculated for the entire area using the features derived from ALOS and Landsat imagery. The surrogate plot sample was taken to train the model.

Furthermore, micro-segments with mean area of 0.4 and 1 hectare were generated for the study area. Vegetation height and density features derived from LiDAR were used in the segmentation process (Figure 2). The satellite-based estimates were aggregated into the segments and validated against the accurate biomass map.

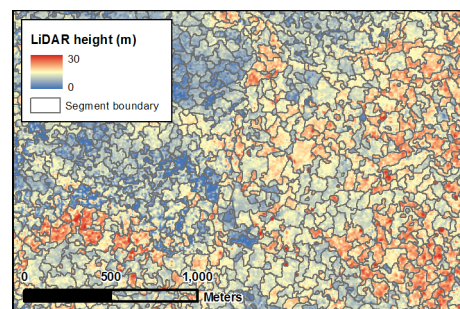


Figure 2: Micro-segments derived from LiDAR. The mean segment area is 0.4 ha.

Results

The biomass estimation using both LiDAR and satellite data resulted in 6.2% Root Mean Square Error (RMSE) for dominant diameter, 7.6% RMSE

for dominant height, 16.4% for basal area and 23.3% for total volume (Figure 3 and 4). RMSE for the remaining target variables is not reported as the variables were modelled from the volume and the error is therefore identical.

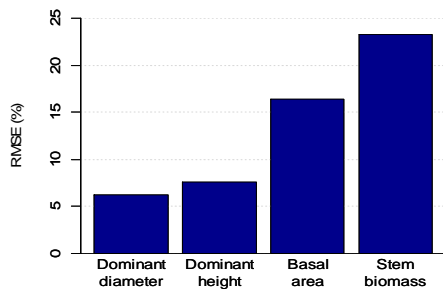


Figure 3: Root Mean Square Error of biomass estimates validated against the original field plots.

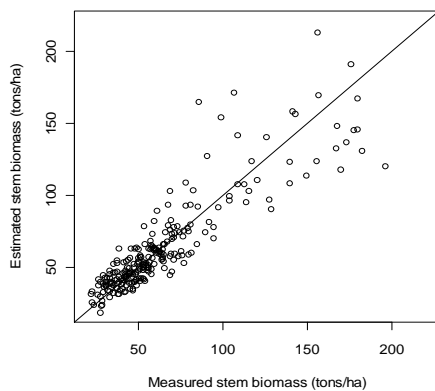


Figure 4: Scattergram of measured and estimated stem biomass.

The accurate biomass map was used for creating a surrogate sample that better describes the variation in the biomass. Most notably, areas with high biomass concentration were not well represented by the original sample plots, while the surrogate sample adds a significant amount of plots with total volume above 300 m³/ha (Figure 5).

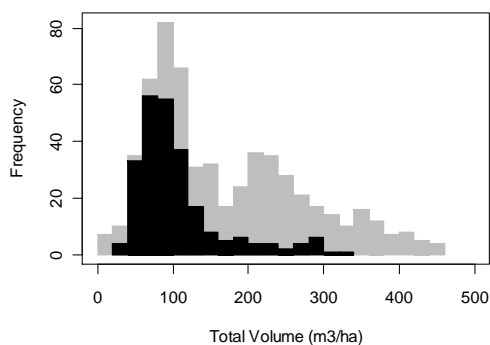


Figure 5: Total volume histograms of original sample plots (black bars) and the surrogate sample (gray bars).

At the plot level (the area of a single plot was 400 m²), the satellite-based estimation of the entire area yielded RMSE of 19.2% for dominant diameter, 18.5% for dominant height, 26.5% for basal area and 34.4% for stem biomass when using the surrogate plots as a teaching set. This is in contrast to using the original field plots which do not fully describe biomass within the LiDAR transects and the RMSE of stem biomass reached 39.7%. The estimation error drops dramatically when calculated at the scale of 0.4 and 1 hectare. For the 1-hectare segments, RMSE of dominant diameter was 8.3%, RMSE of dominant height 9.21%, RMSE of basal area 19.5% and RMSE of stem biomass 23.9% (Figure 6).

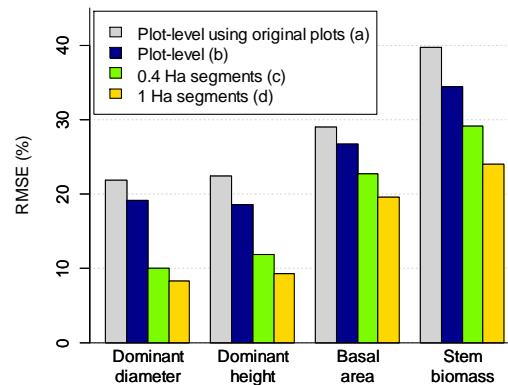


Figure 6: Relative RMSE for satellite-based estimation. a) Plot-level, models trained using the original field sample. b) Plot-level, models trained on the surrogate plots, c) Segments with mean area 0.4 ha, models trained on the surrogate plots. d) Segments with mean area 1 ha, models trained on the surrogate plots.

The costs of forest inventory relying on two-phase sampling are strongly scale-dependent and drop down to and below 10 US cents per hectare for national-scale projects (Figure 7).

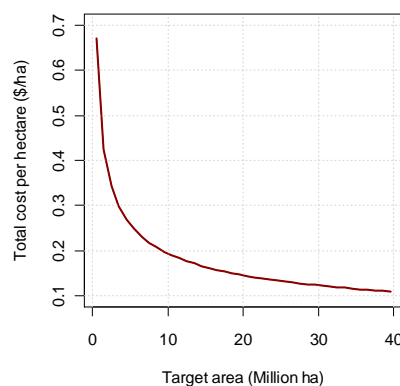


Figure 7: The effect of scale-to-unit cost of biomass inventory when integrating a 10% LiDAR sample, field plot measurements, and optical satellite imagery.

Discussion and conclusions

Various studies have demonstrated that LiDAR technology has great potential for carbon stock estimation in tropical forests. LiDAR technology is well-suited to monitor forest degradation, as LiDAR helps to detect changes in forest canopy structure (Saramaki 2010). The integrated approach makes optimal use of different datasets in estimating tropical forest biomass (Hudak et al. 2002; Næsset 2002; Asner 2009). The possibility of correlating any signal (satellite, aerial image, or LiDAR) with timber volume or carbon stock depends on a strong correlation between the signal and biomass. Peuhkurinen et al. (2010) found a strong correlation between measured and predicted values when stem volume was estimated with LiDAR. The only prerequisite is that some pulses do reach the ground (Gonzalez et al. 2010), which is possible in closed canopy forests up to 50 metres tall, and perhaps taller. But with both optical and SAR (Synthetic Aperture Radar) satellite instruments, this correlation becomes flat when stem volume reaches a certain threshold (Næsset 2009; Gracia et al. 2010). Due to this saturation effect, estimation based on optical imagery may lead to significant underestimation of carbon stock in areas with high above-ground biomass concentrations. In this study, the poor correlation of the satellite data caused underestimation of biomass in areas with total volume above $250\text{m}^3/\text{ha}$ (Figure 8).

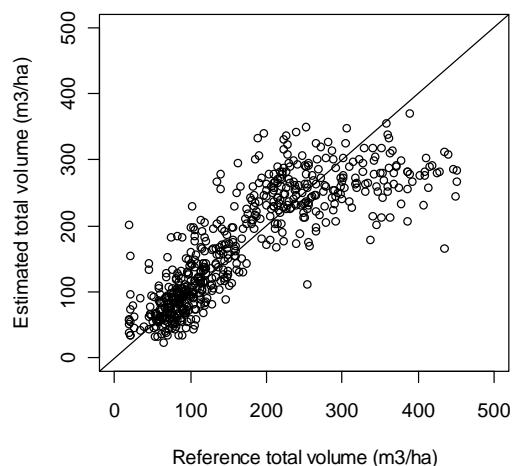


Figure 8: Total volume estimated using satellite data and surrogate plots. The ability of satellite signal to predict biomass saturated at about $250\text{m}^3/\text{ha}$.

Two-phase sampling with LiDAR attains a relative RMSE of 25 to 35 percent in above-ground biomass already on an area of 0.5 ha, whereas alternative methods, such as National Forest Inventories based on permanent

sample plots, optical satellite imagery, and k-NN estimation or visual inspection attain this error level only on areas of 100 ha or even more. Such high spatial resolution is crucial for awarding REDD credits to local forest owners. The integrated methodology provides an opportunity to estimate change in carbon stock at greater aerial extent, and with high spatial resolution, high accuracy, and at relatively low cost (Asner et al. 2010). Accurate, high-resolution, and temporally consistent biomass and carbon stock estimates are needed for REDD-plus monitoring, reporting and verification in line with Tier 3 requirements. Thus, the approach integrating LiDAR, field plot, and satellite data will play a significant role in the future (McNally et al. 2009).

References

- Achard, F; DeFries, R; Eva, H; Hansen, M; Mayaux, P; Stibig, H-J (2007) Pan-tropical monitoring of deforestation. *Environ. Res. Lett.* 2: 045022 (11pp).
- Angelsen, A (ed.) (2008) Moving ahead with REDD: Issues, options and implications. CIFOR, Bogor, Indonesia.
- Asner, GP (2009) Tropical forest carbon assessment: integrating satellite and airborne mapping approaches. *Environ. Res. Lett.* 4: 034009 (11pp).
- Asner, GP; Powell, GVN; Mascaro, J; Knapp, DE; Clark, JK; Jacobson, J; Kennedy-Bowdoin, T; Balaji, A; Paez-Acosta, G; Victoria, E; Secada, L; Valqui, M; Hughes, RF (2010) High-resolution forest carbon stocks and emissions in the Amazon. <http://www.pnas.org/content/early/2010/08/30/1004875107.short> (accessed 8 September 2010).
- Blom, B; Sunderland, TCH; Murdiyarso, D (2010) Getting REDD to work locally: lessons learned from integrated conservation and development projects. *Environmental Science and Policy* 13: 164–172.
- Bottcher, H; Eisbrenner, K; Fritz, S; Kindermann, G; Kraxner, F; McCallum, I; Obersteiner, M (2009) An assessment of monitoring requirements and costs of 'Reduced Emissions from Deforestation and Degradation'. *Carbon Balance and Management*, 4: 7. <http://www.cbmj.com/content/4/1/7> (accessed 1 August 2010).

Falkowski, MJ; Evans, JS; Martinuzzi, S; Gessler, PE; Hudak, AT (2009) Cha

racterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. Remote

Sensing of Environment, 113: 946–956.

Garcia, M; Riano, R; Chuvieco, E; Danson, FM (2010) Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment* 114: 816–830.

Gautam, BR; Kandel, PN (2010) Working paper on LiDAR mapping in Nepal. Submitted to the Ministry of Forests and Soil Conservation on 26 March 2010.
<http://www.forestrynepal.org/publications/article/4771> (accessed 15 July 2010).

Gonzalez, P; Asner, GP; Battles, JJ; Lefsky, MA; Waring, KM; Palace, M (2010) Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. *Remote Sensing of Environment* 114: 1561–1575.

Hollaus, M; Wagner, W; Eberhofer, C; Karel, W (2006) Accuracy of large-scale canopy heights derived from LiDAR data under operational constraints in a complex alpine environment. *ISPRS Journal of Photogrammetry & Remote Sensing* 60: 323–338.

Hudak, AT; Lefsky, MA; Cohen, WB; Berterretche, M (2002) Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment* 82: 397–416.

Junttila, V; Maltamo, M; Kauranne, T (2008) Sparse Bayesian Estimation of Forest Stand Characteristics from Airborne Laser Scanning. *Forest Science*, 54(5): 543–552.

Kellndorfer, J (2007) New eyes in the sky: cloud-free tropical forest monitoring for REDD with the Japanese advanced land observation satellite (ALOS). Presented at United Nations Framework Convention on Climate Change (UNFCCC) Conference of the Parties (COP), Thirteenth sessions, Bali, Indonesia.

LAPAN and JAXA (2009) Indonesia ALOS Pilot Project Cooperation between LAPAN and JAXA, for Some Application. Available at http://www.sapc.jaxa.jp/about/data/indonesia_pilotproject_e.pdf (accessed 2 September 2010).

McNally, R; Sage, N; Holland, T (2009) Understanding REDD: Implications for Lao PDR, Nepal and Vietnam (pp 120). http://www.snvworld.org/en/regions/asia/ourwork/Documents/REDD%20final_for%20printing.pdf (accessed 5 August 2010).

Næsset, E (2002) Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80: 88–99.

Næsset, E (2009) Some challenges in forest monitoring. Proposing a “research platform” for development, training, and validation of spaceborne technologies in Brazil. A Presentation. Brazilian Norwegian Workshop on Forest and Marine Monitoring, Ilhabela, SP, Brazil, March 9–12, 2009. 24 slides.

Næsset, E; Gobakken, T; Nelson, R (2009): Sampling and mapping forest volume and biomass using airborne LIDARs. In McRoberts et al (2009): Proceedings of the eighth annual forest inventory and analysis symposium; 2006 October 16–19; Monterey, CA. Gen. Tech. Report WO-79. Washington, DC: U.S. Department of Agriculture, Forest Service. 297–301.

Peuhkurinen, J; Gonzalez, AS; Ketola, J (2010) Using aerial photographs and VHR satellite data as an auxiliary information in airborne LiDAR-based estimation of species-specific forest characteristics. Paper presented at Forestsat 2010: Operational tools in forestry using remote sensing techniques. September 7–10, Lugo and Santiago de Compostela, Spain.

Phelps, J; Webb, EL; Agrawal, A (2010) Does REDD+ Threaten to Recentralize Forest Governance?
http://www.rightsandresources.org/documents/files/doc_1461.pdf (accessed 15 July 2010).

Saramaki, T (2010) REDD assessment: modelling canopy cover from ground measurements and lidar variables. Bachelor thesis submitted to the University of Eastern Finland (pp 1–24).

Seager, S; Turner, EL; Schafer, J; Ford, EB (2005) Vegetation’s Red Edge: A Possible Spectroscopic Biosignature of Extraterrestrial Plants *Astrobiology*, Volume 5, Issue 3 (pp 173–194).

Springate-Baginski, O; Wollenberg, E (eds.) (2010) REDD, forest governance and rural livelihoods: the emerging agenda. CIFOR, Bogor, Indonesia. (pp 173–194).

Tipping, ME (2001) Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research* 1: 211–244.

Trigg, SN; Curran, LM; McDonald, AK (2006) Testing the utility of Landsat 7 satellite data for continued monitoring of forest cover change in protected areas in Southeast Asia. *Singapore Journal of Tropical Geography* 27: 49–66.