International Expert Group Meeting on Geospatial Information Systems for Multi-scale Forest Biomass Assessment and Monitoring in the Hindu Kush-Himalayan Region ICIMOD Headquarter, Kathmandu, Nepal.09-10 December 2013.

ESTIMATION OF FOREST BIOMASS USING LIDAR-ASSISTED MULTI-SOURCE PROGRAMME (LAMP)

J. Peuhkurinen [a,*], T. Kauranne [a], J. Hämäläinen [a], B. Gautam [a]

[a] Arbonaut Ltd, Kaislakatu 2 80130 Joensuu, Finland - (jussi.peuhkurinen, tuomo.kauranne, jarno.hamalainen, basanta.gautam)@arbonaut.com

KEY WORDS: Light Detection and Ranging, satellite imagery, REDD, Bayesian interpretation, forest inventory

ABSTRACT:

LiDAR (Light Detection and Ranging) is an active remote sensing technology which provides 3D information of the terrain and vegetation. LiDAR data with pulse density ~1 return per square meter is sufficient for forest inventory applications in general cases. It is also possible to assess vegetation height and density directly from LiDAR data. However, compared to optical or radar satellite data airborne LiDAR data is not efficient tool when high temporal resolution and very large area mapping is required. LiDAR-Assisted Multi-source Programme (LAMP) is a forest inventory methodology that integrates LiDAR data with satellite data and field data for estimating forest characteristics, such as biomass and carbon stocks for large areas. It takes advantage of LiDAR's high precision and satellite data's good temporal and spatial coverage. LAMP methodology was applied in three case studies in tropical countries, namely LAO PDR, Nepal and Ghana. Wall-to-wall LAMP biomass estimates were produced for a grid with a cell size of maximum 1 hectare and verified against field data. The case studies proofed that LAMP is a scalable, fast, robust and cost-efficient approach for estimating forest carbon and biomass. The achieved results indicated that LAMP methodology is a promising approach for achieving Tier 3 requirements of REDD+ in measuring, reporting and verification (MRV) at the national and sub-national scales.

1. INTRODUCTION

Tropical deforestation and forest degradation account for about 20% of annual greenhouse gas (GHG) emissions, thus being the second largest source of GHG emissions globally (IPCC 2007). The Reducing Emissions from Deforestation and forest Degradation (REDD+) scheme may provide sustained incentives for developing countries in the future to reduce emissions from forested lands and invest in sustainable development by providing a financial value for the amount of carbon stored in forests (Angelsen et al. 2009). REDD+ also includes the role of forest conservation, sustainable management of forests, and enhancement of forest carbon stocks in a financing mechanism (Angelsen et al. 2011). A successful REDD+ mechanism will require the design and implementation of operational forest monitoring, reporting and verification systems that are transparent, complete, consistent, comparable, and accurate at national and sub-national scales (Walker et al. 2010, IPCC 2003).

An integrated system of unbiased geospatial and statistical estimators of sequestered carbon amounts across forest land is highly important for REDD+. Combining remotely sensed data with a forest resource inventory provides practical means to generate such information. Remote sensing collects and interprets information about features from a distant location and obtains continuous data over large areas in the forms of continuous thematic maps (e.g., forest biomass). There is tremendous diversity in the number and properties of the sensors and imagery available today ranging from space-borne to airborne to ground-based systems. Each system has different properties with spatial resolution, number of spectral and radiometric bands, temporal frequency and the cost of acquisition. Despite this diversity of sensors, no current remote sensing system directly measures forest biomass and sequestered carbon. Thus, remote sensing is effective at indicating where specific features are and how they are distributed but cannot provide an accurate estimate of how much of that feature is in the mapped area without an integrated resource inventory.

In recent years, airborne LiDAR has become an integral part of operational forest inventory in Scandinavian countries (Næsset 2007). Its high potential for REDD+-related biomass inventories in tropical countries has been well demonstrated (Asner et al. 2009, Gautam et al. 2010, Asner et al. 2012, Asner et al. 2013). Vegetation heights can be acquired with high accuracy using LiDAR height metrics. Since tree height is strongly correlated with tree volume, forest biomass can be predicted with high accuracy when regressing LiDAR metrics

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

with data from field measured plots. Wall-to-wall covering of area of interest with LiDAR is relatively expensive, thus a two-phase estimation approach has been proposed and it only requires LiDAR data from a sample of the study area. This methodology is referred as "LiDAR-Assisted Multi-source Programme" (LAMP) that combines a coverage of LiDAR sample with field plots, and wall-to-wall satellite data to develop forest biomass statistics and map of up to one hectare spatial resolution (Gautam et al. 2013, 2010). The method can be applied in many areas when adjusted to local biophysical conditions.

In this paper we review the LAMP method in a context of three different case studies. The case studies are from Lao PDR (Gautam et al. 2010), Nepal (Gautam et al. 2013) and Ghana (Sah et al. 2012). At first, the general LAMP process is described briefly, then all three studies are presented with the main results and lastly, in discussion chapter we discuss about common findings we have made in the case studies.

2. MATERIALS AND METHODS

2.1 Lidar-Assisted Multi-source Programme (LAMP)

The LAMP method follows a two-phase estimation approach. In the first phase, forest variables related to biomass are estimated with high accuracy from LiDAR information in selected sample areas where full-coverage LiDAR data and ground-truth plots are collected. A rectangular sample block or strip sample design is applied to sample LiDAR data over the area of interest. The field plots are used as training data set for the first-phase biomass estimation. In the second phase, these highly accurate estimates in the LiDAR sample area are used as surrogate plots (simulated field plots) in the interpretation of medium-resolution satellite scenes for the entire study area (Gautam et al. 2010).

LAMP phase 1: Estimating forest parameters for LiDAR coverage area

In the first phase of the LAMP approach, a regression model is generated based on the relationship between LiDAR metrics (height and density) and field measurements. It has been shown that Sparse Bayesian methods offer a flexible and robust tool for regressing LiDAR echo histograms with forest parameters. While performing comparably to traditional regression methods, they are computationally more efficient and allow better flexibility than step-wise regression (Junttila et al. 2008, Junttila et al. 2010). The Sparse-Bayesian regression model is applied to predict forest characteristics for a set of thousands of surrogate plots of about 1-hectare size within the forested area of the LiDAR coverage.

LAMP phase 2: Expanding the estimates to the entire area of interest (AOI) using satellite data

In the second phase of the LAMP approach, the forest characteristics that are estimated for the surrogate plots from LiDAR data are applied as simulated ground-truth to generate a regression model between bio-physical forest parameters and features derived from satellite imagery. Again, we use the Sparse Bayesian method to regress satellite-derived variables with forest characteristics for the locations of the surrogate plots. The satellite-based variables are derived from the satellite data's spectral and textural features and vegetation indices as zonal mean values for the area within each surrogate plot.

In the second phase it is possible to produce Tier 2 level output for forest classes by using surrogate plot estimates to derive forest class specific estimates for mean and variance. This can be done if the inventory area is classified to meaningful forest classes using, for example, satellite data.

The final Tier 3 level output includes biomass and carbon estimates for higher spatial resolution. The spatial resolution of the Tier 3 level outputs is up to 1 hectare pixel size. The mean value biomass/carbon calculated from the forest class mean values (Tier 2 level output) and the mean from Tier 3 level output (1 hectare grid) are the same and both are unbiased estimates of the mean biomass/carbon within the area of interest.

2.2 Case studies

The LAMP method was used in three case studies. In this section we will briefly describe the main features of each case study. The summary of materials of each case study are presented in Table 1.

	Lao PDR	Nepal	Ghana
Inventory area, km^2	250	23300	15153
LiDAR sample area, % of inventory area	10	5	5
Laser scanner	Leica ALS 40	Leica ALS 50-II	Leica ALS 50-II
Point density, points/m^2	~1	~0.8	~2.0
Satellite data	Alos AVNIR-2	Landsat 5 TM	Alos AVNIR-2/ DMC
Number of images	1	5	7/1
Field plot type	rectangular	circle	rectangular
Field plot area, m^2	400	500	400
Number of field plots	328	738	254

Table 1. Summary of the materials of the case studies.

2.2.1 Lao PDR

The study area was situated in Vientiane, Lao PDR. LiDAR survey covered the whole study area, total of 25 000 hectares. The LiDAR survey and field campaign were carried out in 2009 and the satellite imagery (ALOS AVNIR-2 and Landsat 7) were from years 2006 and 2000, respectively. To test the two-phase LAMP approach, every tenth LiDAR flight line (strip) was used as a sample. In order to acquire plot sample that well represent the variation in biomass in the study area, surrogate plots were placed randomly with probability proportional to the estimated biomass. 90 % of the area was interpreted using satellite data with the regression models based on the LiDAR estimate of the surrogate plot data. (Gautam et al. 2010)

2.2.2 Nepal

Nepal study area was located in Terai Arc Landscape in southern Nepal. The LiDAR survey and field campaign were carried out in 2011 and the satellite imagery, total of 5 Landsat 5 TM images, were from years 2010 and 2011. The applied LiDAR sample was a weighted random block sample, with the block size of 5 km by 10 km. The weights were decided by expert judgement of forest types' variation. The LiDAR sample covered 5 % of the total of over 23 300 km^2 study area. Total of 738 systematically located field plots measured inside the LiDAR blocks were used in phase 1. The images were radiometrically normalized and mosaicked. In LAMP phase 1 biomass models were estimated on LiDAR features and field measured biomass. In LAMP phase 2 the phase 1 models were used to generate 10 000 surrogate plots of 1 hectare size and the surrogate plot estimates were used to generate phase 2 model. The final result was a grid level estimates (cell size 1 hectare) for the whole study area. (Gautam et al. 2013)

2.2.3 Ghana

The study area was located in the western border of Ghana. The total area was 15153 km^2 of which 5 % was covered with a systematic LiDAR strip sample. Seven ALOS AVNIR-2 and one DMC satellite scenes were used for producing Land use classification for the study area. (Sah et al. 2012)

Weighted cluster plot sample of 254 field plots was used to generate phase 1 regression models for total of 4 forest strata (closed canopy forests, open forests and croplands within wet, moist and dry zones). In phase 2 the phase 1 models were used to estimate mean and variance for each forest zone by using the whole LiDAR sampled area. The final result of the LAMP in this case was expected to be a tier 2 level output, i.e. LiDAR-model derived means and variances per each ecological forest zone.

3. RESULTS

3.1 LAMP phase 1 estimation results

The phase 1 biomass estimates were validated against field plots. In Lao PDR the relative root mean squared error (RMSE%) value for mean aboveground biomass estimate was 23.3 % when the LiDAR estimates were validated against the original field plots of size 400 m^2 (Gautam et al. 2010). In Nepal the phase 1 LiDAR models were validated against independently sampled larger validation plots of size 2826 m^2. The RMSE% was 17.0 (Gautam et al. 2013). In Ghana the LiDAR model results were validated in a similar way as in Lao PDR.

3.2 LAMP phase 2 estimation results

In Lao PDR and Nepal case studies, where tier 3 level outputs were produced, the phase 2 outputs were validated against LiDAR estimates of surrogate plots. In Lao PDR RMSE% of mean aboveground biomass estimate was 23.9 at 1 hectare level (Gautam et al. 2010). In Nepal the corresponding RMSE% value was 42.1 (Gautam et al. 2013). The reported RMSE% values do not include the phase 1 model bias, and therefore, may underestimate the true error slightly. The saturation of satellite imagery's signal is a common problem when optical satellite data is applied in biomass estimation. In Lao PDR the RMSE% value indicated better estimation results than in Nepal. By comparing the scatterplots of the surrogate plot biomass and LAMP phase 2 output it can be noted that in Lao PDR the satellite signal saturates at about 200 tons/ha, whereas in Nepal case study there is no such an effect visible (Figure 1). The Nepal estimates were produced by retaining the variation of surrogate plots. This decreases the saturation effect, but increases the RMSE%. In Ghana tier 2 level output was provided, and therefore, no RMSE% values for 1 hectare level were available.

3.3 Findings concerning materials and costs

In the following chapters we present some of the key findings of the three case studies concerning the specifications of the input data materials and the reference data costs.

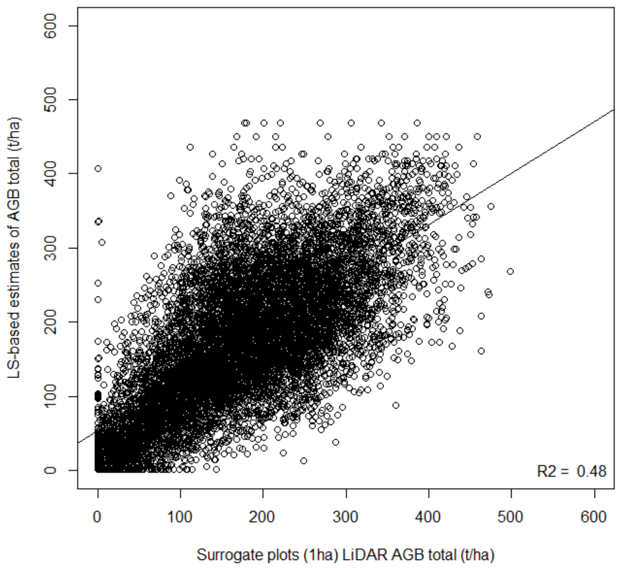

Figure 1. The scatterogram of Nepal case study LAMP pahse 2 estimates and surrogate plot LiDAR estimates at 1 hectare level (Gautam et al. 2013).

3.3.1 Field plots

Field plots used in LAMP have some requirements that differ from field based surveys. These are plot positioning and plot size. To overlap field data with LiDAR data accurately, the plot positioning error should be minimized and the plot size maximized. According to the experiences from the three case studies field plots should be positioned with differentially corrected GNSS to achieve sub-meter accuracy when plot-size is relatively small.

Optimal field plot size is related to plot positioning accuracy, spatial pattern of tree locations and the size of the tree crowns. In field sample the tree is decided to be in plot, if its trunk centre from the height of 1.3 meters is inside the plot. If the tree is close to the plot border, great part of the crown can be outside the plot. LiDAR sees the crown and the trees, which trunk is inside the plot, but at least part of the canopy outside the plot are differently presented in field data and LiDAR. This border effect increases as a function of inverse plot size and mean tree crown size. The spatial pattern of tree locations affects so that more clustered it is the more significant the border effect can be. In regular patterns the effect can be large, too, if the plot shape and location does not take into account the spatial trend. In natural tropical forests the tree crowns can be very large, tree sizes can have a lot of variation and spatial pattern is random or clustered. Thus the optimal plot size is usually larger in natural tropical forest than, for example, in plantations.

3.3.2 LiDAR sample design

LiDAR sample design is a trade of between costs and accuracy of the estimates. In the most straight forward approach the whole area is scanned with wall-to-wall LiDAR. Usually, this is not a feasible solution and some kind of sampling strategy should be used instead. 2-10 % sample rate is sufficient for LAMP. A systematic or random strip sample allows a good presentation of the whole AOI and gives a good starting point to estimate unbiased biomass/carbon estimates. However, using strip sampling may give poor presentation for forest types which present only minority of the whole inventory area. Block sampling can be more efficient in large and fragmented forest areas, since blocks can be designed so that large enough sample is collected from each forest type. The costs of LiDAR collection are very much dependent on LiDAR sample design, and therefore, the sample design should be optimized for each project individually taking into account the representativeness of the sample and issues considering the total flight time.

In all three case studies discrete return LiDAR was used. There were some indications that if the vegetation structure is very dense the pulse penetration to the ground can be a problem. Thus, to provide enough ground observations full waveform LiDAR could give more reliable data.

3.3.3 Satellite data

LAMP is not dependent on any particular satellite data. Requirements are that the signal in satellite data should correlate with the biomass or carbon and that the geometric accuracy is good compared to pixel size. Using the LiDAR estimated surrogate plots allows the use of large basic estimation unit, for example 1 hectare pixel size, which is not feasible if field plots are used directly. This feature of LAMP gives excellent possibilities to use low or medium resolution satellite data. The challenge of satellite data is the image normalization without losing the signal. The correlation between the optical satellite data features and the amount of biomass is low compared to correlation of LiDAR features and biomass. Further processing the imagery so that the images are spectrally equivalent does not usually improve the correlation.

3.3.4 Reference data cost

LiDAR provides cost-efficient means of acquiring reference data for above ground biomass inventories and distribution mapping. The cost benefits are the highest in the conditions where the rate of above ground biomass variation within forest strata and field measurement costs are high. In case of the field plot based approach more plots have to be collected when there is more above ground biomass variation to achieve the targeted estimation accuracy. Also, digital terrain model reference is necessary when integrating field reference and radar datasets and validating biomass maps with high spatial resolution and accuracy (Mitchard et. al 2009).

The figures 2, 3, 4 and 5 illustrate the scale and conditions when the LiDAR acquisition with a limited number of modelling plots (50) per stratum generates cost savings. The LiDAR acquisition reference costs are modelled based on the operational costs from various country cases in Asia, Africa and Latin America. To provide a realistic view the sensor and operators are mobilised from abroad, even though local service providers and LiDAR data archives are available in many countries. The estimation model also considers mobilisation, acquisition and pre-processing costs. In addition, 2-% LiDAR coverage per stratum expected to be sufficient and field measurement costs per plot (USD 100 – 2500) to be equal in both LiDAR-assisted and plot-based cases.

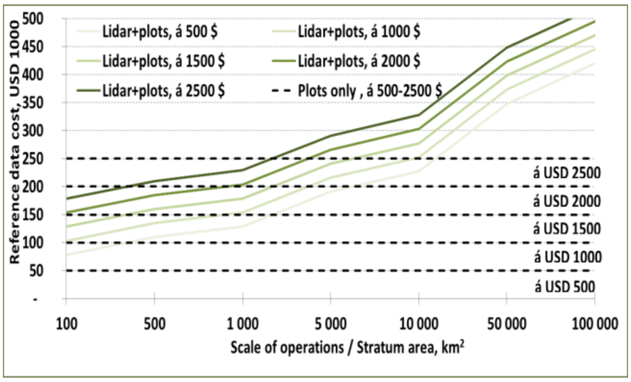

Figure 2. The scale of operations and reference data acquisition costs for LiDAR-assisted and plot-based approaches. The required number of measured plots is assumed to be 100 for the plot-based approach.

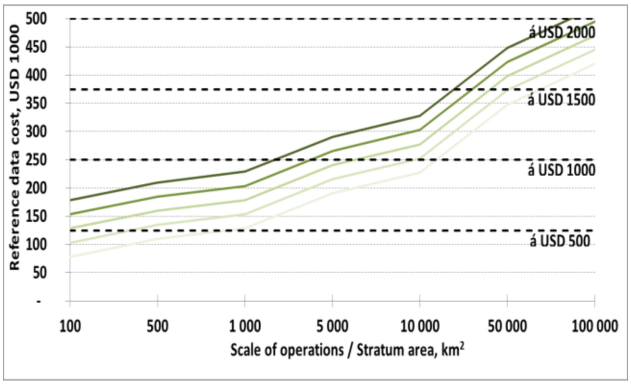

Figure 3. The scale of operations and reference data acquisition costs for LiDAR-assisted and plot-based approaches. The required number of measured plots is assumed to be 250 for the plot-based approach.

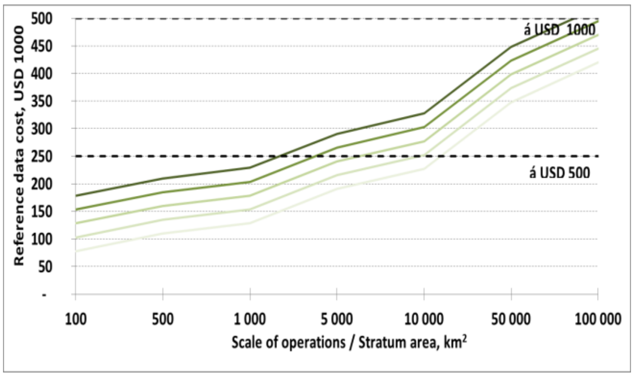

Figure 4. The scale of operations and reference data acquisition costs for LiDAR-assisted and plot-based approaches. The required number of measured plots is assumed to be 500 here for the plot-based approach.

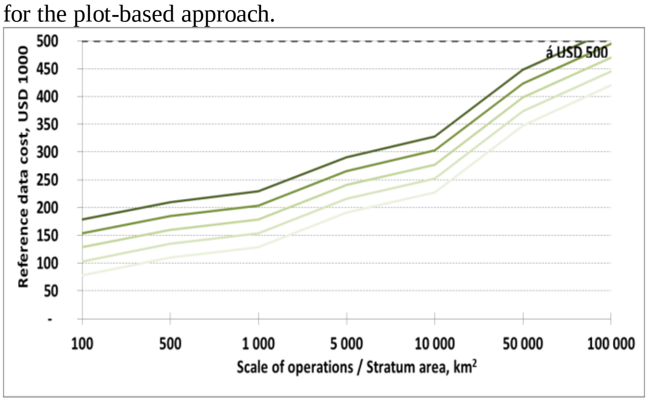

Figure 5. The scale of operations and reference data acquisition costs for LiDAR-assisted and plot-based approaches. The required number of measured plots is assumed 1000 here for the plot-based approach.

4. CONCLUSIONS

LAMP is an agile, scalable and reproducible approach for large area biomass and carbon inventories. It can be applied for several forest inventory problems and it is not dependent on some certain input data. However, this gives a user a variety of parameters, which affect to the end result quality. The reference data costs can be significantly lower and the output data more valuable than in traditional field-based inventories or in satellite based inventories not applying LiDAR.

References

Angelsen, A., Boucher, D., Brown, S., Merckx, V., Streck, C., & Zarin, D. (2011). *Guidelines for REDD+ reference levels: Principles and recommendations.* Prepared for the Government of Norway by the Meridian Institute. http://www.REDD-OAR.org (20.11.2013).

Angelsen, A., Brown, S., Loisel, C., Peskett, L., Streck, C., & Zarin, D. (2009). *Reducing Emissions from Deforestation and Forest Degradation (REDD): An options assessment report.* Prepared for the Government of Norway by the Meridian Institute http://www.REDD-OAR.org (20.11.2013).

Asner, G.P., Mascaro, J., Anderson, C., Knapp, D.E., Martin, R.E., Kennedy-Bowdoin, T., Breugel, M.V., Davies, S., Hall, J.S., Muller-Landau, H.C., Potvin, C., Sousa, W., Wright, J. and Bermingham, E. (2013). High-fidelity national carbon mapping for resource management and REDD+. Carbon Balance and Management, 8:7. http://www.cbmjournal.com/content/8/1/7 (20.11.2013).

Asner, G.P., Clark, J.K., Mascaro, J., Vaudry, R., Chadwick, K.D., Vieilledent, G. & Knapp, D.E. (2012). Human and environmental controls over aboveground carbon storage in Madagascar. *Carbon Balance and Management*, 7, 2. doi:10.1186/1750-0680-7-2

Asner, G.P., Hughes, R.F., Varga, T.A., Knapp, D.E., & Kennedy-Bowdoin, T. (2009). Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems*, 12, 261-278.

Gautam, B., Peuhkurinen, J., Kauranne, T., Gunia, K., Tegel, K., Latva-Käyrä, P., Rana, P., Eivazi, A., Kolesnikov, A., Hämäläinen, J., Shrestha, S.M., Gautam, S. K., Hawkes, M., Nocker, U., Joshi, A., Suihkonen, T., Kandel, P., Lohani, S., Powell, G., Dinerstein, E., Hall, D., Niles, J., Joshi, A., Nepal, S., Manandhar, Kandel, U. Y. & Joshi, C. (2013). Estimation of Forest Carbon Using LiDAR-Assisted Multi-source Programme (LAMP) in Nepal. Presented at International Conference on Advanced Geospatial Technologies for Sustainable Environment and Culture, 12-13 September, Pokhara, Nepal (an event of ISPRS, Technical Commission VI, Education and Outreach, Working Group 6.) 7 p.

Gautam, B., Tokola, T., Hämäläinen, J., Gunia, M., Peuhkurinen, J., Parviainen, H. & Sah, B. (2010). Integration of airborne LiDAR, satellite imagery, and field measurements using a two-phase sampling method for forest biomass estimation in tropical forests. In B. Gautam (Ed.), *International Symposium on "Benefiting from Earth Observation".* 4 -6 October 2010, Kathmandu, Nepal. (pp. 1-7).

Intergovernmental Panel on Climate Change. (2007). *Climate Change 2007: Synthesis Report.* The Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge and New York: Cambridge University Press. http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_synthesis_report.htm (20.11.2013).

Intergovernmental Panel on Climate Change. (2003). Good practice guidance for land use, land-use change and forestry. Hayama, Kanagawa, Japan: Institute for Global Environmental Strategies. http://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_contents.html (20.11.2013).

Junttila, V., Kauranne, T., & Leppänen, V. (2010). Estimation of Forest Stand Parameters from Airborne Laser Scanning Using Calibrated Plot Databases. *Forest Science*, 56, 257-270.

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

Mitchard, E., Saatchi, I., Woodhouse, I., Nangendo, G.,Ribeiro, N., Williams, M., Ryan, C., Lewis, S., Feldpausch, T. & Meir, P. (2009). Using satellite radar backscatter to predict aboveground woody biomass: a consistent relationship across four different African landscapes. *Geophysical Research Letters*, 36, L23401.

Næsset, E. (2007). Airborne laser scanning as a method in operational forest inventory: status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research*, 22, 433-442.

Sah, B. P., Hämäläinen, J. M., Sah, A. K., Honji, K., Foli, E. G., and Awudi, C. (2012). The Use of Satellite Imagery to Guide Field Plot Sampling Scheme for Biomass Estimation in Ghanaian forest, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, I-4, 221-226.

Walker, W.S., Stickler, C.M., Kellndorfer, J.M., Kirsch, K.M., & Nepstad, D.C. 2010. Large-Area Classification and Mapping of Forest and Land Cover in the Brazilian Amazon: A Comparative Analysis of ALOS/PALSAR and Landsat Data Sources. *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, 3, 594-604.