THE POTENTIAL OF DISCRETE RETURN, SMALL FOOTPRINT AIRBORNE LASER SCANNING DATA FOR VEGETATION DENSITY ESTIMATION

Felix Morsdorf[1], Benjamin Koetz[1], Erich Meier[1], K.I. Itten[1] and Britta Allgöwer [2]
1) Remote Sensing Laboratories
2) Geographic Information Systems
Department of Geography
University of Zurich
email:morsdorf@geo.unizh.ch

Working Group III/3

KEY WORDS: laser scanning, correlation, point cloud, forestry, land cover

ABSTRACT

We evaluate the potential of deriving a vegetation leaf area index (LAI) from small footprint airborne laser scanning data. Based on findings from large area histograms of discrete laser returns for two contrasting plots, LAI is estimated from the fraction of first to last and single returns inside the canopy. The canopy returns are classified using thresholding of LIDAR raw data heights subtracted by interpolated digital terrain model heights. This should yield LAI estimates being independent of fractional vegetation cover, an ambiguity many passive optical approaches suffer from. Validation is carried out using 78 georeferenced hemispherical photographs, with LAI and gap fractions for a range of zenith angles being computed using the Gap Light Analyzer (GLA, Frazer et al. [1997]). Since the range sensitivity of the hemispherical photographs is not a priori known, we use variable LIDAR data trap sizes to find a suitable diameter. This is achieved searching the maximum R^2 value of the regression for the trap size range from 5 to 50 m diameter. Larger diameters (30 m) provide best results for our canopy types. Regressions of LIDAR estimates shows a moderate agreement with field data based on hemispherical photographs, with R^2 0.6 for LAI. Due to either heterogenity of the canopy or geolocation errors, a quite large amount of noise seems to be attributed to the regression, explaining the somewhat low values of R^2.

1 INTRODUCTION

The availability of robust estimations of vegetation density measures such as LAI is critical for a number of applications. Passive optical remote sensing systems are most popular for the indirect retrieval of LAI, due to their availability and relatively low costs. Most often regression methods using some kind of band ratio are applied [Cohen et al., 2003; Colombo et al., 2003], and in some cases radiative transfer modeling [Koetz et al., 2004; Atzberger, 2004] is used. One of the largest problems that these approaches face is the unknown or not well known vertical structure of the canopy. LIDAR (**LI**ght **D**etection **A**nd **R**anging) systems can overcome this short come by penetrating the canopy and revealing the vertical stratification of the canopy. Thus, LIDAR systems have been widely used for stand wise derivation of structural parameters [Lovell et al., 2003; Means et al., 2000; Lefsky et al., 1999], often by means of regression methods choosing some LIDAR predictor variables (e.g. height percentiles) for ground based measures of structural information [Naesset, 2002, 2004; Cohen et al., 2003; Andersen et al., 2005]. With high point density, the derivation of single tree metrics becomes possible, it's feasibility has been shown by Hyyppae et al. [2001]; Andersen et al. [2002]; Morsdorf et al. [2004]. Some researchers have already assessed the potential of deriving LAI from laser scanning data. Riano et al. [2004] used a relation from Gower et al. [1999] to compute LAI from the gap fraction distribution derived by means of airborne laser scanning, whereas Lovell et al. [2003] used ground based laser range finder information to model LAI. Radiative transfer modeling of LIDAR waveforms for different canopy types has also been accomplished [Sun and Ranson, 2000; Ni-Meister et al., 2001], but so far only for large footprint sensors. Small footprint sensors are in most cases only capable of recording discrete returns (e.g. first and last return), but not the entire waveform. Still, these discrete returns contain valuable information about the vegetation density and structure at a much higher spatial resolution (O $\sim$ 1 m) that large footprint sensors can not provide. It has been shown that first and last returns can be used to model stand properties such as basal area and LAI [Lim et al., 2003], but first and last pulse information have been put into a single vegetation class in this study. Our objective is to evaluate the potential of estimating LAI from small footprint laser data exploiting the information contained in multiple return types, using field measured LAI values as validation.

2 DATA

2.1 Site description and Field Data

The study area for the acquisition of the field data is located in the eastern Ofenpass valley, which is part of the Swiss National Park (SNP). The Ofenpass represents a dry inner-alpine valley with rather little precipitation (900-1100 mm/a). Surrounded by 3000 meter peaks, the Ofenpass valley starts at about 1500 m a.s.l. in the west and quickly reaches an average altitude of about 1900 m a.s.l towards the east. The south-facing Ofenpass forests, the location of the field measurement, are largely dominated by mountain pine (Pinus montana ssp. arborea) and some stone pine (Pinus cembra) as a second tree species, being of interest for natural succession. These forest stands can be classified as

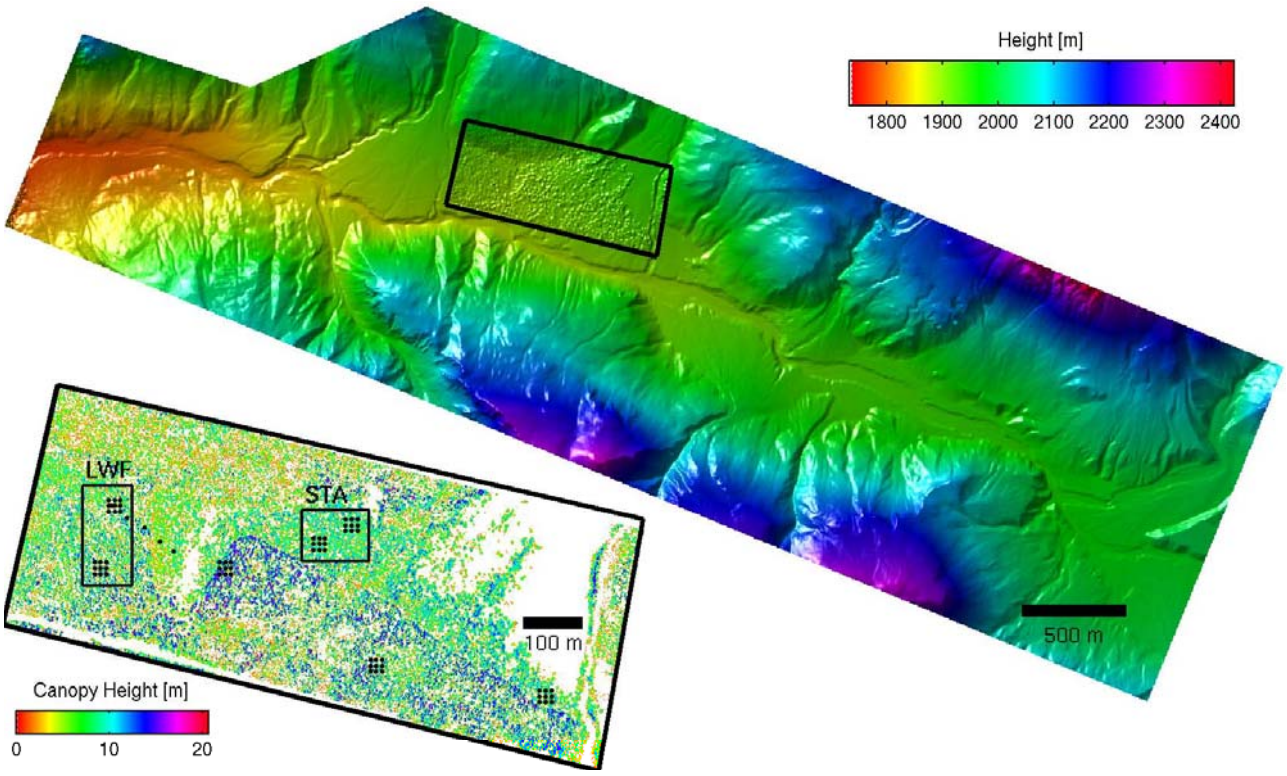

Figure 1: The Digital Terrain Model (DTM) of the Ofenpass area in the Swiss National Park. The smaller area marked by the black box was sampled with higher point density due to the lower flying height of 500 m above ground. A canopy height map of that area is displayed in the lower left. Black dots mark positions of hemispherical photographs. Black squares mark areas where the histograms in Figures 2 and 3 were computed from.

woodland associations of Erico-Pinetum mugo. The understory is characterized by low and dense vegetation composed mainly of Vaccinium, Ericaceae, and Seslaria species. In Figure 1 an overview of the test site is given. More than 20 % of the stand are upright standing dead trees, with the minimum tree age being 90 years, the mean and maximum being 150 and 200 years respectively. The whole stand has regenerated after a period of clear cutting in the 18th and 19th century, and has been without any management since the foundation of the Swiss National Park in 1914. The main cause for dying of the trees is the root rot fungi, as described in Dobbertin et al. [2001]. The small plot in Figure 1 shows a canopy height model (CHM) of the area over flown with the lower altitude. Black dots indicate positions where hemispherical photographs were taken. 73 of these plots were geolocated by means of handheld GPS and five using a differential GPS.

2.2 Deriving field estimates of LAI

LAI was first defined as the total one-sided area of photo synthetic tissue per unit ground surface area Watson [1947]. This definition is only valid for broad leaf forests though, and consequently Myneni et al. [1997] defined the LAI as the maximum projected leaf area per unit ground surface area. There are various ways of determining LAI in the field, a contemplative summary is given by Jonckheere et al. [2004]. Methods can be categorized in two classes, direct and indirect. Direct methods generally use destructive sampling to estimate the total number of leaves on a tree and their included angles and distribution to estimate LAI. Indirect methods mostly measure some aspect of the radiative regime and infer the LAI from the distribution of light inside the canopy. These approaches rely on some homogeneity inside the canopy (e.g. uniform leaf angle distribution) and tend to produce errors if these constraints are not fulfilled. Consequently, a lot of samples are needed in the case of heterogene canopy to get stable and reliable estimates of these parameters [Weiss et al., 2004]. Using a Nikon 4500 digital camera together with fish-eye lens and a tripod, we took 73 hemispherical photographs at plots located by means of a handheld GPS, thus introducing a positional uncertainty of a about 5 m. Five were were geolocated using differential GPS equipment, thus reducing the positional uncertainty to some centimeters. The hemispherical photographs were analyzed using the Gap Light Analyzer (GLA, [Frazer et al., 1997]), by means of manual thresholding.

2.3 Laser Scanning Data

In October 2002 a helicopter based LIDAR flight was carried out over the test area, covering a total area of about 14 km^2. The LIDAR system used was the Falcon II Sensor developed and maintained by the German company TopoSys. The system is a push-broom laser altimeter recording both first and last reflection from the laser signal (first/last pulse). The flight was conducted with a nominal height over ground of 850 m, leading to an average point density of more than 10 points per square meter (p/m^2). A

smaller subset of the area (0.6 km^2) was over flown with a height of 500 m above ground, resulting in a point density of more than 20 p/m^2, thus combining the two datasets would yield to a point density of more than 30 p/m^2 for both first and last pulse. We only used data from the lower flight in this study. The footprint sizes were about 0.9 m in diameter for 850 m flight altitude and about 0.5 m in diameter for 500 m altitude. The raw data delivered by the sensor (x,y,z - triples) was processed into gridded elevation models by TopoSys using the company's own processing software. The Digital Surface Model (DSM) was processed using the first pulse reflections, the Digital Terrain Model (DTM) was constructed using the last returns and filtering algorithms. The grid spacing was 1 m for the large area and 0.5 m for the smaller one, with a height resolution of 0.1 m in both cases. A quality analysis of the raw data was done using six artificial reference targets and is described in detail in Morsdorf et al. [2004]. Standard deviation of height estimates based on raw echos on these targets were as low as 6 cm, with the internal accuracy of the LIDAR data being well below pixel size of the raster models.

3 METHODS

The triggering of a laser echo from an return signal depends on both the size and the reflectivity of the illuminated target, with the latter having a larger influence on a target's visibility than the first [Baltsavias, 1999]. In case of vegetation, it will depend as well on the distribution of reflecting elements such as leaf or needle shoots. In most vegetation cases, a first return is triggered close to the top of the canopy, with an systematic underestimation due to vegetation density issues and potential undersampling (see Gaveau and Hill [2003] for details). If the vegetation is not too dense, a part of the beam can further penetrate the canopy, until the threshold for the intensity is surpassed a second time and the so called last return is triggered. Depending on the vegetation openness and density this can be on the ground or inside the vegetation. A minimum distance needs to be in between first and last return for their separation, which depends primarily on the pulse duration of the laser emitter. For the Falcon II systems this minimum distance is a about a meter, which can be seen in Figures 2 (a) and 3 (a). In the two figures a height distribution of the difference of first and last return is shown, and one can clearly notice the gap from zero to about 1.2 m. The heights in the lower panels, namely Figures 2 (b) and 3 (b), have been subtracted by terrain heights interpolated to raw data coordinates from the DTM Toposys provided. With a system recording first and last pulse, we can classify three different types of returns scenarios :

- first echo
- last echo
- single echo, first return = last return

The term *single echo* describes the case where only one return is triggered from a return signal, resulting in both values having the same height. Most single returns will come from plain surfaces such as roads or generally from the ground, but there are some in the vegetation, as can be seen in the lower panels of Figures 2 and 3. In these figures are a illustration of these three return types. Colored in medium and darker gray, the last and single returns are (if not on ground) mostly concentrated in the upper canopy, with their maximum just before the maximum of total returns. Note that limits of the y-axis have been lowered for better visibility of the vegetation part of the histogram, and thus, the percentage of last/single returns to first return and the absolute number of ground returns can not be drawn from this graph. Histograms have been derived from areas with different vegetation densities called LWF (Fig. 2) and STA (Fig. 3). These areas are marked by black squares in Fig. 1.

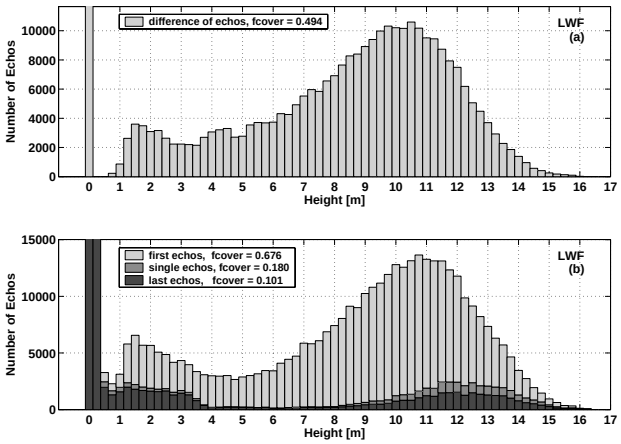

Figure 2: Histogram of difference of first and last pulse (a), histogram of first, last and single echos (b) for site LWF.

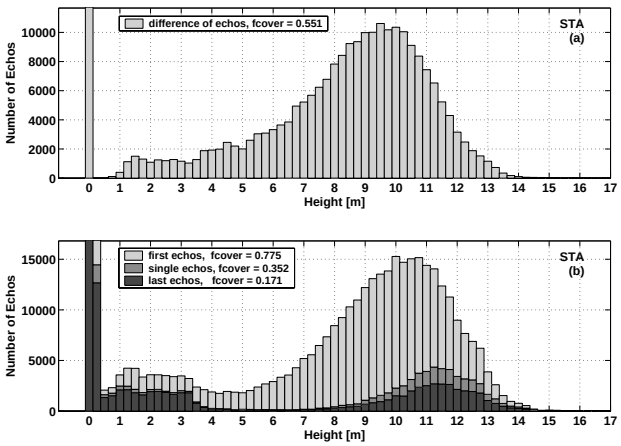

Figure 3: Histogram of difference of first and last pulse (a), histogram of first, last and single echos (b) for site STA.

Our hypothesis is that the probability of either of these three return types will depend on vegetation density, if vegetation reflectance can be assumed to be constant. Whether this assumption is valid for our study, will be discussed below. When comparing the two sites LWF and STA, one can note that the fraction of last and single returns in the upper canopy and lower canopy is higher for STA then for LWF, with the mean LAI from field measurements being about

1 at LWF and about 2 at STA. We concluded that the vegetation density could be derived by using this fraction as a proxy. Thus, we propose to compute a LAI proxy from LIDAR data by:

$$LAI_{prox} = \frac{\sum E_{first}}{\sum E_{single} + \sum E_{last}} \quad (1)$$

E_{first}, E_{single} and E_{last} denote the three types of returns described in detail above, but only for crown returns. The observation of different return fractions for different vegetation densities from Figures 2 and 3 leads to the verbalisation of Equation 1, which is a quantitive transcription of our hypothesized relationship. The vegetation returns are classified by thresholding the height over terrain of the raw laser hits with a value of 1 m. This distinguishability is an advantage of small footprint laser scanners raw data combined with a DTM over large footprint systems and other passive optical approaches.

One of the largest problems in validating the LAI by indirect methods arises from the fact that needles are clumped into shoots and the dispersion of crowns in the canopy, which manifests a discontinous plant canopy. These two effects or better, generally the inhomogenous distribution of canopy elements (crowns,branches,twigs,shoots and leafs) on all possible scales need to be adressed. For many canopy types, clumping at shoot scale and inhomogenous crown dispersion are the dominant effects Ni-Meister et al. [2001]. Clumping at shoot scale can be adressed by correcting the indirect LAI estimates (often called effective LAI, LAI_{eff} with a factor depending on the projection function of canopy elements [Weiss et al., 2004]. We decided to derive only LAI_{eff}, since a simple coefficients does not alter the quality of our regressions. If one needs values for true LAI, one would have to multiply our values by 1.75 [Koetz et al., 2004]. Dispersion of crowns inside the canopy is not an issue for our approach, since we compute the fraction of returns only from returns being off ground more than a meter, hence trees, since no shrubs are present on our study site. Thus, LAI proxy is only derived from crowns, due to the advantage of high point density, small footprint scanner being able to discrimate vegetation from ground returns based on the Digital Terrain Model. In order to assess the feasibility of our approach regarding the assumption on uniform reflectance of the canopy, we conducted some tests using PROSPECT [Jacquemoud and Baret, 1990], modeling the reflectance of the green canopy elements in our study site using average leaf parameters from field measurements [Koetz et al., 2004]. We varied the moisture content, as at 1560 nm absorbtion due to moisture is the dominant effect, in a range observed in the field and calculated the leaf reflectance. This yielded 20.8 % reflectance for the lowest of moisture observed and 19.2 % for the highest value observed, making up for an absolute difference of only 1.6 %. This is small enough to be neglected, considering results from practical test using artifical targets on target visibility using different reflectances [Wotruba et al., 2005]. All other parameters of PROSPECT were left constant, since our test site is predominantly covered by only one tree type, mountain pine.

4 RESULTS

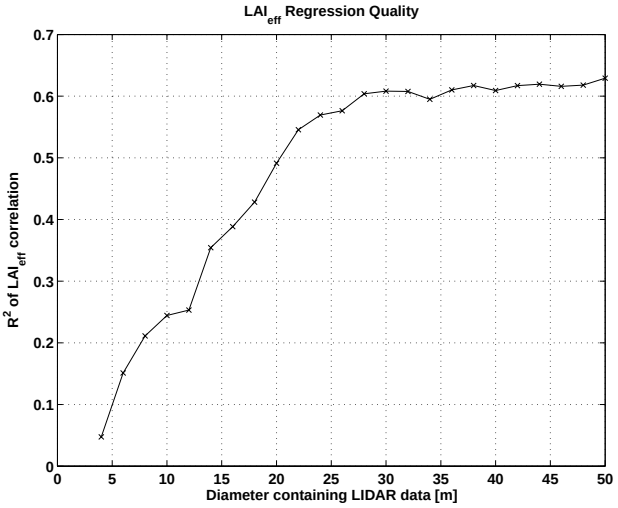

Figure 4: Coefficient of determination (R^2) for LAI regression for a range of LIDAR raw data patches. Field measured LAI was computed from hemispherical photographs.

Since we did not know how large the LIDAR data trap size needs to be in order to capture the same area as the hemispherical photographs, we computed our LAI proxy for a range of LIDAR data trap sizes. Then we did a regression analys and computed the R^2 as a measure of regression quality. The outcome is depicted in Fig. 4. The values of R^2 are as low as 0.1 for patches being 5 m in diameter and reaches a local maximum of about 0.6 at 30 m in diameter. For values larger than that, there is no significant rise of R^2. Thus, we chose a diameter of 30 m for the LAI regressions in Fig. 5. The lack of a clearly visible inflection point for the values of R^2 in Fig. 4 is probably due to the fact that for higher data trap sizes the values of LAI become more similar to each other. This yields a flatter regression curve with lower noise, and thus still a high R^2, even though the meaningfulness of the regression is reduced. The quite large size of the LIDAR raw data trap could lower the effect of inprecise georeferencing to some extent, even though the inner few meters of a hemispherical photo should contribute more to field measured LAI values.

In Fig. 5 the regression of LIDAR derived LAI_{eff} and field measured (hemispherical photographs) is depicted. R^2 is 0.6, with the RMS being 0.03. The regression models coefficients were used to compute LAI_{eff}, the original values of our LAI_{eff} proxy (fraction of first to single/last pulse in vegetation) were in the range of 0.1 to 0.3. The black line denotes the one on one relationship for computed LAI_{eff} values. It is also visible that the range of values (hence noise) around the one on one relation is larger for higher LAI_{eff} values.

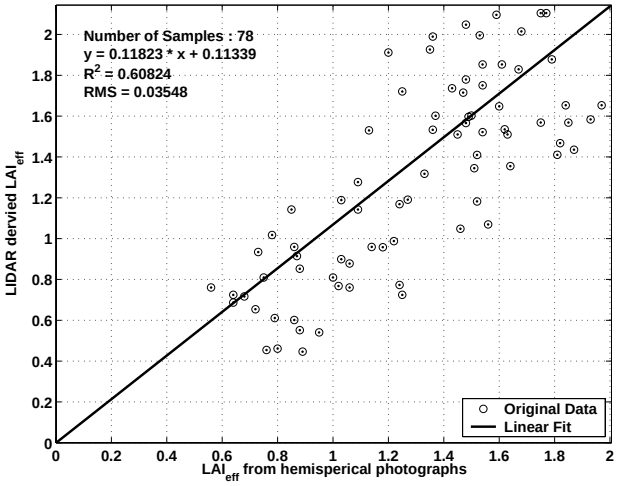

Figure 5: Regression of LIDAR derived LAI_{eff} with LAI_{eff} from hemispherical measurements. Processing of hemispherical photgraphs was done using GLA. A LIDAR data trap size of 28 m was used.

5 DISCUSSION AND CONCLUSIONS

In the past years, airborne laser scanning has been established as a valuable tool for forest structural analyses. We have tried to establish a new predictor variable for vegetation density, namely LAI, by using the fraction of different returns types inside the canopy. Regression of the LIDAR estimates with field data from hemispherical photographs showed moderate agreement, with the R^2 being 0.6 and the RMS being 0.035. Quite an amount of noise was visible in our regression, leading to somewhat lower R^2 values. Possible cures could be the use of more samples and even more precise georeferencing, e.g. differential GPS measurements for all sampling locations. For portability of this approach to other study sites, a vegetation specific correction parameter could be derived. This parameter will probably be influenced by two factors: reflectance differences and clumping of canopy elements at smaller scales than footprint size. Clumping is a well studied phenomenon, since it needs to be corrected for most field based methods as well (e.g. LAI2000, hemispherical photographs, ceptometer, [Jonckheere et al., 2004]). Correction factors have been derived for various canopy types, that are used in day to day field work. Hence, it would not be too big of a problem to extent their use to airborne LIDAR data. Canopy reflectance, though, is an issue that needs to a more complicated procedure. For LIDAR - LAI_{eff} estimates to be inter-comparable, the reflectance of canopy elements in the wavelength of the laser (our sensor is at 1560 nm) should be roughly the same. If not, a correction factor needs to be applied. Modern LIDAR systems allow as well the recording of the intensity of the returned signal, so this information could be used for assessing effects of canopy reflectance differences. Unfortunately, this option was not yet available in 2002 when our flight campaign was carried out. A more sophisticated treatment of the three return classes to retrieve LAI, instead than just taking the fraction of first to single and last returns could also improve the results. It might be that the single returns a more sensitive to LAI changes and vice versa. This will be looked at in future studies. It would be a benefit to have instruments that deliver inter-comparable results for LAI from very different areas. This can only be achieved by taking the radiative regime and it's involved physical processes into account. We have tried to take a step into that direction by choosing direct, physically meaningful metrics of LIDAR raw data for LAI opposed to regression or multiple regression with a set of statistical predictor variables as in Naesset [2004] or Lim et al. [2003]. Combined with the method for single tree extraction [Morsdorf et al., 2004], the work we presented would allow the direct retrieval of an close to "actual" foliage profile from airborne laser scanning data [Ni-Meister et al., 2001; Lovell et al., 2003]. This is due to the fact that only the small scale clumping (smaller than branch level) needs to be corrected, whereas crown dispersion is actually known due to high point density and small footprint size of the laser scanner used in this study.

6 ACKNOWLEDGMENTS

This project is funded by the EC project "Forest Fire Spread and Migitation" (SPREAD), EC-Contract Nr. EVG1-CT-2001-00027 and the Federal Office for Education and Science of Switzerland (BBW), BBW-Contract Nr. 01.0138.

References

Andersen, H.-E., McGaughey, R. J. and Reutebuch, S. E., 2005. Estimating forest canopy fuel parameters using lidar data. Remote Sensing of Environment 94(4), pp. 441–449.

Andersen, H.-E., Reutebuch, S. E. and Schreuder, G. F., 2002. Bayesian object recognition for the analysis of complex forest scenes in airborne laser scanner data. In: ISPRS Commission III, Symposium 2002 September 9 - 13, 2002, Graz, Austria, pp. A–035 ff (7 pages).

Atzberger, C., 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. Remote Sensing of Environment 93(1-2), pp. 53–67.

Baltsavias, E. P., 1999. Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry and Remote Sensing 54(2-3), pp. 199–214.

Cohen, W. B., Maiersperger, T. K., Gower, S. T. and Turner, D. P., 2003. An improved strategy for regression of biophysical variables and landsat etm+ data. Remote Sensing of Environment 84(4), pp. 561–571.

Colombo, R., Bellingeri, D., Fasolini, D. and Marino, C. M., 2003. Retrieval of leaf area index in different vegetation types using high resolution satellite data. Remote Sensing of Environment 86(1), pp. 120–131.

Dobbertin, M., Baltensweiler, A. and Rigling, D., 2001. Tree mortality in an unmanaged mountain pine (*pinus mugo var. uncinata*) stand in the swiss national park impacted by root rot fungi. Forest Ecology and Management 145, pp. 79–89.

Frazer, G., Trofymow, J. and Lertzman, K., 1997. A method for estimating canopy openess, effective leaf area index and photosynthetically active photon flux density using hemispherical photography and computerized image analysis techniques. Technical report, Information Report BC-X-373, Natural Resources Canada, Canadian Forest Service, Pacific Forestry Centre, Victoria, BC.

Gaveau, D. and Hill, R., 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. Canadian Journal of Remote Sensing 29, pp. 650–657.

Gower, S. T., Kucharik, C. J. and Norman, J. M., 1999. Direct and indirect estimation of leaf area index, fapar, and net primary production of terrestrial ecosystems. Remote Sensing of Environment 70, pp. 29–51.

Hyyppae, J., Kelle, O., Lehikoinen, M. and Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing 39, pp. 969–975.

Jacquemoud, S. and Baret, F., 1990. Prospect: A model of leaf optical properties spectra. Remote Sensing of Environment 34(2), pp. 75–91.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M. and Baret, F., 2004. Review of methods for in situ leaf area index determination: Part i. theories, sensors and hemispherical photography. Agricultural and Forest Meteorology 121(1-2), pp. 19–35.

Koetz, B., Schaepman, M., Morsdorf, F., Itten, K. and Allgöwer, B., 2004. Radiative transfer modeling within a heterogeneous canopy for estimation of forest fire fuel properties. Remote Sensing of Environment 92(3), pp. 332–344.

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A. and Harding, D., 1999. Lidar remote sensing of the canopy structure and biophysical properties of douglas-fir western hemlock forests. Remote Sens. Environ. 70, pp. 339–361.

Lim, K., Treitz, P., Baldwin, K., Morrison, I. and Green, J., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing 29(5), pp. 658–678.

Lovell, J., Jupp, D., Culvenor, D. and Coops, N., 2003. Using airborne and ground-based ranging lidar to measure canopy structure in australian forests. Can. J. Remote Sensing 29(5), pp. 607–622.

Means, J. E., Acker, S. A., Fitt, B. J., Renslow, M., Emerson, L. and Hendrix, C., 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogrammetric Engineering & Remote Sensing 66(11), pp. 1367–1371.

Morsdorf, F., Meier, E., Kötz, B., Itten, K. I., Dobbertin, M. and Allgöwer, B., 2004. Lidar-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. Remote Sensing of Environment 3(92), pp. 353–362.

Myneni, R., Nemani, R. and Running, S., 1997. Estimation of global leaf area index and absorbed par using radiative transfer models. IEEE Transactions on Geoscience and Remote Sensing 35, pp. 1380–1393.

Naesset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80(1), pp. 88–99.

Naesset, E., 2004. Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. Remote Sensing of Environment 91(2), pp. 243–255.

Ni-Meister, W., Jupp, D. L. B. and Dubayah, R., 2001. Modeling lidar waveforms in heterogeneous and discrete canopies. IEEE Transactions on Geoscience and Remote Sensing 39(9), pp. 1943–1958.

Riano, D., Valladares, F., Condes, S. and Chuvieco, E., 2004. Estimation of leaf area index and covered ground from airborne laser scanner (lidar) in two contrasting forests. Agricultural and Forest Meteorology 124(3-4), pp. 269–275.

Sun, G. and Ranson, K., 2000. Modeling lidar returns from forest canopies. IEEE Transactions on Geoscience and Remote Sensing 38(6), pp. 2617–2626.

Watson, D. J., 1947. Comparative physiological studies in the growth of field crops. i. variation in net assimilation rate and leaf area between species and varieties, and within and between years. Ann. Bot. 11, pp. 41–76.

Weiss, M., Baret, F., Smith, G. J., Jonckheere, I. and Coppin, P., 2004. Review of methods for in situ leaf area index (lai) determination: Part ii. estimation of lai, errors and sampling. Agricultural and Forest Meteorology 121(1-2), pp. 37–53.

Wotruba, L., Morsdorf, F., Meier, E. and Nüesch, D., 2005. Assessment of sensor characteristics of an airborne laser scanner using geometric reference targets. subm.