Sampling strategies for modelling aboveground biomass using optimal imagery and LiDAR metrics

Seamus Murphy[1, *]

[1]Corresponding Author

[*]Bangor University

Word Count:

Abstract

Estimation of forest above ground biomass has been conducted with a range of data sources and through a variety of prediction methods. Despite notable advancements in this field of remote sensing studies, statistical uncertainties still exist. In this study, we compared seven prediction methods according to simple random sampling and stratified sampling. The forest site was classified into nine land cover classes based on three different machine learning algorithms. We externally validate predictions against biomass measurements from 50 concentric field plots, obtaining a RMSE of 5.5 – 7.5%. Accuracy increased with use of support vector machine classifier. Although M- and S-estimators produced lowest RMSE scores, cross-validation suggested models were over-fitting to the data. Highly variable diagnostic results suggest model identification requires comprehensive screening approaches that address not only variance inflation and bayesian information but informs model selection based on uncertainty scoring beyond internal error rates.

Keywords: aboveground biomass; LiDAR; optical imagery; machine learning; prediction model; cross-validation; forest inventory; accuracy assessment

1. Introduction

In the Clocaenog Forest of Denbighshire, North Wales, an integrated, multi-phase sampling was employed to estimate aboveground biomass, using simple random sampling of field plot data, stratified random sampling using image classification of vegetation, and LiDAR-based prediction of forest biomass. This was done by combining data from Airborne Laser Scanning and field measurements. In this study, we compared prediction methods and conservative estimates based on various sampling strategies and cross-validation settings. Aboveground biomass was predicted using common modelling algorithms, including simple linear regression, stepwise linear regression, nonlinear least square regression, iterated re-weighted least square regression based on the Huber-loss estimator, the Tukey estimator, and the Hampel estimator, as well as maximum likelihood classifier, random tree classifier, and the support vector machine methods. Integrating remote-sensing data into a multi-phase forestry sampling requires a statistical framework that is based on unbiased model-assisted estimators.

2.1. Methods; Field measurements

In the initial phase of sampling, allometric models were validated based on the internal rate of error. Candidate equations were identified using a system of screening according to three requirements: 1) proximity of original study to the current field site, 2) use of 'dbh' as the predictor in equation, 3) and appropriate range of 'dbh' values assessed when fitting the biomass equation. The following table presents eight candidate equations and their summary information:

Table 1. Aboveground biomass equations by tree species. Format and parameter values provided under "Equation" as a, b, D. #Index number of original publication citing equation.

Species (Index[#])	Range of D	Equation	α	β	Country
M1_*Betula pendula* (32)	/	α*Dβ	0.2511	2.29	UK
M2_*Betula pendula* (10)	2.9-30	α+β*ln(D)	-2.4166	2.4227	UK
M3_*Betula pendula* (10)	2.9-26	α+β*ln(D)	-2.7584	2.6134	UK
M4_*Betula pendula* (10)	3.3-16	α+β*ln(D)	-2.1625	2.3078	UK
M5_*Betula pendula* (10)	3.5-23	α+β*ln(D)	-2.6423	2.4678	UK
M6_*Betula pubescens* (35)	10-90	α*Dβ	-2.162	2.3078	UK
M7_*Betula pubescens* (35)	0.8-8.5	α*Dβ	0.00029	2.50038	Sweden
M8_*Picea sitchensis* (B04)	2-40	α*Dβ	0.028	2.710	Ireland

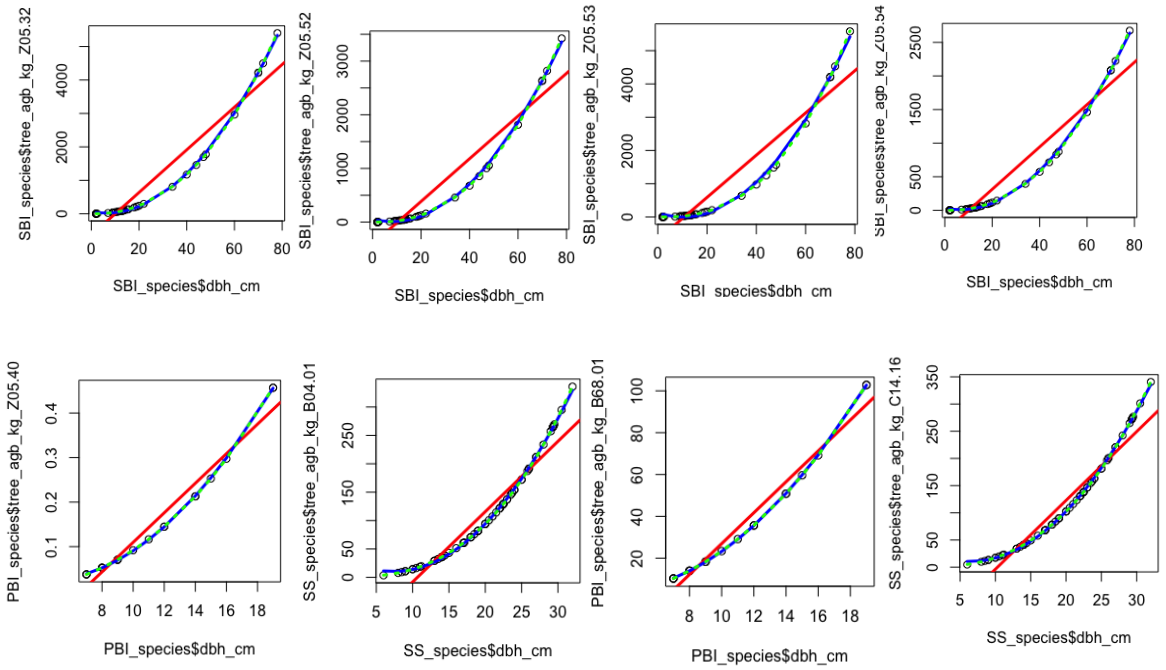
Table 2 Descriptive statistics of field plotes

Species	Mean dhb	SD	Median dbh	SE	Range
Betula pendula	23.7	23.1	13.5	3.6	76.0
Betula pubescens	11.1	3.9	9.5	0.9	12.0
Picea sitchensis	19.7	6.5	20.0	0.8	26.0

Descriptive analysis found highly dispsersed data, violation of homogeneity and non-normal distribution (Table 3). As a result, models were compared by assessing the bias between predicted and observed values, which were plotted in the following 'residuals versus fitted values' diagrams (Fig.1-8).

Table 3. Descriptive statistics of response variables based on 9 biomass equations. #Shaprio-Wilks's p-values reporting significant deviation from normal distribution at 0.05, 0.01, 0.001 levels as *, **, ***

Variable	Mean	SD	SE	Skewnes s	Kurtosis	p-value#
M1: SBI_agb_Z05.32	872.4	1533.7	242.5	1.73	1.53	***
M2: SBI_agb_Z05.52	531.8	962.0	152.1	1.77	1.68	***
M3: SBI_agb_Z05.53	824.5	1546.3	244.5	1.82	1.90	***
M4: SBI_agb_Z05.54	429.6	758.3	119.9	1.73	1.55	***
M6: PBI_agb_B68.01	35.1	29.1	6.5	1.19	0.22	***
M7: PBI_agb_Z05.40	0.2	0.1	0.0	1.25	0.39	***
M8: SS_agb_C14.16	118.3	85.7	10.4	0.65	-0.48	**



Figures 1-8 Scatterplot and regression lines of the relation between species dbh and the calculated aboveground biomass; Linear (red curve) vs nonlinear algorithmic regression lines (blue curve).

In table 4 below, the root mean square error of model residuals (RMSE) was presented in absolute terms. Models with highest RMSE values were omitted, providing the final cut of eight candidate models. For silver birch, M3 model was best fit to the current dataset. Bunce, who carried out an inventory survey in the nearby Meathop Wood, derived this equation (Bunce 1968). Though, the range of dbh values appears less dispersed than this current dataset. For downy birch, some confusion arose. While model diagnostics indicated M7 was best fit, descriptive results suggested this equation was heavily conservative. Therefore, final equation was chosen from M6 model, which was also highly significant. Interestingly, M6 was also derived from the same study carried out by Bunce (ibid.).

Table 4. Parameter estimates (& standard error for biomass models: NLS-NLS5. SE: standard errors. RMSE: root mean square. Significant estimates reported at the levels of 0.05, 0.01, 0.001 as *, **, ***.

Parameter	M1 (SBI)	M2 (SBI)	M3 (SBI)	M4 (SBI)
α	56.774 (9.901)***	52.076 (9.344)***	121.037 (22.548)***	29.817 (5.220)***
β	1.027 (0.011)***	0.684 (0.010)***	1.181 (0.025)***	0.5122 (0.006)***
RMSE	28.61	27.00	2.38	15.09

Parameter	M6 (PBI)	M7 (PBI)	M8 (SS)	M9 (SS)
α	7.185 (0.346)***	0.055 (0.003)***	38.760 (1.968)***	30.372 (1.467)***
β	-2.200 (0.059)***	0.002 (0.001)***	0.526 (0.006)***	0.495 (0.004)***
RMSE	0.14	0.00	1.98	1.47

2.2. Methods; Image stratification

In Phase II, the Clocaenog Forest site was then classified according to vegetation composition. Using multi-band false colour infrared image of the site, four steps of analysis were taken, including 1) exploration and segmentation of image data, 2) non-supervised classification of extracted image data using the ISODATA algorithm 3) supervised classification of image data using maximum likelihood classifier (MLC), support vector machine methods (SVM) and random tree classifier (RTC), and 4) a post-classification accuracy assessment.

Optical imagery was extracted from WorldView-2 satellite acquired from the following raster files. To check for classification error through unequal spread of means, the distribution of each predictor variable was assessed (Table 5). Applying the ‘SegmentMeanShift’ tool in ArcGIS 10.3, the image data was partitioned into

coarsely homogenous spectral clusters derived using the kernel density function (Jain, Murty, and Flynn 1999).

Table 5. Sources and descriptive statistics of rasterband variables employed in the present study

Satellite data	Acquisition date	Spectral band	Spectral resolution (μm)	Spatial resolution (m)
WorldView-2	2009 July 01	B1. NIR	(0.70-100)	0.46
		B2. Red	(0.50-0.58)	0.46
		B3. Green	(0.62-0.70)	0.46

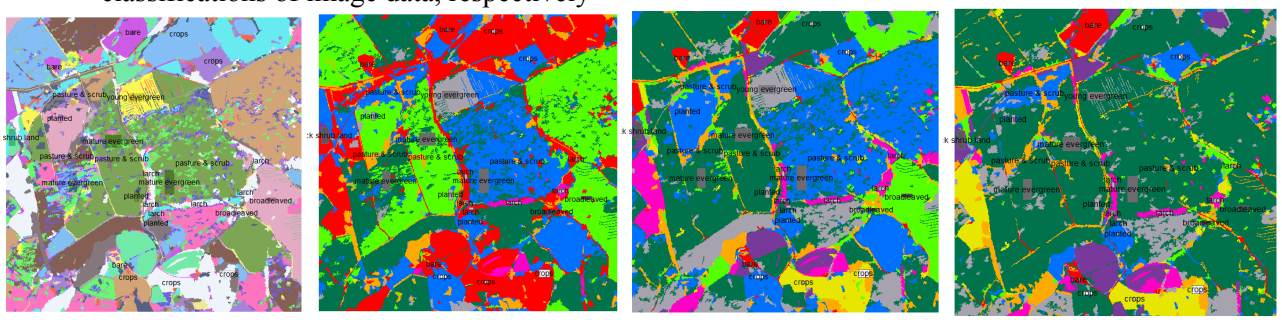
	Mean	SD	Min	Max	Correlation matrix		
					Band 1	Band 2	Band 3
Band 1	115.63	38.11	0.00	255.00	1.0000	0.7225	0.7879
Band 2	106.28	35.29	3.00	255.00	0.7225	1.0000	0.9738
Band 3	90.37	29.00	0.00	255.00	0.7879	0.9738	1.0000

Greatest variance was observed in NIR band (SD = 38.1116). Highly significant correlation was observed between band 2 and band 3 (R = 0.974). Due to the different characteristics of electromagnetic energies and the different interactions they produce with surface targets and atmospheric interference, these bands provided different information regarding land cover (Elvidge and Chen 1995; Milchunas 1998; Treitz and Howarth 1999). If considering grayscale data ranges, a possible reason for red band having shorter range could relate to the fine resolution of this image (0.5m), which means that wavelengths are narrower. This effect can limit spatial resolution and texture relations between pixel values, which may limit classification algorithms (Chakraborty et al. 2013) or obscure lower values (black pixels) from red band wavelengths.

Using segmented data from above, pixels were clustered using the isodata algorithm (interactive self-organising data analysis technique, Chakraborty et al. 2015; 2013; 2012; Chakraborty, Singh, and Dutta 2017; Thakur and Chakraborty 2019). Based on the k-means algorithm or the minimum Euclidean distance, isoclustering process iteratively assigns pixels to nearest classes that are recalculated according to their members locations (Gordon and Henderson 1977; Mao and Jain 1996). Class mean vectors are recalculated and pixels reassigned until there is limited change between iterations (Jain, Murty, and Flynn 1999), and end-point class signatures are defined and pixels are classified according to the minimum distance classifier.

Three classification methods were applied. Each of the MLC, RTC, and SVM classifiers were trained to 9 training samples, before the three algorithms were run across through spectral data. The MLC tool calculated probabilities of pixels belonging to each training sample before reassigning them to new classes. Similar to pixel-based classification, the RTC, and SVM methods used object-oriented feature extractions to classify segments or super pixels within the data. Visually this represented a coarser classification that MLC (Figure 8-11). To review the distribution of these classes per tool, spectral signatures are provided in appendix.

Figs. 8-11. Left to right: Unsupervised ISO and supervised MLC, RTC and SVM classifications of image data, respectively



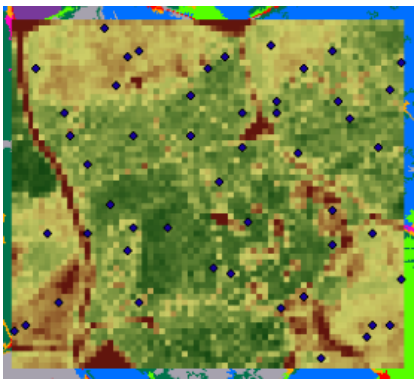
SVM (64%) and RTC (56%) proved more accurate than MLS (43%). Research highlights machine learning algorithms for land-cover classifications (Xie et al. 2019), though choice of tool depends on site and classification system put in place (Lu and Weng 2007).

Table 6. Accuracy estimates of land-cover classes of MLC, SVM, RTC

Classifier	# Classes	User accuracy	Kappa agreement	Significant class
MLC	9	0.8876	0.4307	C6 (0.9090)
RTC	9	0.9375	0.5612	C8 (0.8732)
SVM	9	0.9546	0.6380	C3 (0.1000)

2.3. Methods; Building a LiDAR model for predicting above groundbiomass

In phase III, candidate LiDAR models were evaluated using accuracy assessments. Efforts are made to check all four types regression violations and to evaluate if models are over-fitted to the data based on k-fold validation techniques. Models were calibrated between field plot data and remote sensing statistics, which were linked using global navigation satellite systems (Ruben Valbuena et al. 2012).



Remote sensing LiDAR data was generated from an airborne laser scanning survey (ALS), from which three predictor variables were generated: mean tree height (MeanH) (Hopkinson, Chasmer, and Hall 2008; Magnussen and Boudewyn 1998; Means et al. 2000; Naesset 1997; Næsset 2002; Næsset and Økland 2002; Popescu 2007), standard deviation (SD), and fractional canopy height (Cover) based on ratio of height returns above 1.3m to total returns (Hill and Broughton 2009; Hopkinson et al. 2006; Hopkinson and Chasmer 2009; Morsdorf et al. 2005, 2006; Wasser et al. 2013). Descriptive statistics of these predictor variables are presented below in Table 7.

Table 7. Descriptive statistics of predictors: Shapiro-Wilk (SW) score of normality.

Variable	Mean	SD	SE	Skewness	Kurtosis	SW
MeanH	13.05	5.77	0.82	-0.32	-1.40	***
Cover	0.92	0.08	0.01	-1.76	4.03	***
SD	5.41	2.84	0.40	-0.15	-1.47	***

A simple linear regression was conducted (M1) and its residuals were explored using regression diagnostics. These results are presented below in Table 8. Results indicated signs of heteroscedasticity, multi-collinearity and influential outliers. Heteroscedasticity was observed by the Breusch-Pagan test ($p = 0.002$), and confirmed by the "Residuals vs Fitted" and "Scale Location". Although no collinearity problems were detected by Durban-Watson test, the Variance-Inflation-Factor identified Standard Deviation as a problematic predictor. The existence of influential outliers was observed using the Bonferroni test ($p = 0.025$). To explore measures of influence, analyses ran a Cook's Distance graph, a DFBETA panel, and a "Studentized Residuals vs Leverage" plot using the *lmtest* and *olsrr* packages in R. Two observations (obs 9 &18) were removed and a new dataset was developed "Clocaenog_Cleaned". To compare models, both datasets were used, and models corresponding to these are hereafter labeled "-comp" and "-clean".

Table 8. Hypothesis tests of linear regression: Breusch-Pagan (BP) test of heteroskedasticity of residuals, Durbin-Watson (DW) test to search for multi-collinearity of data, Variance Inflation Factor (VIF) test to understand collinearity problems, Bonferroni (BF) test to search for outliers.

Model	SW	BP	DW	BF	Variance Inflation Factor		
					MeanH	SD	Cover
M1 (LM.comp)	0.131	0.002	0.345	0.025	3.234	5.451	2.886

Based on these results, 7 new models were generated. M2 included a simple linear

regression based on the cleaned data. M3 included a stepwise linear regression using two predictor variables (MeanH and SD) and complete dataset, which were selected from a stepwise backwards 10-kfold validation and the *caret* R package. M4 included a nonlinear least squares regression based on complete dataset and M5 included a nonlinear least squares regression using cleaned dataset. Using the *mass* R package, M6-8 models included iterated re-weighted least square regressions based on the Huber-loss estimator, the Tukey estimator and the Hampel estimator. Residuals of these 7 additional models were examined using same hypothesis tests as above. In Table 9 below, similar violations were observed. Due to its nonparametric nature, the model that showed fewer problems was M4.

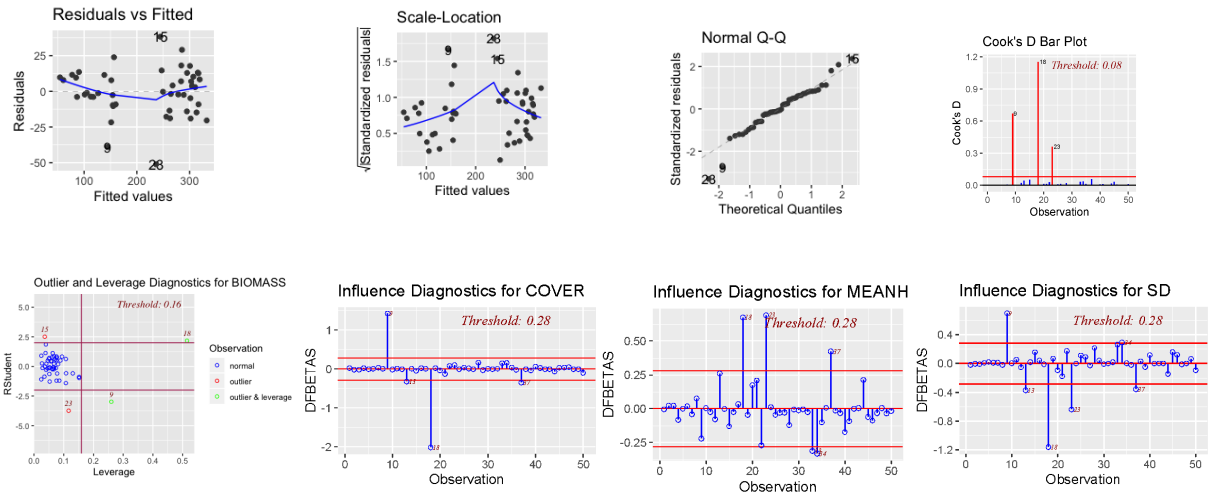


Table 9. Regression diagnostics.

Model	SW	BP	DW	BF	Variance Inflation Factor		
					MeanH	SD	Cover
M2.LM.clean	0.949	0.187	0.303	0.001	3.743	4.951	9.493
M3.LM.stepwise	0.916	0.586	0.437	0.010	1.903	1.903	/
M4.NLS.comp	0.968	0.417	0.344	0.049	28.554	12.128	2.959
M5.NLS.clean	0.961	0.110	0.287	0.002	34.967	13.391	5.221
M6.RR.huber	0.002	0.002	0.345	0.025	3.234	5.451	2.886
M7.RR.tukey	0.000	0.002	0.345	0.025	3.234	5.451	2.886
M8.RR.hampel	0.005	0.002	0.345	0.025	3.234	5.451	2.886

As results above were not conclusive with regards to regression violations, all models were analysed using a resampling validation technique: the 10-Kfold cross validation (3-repeat). This method was used to provide estimates of the bias between resampling procedures and estimates of each model. It was also used to explore levels of

precision in terms of the variation of results (RMSE). Using *DescTools* package in R, Theil's U estimate of error was provided to examine the level of unexplained variance in each model (Table 10.).

Table 10. Uncertainty estimates of LiDAR regression models

	Full Models					Cross Validation	
Model	RMSE %	MAE %	Theil U	AIC	BIC	RMSE %	MAE %
M1.LM.comp	16.44 (7.5)	11.95 (5.4)	0.998	427.727	437.286	16.63 (7.6)	13.67 (6.2)
M2.LM.clean	15.30 (6.9)	10.95 (4.9)	0.998	403.926	413.282	15.92 (7.1)	13.15 (5.9)
M3.LM.stepw	17.86 (8.1)	12.47 (5.7)	0.997	435.041	442.689	16.44 (7.5)	13.34 (6.1)
M4.NLS.comp	16.30 (7.4)	11.19 (5.1)	0.998	427.766	439.239	16.34 (7.4)	13.26 (6.0)
M5.NLS.clean	15.18 (6.8)	10.37 (4.6)	0.998	404.083	415.310	15.90 (7.1)	13.91 (6.2)
M6.R.Huber	13.11 (5.9)	11.17 (5.1)	1.001	430.087	429.647	17.19 (7.8)	14.08 (6.4)
M7.R.Tukey	12.14 (5.5)	11.17 (5.1)	1.006	435.415	444.975	17.31 (7.9)	13.94 (6.3)
M8.R.Hampel	13.34 (6.1)	11.17 (5.1)	1.001	428.938	438.499	17.48 (7.9)	14.08 (6.4)

Based on these results, the Tukey robust regression models provided highest accuracy based on its lowest RMSE score. However, all robust regressions showed signs of over-fitting the data. This was evident from the substantial difference observed between full model and cross validation error rates, as was as their higher AIC and BIC scores. The next best model M5 included the nonlinear least squared model based on cleaned dataset (RMSE = 15.18).

3. Results; Conservative LiDAR-based estimations of aboveground biomass based on simple random sampling and stratified random sampling

Estimates of aboveground biomass across the site were generated using a regression analysis of 50 randomly sampled field plots and linked LiDAR data points from phase III. From 8 regression analyses, two candidate regression models were selected from previous exercise (Table 11). However, since image classifications and LiDAR estimates were based on the full dataset of 50 observations, the second candidate model was omitted. Therefore, parameters from the nonlinear least squares regression model (M4) were fitted to the biomass formula below (1).

7.03 * “MEANH” + 146.78 * “COVER” + 12.72 * “SD” (1)

Table 11 Regression models and parameters used for calculating above ground-biomass. AIC: Akaike Criterion Information, BIC: Bayesian Information Criterion.

	RMSE(%)	AIC	BIC	Theil’s U	Paramter estimates (SE)		
					MeanH	SD	Cover
M4.NLS.comp	16.30(7.4)	427.77	439.24	0.998	7.03(2.16)**	12.72(2.16)***	146.78(51.59)**
M5.NLS.clean	15.18(6.8)	404.08	415.31	0.998	6.95(2.27)**	13.23(2.82)***	159.11(91.51)

Based on simple random sampling, mean biomass density was estimated as 220.63Mt/ha (95% CI 195.64, 245.62). Confidence intervals were also generated using the *DescTools* package in R studio (95% CI 195.01, 246.25). As *DescTools* estimates presented a higher certainty than the first estimates, the manually generated confidence intervals were selected for computing the following estimations. Calculating the sample area from pixel dimensions in the ‘property>source’ tab in ArcGis, the total area of the site was 107.75ha. Using this value, the total sum of biomass was equal to 23,772.76Mt (95% CI 21,080.21, 26,465.08). Based on market price of £50/Mt, total forest biomass was worth £1,118,638 (95% CI 1,054,010.5, 1,323,254).

Based on image classification of remote imagery, which was generated using the maximum likelihood algorithm, weighted statistics of biomass were calculated for the forest classes (Table 2). Post-stratification of classes was conducted using a weighting system based on the sampling derivative of the proportion of class areas to the total area of forested land. The total area of forested land was 97.8ha. The largest forest class was ‘Evergreen’ (83.4ha), which accounted for 85.2% of the forested area, followed by ‘Mixed Broadleaved’ (7.1%, 6.9ha), ‘Larch’ (6.1%, 6.0ha), and ‘Planted’ (1.6%, 1.6ha).

Weighted estimates were generated twice), first using the *Hmisc* package in Rstudio (225.07Mt/ha, 95% CI 50.71, 399.43), and second using manual calculations

(226.40Mt/ha, 95% CI 221.52, 231.29). Due to higher uncertainty generated by *Hmisc* package, manually generated results were used in the following estimations. Based on weighted estimates, the total sum of biomass was equal to 21,141.92Mt (95% CI 21.664.66, 22,620.16).

Table 12. Biomass estimates of forest classes using random sampling and stratified sampling.

	Mean (Mt/ha)	SD	95% CI		Tot Vol (Mt)	Tot Val (£)
			Lower	Upper		
1. C4 Planted	159.30	/	/	/	248.8	12,440
2. C5 Mix_Broad	114.58	22.88	90.38	138.60	794.9	37,495
3. C6 Larch	209.25	40.66	156.05	574.55	1,253.4	62,670
4. C7 Evergreen	**238.20**	88.00	210.42	265.98	19,854.8	979 ,200
5. All_Field_Plots	220.63	90.16	195.01	246.25	21,577.4	1,078,870
6. All_Image_Class	226.40	88.96	221.52	231.29	22,141.9	1,107,095

Using image classifications and weighting system as above (proportionate to total forest area), biomass estimation was carried our using LiDAR datapoints for the same four forested classes (Table 12). These LiDAR values were generated using 'zonal statistics' in the 'spatial analyst' toolbox of ArcGis software. Using similar formula of variance, standard error and Gaussian distribution, this produced a mean biomass density of 247.55Mt/ha and lower and upper bounds of 247.17 and 247.93, respectively. From this, the total forest biomass was estimated to be 24,210.39Mt (95% CI 24,173.23, 24,247.55).

Table 13. Aboveground biomass estimates based on simple random and stratified sampling

	Mean (Mt/ha)	SD	95% CI		Tot Vol (Mt)	Tot Val (£)
			Lower	Upper		
5. All_Field_Plots	220.63	90.16	195.01	246.25	21,577.4	1,078,870
6. All_Image_Class	226.40	88.96	221.52	231.29	22,141.9	1,107,095
7. All_LiDAR_Im_Class	247.55	0.19	247.17	247.93	24,210.4	1,210,520

Comparing between the aggregated results above ("5. All - Field Plots", "6. All – Image Classification", "7. All – LiDAR & Image Classes"), we find that values generated from the LiDAR data were higher than from other methods. With all estimates now available, results confirmed a lack of bias as no methods produced a mean beyond the range of confidence intervals of other methods. Applying the

principle of conservativeness, biomass sales were compared using lower bounds of estimates set at the level of 95% confidence. Accordingly, conservative estimates generated using LiDAR data points and weighted stratification from the image classification (7.All – LiDAR & Image Classification), total biomass sales were worth £1,208,661.3. Therefore, the LiDAR survey provided an increase in estimated sales of £125,428.5 compared to stratified image classification. Compared to the initial, simple random sampling of field plots, LiDAR estimates increased even further by £154,650.8.

If we take into account the extra cost of LiDAR surveying, we can evaluate the value of investing in this method. If the cost of LiDAR survey was £10 per hectare and the total area of forest was 97.8ha, the LiDAR survey required a total budget of £970.80. Compared to image classification, LiDAR surveying provided the added value of £124,457.7. Compared to simple random sampling of field plots, LiDAR survey provided the added value of £153,680.

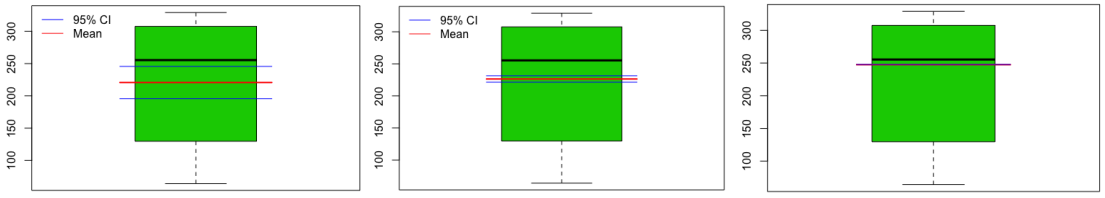


Figure 12. Boxplot graph comparing means and confidence intervals from the three estimation methods. Left: Simple random sampling and unweighted estimations. Centre: Image classification and weighted estimations. Right: LiDAR estimations using weighted classification from optical imagery.

4. Discussion

In order to provide a comprehensive cost-efficiency analysis of remote sensing biomass estimation tools, it is worthwhile comparing between classification methods, particularly if the forested site is significantly large in extent. Accuracy assessments in phase II indicated that the Support Vector Machine algorithm produced higher levels of accuracy than the Maximum Likelihood algorithm and the Random Tree classifiers. However, no significant differences were observed between their stratification of forest sample site. This is likely the result of the small land areas of the four forested classes.

In phase III, regression analyses were assessed before fitting them to LiDAR biomass estimation models. It might be useful for statistics newcomers like myself to follow a recommended workflow of regression diagnostics? In addition, since some tests showed certain violations more significantly than others, it seems that such model assessment toolkit or protocol would be highly beneficial. As a suggestion, would ether the Bonferroni correction test or the Holm-Bonferroni step-by-step methods seem relevant here? In particular, there a number of common pitfalls to look out for such as 1) multi-collinearity and 2) influence and leverage of outliers. Both were identified in phase III.

Regarding the former, studies have shown that collinear variables can give the appearance of a highly significant model (Graham 2003; Sawadogo et al. 2010), which can lead to over-estimates of biomass (Basuki et al. 2009; Colgan, Asner, and Swemmer 2013). To avoid such errors, some have recommended the application of variance inflation factor metrics (Chave et al. 2005; Sileshi 2014, 249), or the use of step-wise regression methods (Carrascal, Galván, and Gordo 2009; Ruben Valbuena et al. 2012).

Regarding the latter, the potential effect of outliers has led some to recommend use of robust regression (Sileshi 2014, 248). Therefore, a number of estimators need exploring, such as the M-estimators and S-estimators. Results from phase III indicated that Tukey's robust regression produced lowest RSME with the current dataset, though evidence of over-fitting was also indicated. In order to choose between candidate models, it worthwhile exploring various information criteria and cross validation methods given the nature of the data and model. While some recommend AIC and BIC scoring (Lukacs, Burnham, and Anderson 2010), others assume an optimal number of predictors are in use (Guthery et al. 2005). Alternatively, others recommend the use of Theil's uncertainty scoring (Valbuena et al. 2017). This emerging area of research is nonetheless an important aspect of remote sensing estimations of biomass that requires scientific debate to provide an approved protocol and workflow.

Appendix I Biomass equations

- M1 Hughes, M. K. (1971). Tree biocontent, net production and litter fall in a deciduous woodland. *Oikos*, 62-73
- M2 Bunce, R. G. H. (1968). Biomass and production of trees in a mixed deciduous woodland: I. Girth and height as parameters for the estimation of tree dry weight. *The Journal of Ecology*, 759-775.
- M3 Bunce, R. G. H. (1968). Biomass and production of trees in a mixed deciduous woodland: I. Girth and height as parameters for the estimation of tree dry weight. *The Journal of Ecology*, 759-775.
- M4 Bunce, R. G. H. (1968). Biomass and production of trees in a mixed deciduous woodland: I. Girth and height as parameters for the estimation of tree dry weight. *The Journal of Ecology*, 759-775.
- M5 Bunce, R. G. H. (1968). Biomass and production of trees in a mixed deciduous woodland: I. Girth and height as parameters for the estimation of tree dry weight. *The Journal of Ecology*, 759-775.
- M6 Johansson, T. (1999). Biomass equations for determining fractions of pendula and pubescent birches growing on abandoned farmland and some practical implications. *Biomass and bioenergy*, *16*(3), 223-238.
- M7 Johansson, T. (1999). Biomass equations for determining fractions of pendula and pubescent birches growing on abandoned farmland and some practical implications. *Biomass and bioenergy*, *16*(3), 223-238.
- M8 Chojnacky, D. C., Heath, L. S., & Jenkins, J. C. (2014). Updated generalized biomass equations for North American tree species. *Forestry*, *87*(1), 129-151.
- M9 Black, K., Tobin, B., Saiz, G., Byrne, K. A., & Osborne, B. (2004). Improved estimates of biomass expansion factors for Sitka spruce. *Irish Forestry*.
- M10 Chojnacky, D. C., Heath, L. S., & Jenkins, J. C. (2014). Updated generalized biomass equations for North American tree species. *Forestry*, *87*(1), 129-151.
- M11 Chojnacky, D. C., Heath, L. S., & Jenkins, J. C. (2014). Updated generalized biomass equations for North American tree species. *Forestry*, *87*(1), 129-151.

Bibliography

Basuki, T M, P E Van Laake, A K Skidmore, and Y A Hussin. 2009. "Allometric Equations for Estimating the Above-Ground Biomass in Tropical Lowland Dipterocarp Forests." *Forest ecology and management* 257(8): 1684–94.

Bunce, R G H. 1968. "Biomass and Production of Trees in a Mixed Deciduous Woodland: I. Girth and Height as Parameters for the Estimation of Tree Dry Weight." *The Journal of Ecology*: 759–75.

Carrascal, Luis M, Ismael Galván, and Oscar Gordo. 2009. "Partial Least Squares Regression as an Alternative to Current Regression Methods Used in Ecology." *Oikos* 118(5): 681–90.

Chakraborty, D, V M Chowdhary, D Dutta, and J R Sharma. 2013. "Classification of High Spatial Resolution Image Using Multi Circular Local Binary Pattern and Variance." *International Journal of Electronics and Computer Engineering* 4(6): 1648–53.

Chave, Jérôme et al. 2005. "Tree Allometry and Improved Estimation of Carbon Stocks and Balance in Tropical Forests." *Oecologia* 145(1): 87–99.

Colgan, Matthew S, Gregory P Asner, and Tony Swemmer. 2013. "Harvesting Tree Biomass at the Stand Level to Assess the Accuracy of Field and Airborne Biomass Estimation in Savannas." *Ecological Applications* 23(5): 1170–84.

Gordon, A D, and J T Henderson. 1977. "An Algorithm for Euclidean Sum of Squares Classification." *Biometrics*: 355–62.

Graham, Michael H. 2003. "Confronting Multicollinearity in Ecological Multiple Regression." *Ecology* 84(11): 2809–15.

Guthery, Fred S, Leonard A Brennan, Markus J Peterson, and Jeffrey J Lusk. 2005. "Information Theory in Wildlife Science: Critique and Viewpoint." *The Journal of Wildlife Management* 69(2): 457–65.

Hill, R A, and Richard K Broughton. 2009. "Mapping the Understorey of Deciduous

Woodland from Leaf-on and Leaf-off Airborne LiDAR Data: A Case Study in Lowland Britain." *ISPRS Journal of Photogrammetry and Remote Sensing* 64(2): 223–33.

Hopkinson, Chris et al. 2006. "Towards a Universal Lidar Canopy Height Indicator." *Canadian Journal of Remote Sensing* 32(2): 139–52.

Hopkinson, Chris, and Laura Chasmer. 2009. "Testing LiDAR Models of Fractional Cover across Multiple Forest Ecozones." *Remote Sensing of Environment* 113(1): 275–88.

Hopkinson, Chris, Laura Chasmer, and R J Hall. 2008. "The Uncertainty in Conifer Plantation Growth Prediction from Multi-Temporal Lidar Datasets." *Remote Sensing of Environment* 112(3): 1168–80.

Jain, Anil K, M Narasimha Murty, and Patrick J Flynn. 1999. "Data Clustering: A Review." *ACM computing surveys (CSUR)* 31(3): 264–323.

Lu, Dengsheng, and Qihao Weng. 2007. "A Survey of Image Classification Methods and Techniques for Improving Classification Performance." *International journal of Remote sensing* 28(5): 823–70.

Lukacs, Paul M, Kenneth P Burnham, and David R Anderson. 2010. "Model Selection Bias and Freedman's Paradox." *Annals of the Institute of Statistical Mathematics* 62(1): 117.

Magnussen, S, and P Boudewyn. 1998. "Derivations of Stand Heights from Airborne Laser Scanner Data with Canopy-Based Quantile Estimators." *Canadian journal of forest research* 28(7): 1016–31.

Mao, Jianchang, and Anil K Jain. 1996. "A Self-Organizing Network for Hyperellipsoidal Clustering (HEC)." *Ieee transactions on neural networks* 7(1): 16–29.

Means, Joseph E et al. 2000. "Predicting Forest Stand Characteristics with Airborne Scanning Lidar." *Photogrammetric Engineering and Remote Sensing* 66(11): 1367–72.

Morsdorf, Felix et al. 2005. "The Potential of Discrete Return, Small Footprint Airborne Laser Scanning Data for Vegetation Density Estimation." *Proceedings of ISPRS WG III/3* 3(4): 3.

———. 2006. "Estimation of LAI and Fractional Cover from Small Footprint Airborne Laser Scanning Data Based on Gap Fraction." *Remote Sensing of Environment* 104(1): 50–61.

Naesset, Erik. 1997. "Determination of Mean Tree Height of Forest Stands Using Airborne Laser Scanner Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 52(2): 49–56.

Næsset, Erik. 2002. “Predicting Forest Stand Characteristics with Airborne Scanning Laser Using a Practical Two-Stage Procedure and Field Data.” *Remote sensing of environment* 80(1): 88–99.

Næsset, Erik, and Tonje Økland. 2002. “Estimating Tree Height and Tree Crown Properties Using Airborne Scanning Laser in a Boreal Nature Reserve.” *Remote Sensing of Environment* 79(1): 105–15.

Popescu, Sorin C. 2007. “Estimating Biomass of Individual Pine Trees Using Airborne Lidar.” *Biomass and Bioenergy* 31(9): 646–55.

Sawadogo, Louis et al. 2010. “Allometric Prediction of Above-Ground Biomass of Eleven Woody Tree Species in the Sudanian Savanna-Woodland of West Africa.” *Journal of Forestry Research* 21(4): 475–81.

Sileshi, Gudeta W. 2014. “A Critical Review of Forest Biomass Estimation Models, Common Mistakes and Corrective Measures.” *Forest Ecology and Management* 329: 237–54.

Treitz, Paul M, and Philip J Howarth. 1999. “Hyperspectral Remote Sensing for Estimating Biophysical Parameters of Forest Ecosystems.” *Progress in Physical Geography* 23(3): 359–90.

Valbuena, R et al. 2017. “Enhancing of Accuracy Assessment for Forest Above-Ground Biomass Estimates Obtained from Remote Sensing via Hypothesis Testing and Overfitting Evaluation.” *Ecological modelling* 366: 15–26.

Valbuena, Ruben, Francisco Mauro, Roberto Rodriguez-Solano, and Jose Antonio Manzanera. 2012. “Partial Least Squares for Discriminating Variance Components in Global Navigation Satellite Systems Accuracy Obtained under Scots Pine Canopies.” *Forest Science* 58(2): 139–53.

Wasser, Leah, Rick Day, Laura Chasmer, and Alan Taylor. 2013. “Influence of Vegetation Structure on Lidar-Derived Canopy Height and Fractional Cover in Forested Riparian Buffers during Leaf-off and Leaf-on Conditions.” *PLoS One* 8(1).

Xie, Zhuli et al. 2019. “Classification of Land Cover, Forest, and Tree Species Classes with ZiYuan-3 Multispectral and Stereo Data.” *Remote Sensing* 11(2): 164.