# European Journal of Forest Research

Combining canopy height and tree species map information for large scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes

--Manuscript Draft--

Manuscript Number:	EFOR-D-17-00182R1
Full Title:	Combining canopy height and tree species map information for large scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes
Article Type:	Original Research Paper
Keywords:	OLS Regression; standing timber volume; LiDAR canopy height model; satellite-based tree species classification; calibration; Forest Inventory; angle count sampling
Corresponding Author:	Andreas Hill Eidgenossische Technische Hochschule Zurich Zurich, SWITZERLAND
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Eidgenossische Technische Hochschule Zurich
Corresponding Author's Secondary Institution:	
First Author:	Andreas Hill, Msc.
First Author Secondary Information:	
Order of Authors:	Andreas Hill, Msc.
	Henning Buddenbaum, Dr.
	Daniel Mandallaz, Dr. habil.
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	A timber-volume regression model applicable to the entire forest area of the federal German state of Rhineland-Palatinate is identified using a combination of airborne laser scanning (ALS)-derived metrics and information from a satellite-based tree species classification map available on the federal state level. As is common in many forest inventory datasets, strong heterogeneity in the ALS data due to different acquisition dates and misclassifications in the tree species classification map had noticeable effects on the regression model's performance. This article specifically addresses techniques that improve the performance of ordinary least square regression models under such restricting conditions. We introduce a calibration technique to neutralize the effect of misclassifications in the tree species variable that originally caused a residual inflation of 0.05 in adjusted R2. Incorporating the calibrated tree species information improved the model accuracy by up to 0.07 in adjusted R2 and suggests the use of such information in forthcoming inventories. We also found that including ALS quality information as categorical variables within the regression model considerably mitigates issues with time lags between the ALS and terrestrial data acquisition and ALS quality variations (increase of 0.09 in adjusted R2). The model achieved an adjusted R2 of 0.48 and a cross-validated root mean square error (RMSEcv) of 46.7% under incorporation of the tree species and ALS quality information, and was thus improved by 0.12 in adjusted R2 (5% in RMSEcv) compared to the simple model only containing ALS height metrics (adjusted R2=0.36, RMSEcv=51.7%)

Review

Combining canopy height and tree species map information for large scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes

Andreas Hill, Henning Buddenbaum, Daniel Mandallaz

20. Januar 2018

Reviewer 1:

This relevant and interesting study describes the development of a working/linking model to be used with model-supported estimators. The challenge of the study was the large spatial extent in combination with high resolution auxiliary variables and field data that resulted in severe inconsistencies that had to be handled. The applied approach is statistically rigorous and adequately described; the text is well written and clear.

Comments:

1. *Even though the development of the working/linking model is important, I would have appreciated if it was also presented what the effect of using the model is for estimates. How big is the relative efficiency for estimates in RLP? Consider discussing the impacts of using an internal model (as here) vs. an external model.*

   We are currently writing a follow-up article in which the 'final' regression model presented in this article is in fact used as an internal model for model-assisted small area estimations of standing timber volume within 390 forest management units in RLP. In this framework, we also discuss the pro and cons of internal vs. external models. However, while the demonstration of a small area double sampling procedure for the German NFI has actually been the underlying overall objective, we decided to publish the findings in two separated articles. Our motivation to address the model building in a separate first article as a pre-study has been that the identification of the 'best possible' regression model turned out to be a major issue of the entire study. In particular, we came to the conclusion that the three major issues we had to deal with (i.e. heterogeneity in the remote sensing data, identification of the optimal support under angle count sampling, incorporating tree species information including the handling of misclassifications) have not yet been dealt and linked to each other in this detail.

2. *Please discuss the results with the study by Kirchhoefer, et al. (2017) Considerations towards a Novel Approach for Integrating Angle-Count Sampling Data in Remote Sensing Based Forest Inventories, Forests*

   Thank you very much for the hint to this study. We discussed the findings of the study at multiple occasions in the 'Discussion' (section 4), i.e. at the end of section 4.2 'Stratification according to ALS Years and Tree Species' (page 13, RHS), and in section 4.3 'Choice of Support under Angle Count Sampling' (page 14, RHS).

3. *Despite of the large number of observations, the number of 39 explanatory variables is quite large for a linear model. Could you discuss implications of this? Except for 2007, the model parameters do not seem to be all too different (Fig 7). Consider merging several years (all except for 2007?) into one factor as a simple means to reduce the number of model parameters.*

   1) Concerning the number of parameters, we here considered the often cited rule of thumb by Draper and Smith (2014) {Applied regression analysis, Chapter 15.1}, i.e. one should at least have 5-10 observations per parameter in the regression model in order to avoid the issue of overfitting. For our final model (39 parameters), that would have implied a sample size of at least 390 observations. Since the actual number of observations used for model fitting was 5171 and hence considerably beyond the threshold suggested by Draper and Smith, the number of 39 parameters was not regarded to be critical. We added a comment on this in section 4.1 'Stratification according to ALS Years and Tree Species', page 13 (RHS), line 825-830.

   2) Concerning merging several years (all except for 2007) into one factor: As described in section 2.3.1, we already merged the ALSyears 2006 and 2007, and 2012 and 2013. However, as emphasized in Fig.8 and Table 2, the purpose of using the ALSyears factor-levels was to allow for higher model accuracies in ALSyear-strata in which the height data showed less noise or were in better temporal alignment with the terrestrial inventory than in other ALSyears. The so improved model accuracies were however substantially reduced or even removed when merging several ALSyears-levels, since this increased the noise in the data and lowered the overall R^2 values as well as the R^2-values within the original ALSyear strata. We show this in the table below, comparing the R^2-values of the final model (as given in the article) and those achieved when merging the factor-levels except '2007' (called the 'merged model'). The same was investigated when merging other levels. As depicted in section 'Discussion' page 13/14 line 862-869, achieving those higher model accuracies within domains that are located in the respective ALSyear-strata are considered to improve the estimation errors in the frame of our upcoming small area estimation analysis. We will particularly investigate this issue in the upcoming article. For the mentioned reasons, we thus decided not to merge any ALSyear factor-levels. We added a comment on this in section 4 'Discussion', page 13 (RHS) line 840-853.

$ALSyear$	$Area_{ALSyear}$	R^2 final model	R^2 merged model	n
2012	2807	0.61	0.53	408
2011	4361	0.57	0.45	883
2010	4182	0.51	0.43	1171
2009	2100	0.42	0.34	559
2008	2968	0.48	0.41	701
2008_1	2116	0.33	0.33	394
2007	3498	0.46	0.37	418
2003	602	0.27	0.23	529
2002	775	0.44	0.40	314

4. *The use of square supports is at least uncommon. Basically all studies I am aware of use circular supports (in the case of circular sample plots as here). Therefore, this choice should be justified a bit more or be revised. The reason given that the support should allow for a potential tessellation does not seem to hold because the exact plot location will not allow for an alignment with a potential tessellation grid anyways. To my understanding, the area (size) of the support needs to fit to the field data and the tessellation grid, especially if scale-dependent explanatory variables are used. The shape of the support should resemble the field data. I think the model variance is artificially increased by selecting a support that does not fit (i.e. is not circular) with the field data.*

   1) Thank you very much for the hint. We recalculated the entire analysis under the use of circular supports. It turned out that this only had minor effects on all accuracy metrics that were derived in our study. The results were almost identical to the previous ones and did not affect the major findings of the article. It however led to a further improvement of the model accuracy in the ALSyear-strata 2012, 2011 and 2010 (Table 2). We thus decided to switch to the use of circular supports with respect to your comments and changed all respective text-sections, numbers, Figures and Tables accordingly.

   2) Rational for rectangular supports: After reconsideration, we became aware that the reason given in the article for using rectangular supports (i.e. theoretical tessellation of the forest area) was indeed not correct. In the *infinite population approach* used by the model-assisted estimators by Mandallaz (and also those proposed by Saborowski 2010 {Double sampling for stratification in periodic inventories-Infinite population approach.*Forest ecology and management*}.), the estimators do in fact not impose any assumptions on the geometry and size of the supports. This is because the explanatory variables are sampled from an infinite population of points defining the continuous distribution (i.e. surface) of the explanatory variable. We rephrased the respective text in section 2.4, page 7, line 515-518.

5. *Was it not necessary to remove outliers or other influential observations from the data set? It sounds almost too good to be true, if that was not necessary.*

   We conducted an Influence Analysis, i.e. leverage / outlier detection, for the 'final model'. We here considered the usually applied criteria of *Leverages* and *Cook's Distance* as amongst others described in Fahrmeir 2013 {Regression: models, methods and applications. *Springer Science & Business Media*. page 164-167}. The critical threshold of $2p/n$ (i.e. twice the average of the hat matrix' diagonal entries), was exceeded by 10% of the observations. However, only 3% of these leverage points were assigned to studentized residuals > 1 or $<$ -1. Leaving these 3% of points out and recalculating the final model lead to a R^2_{adj} of 0.494 compared to 0.485 when including them in the regression. The Cook's Distance values D_i did not exceed a value of 0.019, and thus were far apart from the often cited critical threshold of $D_i > 0.5$. Based on these findings, we decided not to remove observations from the modelling data set. We added a paragraph commenting on this issue of influential data points in section 3.3 'Final Regression Model', page 9 RHS line 690-707.

   Especially if the regression model is used as an internal model for design-based estimations (which will be the case in our follow-up study), we generally consider removing potential outliers or leverage points an issue that has to be handled with extreme caution. In the design-based framework, removing data points from the modelling/sampling frame is only valid if the respective observations turns out to be truly erroneous. If this is not the case, the removal of outliers or influential data points might increase the model fit, but to the cost of possible bias for the estimates (bias-variance trade-off). This is because excluding an observation from the model fit has in the first instance to be regarded as an interference with the random sampling process. We will comment on this issue in our upcoming study.

6. *It is also somewhat uncommon to derive explanatory variables for CHMs instead from ALS raw data. A lot of information seems to be lost that way. Please justify or revise. Differing pulse densities usually do not have much influence on working models and can easily be considered in the model. This has probably a technical reason?*

Although there are indeed many studies where explanatory variables are derived from the ALS raw data, using a rasterized ALS CHM in this kind of studies is up to our knowledge not at all uncommon. In our particular case however, the main reason was that the available ALS data had very low point densities in many parts of the study area (i.e. less than 1 point per m^2). Owing to the low point density it is difficult to extract meaningful statistical descriptors which might significantly improve the model accuracies. This is especially because the 'mean canopy height' as a predictor variable already has the highest predictive power for our plot-based timber volume predictions. We expect that the derivation of this variable from the rasterized representation is well sufficient and that its derivation from the ALS raw data will not increase its explanatory power. We had also tested additional variables derived from the CHM, such as 'height-percentiles' (25%, 50%, 75%) and the 'maximum height value'. None of them added any explanatory power to the model. We are thus convinced that deriving the selected predictor variable directly from the rasterized CHM is justified. Since two years the ALS data acquired by the Geodetic Survey provide substantially higher point densities, which is more promising with respect to deriving ALS echo statistics directly from the point cloud in future studies.

Another more long-term oriented reason for using the rasterized ALS height information was the transferability of the algorithms to the photogrammetric canopy height model (mentioned in section 4 'Discussion'), which will be updated in much shorter terms as future ALS campaigns, and thus provides a much better temporal alignment with the terrestrial inventories. An example from Norway where such information is used can be found in Breidenbach & Astrup 2012. {Small area estimation of forest attributes in the Norwegian National Forest Inventory}, *European Journal of Forest Research*.

With respect to the reasons explained we believe that revising the approach is not justified, considering the trade-off between potential improvements and the required processing efforts (our entire analysis algorithms have been set up and optimized for 'raster'-processing operations in a PostgreSQL database which holds the rasterized CHM as well as the tree species classification map).

7. *Why and how were the ALS raw data thinned before interpolating them to grids? (Consider giving point densities in the more common unit point per m2.)*

   The data was delivered as two separate datasets comprising the Vegetation First Pulse (VEF) and Ground (GRD) points. In order to create a surface model (DSM) in a given raster resolution, the highest point of the combined VEF and GRD dataset was identified in each raster cell and saved as a thinned surface point cloud. For the elevation model (DEM), the mean of all GRD points in the cell was calculated, and the result was saved as a thinned ground point cloud. The thinned point clouds were then aggregated to larger tiles and interpolated to raster images using a Delauney interpolation in the Matlab software. The resulting DSM and DEM raster sets were then subtracted from each other to calculate a canopy height model (CHM) in raster format, providing discrete information about the canopy surface height of the entire forest area of RLP in a spatial resolution of 5 meters. The thinning process led to much smaller datasets that could be processed in larger tiles and considerably lowered processing times than the original dense point clouds. Since the data was recorded in leaf-off condition, the original point clouds contained many returns from within the crowns of deciduous trees. The thinned dataset here also provided the advantage that those measurements did not skew the vegetation height estimate in the final CHM. We rephrased the paragraph in Section 2.3.1 'ALS Canopy Height Model' page 4 RHS accordingly to be more clear about this issue. We also changed the units to point per m^2.

8. *Edge correction. Figure 3b suggests that supports were clipped at forest boundaries. 1) There is an additional data set for forest extent of public forests. Could it be described a bit more? Is the model, strictly speaking, only valid for public forests? How were plots on private forests treated? 2) Does clipping of supports fit to the type of edge correction in angle count sampling used in BWI3?*

   1) <u>Restriction to state and communal forest area:</u> We indicated the reason for this in the Introduction (page 2 RHS, last paragraph): 'Our study is embedded in the current implementation of model-assisted regression estimators (Mandallaz, 2013a,b; Mandallaz et al, 2013) for estimating the standing timber volume within the state and communal forest management units over the entire state of Rhineland-Palatinate'. Actually, we use the 'final' regression model presented in this article as an internal model for model-assisted small area estimations of standing timber volume within 390 the state and communal forest management units in RLP (to be presented in our upcoming follow-up article). The state and communal forest area thus constitutes the sampling frame on which the regression model identified in this article is subsequently applied to. Already restricting the set of sample plots used for modelling to this sampling frame in this article provides the advantage that when used as an *internal model* in design-based estimators, the regression model predictions already hold the assumption on the residuals to be zero on average for state and communal forest by construction of OLS technique (see amongst others Mandallaz 2013 {Design-based properties of some small-area estimators in forest inventory with two-phase sampling. *Canadian Journal of Forest Research*}).

We added this rational to the article (page 4, Line 254-264) to make the reason for restricting to state and communal forest more transparent. We also hope that this addresses your question regarding the 'validity' of the model. It also has the huge advantage that we can refer to this article for details of the used regression model in our upcoming follow-up article that focuses much more on the application of the double-sampling estimation techniques.

2) Forest ownerships in RLP: We are not sure what the reviewer refers to as 'public' forest. We added some sentences in section 2.1 'Study Area', LHS line 184-186 describing the three forest ownership classes in RLP, i.e. state forest, communal forest and private forest.

3) Clipping of supports: The clipping of supports at the forest boundary is a means to optimize the coherence between explanatory variables computed at the forest boundary and the corresponding terrestrial response variable, thereby optimizing the model fits for such observations (see Mandallaz et al. 2013 {New regression estimators in forest inventories with two-phase sampling and partially exhaustive information: a design-based monte carlo approach with applications to small-area estimation}). In the BWI3 survey, edge correction is applied at the forest border at the individual tree level. This means that sample trees whose inclusion circles are intersected with the forest border are assigned with a corrected (*increased*) counting factor. This method is used to compensate for the fact that part of the trees inclusion circle is outside the forest area. Consequently, the terrestrially determined timber volume value of a sample plot with existing boundary effects would be underestimated if the edge correction was neglected. Now, as obvious from Figure 3b) (upper left support), the ALS mean canopy height will drop to around zero outside the forest area (i.e. beyond the forest border). Including these 'zero' height pixels when calculating the value of the mean canopy height for this plot will severely attenuate the mean canopy height value towards zero (this effect will increase with increasing proportion of the support lying outside the forest border). However, the terrestrially recorded timber volume value has been compensated for the edge effect by increased counting factors for the affected sample trees. Neglecting the boundary correction of the support would thus increase the discrepancy between the value of the explanatory variable (attenuated towards zero) and the terrestrial timber volume value (increased by corrected counting factors). An optimized comparability can thus be realized if we restrict the calculation of the mean canopy height to those height pixels lying within the forest border, thereby avoiding the attenuation towards zero. In our opinion, the proposed clipping-method thus indeed accounts very well for the edge correction in angle count sampling used in BWI3. We added a comment on this in section 2.3.1 'ALS Canopy Height Model', page 4 RHS, last paragraph line 315 - 322.

9. *Calibration. Did this study really ïntroduce a calibration technology"? Consider revising. Why is the tree-species model a "calibration model"? Calibration model sounds like the parameters of the original model were adjusted. Calibration is also an estimation technique and could be misunderstood in this context. Would it be a calibration model if a traditional tree species map was used as explanatory variables? Is it not simply a model with some categorical explanatory variables?*

   Our proposed method is indeed a true calibration as known from regression-calibration approaches in the sense that 'parameters of the original model are adjusted'. Classical statistical calibration models are used to calibrate an error-prone variable that can cheaply be measured in high quantity on its corresponding exact, i.e. error-free variable whose recording is however very cost intensive. We exactly used this statistical framework to calibrate the estimated main tree species from the classification map (which revealed the quantified misclassification errors shown in Fig.4) on the error-free (i.e. exact) main tree species calculated from the set of sample trees in each terrestrial plot. This calibration of the tree species variable led to an increase in the classification accuracies and more importantly, considerably reduced the effect of the misclassification errors on the regression coefficients and thus increased the model accuracy when using the *calibrated* main plot tree species as a categorical variable in the regression model. Concerning your question, the term 'calibration model' refers to the random forest algorithm that is used to *calibrate* the error-prone tree species variable on the error-free determined main plot tree species (determined from the set of sample trees at each plot location). The regression model using this calibrated categorical variable is indeed 'simply a model with some categorical explanatory variables'. The proposed calibration method was also well received by the other reviewer and pronounced to be an 'important' part of our work. On suggestion of the other Reviewer to 'put the calibration more popular' in our article, we rephrased and expanded Section 2.3.2 'Calibration', page 5-6. We here particularly considered your input and tried to be much more clear about the basic idea of calibration in measurement error statistics, and differentiate it from the topic of calibration estimation.

10. *The issue of using y-transformed models or not is of high importance. By discussing why g-weight variance is of important, this paper could contribute considerably to the discussion. In addition: How are negative predictions dealt with in practice? Set to 0?*

    1) g-weight variance: This is a very interesting hint and idea for our currently written follow-up article that actually deals with the design-based small area estimations in RLP. In fact, the application of the g-weight

variances is a fundamental part of the upcoming study and we will incorporate your suggestion in this study. We thus only added a small comment in this article in section 3.3 'Interpretation of Final Regression Model' page 11 (LHS) line 736-747, and will discuss this issue in more detail in the follow-up article.

2) negative predictions: The purpose of this study was the identification of a best possible regression model to be integrated in the model-assisted estimators. We added a respective comment in section 3.3 'Interpretation of Final Regression Model' that rarely occurring negative predictions will most likely not have any influence on model-assisted estimates when multiple predictions are averaged over spatial domains.

11. *P1,L44 RHS: Is Beaudoin et al really the right reference here? The concept is anyways much older and Næsset, E. (1997) Estimating Timber Volume of Forest Stands Using Airborne Laser Scanner Data. Remote Sensing of Environment should be considered.*

    We revised the given reference of Beaudoin and now refer to the article of Brosofske et al. 2014 {A review of methods for mapping and prediction of inventory attributes for operational forest management. *Forest Science, Society of American Foresters*} who give a very nice and extensive review of applied mapping techniques in forestry, including various references to applications in the Nordic countries (page 1, 'Introduction', Line 42). We also added a sentence particularly emphasizing the long history of timber volume prediction models with reference to Naesset 1997 {Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*} (page 1 RHS, 'Introduction', Line 49-52).

12. *P2,L15: Does Van Aardt et al. fit into the list of references?*

    We removed Van Aardt from the list of the given references and replaced it with Bohlin et al. 2017 {Mapping forest attributes using data from stereophotogrammetry of aerial images and field data from the national forest inventory. *SILVA FENNICA*} as an up-to-date reference for mapping approaches applied in Sweden using data from the national forest inventory (page 2, LHS, 'Introduction', line 71).

13. *P2,L59: It may be of interest around that line that a variable describing deciduous proportion derived from leaf-off ALS data was used by Breidenbach et al. (2008) Mixed-effects models for estimating stand volume by means of small footprint airborne laser scanner data. Photogrammetric Journal of Finland to improve a model for timber volume.*

    Thank you very much for drawing our attention to this article. We rephrased the respective section in the 'Introduction', page 2 RHS line 112-117 accordingly and now mention the study of Breidenbach et al. (2008). We also reference to this article in our 'Discussion' page 13 RHS (line 811-816) and page 14 LHS (line 877-886), as it nicely supports and complements several findings of our article.

14. *P2,L5-8 RHS: 'One of the rare examples... ' I cannot follow here. The way it is described, it is exactly the approach commonly used in the Nordic countries. There must be hundreds of studies one could cite. Probably a misunderstanding?*

    That was indeed not well formulated. Our intention of this section was actually to identify the gap of knowledge that we tried to address by dealing with tree species information in our study. We particularly wanted to point out that - up to our knowledge - there has not been a study that investigated the integration of this amount of tree species categories (i.e. 5 categories: beech, oak, spruce, douglas fir, scots pine) as explanatory variables in a prediction model. However, we unfortunately had difficulties to find a lot of studies that use similar approaches than described in our article, i.e. using estimated tree species information (categorical or continuous) as explanatory variables in prediction models which go beyond the distinction of 'coniferous/broadleaf'. This lack of knowledge is amongst others addressed in White et al. 2016 and Koch 2010. We rephrased the respective section 'Introduction' page 2 RHS line 112-125 accordingly and decided not to mention Packalen et al. 2006 since their approach seemed in fact to differ too much from our approach.

15. *P13,L31-37: Why blending ITC into this article? Does the concept of supports fit there at all? Consider removing.*

    The sentence has been deleted from the article.

16. *Consider discussing the model a bit more with other models published in central Europe (maybe based on smaller study sites) and other studies based on NFI data. Why do we see differences?*

    We now discuss our results with 4 similar studies that have been carried out in central Europe and address different aspects of our article. These studies comprise: Latifi et al. (2012), Breidenbach et al. 2008, Maack et al. 2016 and Kirchhoefer et al. 2017. The added discussion can be found in Section 4 'Discussion', 'Stratification according to ALS Acquisition Years and Tree Species,' page 13.

17. *P7,L30-31: Ambiguous. Please rephrase. Is dominant species here decided based on field data or the map?*

    We rephrased this sentence accordingly. We also rephrased the sentence about how the accuracy assessment was made at the beginning of section 2.5 'Model Building and Evaluation', page 7 line 523-530: *'In order to judge the quality of the treespecies variable, the user's accuracy for each classified species category and the overall accuracy of the classification scheme was calculated based on the confusion matrix (Congalton and Green 2008). As reference data, we calculated the actual main plot tree species by applying the respective threshold to the sample trees of each sample plot'.*

18. *P7,L7-32: The model interpretation is a bit lengthy and may be best described by a reference to Fig 7.*

    We think the Reviewer is referring to P10,L7-32 (section 'Interpretation of Final Regression Model'). In fact, the model interpretation is descibed by a reference to Fig. 7 (first sentence). We do not share the Reviewer's opinion that the model interpretation is to lengthy, considering that it is - up to our knowledge - the first study where such detailed tree species information (5 tree species) has been used in a prediction model. We therefore consider that providing a detailed interpretation of the model is an important part of the article. With respect to the reasons explained we believe that revising the section is not justified.

19. *Please add a table on field measured values and relative RMSE values (%).*

    We added a table on the field measured values within the state and communal forest area at the end of section 2.2 'Terrestrial Inventory Data', page 4. RMSE-values in % have also been calculated for the model accuracies and added in the article. There are however no RMSE-values for the field measured values.

20. *Please consider that the Conclusion is not an extension of the Discussion. Consider revising to shorten and to avoid references in the Conclusion*

    We have completely revised the Conclusion by transferring some aspects (such as the photogrammetric CHM in RLP) to the section 'Discussion'. We significantly shortened the 'Conclusion' (section 5, page 15) and restrict to shortly present the three major conclusions of our article.

21. *Fig 4: Please define 'ind'.*

    'ind' is now defined in the caption of Fig. 4. It was also already defined in section 2.4. 'Choice of Support under Angle Count Sampling'.

22. *Is Fig 5 really needed?*

    We consider Figure 5 as one of the most important graphics in our article. It gives the visual overview about all model evaluations carried out under the different parametrizations (threshold and support choices, calibration). It addresses all 3 objectives of the article that are defined at the end of the 'Introduction'. We thus decided to keep the graphic. On suggestion of the other Reviewer, we changed the layout of Fig.5 in order to provide a better visual distinction.

23. *Figure 6: Could you expand in the text why two distinct clusters are visible? This remained a bit unclear. The dashed lines are hardly visible.*

    The differentiation into two distinct point clouds results from the poor model performance under support size $q100$ for the CHM variables (i.e. the lower point cloud). We added this information in Section 3.2 page 9 RHS 'Effect of Missclassifications' (line 657-660). We also changed the background to white in order to improve the visibility of the dashed lines.

24. *Figure 8: Consider showing only the simple model and the final model. The others seem to be of little relevance in this context and just make the graphic difficult to read.*

    We changed the graphic with respect to suggestions of the other Reviewer in order to provide a better visual distinction. We would however very much like to keep all visualized models in the graphic. It was our particular intention to show the differences in model accuracy between these models within the different ALSyears. Visualizing submodel 2 (use of ALSyears as predictor variable) is of importance since one can see the improvement of the model accuracy when using the treespecies predictor variable additional to the ALSyear variable.

25. *Figure 7: (The fig appears after Fig 8.) The single panels are very small and difficult to read. Consider just showing one panel to exemplify.*

    The order of the Figure 7 and Figure 8 has been corrected. We however would very much like to keep Figure 7 as is, because our intention was to show the different observation point clouds per ALS acquisition year at least once in the article (plotted in the background of each panel). We also wanted give the reader a visual impression of how well the tree species specific regression lines fit to the data in each ALS year. For a better readability, we adjusted the graphic by removing the grey background and by increasing the contrast of the underlying data points. We hope the Reviewer can agree with our decision not to further revise this Figure.

26. *Consider revising the use of percent (%) vs. percentage points. My impression is that the terminology could have been used in the wrong way. (And R2 values are proportions, not percentages.)*

    We revised the notation of % and changed the R^2-values into percentage points (i.e. 0.5 instead of 50%).

27. *Terminology (p3, l16): Consider using ALS throughout the paper as the correct acronym for airborne laser scanning. (Or lidar which is consistent with acronyms like radar or laser.)*

    Has been changed accordingly.

28. *Please add page range when citing books.*

    We used the article of Breidenbach and Astrup 2012 {Small area estimation of forest attributes in the Norwegian National Forest Inventory}, *European Journal of Forest Research*) as example and added the book-chapters to the book-references of Mandallaz 2008, Gustafson 2003, Caroll et al. 2006 as well as the page range of Draper and Smith 2014 in our article accordingly.

29. *Eq1: are $Y(x)$ and $e(x)$ defined? Consider presenting table 2 close to eq 1 as the submodel names were undefined until table 2 was presented.*

    The timber volume density Y(x) has been defined on page 4 LHS line 265-269, section 2.2 'Terrestrial Inventory data'. Additionally, the submodels are now described *before* table 2 (now table 3) and Fig.8. We changed the letter e to ϵ to use a common way to notate the error term of a OLS regression model.

30. *P5L37: Is the threshold defined?*

    Thank you very much for the hint. We have not been aware that this information was missing. We added this to the paragraph accordingly in section 2.3.2. 'Prediction of main tree species' page 5 line 386-389.

31. *P8L17-21, RHS: Discussion?*

    We moved this sentence to the Discussion, section 4.2 'Calibration of Tree Sprecies Map Information', page 14 line 891-895.

32. *Format: Consider submitting a one-column, two line spacing manuscript, such that line numbers are available to all lines.*

    The editing was done by the editorial office and we did not have influence on it. The version we prepared for the reviewers did in fact have line numbers on both sides, and we will ask the editor to provide this version to the Reviewers.

Reviewer 2:

It is a very interesting and useful approach and paper. Your research question is well defined, extensively investigated and explained. Regarding your figures please be aware, that printing or digital presentation will be very small. Try to use clear distinguishable colors and markers. I recommend removing the grey background in all figures. Some phrases are quiet long and contain several aspects. Try to shorten your sentences, by separating them according to the different aspects or contents.

Comments:

1. *P3 line 46 Column 2: Did you measure the absolute tree height of every sample tree, or was it a estimation according to the measurements of a sub-sample of tree?*

   In fact, the height is *measured* only for a subset of the sample trees at each plot. The height-values for the remaining sample trees are then *estimated*. The taper functions however use the height-value of a sample tree without distinction between 'measured' and 'estimated' as one of the explanatory variables. We rephrased the sentence accordingly (page 3-4, section 'Terrestrial Inventory Data', line 230-237).

2. *P3 Line 48 Column 2: Do you have any information about the type and positional accuracy of the used GPS or GNSS-Techniques*

   Yes, we in fact looked at this issue very closely in an individual study of Lambrecht et al. 2017 {A Machine Learning Method for Co-Registration and Individual Tree Matching of Forest Inventory and Airborne Laser Scanning Data. *Remote Sensing*}. The analysis was carried out for a subsample of all sample plots in RLP and indicated that horizontal DGPS errors do not exceed 8 meters for 80% of all plots. We were unfortunately not provided with estimated accuracy metrics from the actual DGPS acquisition software that was used by the field crews. The dataset provided from the authorities only contained information about the number of DGPS measurements used to average the final plot position coordinate (100 measurements at each plot center) and the PDOP-value. Both seemed unsuitable to give a reliable accuracy for the plot positions. The study of Lambrecht et al. is the first study in RLP that have provided an idea about the actual positional accuracy. We added a respective comment in the article (page 4, LHS, line 237-246).

3. *P3 Line 57/58: Why did you restrict to stat and communal forests?*

   We indicated the reason for this in the Introduction (page 2), last paragraph: 'Our study is embedded in the current implementation of model-assisted regression estimators (Mandallaz, 2013a,b; Mandallaz et al, 2013) for estimating the standing timber volume within the state and communal forest management units over the entire state of Rhineland-Palatinate'. We are currently writing a follow-up article, where the 'final' regression model presented in this article is in fact used as an internal model for model-assisted small area estimations of standing timber volume within 390 the state and communal forest management units in RLP. The state and communal forest area thus constitutes the sampling frame on which the regression model identified in this article is subsequently applied. Already restricting the set of sample plots used for modelling in this article provides the advantage that when used as an *internal model* in design-based estimators, the regression model predictions already hold the assumption on the residuals to be zero on average for state and communal forest by construction of OLS technique (see amongst others Mandallaz 2013 {Design-based properties of some small-area estimators in forest inventory with two-phase sampling. *Canadian Journal of Forest Research*}). We added this rational to the article (section 'Introduction', page 4, LHS, line 254-264) to make the reason for restricting to state and communal forest more transparent. It also has the huge advantage that we can refer to this article for details of the used regression model in our upcoming follow-up article that focuses much more on the application of the double-sampling estimation techniques.

4. *P4 Line 24-26: In general the point density is given in points per m^2.*

   Values have been changed accordingly.

5. *P4 Line 47 -51: How did you calculate the canopy height model, by subtraction?*

   We forgot to mention this. Yes, the CHM was calculated by subtracting the DEM from the DSM. Based on suggestions of the other Reviewer, we rephrased this section 2.3.1 'ALS Canopy Height Model' (page 4, first paragraph), and also added the missing information about the CHM calculation.

6. *P4 L1 Column 2: How did you define the forest borders?*

   As mentioned on page 4, section 2.3.1 'ALS Canopy Height Model': The support *'was previously intersected with the state and communal forest area defined by a polygon mask and thereby corrected for edge effects at the forest border'*. So referring to your question, the forest borders of the state and communal forests were defined by the mentioned polygon mask, which was provided by the forest service of RLP. We tried to rephrase the sentence to make this more clear to the reader. We also added a few sentences commenting on the purpose of this method in section 2.3.1 (page 4, 2nd paragraph), as requested by the other Reviewer.

7. *P5 Line 34 – 40: Mention here, that you investigate this threshold in your work.*

   Thank you very much for the hint. We have not been aware that this information was missing. We added this to the paragraph in section 2.3.2 'Prediction of main tree species' (page 5) accordingly.

8. *P5 Line 11-36 Column 2: Why did you use these variables and not others? I would recommend putting the explanation of your calibration method more popular in your article. It is an important part of your work.*

   An advantage for using those explanatory variables in the calibration model was that they also provided explanatory power in the regression model, so they could be used for both, the calibration *and* the regression model. Using these variables thus considerably saved computation time as well as data storage space. We added this comment in Section 'Calibration' (page 5 and 6). Triggered by your suggestion above, we rewrote the entire Section 'Calibration' in order to be more clear about the technique of calibration in measurement error statistics, and to explain the idea of transferring this method to our misclassification problem in more detail.

9. *P6 Line 46: the title should be "model building and evaluation" instead of model validation*

   Has been changed accordingly.

10. *P7 Line 4-18: What are the values for lidar year, 0 and 1 or N/A and 1 or anything else?*

    Categorical variables are always recoded by a unique combination of '0' and '1'- values in the design matrix of a linear regression model. We however think that the recoding is rather theoretical knowledge which a reader familiar with regression techniques might have. However, for the reader it is just important to see that each factor level (i.e. each tree species and each ALS year category) has its own regression coefficient to be fitted. This leads to the amount of parameters used in a respective linear model.

11. *P7 figure4: the figure is overloaded. Try to use colored lines instead of colors and different markers. Is it really necessary to show both n and threshold? If yes, put them into two separate figures. Or could you explain*

    We removed the 'n' from the figure as it did in fact not add any valuable information. We also changed the layout of Fig. 5 (and all remaining figures) based on your suggestions and think it really improved the graphic (now on page 8). The changes comprised: a) using flexible scales among the tree species groups in order to 'zoom' in and make the displayed information better to distinguish from each other. We also placed the legend on the bottom like you suggested and adapted the colour-scheme to support the visual distinction. The background has been changed to white.

12. *P8 Line 1: Title should be "Calibration" and not only "Effect of Calibration".*

    Has been changed accordingly.

13. *P8 L55-59: Separate it in at least two sentences.*

    We rephrased the section accordingly (now on page 7 LHS, line 523-529, section 2.5 'Model Building and Evaluation').

14. *P9 figure5: the figure is hard to differentiate: Use color instead of different markers, put the legend below or above could already help.*

    We changed the layout of Fig.5 on page 10 (and also Fig.4 on page 8) based on your suggestions and think it really improved the graphic. The changes comprised: a) using flexible scales among the tree species groups in order to 'zoom' in and make the displayed information better to distinguish. We also placed the legend on the bottom like you suggested. We also adapted the colour-scheme to support the visual distinction.

15. *P9 figure 6: Where do the different grey values refer to?*

    We missed to mention that we used semitransparent colour for the data points to visualize overlap, and to indicate where most of the calculated model accuracies are located in the plot. We added this information to the caption of Figure 6 page 10.

16. *P10 figure 8: I recommend limiting the y-axis between 0.2 and 0.6. A legend is missing and should be added. The long title of the figure contains text, which should be included in the text and not in the title.*

    We changed the figure 8 (now page 13) with respect to your suggestions, removed the grey background and changed the colour scheme for a better visual distinction. We also moved the description of the displayed models into the text (section 3.3, 'Effect of Time-Lags and Heterogeneity in ALS Data', page 11 RHS, line 754-761).

17. *P10 Table 1: I recommend mentioning RMSE% instead of SSE*

    Additional to $RMSE_{cv}$ or $RMSE$, we now also calculated $RMSE_{cv}[\%]$ and $RMSE[\%]$. These values have respectively been added to the text-sections, the supplementary data tables (Table 2 and 3) as well as Table 1 in the article instead of SSE. We also decided to map the $RMSE_{cv}[\%]$ instead of $RMSE_{cv}$ in Fig. 5 page 10 (right).

18. *P12 Table2: 1) submodel 3 isn't mentioned and explained in the text. 2) It isn't possible to reconstruct how you get your amount of parameters for each model. In addition the amount of parameters isn't discussed in the text. So I recommend either to explain (I would prefer) or to leave it out.*

    1) We now explicitly mention submodel 3 in section 'Effect of Time-Lags and Heterogeneity in ALS data' (page 11 RHS, line 754-761).

    2) With respect to the amount of model parameters, the intention of giving the number of parameters in Table 2 (now Table 3) was to show the reader that expanding the model-terms comes with an increase in parameters to be fitted. We used this information to add a comment to overfitting-problems in section 4

'Discussion', page 13 RHS line 821-826.

In general, one gets the number of parameters of an OLS model by summing over the number of continuous variables, the number of categories of the continuous variables and the number of interaction variables. For the final model, this is 1 intercept + 3 main effects for continuous variables + 13 main effects for categorical variables + 22 interaction variables = 39 regression coefficients (parameters) to be fitted. We added this information in section 3.3 'Final Regression Model' RHS line 686-689. We however decided not to further explain how to reconstruct the number of parameters. This is because when reconstructing the number of parameters, one has to consider that the respective reference category is always incorporated in the respective main effect. This is due to the recoding-method used for setting up the design-matrix of an OLS regression model. We think that a detailed explanation of this would go beyond the scope of the article and in fact lead to a repetition of OLS regression theory, which is not the intention of the article. We assume that an interested reader should be able to reconstruct the amount of parameters after consulting one of the many available books on OLS regression models. For reconstruction, the number of categories (factor levels) for the $ALSyear$-variable and the $treespecies$-variable are repeatedly provided in section 2.3.1, section 2.3.2, Figure 4, Figure 7 and Table 2). We thus think that a revision of table 3 is not justified.

19. *P12 Line35: This is the wrong reference: Table1 instead of Table2.*

    Has been corrected.

20. *P13 Line 43 Column 2: aerial instead of areal.*

    Typo has been corrected.

Documentation of changes:

Note: In order to facilitate the identification of changes to the previous manuscript, we provided a version of the revised manuscript to the editor and the reviewers in which *changes* to the previous text have been colored in *red*, and *added* text sections have been colored in *green*.

The following changes affect the entire article:

1. Triggered by the suggestion of Reviewer#1, we recalculated the entire analysis based on *circular* rather than *rectangular* supports. The results were almost identical to the previous ones and did not affect the major findings of the article. It however led to a further improvement of model accuracy in the ALSyear-strata 2012, 2011 and 2010 (Table 1). We changed the respective text-sections, tables and figures accordingly.

2. At the suggestion of Reviewer#1, we changed the terminology of LiDAR into ALS.

3. At the suggestion of Reviewer#2, we removed the grey background in all Figures. We also changed the layout of Fig.4 and Fig.5 based on the suggestions of Reviewer#2. These changes comprised: a) using flexible scales among the tree species groups in order to ßoomïn and make the displayed information better to distinguish. We placed the legend on the bottom like suggested by Reviewer#2. We also adapted the color-scheme to support the visual distinction.

4. Additional to $RMSE_{cv}$ or $RMSE$, we now also calculated $RMSE_{cv}[\%]$ and $RMSE[\%]$. These values have respectively been added in the supplementary data tables (Table 2 and 3) as well as given in Table 1 in the article instead of SSE (requested by Reviewer#2). We also decided to map the $RMSE_{cv}[\%]$ instead of $RMSE_{cv}$ in Fig. 5 (right). We also added the RMSE-formulas in section 2.5 'Model Building and Evaluation'.

The following text-sections have been subject to changes and/or adjustments:

1. Page 1, RHS, line 42: We added the Chapter-information to reference 'Mandallaz 2008', as suggested by Reviewer#1.

2. Page 1, RHS, line 42: We substituted the reference of Brosofske et al. 2014 for Beaudoin et al. 2014 as suggested by Reviewer#1.

3. Page 1, RHS, line 49-51: We added a reference to Naesset 1997 as suggested by Reviewer#1.

4. Page 2, RHS, line 112-126: We added a reference to Breidenbach et al. 2008 and removed the reference to Packalen and Maltamo 2006, as suggested by Reviewer#1.

5. Page 3, LHS, line 184-186: We added a sentence on the 3 forest ownership classes in Rhineland-Palatinate.

6. Page 3/4, RHS/LHS, line 230-246: Based on questions of Reviewer#2, we added a few sentences on a) the measurements of tree heights in the German NFI, and b) accuracy information on the DGPS-determined field plot coordinates.

7. Page 4, LHS, line 254-265: Based on questions of both Reviewers, we develop the rational for restricting the dataset to state-and communal forest in more detail.

8. Page 4, LHS, Table 1: We added a table on the field measured values as suggested by Reviewer#1.

9. Page 4, RHS, line 282-309: Based on recommendations of both Reviewers, we rephrased the calculation of the CHM. The units are now also given only in points per m^2.

10. Page 4, RHS, line 315-322: Based on questions of both Reviewers, we explain the support-adjustment at forest borders in more detail.

11. Page 5, Fig.2, caption: We added a remark that the transparent blue layer is the polygon mask defining the state-and communal forest area.

12. Page 5, RHS, line 366: As a consequence of using circular supports, the missing tree species information changed to 411 sample plots.

13. Page 5, RHS, 385-389: Suggested by both Reviewers, we defined the tested thresholds.

14. Page 5 line 391 - Page 6 line 465: We completely rephrased the section about the Calibration of the treespecies variable, based on suggestion of Reviewer#2 and questions of Reviewer#1.

15. Page 7, LHS, line 522: We changed the title of the section according to the recommendation of Reviewer#2.

16. Page 7, LHS, line 523-529: We rephrased the sentence as suggested by Reviewer#2.

17. Page 7, RHS, Figure 3 b): Now showing the *circular* supports.

18. Page 7, RHS, line 558: due to the use of circular supports, the sample size of the modelling dataset changed to 5171 as the number of missing treespecies plots changed to 411.

19. Page 7, RHS, bottom: We added the formulas for RMSE.

20. Page 8, LHS, top: We changed the error term of the OLS regression model from e to ε in order to stick to a more common notation.

21. Page 8, LHS, line 563-565: Suggested by Reviewer#1, we rephrased the sentence to avoid ambiguities.

22. Page 8, Fig.4: We updated the figure to the circular-support evaluation and adapted the layout as suggested by Reviewer#2.

23. Page 9, LHS, line 623-627: Now, the best model we found overall in terms of adjusted R^2 and cross-validated RMSE values (now also given in %) was the model later used as the final model ($q50$ support for both CHM and *treespecies*-variables). In the previous version, the best model was found with the support settings $q50$ for the CHM-variables and the $q100$ for the *treespecies*-variable

24. Page 9, LHS, line 633-644: We rephrased this section as recommended by Reviewer#2.

25. Page 9, RHS, line 658-660: We moved the explanation why two distinct point clouds are visible in Fig.6 (page 10) to this section, as suggested by Reviewer#1.

26. Page 9, RHS, line 662-663: residual inflation of the R^2 is now given in percentage points rather than in %.

27. Page 9, RHS, line 686-689: Recommended by Reviewer#2, we reconstruct the number of parameters used in the final model (table 3).

28. Page 9/11, RHS/LHS line 690-707: Based on questions of Reviewer#1, we provide details of an leverage/outlier analysis of the final regression model and provide the reasons why no plots except the 'zero-plots' were removed from the modeling dataset.

29. Page 10, Fig 5 and Fig 6: The layout was changed according to recommendations of both Reviewers.

30. Page 11, LHS, line 736-747: Based on comments of Reviewer#1, we extended the reason why no log-transformation of the response variable was applied.

31. Page 11, RHS, line 754-761: Comments on each submodel given in Table 3 (page 12) were added based on comments of Reviewer#2.

32. Page 11, RHS, Table 2: We added the area of each ALS acquisition for which the derived model accuracies are valid. We also added the RMSE% as suggested by Reviewer#2.

33. Page 11, RHS, line 775-782: Updated based on recalculations (circular supports).

34. Page 12, Fig 7: Updated based on recalculations (circular supports). Layout changed based on recommendations of Reviewer#2.

35. Page 12, Table 3: We substituted RMSE% for SSE, as recommended by Reviewer#2.
36. Page 13, Fig 8: We changed the Layout according to suggestions of both Reviewers.
37. Page 13 & 14, LHS/RHS, line 800-819, 834-837, 840-853, 877-886: As requested by Reviewer#1, we compare and discuss our findings with 4 similar studies carried out in central Europe, i.e. Latifi et al. 2012, Breidenbach et al. 2008, Maak et al. 2016, and Kirchhoefer et al. 2017.
38. Page 13, RHS, line 825-829: Based on questions of Reviewer#1, we comment on potential overfitting issues.
39. Page 14, LHS, line 891-895: As recommended by Reviewer#1, we moved this sentence from the 'Results'-section to the 'Discussion'.
40. Page 14, RHS, line 929-972: We rephrased this section based on suggestions of Reviewer#1 (we now also discuss our findings with the study of Kirchhoefer et al.).
41. Page 15, Conclusion: We completely rephrased the section 'Conclusion' based on the recommendation of Reviewer#1. Basically, we tried to avoid the conclusion to be an extension of the Discussion.

European Journal of Forest Research manuscript No.
(will be inserted by the editor)

Combining canopy height and tree species map information for large scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes

Andreas Hill · Henning Buddenbaum · Daniel Mandallaz

Received: date / Accepted: date

Abstract A timber-volume regression model applicable to the entire forest area of the federal German state of Rhineland-Palatinate is identified using a combination of airborne laser scanning (ALS)-derived metrics and information from a satellite-based tree species classification map available on the federal state level. As is common in many forest inventory datasets, strong heterogeneity in the ALS data due to different acquisition dates and misclassifications in the tree species classification map had noticeable effects on the regression model's performance. This article specifically addresses techniques that improve the performance of ordinary least square regression models under such restricting conditions. We introduce a calibration technique to neutralize the effect of misclassifications in the tree species variable that originally caused a residual inflation of 0.05 in adjusted R^2. Incorporating the calibrated tree species information improved the model accuracy by up to 0.07 in adjusted R^2 and suggests the use of such information in forthcoming inventories. We also found that including ALS quality information as categorical variables within the regression model considerably mitigates issues with time lags between the ALS and terrestrial data acquisition and ALS quality variations (increase of 0.09 in adjusted R^2). The model achieved an adjusted R^2 of 0.48 and a cross-validated root mean square error (RMSE$_{cv}$) of 46.7% under incorporation of the tree species and ALS quality information, and was thus improved by 0.12 in adjusted R^2 (5% in RMSE$_{cv}$) compared to the simple model only containing ALS height metrics (adjusted R^2=0.36, RMSE$_{cv}$=51.7%).

Andreas Hill
Department of Environmental Systems Science, ETH Zurich, Universitaetstrasse 22, 8092 Zurich, Switzerland
Tel.: +41 44 632 32 36
E-mail: andreas.hill@usys.ethz.ch

Henning Buddenbaum
Environmental Remote Sensing and Geoinformatics Department, Trier University, 54286 Trier, Germany
Tel.: +49 651 201 4729
E-mail: buddenbaum@uni-trier.de

Daniel Mandallaz
Department of Environmental Systems Science, ETH Zurich, Universitaetstrasse 22, 8092 Zurich, Switzerland
Tel.: +41 44 632 88 35
E-mail: daniel.mandallaz@usys.ethz.ch

Keywords OLS Regression · standing timber volume · ALS canopy height model · satellite-based tree species classification · calibration · forest inventory · angle count sampling

1 Introduction

Forest inventory methods are the primary tools used to assess the current state and development of forests over time. They provide reliable evidence-based information that is used to define and identify management actions as well as to adapt forest management strategies to both national and international guidelines. Two methods that have become particularly attractive are so-called *double-sampling* (Mandallaz, 2008, Ch. 5) and *mapping* (Brosofske et al, 2014) procedures. The core concept of these methods is to use predictions of the terrestrial target variable at additional sample locations where the terrestrial information has not been gathered. These predictions are produced by models that use explanatory variables derived from *auxiliary data*, commonly in the form of spatially exhaustive remote sensing data in the inventory area. Especially models to predict timber volume based on airborne laser scanning (ALS) have been extensively investigated for a long time (Næsset, 1997). The specific scope of double-sampling is to enlarge the terrestrial sample size by a much larger sample of predictions of the target variable in order to gain higher estimation precision without performing additional expensive terrestrial measurements. Model-dependent and design-based regression estimators are used in a broad range of double sampling concepts and methods (Gregoire and Valentine, 2007; Köhl et al, 2006; Mandallaz, 2013a,b; Saborowski et al, 2010; Schreuder et al, 1993) and have been applied to existing inventory systems (Breidenbach and Astrup, 2012; von Lüpke and Saborowski, 2014; Magnussen et al, 2014; Mandallaz et al, 2013; Massey et al, 2014). While double-sampling methods provide reliable estimates for a given spatial unit, e.g. a forest district, they do not provide information about the spatial distribution of the estimated quantity within this area. For this reason, the same modeling technique used in double-sampling procedures has also been intensively used to produce exhaustive prediction maps that provide pixel-wise estimations of a target variable in high spatial resolution (Bohlin et al, 2017; Hill et al, 2014; Latifi et al, 2010; Nink et al, 2015; Tonolli et al, 2011).

To allow for an area-wide application of the prediction model, both double sampling and mapping methods require that the remote sensing data are available over the entire inventory area. This is usually not a limiting factor in *small-scale* applications. In the optimal case, the remote sensing data are in principle collected in accordance to the specific study objective. Quality standards that have often been addressed are that *a)* the remote sensing data should be acquired close to or even at the time of the terrestrial inventory in order to ensure best possible comparability between the target variable on the ground and the remote sensing derived variables (McRoberts et al, 2015); *b)* the remote sensing technology and its spectral and spatial resolution should be chosen according to the modelling purpose (Köhl et al, 2006); and *c)* the variation in quality of the remote sensing data over the inventory area should be minimized in order to avoid artificial noise in the data (Naesset, 2014). Despite the increasing availability and decreasing costs of remote sensing data (White et al, 2016), these quality standards of the remote sensing data can often not be guaranteed for *large-scale* applications (Maack et al, 2016), and trade-offs must be accepted (Jakubowski et al, 2013). The prime objective is then to produce the best possible prediction model given the restrictions imposed by the available remote sensing information. The exploration of scarcely used remote sensing products and the optimization of prediction models under severe quality restrictions in the remote sensing data are thus one of the challenges in large-scale model-supported inventory applications.

Among the still rarely used remote sensing data in large scale applications, the integration of tree species information in prediction models - especially for timber volume estimation - has been stated as some of the most promising but often missing information (Koch, 2010; White et al, 2016). As timber volume estimations on the single tree level in forest inventories are often based on species-specific biomass and volume equations (Husmann et al, 2017; Zianis et al, 2005), the application of species-specific models is expected to be a key factor for improving estimation precision (White et al, 2016). This has been supported by studies from Breidenbach et al (2008) who achieved a substantial improvement in accuracy of their timber volume prediction model when including a variable estimating the deciduous proportion derived from leaf-off ALS data. Similar gains in model performance were also reported by Straub et al (2009) and Latifi et al (2012) who used broadleaf and coniferous information based on color infrared orthophotos as a categorical explanatory variable. However, studies that explore the use of more species-specific information (i.e. a further discrimination of tree species) as explanatory variables have been rare. Further investigations are thus necessary especially in countries whose forests are characterized by a larger variety of tree species that may also occur in mixed and uneven-aged stands (McRoberts et al, 2010). The area-wide tree species information in most studies was obtained from satellite and airborne remote sensing sensors based on automatic classification methods. Whereas the presence of misclassifications has already been addressed (Latifi et al, 2012), an issue that has so far been neglected is how misclassifications actually affect the prediction model (Gustafson, 2003).

A frequently encountered problem in large scale forest inventories is the lack of temporal synchronicity between the remote sensing acquisition and the terrestrial survey. As a result, the available remote sensing data often exhibit notable time-lags with respect to the date of the terrestrial inventory. This has often been addressed as a major drawback, especially for the application of design-based change estimation (Massey and Mandallaz, 2015).

Our study is embedded in the current implementation of design-based regression estimators (Mandallaz, 2013a,b; Mandallaz et al, 2013) for estimating the standing timber volume within the state and communal forest management units over the entire state of Rhineland-Palatinate (RLP, Germany). With respect to this overall objective, the aim of this study was to derive an ordinary least square (OLS) regression model to generate predictions of the standing timber volume associated with a sample location of the Third German National Forest Inventory (BWI3) over the entire state and communal forest area (6155 km^2). A merged ALS dataset from different acquisition years and a satellite-based tree species classification map for the five main tree species in RLP was available for the entire inventory area and consequently used to derive predictor variables. The major limiting factors for using these data in a regression analysis are **(i)** variation in the ALS data quality as well as time-lags of up to 10 years between the ALS acquisitions and the terrestrial survey, **(ii)** misclassifications in the tree species classification map and **(iii)** the ambiguous choice of a suitable extraction area (*support*) for all remote sensing information under angle count sampling in the terrestrial survey (variable sample plot sizes). For this reason, we address the following specific research questions:

1. How can tree species map information be optimally used within a regression model that predicts timber volume? What effects do misclassifications have on the predictions and how can these effects be minimized?
2. What are the effects of quality restrictions and substantial time lags between the ALS- and terrestrial data acquisition on the regression model and how can these effects be mitigated?
3. Does support size influence model accuracy? What is the optimal support size and what are the determining factors?

2 Materials and Methods

2.1 Study Area

The German federal state Rhineland-Palatinate (RLP) is located in the western part of Germany and borders Luxembourg, France and Belgium (figure 1). With 42.3% (appr. 8400 km^2) of the entire state area (19850 km^2) covered by forest, RLP is one of the two states with the highest forest coverage among all federal states of Germany (von Thünen-Institut, 2014). The forest area of RLP is divided into three ownership classes, i.e. state forest (27%), communal forest (46%) and privately owned forest (27%). The most frequent tree species in RLP are European beech (*Fagus sylvatica*, 21.8%), oak (*Quercus petrea* and *Quercus robur*, 20.2%), Norway spruce (*Picea abies*, 19.5%), Scots pine (*Pinus sylvestris*, 9.9%), Douglas fir (*Pseudozuga menziesii*, 6.4%), European larch (*Larix decidua*, 2.4%) and Silver fir (*Abies alba*, 0.7%). The share of broadleaf tree species is 58.7%. The forests of RLP further exhibit heterogeneous structures (von Thünen-Institut, 2014): around 82% of the forest area in RLP are mixed forest stands (i.e. at least two different tree species occur in the same stand) and 69% of the forest area exhibit a multi-layered vertical structure. While the average tree age is around 80 years, most of the forest area (20%) is occupied by trees between 40 and 60 years of age, whereas 27% of the trees are older than 100 years. Spatially variable climate conditions have a strong influence on the local growth dynamics as well as tree species composition and create a large variety of forest structures, ranging from characteristic oak coppices (Moselle valley), pure spruce, beech and Scots pine forests (e.g. Hunsrück and Palatinate forest) to mixed forests comprising variable proportions of oak, larch, spruce, Scots pine and beech. Accordingly, RLP has been divided into 16 bioclimatic growing regions that form homogeneous areas with respect to the afore mentioned characteristics (Gauer and Aldinger, 2005).

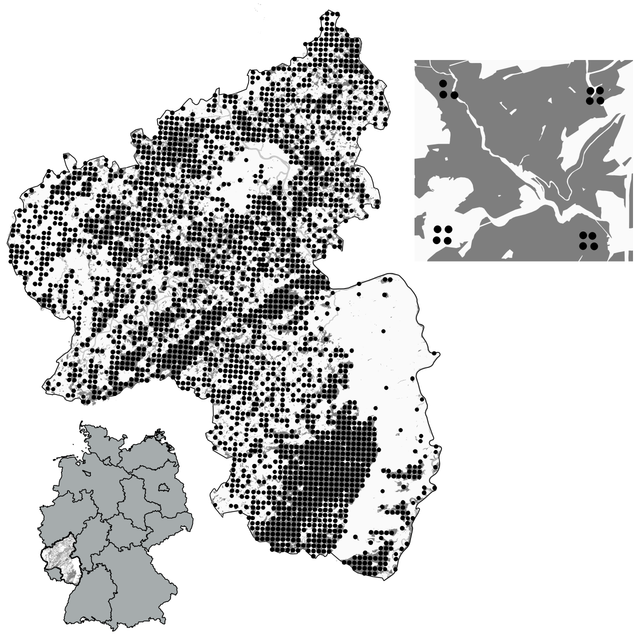


Fig. 1: Spatial distribution of the BWI3 cluster samples over Rhineland-Palatinate

2.2 Terrestrial Inventory Data

The German National Forest Inventory (NFI) is carried out over the entire forest area of Germany in reoccurring time periods of 10 years. The most recent inventory (BWI3) has been conducted in the years 2011 and 2012. In this framework, Rhineland-Palatinate is covered by a 2x2 km grid that defines the sample locations for the terrestrial survey. A sample unit consists of four sample locations (also referred to as *sample plots*) that are arranged in squares (so called *clusters*) with a side length of 150 metres (figure 1). The number of plots per cluster can however vary between 1 and 4 depending on forest/non-forest decisions on the plot level (Bundesministerium für Ernährung, 2011). In the field survey of the BWI3, sample trees for timber volume estimations are selected according to the angle count sampling technique (Bitterlich, 1984), using a basal area factor (*BAF*) of 4 that is respectively adjusted for boundary effects at the forest border (Bundesministerium für Ernährung, 2011). A further selection criterion for a tree to be recorded is a diameter at breast height (*dbh*) of at least 7 cm. This sampling technique was applied to 8092 sample plots (2810 clusters) in RLP, resulting in the collection of 56561 sample trees for which the dbh, the tree diameter at 7 m (*D7*) and the tree species were recorded for all trees. Tree height measurements were conducted only for a subset of all sample trees and used to predict the height for the remaining sample. During the last inventory, all plot center positions were remeasured with differential global positioning system (DGPS) technique. Knowledge about the exact plot positions were considered crucial to provide optimal comparability between the terrestrial observations and the information derived from the auxiliary data. A detailed analysis by Lamprecht et al (2017) indicated that horizontal DGPS errors do not exceed 8 meters for 80% of all plots in RLP. For 162 plots, the DGPS coordinates were replaced by their former target coordinates due to missing or implausible values. In order to derive a volume estimation for each sample tree, the BWI3 estimates a taper curve for each sample tree by calibrating the random effects term of linear mixed-effects taper models with the set of diameters and corresponding height measurements taken from the respective sample tree (Kublin et al, 2013). The integration of the derived taper curves consequently lead to a volume prediction for each sample tree. Since the overall objective of the study was to subsequently use the identified regression model for design-based timber volume estimations of state and communal forest management units, we already restricted the sample plots used for modeling to the state and communal forest area (73% of the entire forest area of RLP). This provides the advantage that when the regression model is used as an *internal model* in design-based estimators, the model predictions hold the assumption on the residuals to be zero on average over the state and communal forest area by construction of OLS technique (Mandallaz, 2013a,b; Mandallaz et al, 2013). The dataset of this study hence comprised 5791 plots (2055 clusters). For this sample, the timber volume density per hectare on plot level, $Y(x)$, was calculated according to the formula of one-phase one-stage sampling (Mandallaz, 2008, Ch. 4.2). The timber volume density per hectare on plot level was used as the response variable in the regression analysis.

Table 1: Descriptive statistics of the forest observed on NFI sample plots located iwithin communal and state forest area (n=5791).

Variable	Mean	SD	Maximum
Timber Volume (m^3/ha)	300.86	195.55	1375.31
Mean DBH (mm)	354.90	137.22	1123.20
Mean height (dm)	239.60	72.43	497.43
Mean stem density per hectare	101.00	114.01	1010.31

2.3 Auxiliary Data

2.3.1 ALS Canopy Height Model

Between 2003 and 2013, the topographic survey institution of RLP acquired airborne laser scanning (ALS) data over the entire state of RLP at leaf-off condition (Figure 2). The objective of this campaign was to derive a countrywide digital terrain and surface model based on the acquired ALS point clouds. During the extended acquisition period, airborne laser scanning technology and data quality evolved significantly. The tiles recorded in 2002 and 2003 have a rather poor quality with about only 0.04 points per m^2 , while more recently acquired datasets contained about 5 points per m^2 . The data was delivered as two separate datasets comprising the Vegetation First Pulse (VEF) and Ground (GRD) points. All point clouds were stored as three-column (easting, northing, and height above sea level) ASCII files in tiles of 1 km^2. In order to create a surface model (DSM) in a given raster resolution, the highest point of the combined VEF and GRD dataset was identified in each raster cell and saved as a thinned surface point cloud. For the elevation model (DEM), the mean of all GRD points in the cell was calculated, and the result was saved as a thinned ground point cloud. The thinned point clouds were then aggregated to larger tiles and interpolated to raster images using a Delauney interpolation in the Matlab software (Mathworks, 2017). The resulting DSM and DEM raster sets were then subtracted from each other to calculate a canopy height model (CHM) in raster format, providing discrete information about the canopy surface height of the entire forest area of RLP in a spatial resolution of 5 meters. The thinning process led to much smaller datasets that could be processed in larger tiles and considerably lowered processing times compared to the original dense point clouds. Since the data was recorded in leaf-off condition, the original point clouds contained many returns from within the crowns of deciduous trees. The thinned dataset provided the advantage that those measurements did not skew the vegetation height estimate in the final CHM.

As explanatory variables, the mean canopy height (*meanheight*) and the standard deviation (*stddev*) were calculated as the mean and standard deviation of all raster values within a predefined circle (i.e. *support* of the explanatory variable, see section 2.4) around each sample plot center. In order to correct for edge effects at the forest border, each support area was previously intersected with the state and communal forest area, which was defined by a polygon mask provided by the forest service (figure 3b). Restricting the support area and thus the evaluation of the auxiliary data to the forest area is a means to optimize the coherence between explanatory variables computed at the forest boundary and the corresponding terrestrial response variable

(Mandallaz et al, 2013). The tree height is one prominent predictor variable in the taper functions of the BWI3 that are used to calculate a timber volume value for each sample tree (Kublin, 2003; Kublin et al, 2013). A visual inspection of the tree volumes of all sample trees collected in the BWI3 within RLP against their tree heights also revealed the characteristic shape of an allometric relationship between these variables (Online Resource 1). It was hypothesized that this relationship on single-tree level is also apparent on the aggregated level of a sample plot and cluster, and can be used within the frame of regression modeling.

The strength of correlation between *meanheight* and timber volume on plot level was expected to show high variation according to the mentioned time-lag up to 10 years between ALS acquisition and terrestrial survey. The quality of the height information was also expected to vary according to changing sensor technologies and different point densities used over the years. For these reasons, the ALS acquisition year (*ALSyear*) for each sample plot was considered as a potential categorical explanatory variable to explain the variation in the data introduced by these factors. For this purpose, the acquisition year *2008* was further divided into *2008* and *2008_1*. In the latter, the data quality turned out to be very poor due to sensor failures during the acquisition. Additionally, the years *2006* and *2007* as well as *2012* and *2013* were pooled in order to increase the number of observations per factor level for modelling reasons. As a result, the *ALSyear* variable comprised nine categories (*2002*, *2003*, *2007*, *2008*, *2008_1*, *2009*, *2010*, *2011* and *2012*).

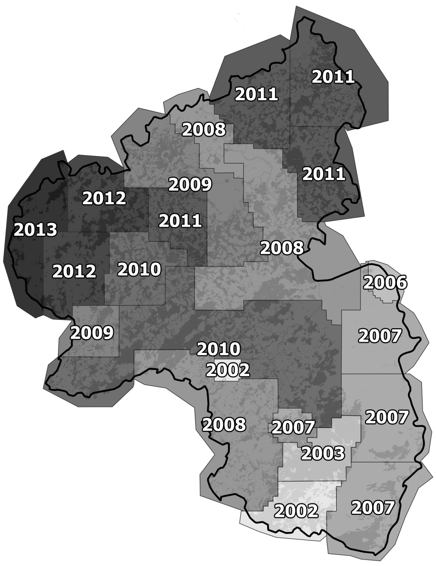


Fig. 2: Separate ALS acquisitions in Rhineland-Palatinate over the years. The colors also indicate the quality of the data: *light*: low point densities ($0.04/m^2$), *dark*: high point densities ($>4/m^2$). Blue semi-transparent layer: state and communal forest area.

2.3.2 Tree Species Classification Map

A countrywide satellite-based classification map of the five main tree species (European beech, Sessile and Pedunculate oak, Norway spruce, Douglas fir, Scots pine) described in Stoffels et al (2015) was used to derive tree species information on sample plot level. The classified tree species map has a grid size of 5 meters and predicts five of the seven tree species that are used in the BWI3 taper functions (Kublin et al, 2013) to calculate the timber volume of a sample tree. Due to unavailable satellite data for the classification, the tree species map excluded one patch with an area of 415 km^2 in the south-west part of RLP, and two further patches with an area of 76 km^2 and 100 km^2 in the northern part (Stoffels et al, 2015). The tree species information was consequently missing for 411 (7%) of the 5791 sample locations.

Prediction of main plot tree species

A visual inspection of all BWI3 sample trees of RLP suggested that a stratification of the relation between tree height and timber volume according to these seven tree species may provide a considerable reduction in variation within the tree species groups (Online Resource 1). This led to the hypothesis that this tree species specific signal might also be apparent on sample plot and cluster level and can consequently be used to increase the accuracy of the prediction model. Based on the tree species classification map, the main tree species of each sample plot was calculated as an additional categorical explanatory variable (*treespecies*) with six categories following a similar approach as Latifi et al (2012): one of the five tree species was assigned as the main plot tree species if its proportion within the edge-corrected support around the sample location exceeded a predefined threshold. If this threshold was not reached by any of the five tree species, the respective sample plot was assigned the category 'Mixed'. We hypothesized that the choice of the threshold-value might have an influence on the resulting classification accuracy and the regression model accuracy (section 2.5). We thus investigated the application of 5 threshold settings, i.e. 0%, 50%, 60%, 80% and 100%.

Calibration

Our analyses revealed that the prediction of the main tree species for a sample plot can be subject to misclassifications (section 3.1). Errors in the explanatory variables of linear regression models can however lead to a bias of the regression coefficients in the direction of zero due to an artificial introduction of noise (Carroll et al, 2006, Ch. 3). This can cause an inflation of the residual variance and a consequent decrease of the model accuracy (Magnussen et al, 2010). In case of classification, the impacts of misclassifications on the model properties are even harder to predict (Gustafson, 2003, Ch. 3). While errors in the explanatory variables do not affect the unbiasedness of the estimators in the design-based framework, a reduction or elimination of the classification errors could provide an improvement of the regression model accuracy and thereby potentially lead to smaller prediction and estimation errors. We therefore addressed the effect of misclassifications in the *treespecies* variable categories as well as means to correct these errors.

We transferred the concept of *regression calibration* as known from classical measurement error statistics (Carroll et al, 2006) to the problem of misclassifications in the *treespecies* variable. In regression calibration, one considers an error-prone explanatory variable W that can be measured in high quantity, whereas X constitutes the same but error-free variable whose determination is however very expensive. In order to yield a corrected or less error-prone version of W, one can define a calibration model $f_{calmod}(X,W)$ that predicts X as a function of W. After calibration on a training set, $f_{calmod}()$ can then be applied to any observed W and yields the corrected, less error-prone variable W_{calib}. Using W_{calib} instead of W in the regression model then asymptotically provides an unbiased estimate of the regression coefficients and thus corrects for the attenuation to zero.

We transferred this concept by using a random forest algorithm (Breiman, 2001) as calibration model. We considered the main tree species of the sample trees at each plot location x as the error-free variable $treespecies_{terr}$, that would also yield the highest model accuracies when used as predictor variable. The objective of the calibration model was thus to provide an improved classification accuracy of each predicted main plot tree species category with respect to $treespecies_{terr}$. The calibration model was considered to correct for potential systematic misclassifications and thus minimize the effect of misclassifications on the regression model when substituting the uncalibrated with the calibrated *treespecies* variable. The random forest algorithm is a machine learning algorithm that grows a large number of decorrelated classification trees by considering only a subset of all provided predictor variables for each split. In the case of classification, new data are thus predicted by aggregating the predictions of all trees using a majority vote. We calibrated the random forest algorithm (f_{RF}) with a set of p predictor variables that comprised the initial prediction of the main plot tree species (*treespecies*), the mean canopy height (*meanheight*) and standard deviation (*stddev*) derived from the CHM, the proportion of coniferous trees estimated from the tree species classification map ($prop.conif$) and the bioclimatic growing region (*wgb*) at the sample location (equation 1). An advantage for using those explanatory variables in the calibration model was that they also provided explanatory power in the regression model. This approach thus saved computation time and minimized data storage. The calibration model was implemented using the random forest algorithm (Liaw and Wiener, 2002) in the statistical software *R* (R Core Team, 2016). The algorithm was grown with 2000 trees, considering $\sqrt{p} \approx 3$ of the predictors for each split.

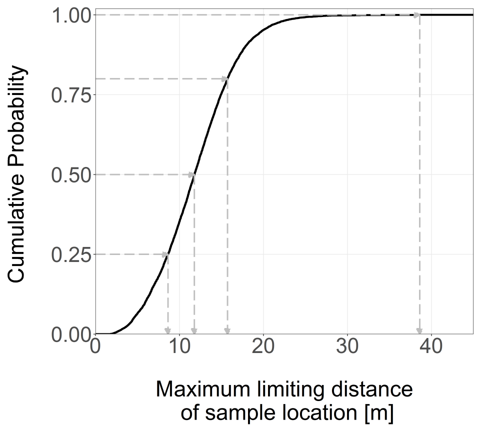
$$treespecies_{terr}(x) = f_{RF}(treespecies, meanheight, stddev, prop.conif, wgb) \tag{1}$$

The calibration model was subsequently applied to the entire dataset. We then investigated the effect on the regression model performance (regression coefficients, model accuracy) when substituting the calibrated (less error-prone) for the uncalibrated (most error-prone) variable, and likewise for the actual (error-free) main plot tree species derived from the sampled trees of the respective sample plot under identical threshold settings.

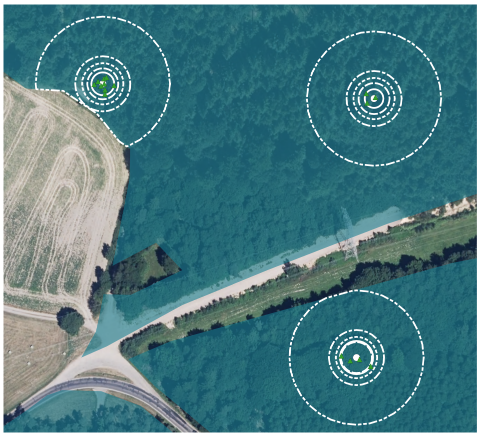
2.4 Choice of Support under Angle Count Sampling

One characteristic of angle count sampling applied in the BWI3 is that a sample plot does not have a fixed radius in which trees are selected (*fixed-radius plot*), but each tree generates an individual radius from the plot center depending on its diameter at breast height (*variable-radius plot*). This tree-individual radius is known as the *limiting distance* from the plot center where the tree would still be included in the sample. A consequence of the absence of a fixed plot radius is the question about the optimal support (Hollaus et al, 2007), i.e. the spatial extent around the plot center in which the auxiliary information is evaluated and transformed into an explanatory variable. It has widely been hypothesized that the best relationship between the target variable on the ground and any explanatory variable derived from the auxiliary information is obtained if the support is spatially identical to the sample plot extent. In case of angle count sampling, an individual extent for each sample plot can be approximated by regarding the maximum limiting distances of its sample trees as the outer plot radius. However, many design-based applications under double-sampling do not allow for a between-plot change of the support for a specific explanatory variable (Mandallaz, 2013a,b).

For this reason, the task is to find a unique support for each auxiliary information that leads to the best overall model accuracy. Deo et al (2016) conducted extensive analysis to identify optimal supports for modelling standing timber volume for *variable-radius plot* designs in conifer forests. They analysed 24 different radii (i.e. circular supports) in which they extracted 57 metrics from a ALS derived point cloud with an average point density of 18 pulses per square meter. They successively evaluated the prediction performance of each support size by using the ALS metrics in a random forest algorithm and comparing the resulting model accuracies. In order to identify the best-performing supports for our explanatory variables, we followed a similar approach. The explanatory variables were calculated using *individual* (i.e. plot-varying) supports (*ind*), i.e. an individual support radius was used for each plot according to the maximum limiting distance of all sample trees associated to the respective sample plot. We then compared the model accuracies achieved by the individual supports against the model accuracies from a set of *fixed* (i.e. non plot-varying) supports. The extents of the fixed supports were chosen from the cumulative distribution function (ECDF) of the maximum limiting distances of all 5791 sample plots of the analysed forest area (Fig. 3a). We considered the 25^{th} ($q25$, 9 meters), 50^{th} ($q50$, 12 meters), 80^{th} ($q80$, 15 meters) and the 100^{th} ($q100$, 38 meters) percentiles, resulting in support radii of 18, 24, 30 and 76 meters (Fig. 3). While in this study we also used circular supports to extract the auxiliary information, also other support-shapes are possible (e.g. rectangles, hexagons). We also want to emphasize that the use of different support sizes for each explanatory variable is perfectly valid in the infinite population framework of design-based estimators (Mandallaz, 2013a,b).



(a) ECDF of maximum limiting distances of all BWI3 sample locations in RLP



(b) Circular supports used to extract explanatory variables around sample locations. *Dash dot dot line*: q100, *dash dot line*: q80, *dot dot line*: q50, *dot line*: q25, *solid line*: individual support, *triangles*: sample trees

Fig. 3: Identification (*a*) and visualization (*b*) of potential supports used for calculating the predictor variables on plot level

2.5 Model Building and Evaluation

In order to judge the quality of the *treespecies* variable, the user's accuracy for each classified species category and the overall accuracy of the classification scheme was calculated based on the confusion matrix (Congalton and Green, 2008). As reference data, we calculated the actual main plot tree species by applying the respective threshold to the sample trees of each sample plot. The classification accuracy was evaluated for all support sizes for both the calibrated and the uncalibrated *treespecies* variables. The measures of the regression model accuracy using both CHM- and *treespecies* variables were defined as the 10-fold cross-validated root mean square error ($RMSE_{cv}$, equation 2) and the adjusted coefficient of determination (adjusted R^2) of the multiple linear regression model defined in equation 3. Additionally, we considered the interaction terms *meanheight*:*treespecies*, *meanheight*^{2}:*treespecies*, *meanheight*:*ALSyear*, *stddev*:*ALSyear* and *meanheight*:*stddev* and performed a variable selection based on the Akaike Information Criterion (AIC) (Akaike, 2011) in order to minimize the number of variables in the model. Due to a pronounced unbalanced design in the *treespecies-ALSyear* strata (Online Resource 2), no interaction between *treespecies* and *ALSyear* was possible. We evaluated the model for all support combinations, considering the use of individual support sizes for each auxiliary information, using both the calibrated and the uncalibrated *treespecies* variable. The calibration model (section 2.3.2) for the *treespecies* variable was recalculated for each respective support-threshold setting.

206 sample plots included no sample trees and the timber volume density $Y(x)$ was thus set to *zero*. These *zero-plots* were removed from the modeling dataset since they acted as leverage points in cases where the ALS height metrics were recorded long before the terrestrial survey. Together with the missing tree species information (section 2.3.2), the modeling dataset s was limited to n=5171 observations.

$$RMSE = \sqrt{\frac{\sum_{x \in s} \left(\hat{Y}(x) - Y(x)\right)^2}{n}} \tag{2a}$$

$$RMSE\% = \frac{RMSE}{\frac{1}{n}\sum_{x \in s} Y(x)} \tag{2b}$$

$$Y(x) = \beta_0 + \beta_1 * meanheight + \beta_2 * meanheight^2 + \\ \beta_3 * stddev + \\ \beta_4 * ALSyear_1 + \ldots + \beta_{12} * ALSyear_9 + \\ \beta_{13} * treespecies_1 + \ldots + \beta_{18} * treespecies_6 + \varepsilon(x) \quad (3)$$

3 Results

3.1 Classification Accuracies

Effect of Support Size and Threshold

The lowest user's accuracies (*UA*) for the uncalibrated tree species variable were mostly realized using high thresholds of 80% and 100% (figure 4). A plausible reason for this is that raising the threshold to higher values (e.g. 80%, 100%) distinctively increases the probability of the reference class (based on the sample trees of the sample location) to be assigned as class 'Mixed', while the much coarser spatial resolution of the tree species map causes the *predicted* class to remain classified as one of the five tree species. However, as the support size is increased, so does the number of tree species raster cells to be evaluated at the sample location, thereby increasing the probability that the predicted class will be 'Mixed'. For this reason, most tree species exhibit an increase in user's accuracy under higher thresholds with higher support sizes. This scale-threshold dependency of the user's accuracy particularly affects tree species that most commonly occur in mixed forest stands in Rhineland-Palatinate (*Scots pine*, *oak* and *beech*), whereas the user's accuracies for tree species that are mostly prominent in pure forest stands (*spruce*, *Douglas fir*) logically turned out to be much more robust to changes in the thresholds and support sizes.

Among the uncalibrated tree species predictions, *beech* and *spruce* produced the best predictions achieving UAs of up to 70% and 80%. Although the predictions for *Douglas fir* and *Scots pine* generally performed less well than *beech* and *spruce*, similar UAs can be produced by adjusting the threshold and support choices. UAs for *oak* never performed better than 50%. A detailed table of the user's and overall accuracies is provided in Online Resource 3.

Calibration

Calibration substantially diminished the effect of the scale-threshold dependency for the five tree species and also increased the UAs for *Scots pine* and *oak*. Whereas the UAs for *beech* and *spruce* were found to be slightly lower after calibration, the overall accuracy under each support choice was always considerably increased by calibrating the tree

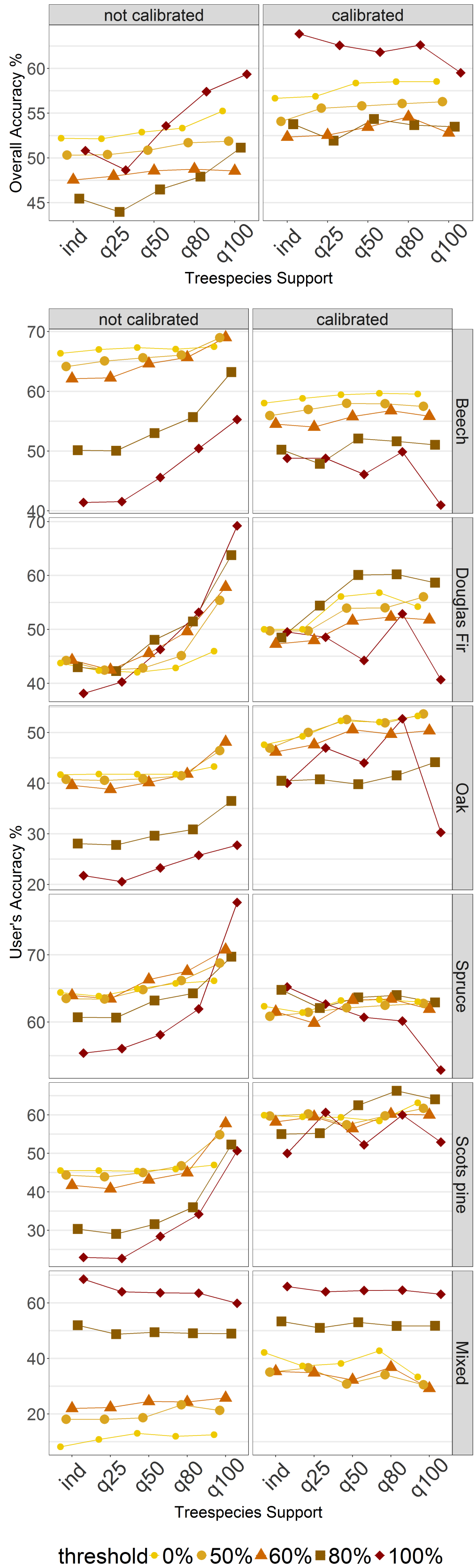


Fig. 4: Classification accuracy for the main tree species of a sample location *before* and *after* calibration: *top)* overall accuracies. *bottom)* user's accuracies. *ind*: plot individual support sizes.

species prediction (figure 4). With respect to the calculated random forest models, the initial tree species prediction (*treespecies*) and the information about the growing region (*wgb*) turned out to be the most valuable information, followed by the estimated proportion of coniferous trees (*prop.conif*) and the mean canopy height (*meanheight*).

3.2 Regression Model Accuracies

Effect of Support Size and Threshold

Figure 5 shows the accuracies of the regression model (equation 3) achieved under all possible combinations of support sizes for the auxiliary data. The stepwise selection procedure always included all considered single and interaction terms. In terms of adjusted R^2 and $RMSE_{cv}$, the analysis revealed that the choice of the CHM support size controls the overall level of the model's accuracy. The information about the main plot tree species can then be used to further improve the model fit under suitable *treespecies* support and threshold settings. When using the uncalibrated *treespecies* variable, an increase of the *treespecies* support size causes an increase in the model performance if low thresholds are used, whereas high thresholds (80%, 100%) cause a decrease in the model performance. This threshold-dependency could be removed by calibrating the *treespecies* variable. The highest adjusted R^2 and the lowest $RMSE_{cv}$ were realized using the $q50$ support for both the CHM and calibrated *treespecies* variables in combination with a *treespecies* threshold of 100%, resulting in (adjusted R^2 of 0.48 and $RMSE_{cv}$ of 136.62 m^2/ha (43.8%). However, various support and threshold combinations for the CHM and *treespecies* variables can be used to yield almost identical $RMSE_{cv}$ and adjusted R^2 values. A detailed table of the model accuracies is given in Online Resource 4.

Effect of Misclassifications

We accessed the magnitude of the misclassification effect for all models that were analysed in section 3.2, i.e. for all possible support and threshold combinations for the CHM and *treespecies* predictor variables. We first compared the adjusted R^2 of each model when using the uncalibrated *treespecies* variable against the adjusted R^2 using the actual, i.e. error-free variable. We then did the same comparison for the model using the calibrated treespecies predictor variable. Figure 6 provides a visualization of this comparison. Note that only the model with the predicted tree species variables can be applied to additional sample locations where no terrestrial survey has been carried out.

As expected, the highest adjusted R^2 for every evaluated model was always achieved using the error-free tree species variable, whereas the missclassifications in the tree species variable led to a systematic decrease of the model accuracy. The calibration of the initially predicted main plot tree species using the random forest classification algorithm (section 2.3.2) turned out to not only improve the classification accuracies (section 3.1), but also to considerably decrease the effect of the missclassifications on the regression model predictions and accuracy. Figure 6 (*right*) shows that the adjusted R^2 under the actual and the calibrated predicted tree species variable are in general much closer to, and in many cases even on the identity line. The differentiation into two distinct point clouds results from the poor model performance under support size $q100$ for the CHM variables (i.e. the lower point cloud). Whereas the misclassifications in the uncalibrated *treespecies* variable led to a residual inflation of 0.01 - 0.05 in adjusted R^2, it was only between 0 and 0.01 after calibration. Further analysis revealed that when using the calibrated *treespecies* variable, the regression coefficients were almost identical to the ones received using the actual main plot tree species.

3.3 Final Regression Model

In order to address research questions 1 and 2 (i.e. the gain in model accuracy by tree species information and effect of heterogeneity in the ALS data), we investigated the model properties in more detail. For this purpose, we decided to use the best found model that was achieved under the support settings of $q50$ for both auxiliary data with a threshold of 100% for the tree species variable as the regression model of choice. The reason for inspecting this model was that *a)* the model provided the highest adjusted R^2 among all validated models while reducing the data handling complexity for upcoming applications (i.e. identical support sizes for all remote sensing data) and *b)* the calibration neutralized the effects of misclassifications on the model predictions. The interaction term between *meanheight*^{2} and *treespecies* (i.e. considering separate curvatures for each tree species) turned out not to have a significant influence on the model accuracy and was dropped, resulting in an adjusted R^2 of 0.48 and a slightly increased $RMSE_{cv}$ of 140.62 m^2/ha (46.7%). The final model thus comprised 39 parameters (regression coefficients), i.e. the intercept, 3 main effects for continuous variables, 13 main effects for categorical variables and 22 interaction parameters (table 3).

We also conducted an analysis for detecting influential data points or outliers for the final regression model. We here considered the commonly applied criteria of leverages and Cook's Distance as amongst others described in Fahrmeir et al (2013, p. 160-167). The critical threshold of $2p/n$ (i.e. twice the average of the hat matrix' diagonal entries) was exceeded by 10% of the observations. However, only 3% of these leverage points were assigned to studentized residuals with values > 1 or < -1. Removing these observations

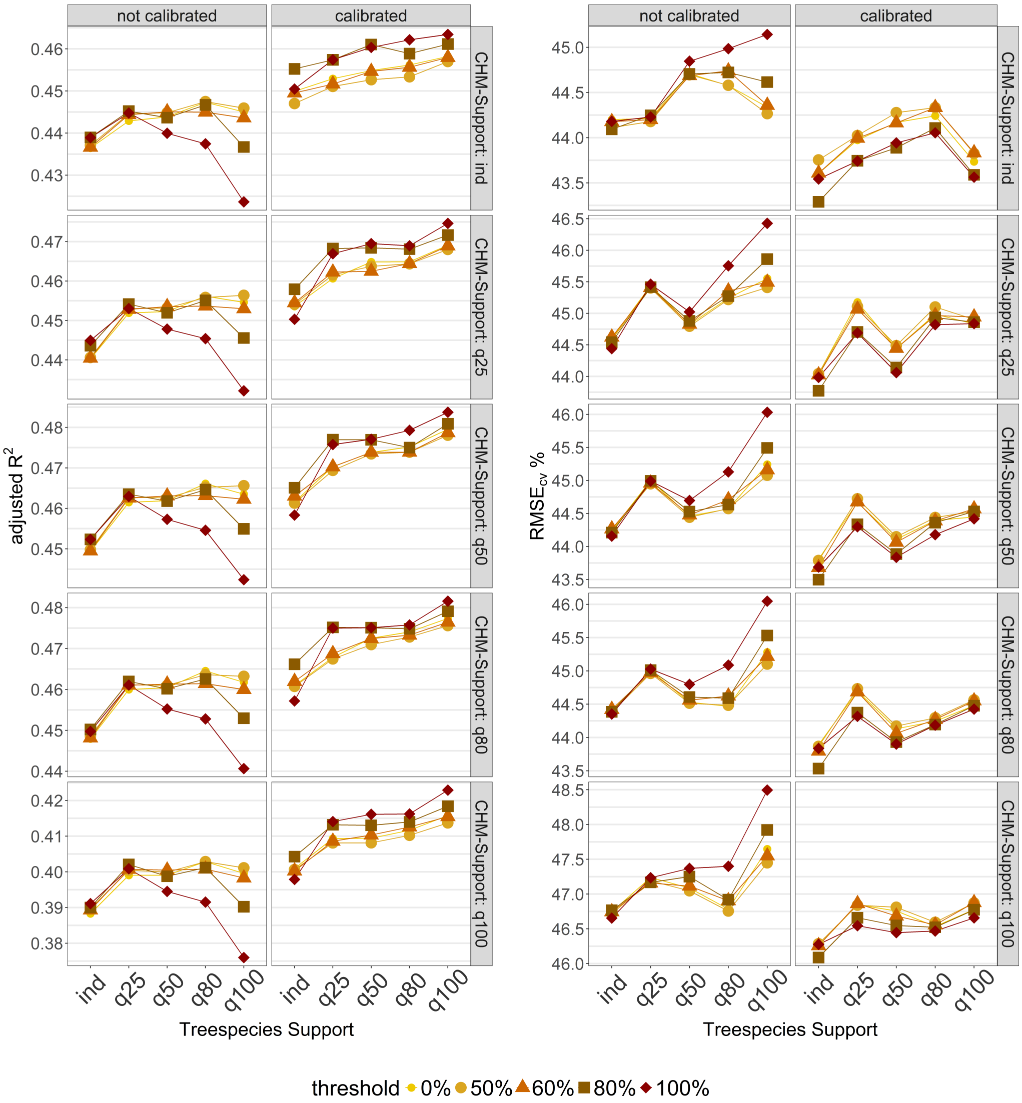


Fig. 5: 10-fold RMSE$_{cv}$[%] and adjusted R^2 realized under various support choices for the CHM and *treespecies* explanatory variables

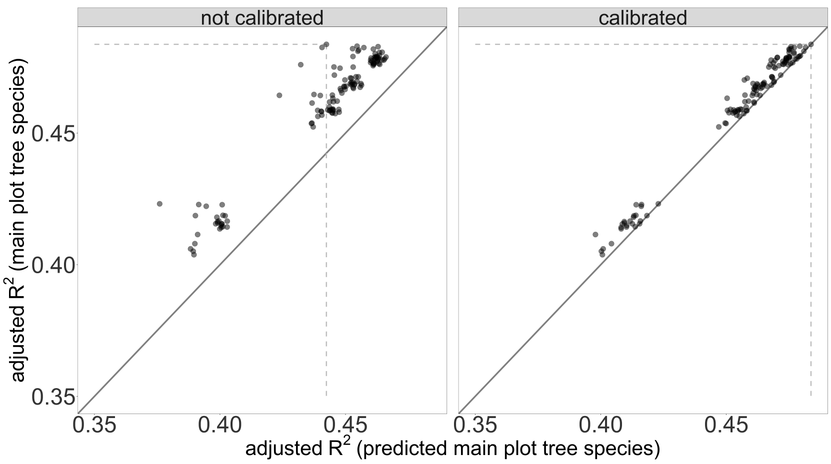


Fig. 6: Effect on the adjusted R^2 when substituting the actual main tree species with the predicted main tree species of a sample plot. The *dotted* line tracks the the model with the highest adjusted R^2 under the use of the error-free *treespecies* variable. Semitransparent colours for the data points are used to visualize overlap.

from the dataset and refitting the model led to an adjusted R^2 of 0.49 compared to 0.48 when including them. Additionally, Cook's Distance values D_i did not exceed a value of 0.019, and were thus far apart from the commonly used critical threshold of $D_i > 0.5$ that indicate a considerably change of the regression model results when omitting them. We thus decided not to remove any observations from the modelling dataset. We thus decided not to remove any observations from the modelling dataset.

Interpretation of Final Regression Model

Figure 7 provides a visualisation of the tree species prediction functions separated by the ALS acquisition years. Sample plots classified as *oak* and *Scots pine* revealed to have an almost identical relationship (nearly identical slopes) for the mean canopy height - timber volume relationship. They only differ by a marginally higher intercept for *Scots pine* plots, meaning that given the same mean canopy height a sample plot dominated by *Scots pine* yields a marginally higher timber volume on the plot level than a plot dominated by *oak*. *Beech*-dominated sample plots tend to achieve a higher timber volume than *oak* and *Scots pine* for canopy heights below 20 meters, but realize the lowest timber volumes for canopy heights above 20 metres. Sample plots dominated by any of the remaining coniferous tree species (*Douglas fir*, *spruce*) revealed to have higher slopes than broadleaf classified plots. This indicates that given the same mean canopy height, sample plots dominated by *Douglas fir* and *spruce* yield higher timber volume values than broadleaf- or *Scots pine* dominated sample plots, and this difference becomes more pronounced with increasing mean canopy heights. Within the group of coniferous-dominated sample plots, *spruce* turned out to have the highest slope, thereby yielding the highest timber volume values for mean canopy heights above 15 meters. An undesired characteristic of the model is that the predicted timber volume can in some cases ($< 1\%$) take negative values for low canopy heights (e.g. for *spruce*-dominated plots with *meanheight* below 5 meters and *stddev* of 4 meters). However, we chose not to use a log-transformation of the response variable. Doing so would have prevented the subsequent calculation of the g-weight variance of the design-based estimators (Mandallaz, 2013a; Mandallaz et al, 2013), which is only possible for response variables on the original scale. The g-weight variance provides the benefit of a better variance estimate for internal models by considering the dependency of the regression coefficients on the realized sample. The rare occurrence of negative predictions were however not considered to have an influence on subsequent design-based estimates when averaging multiple predictions within given spatial domains.

Effect of Time-Lags and Heterogeneity in ALS Data

Incorporating the ALS acquisition year as a categorical variable (*ALSyear*) in the regression model substantially accounted for the variability in the data introduced by *a)* the time-lags between ALS acquisition and terrestrial survey, and *b)* variation in ALS data quality which are due to sensor- and post processing techniques (table 3). Whereas the adjusted R^2 for the regression model without considering the ALS acquisition year as additional predictor variable (*submodel 1*) was 0.36, it could already been increased to 0.40 by including the tree species variable (*submodel 2*). A further stratification by the ALS acquisition year increased the adjusted R^2 of *submodel 1* from 0.36 to 0.45, and the adjusted R^2 of *submodel 3* from 0.40 to 0.48.

Table 2: R^2, RMSE and RMSE% of final regression model within ALS acquisition year strata (*ALSyear*). $Area_{ALSyear}$: Area covered by ALS acquisition given in km^2. *n*: number of validation data.

ALSyear	$Area_{ALSyear}$	R^2	RMSE	RMSE%	n
2012	2807	0.61	135.84	44.87	408
2011	4361	0.57	146.21	48.29	883
2010	4182	0.51	120.90	39.93	1171
2009	2100	0.42	133.42	44.07	559
2008	2968	0.48	130.38	43.06	701
2008_1	2116	0.33	175.43	57.94	394
2007	3498	0.46	136.47	45.08	418
2003	602	0.27	154.48	51.02	529
2002	775	0.44	141.55	46.75	314

We further analysed the model residuals within each ALS acquisition year (within-group variation) for the final model and nested submodels. It turned out that the R^2 values vary distinctly between the ALS acquisition year strata (table 2). More precisely, the within-group R^2 can be higher and lower than the overall R^2 of the respective model. Figure 8 shows that a stratification according to the ALS acquisition years (submodel 2) can already increase the R^2 in most acquisition year strata, compared to the basic model using only the ALS height metrics as predictor variables (submodel 1). In the ALS acquisition year stratum 2007, the increase in R^2 even reached 0.08.

Added Value of Tree Species Map Information

Introducing the predicted main tree species of a sample plot as an additional categorical variable to submodel 2 yielded a further increase in the adjusted R^2 of 0.03 (table 3). However, the improvement was even more pronounced in ALS acquisition years close or identical to the year of the terrestrial inventory (figure 8). We observed an increase of 0.06

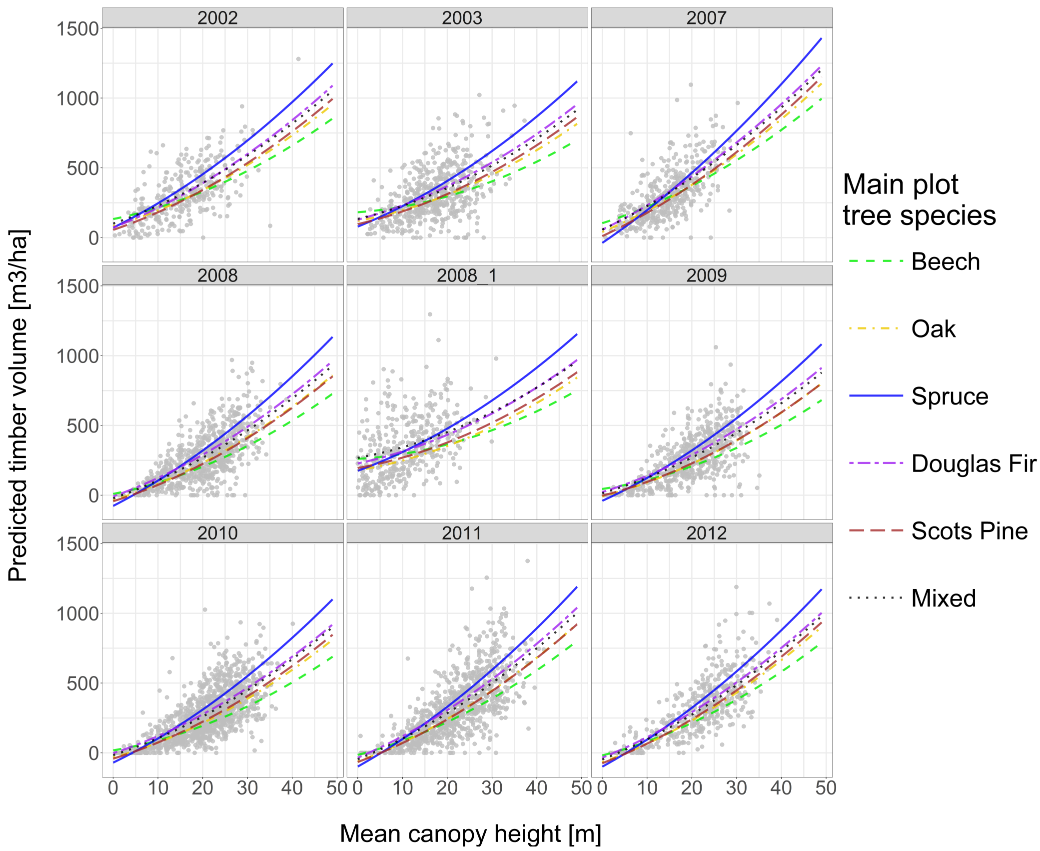


Fig. 7: Visualization of the timber volume prediction function (*final regression model*) on sample plot level for each main plot tree species and ALS acquisition year. For visualization purposes, the predictor variable *stddev* was set to its average value within the respective *treespecies* and *ALSyear* categories. The terrestrially observed timber volume values are plotted in the background.

Table 3: Accuracy metrics for submodels of final OLS regression model. Interaction terms are indicated by ':'.

model terms	model	parameters	R^2_{adj}	$RMSE_{cv}$	$RMSE_{cv}\%$
meanheight + stddev + meanheight^2 + treespecies + ALSyear + meanheight:treespecies + meanheight:ALSyear + meanheight:stddev + stddev:ALSyear	final model	39	0.48	140.62	46.69
meanheight + stddev + meanheight^2 + meanheight:stddev	submodel 1	5	0.36	155.54	51.65
meanheight + stddev + meanheight^2 + ALSyear + meanheight:ALSyear + meanheight:stddev + stddev:ALSyear	submodel 2	29	0.45	145.62	48.35
meanheight + stddev + meanheight^2 + treespecies + meanheight:treespecies + meanheight:stddev	submodel 3	15	0.40	150.32	49.92

in R^2 for ALS acquisition year 2012, and of 0.07 for ALS acquisition year 2011. The analysis illustrated once more that misclassifications in the tree species variable generally reduce model accuracy compared to using error-free tree species information. The residual inflations caused by the misclassifications in the uncalibrated *treespecies* variable within the *ALSyear* strata were up to 0.05 in R^2. However, the calibration was able to substantially decrease or even remove the effects of misclassifications on the model accuracy in all ALS acquisition year strata.

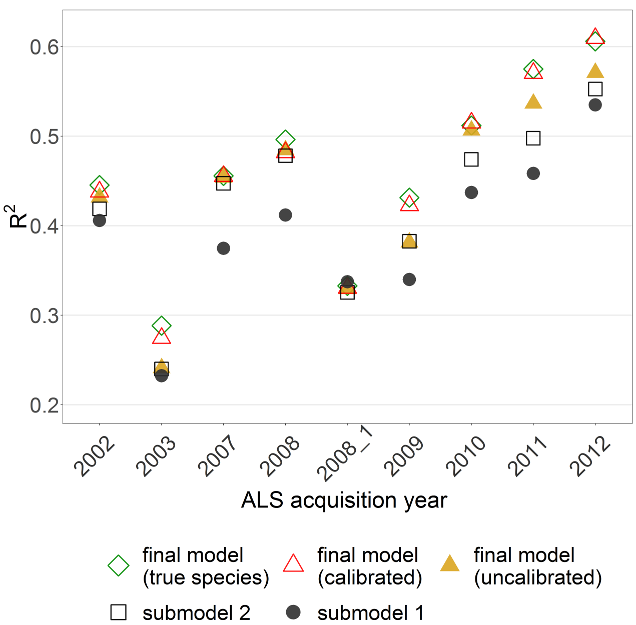


Fig. 8: R^2-values of the final regression model, submodel 1 and submodel 2 achieved *within* the ALS acquisition year strata.

4 Discussion

4.1 Stratification according to ALS Acquisition Years and Tree Species

Incorporating the main tree species of a sample location in the timber volume regression model increased the model accuracy and revealed strong evidence for the existence of a tree species specific behavior concerning timber volume on the plot level. This result seems reasonable regarding the species specific taper functions on single-tree level applied in the BWI3 (Kublin, 2003; Kublin et al, 2013). These findings also agree with those of Latifi et al (2012) who found an almost identical improvement in RMSE of 2% when stratifying to broadleaf and coniferous tree species. The overall RMSE of their model was however 10% smaller than in our study. This might be due to a more heterogeneous dataset of much smaller sample size, but also because the temporal alignment between the auxiliary data acquisition and the terrestrial survey was much better. Additionally, the number of different tree species present in their dataset was lower than in our case and only comprised Scots pine, European beech and oak. The individual effects of spruce and Douglas fir indicated by our model also support the findings of Breidenbach et al (2008), who found a higher percentage of coniferous trees in a sample plot to increase the timber volume predictions. This was not true for Scots pine and oak whose effects turned out to be very similar for our dataset. However, in our study the stratification according to the ALS acquisition years severely limited the flexibility of species-specific prediction functions and model interpretability. In particular, using the ALS acquisition years as categorical variables led to highly unbalanced datasets when stratifying according to the main plot tree species. This prevented the use of further stratification variables such as bioclimatic growing regions due to confounding effects and consequent singularities in the design matrices. It also implied an artificial increase in the number of parameters in the OLS regression model, which was however not regarded as critical with respect to overfitting issues due to the high amount of observations used for fitting the regression coefficients (Draper and Smith, 2014, Ch. 15.1). A stratification to the ALS acquisition years however proved to be an effective means in accounting for the artificially introduced noise in the data caused by quality variations and the large time-lags between the remote sensing and terrestrial data. It allowed for a model accuracy that was very close to those reported by Maack et al (2016) who conducted a very similar study in the German federal state of Baden-Württemberg. Model accuracies were also particularly higher in ALS acquisition year strata in which the data showed considerably less noise or were closer to the date of the terrestrial survey. This effect was significantly reduced or even removed when merging several ALS acquisition year strata. Promising steps with respect to more up-to-date canopy height information have already been made, as the topographic survey institution of RLP is currently processing a canopy height model from aerial imagery acquisitions for 2011 and 2012 covering the entire federal state. These aerial photography acquisitions will in the future be conducted in a two-year period, allowing to derive up-to-date canopy height information in the framework of future forest inventories. For a smaller study area, Kirchhoefer et al (2017) have already demonstrated that similar model accuracies for German NFI data can be achieved using imagery-based canopy height models.

Incorporating the calibrated tree species information further improved the model accuracy by 0.03 in adjusted R^2. Compared to the simple model only containing ALS height metrics, including the ALS quality and calibrated tree species information increased the adjusted R^2 by 0.12 in total. A differentiated evaluation of the final regression model revealed that the highest R^2-values were achieved within ALS acquisitions year strata close or identical with the year of the terrestrial survey, showing differences of up to 0.3 between the R^2s. Also the gain in R^2 by including the tree species information was largest (i.e. 0.07) in combination with ALS information acquired in the year of the terrestrial inventory. These insights were particularly interesting with respect to the further use of the regression model for small area estimations. Small area estimators generally gain modeling strength by defining the prediction model *globally* (i.e. using all data in the inventory area), and then applying the so-derived prediction model to a subset of observations located within the area of interest (Mandallaz, 2013a). Consequently, the proposed stratification technique in the prediction model is expected to yield a gain in model accuracy and a reduction of the small area estimation errors if the small area domain mostly includes data from strata that have high within-strata model accuracies. Findings of Breidenbach et al (2008) indicated that a further increase in model accuracies could possibly be achieved when incorporating these categorical variables as random rather than fixed-effects in linear mixed-effects models (Pinheiro and Bates, 2000). The reason we did not apply this family of models was that small area regression estimators subsequently applied in RLP (Mandallaz, 2013a; Mandallaz et al, 2013) require the internal models to be fitted by OLS technique.

4.2 Calibration of Tree Species Map Information

The accuracy assessment of the initially derived main plot species from the classification map revealed the presence of misclassifications that led to a decrease in model accuracy. This is in agreement with the potential effects of erroneous explanatory variables discussed in Carroll et al (2006) and Gustafson (2003), i.e. an increase of variability (noise) in the data that can increase the amount of unexplainable variance and thereby reduce the model accuracy. One reason for the misclassifications were that the classification algorithm of Stoffels et al (2015) was exclusively trained in pure stands with the objective to predict the *dominant tree species* of a forest stand. Thus, our requirements on the classification map differed considerably from the ones imposed by Stoffels et al (2015) and have to be considered as far more difficult to meet. Firstly, the reference data used in the accuracy assessment also included understory trees that were recorded in the BWI3 sample. Secondly, determining an exact spatial validation unit for a sample location (support) is not possible due to the properties of angle count sampling (section 2.4). Thirdly, distinct discrepancies in the spatial scale between the reference data and the classification map severely hamper exact predictions of the main plot tree species especially in mixed forest stands. The latter issue caused a pronounced dependency of the user's accuracy on the support and threshold choice, particularly for tree species that most commonly occur in mixed forest structures, i.e. *Scots pine* (91%), *oak* (90%) and *beech* (85%) (von Thünen-Institut, 2014). With respect to this set-up, the application of our calibration method proved to be of high value. It led to an increase in the classification accuracies, particularly for those tree species that performed worse in the uncalibrated setup, and thereby successfully minimized and even removed the deleterious effect of misclassifications on model accuracy and regression coefficients. Whereas the extensive analysis in our study deepened the understanding of the afore mentioned scale-effects, an alternative method for future applications could be to use map-derived percentages of each tree species as predictor variables in the random forest algorithm in order to directly predict the terrestrially observed main plot tree species.

4.3 Choice of Support under Angle Count Sampling

The validation of different support sizes underlined that the support choice can impact the accuracy of a prediction model, and thus confirmed the findings of Deo et al (2016). In the present study, differences in the model accuracies however turned out to be small for most support choices. An exception was the choice of the $q100$ support for the CHM derived variables (76 meter radius), where the model accuracy was considerably worse than under the optimal settings. Contrary to our hypothesis, the use of plot-individual supports did not yield the best prediction performance overall. Kirchhoefer et al (2017) recently came to the same result when they transferred the angle-count sampling technique to a pixel-wise selection method of the auxiliary data that resembles the sample tree selection even more precisely. In their study, the application of fixed support sizes did also not perform worse than under variable supports. We consider two plausible reasons for the joint findings: first, the determination of an exact spatial extent that can be transferred to auxiliary data extraction remains technically infeasible under angle count sampling. Thus, angle count sampling does not seem to be adequate when linking inventory information with remote sensing data. Secondly, inaccuracies in the DGPS-measurements of the plot center locations as reported by Lamprecht et al (2017) may have an increased impact on the model accuracy the more exact the auxiliary data derivation spatially corresponds to those of the field survey. However, the extensive analysis carried out in our study also indicated that the optimal support size does not only depend on the spatial extent of the field plots, but also on the spatial resolution of the remote sensing data as well as the context in which the derived information is used in the prediction model. In the case of transforming the tree species information map into a suitable categorical predictor variable, the use of a large support size of 76 meter radius turned out to yield the best model accuracy. However, only few sample locations in the study area were actually characterized by limiting circles of that particular size. An analysis to find the best support settings therefore seems to be advisable prior to further applications of design-based or model-dependent inventory methods so as not to lose model accuracy by unsuitable support choices. The concept of the demonstrated analysis method for identifying suitable supports can be transferred to any kind of auxiliary information, predictor variable and prediction model.

5 Conclusion

We draw three major conclusions from our study: (1) our analyses strongly indicated that the acquisition of auxiliary data close to the date of the terrestrial survey is a key factor to achieve good model accuracies. Particularly for large-scale inventory applications, this requirement is often difficult to meet. In such cases, we consider that the proposed method of including quality information about the auxiliary data in a prediction model can be an effective technique for improving the prediction accuracy. Ongoing studies investigate whether this modelling technique can also lead to smaller estimation errors of design-based estimators. (2) Our study also indicated that the relationship between the field measured timber volume and remote-sensing derived height information is tree species specific. We expect that using the tree species information in a timber volume model would even lead to higher prediction accuracies when combined with explanatory variables that can further explain the variation within each tree species group, such as bioclimatic growing conditions, soil properties and stand density on the plot level. (3) We consider the demonstrated calibration technique to be a valuable method for future studies where an external tree species map (i.e. the map was not created for the specific study objective) is used in prediction models. The application of a calibration model can also be transferred to any error-prone explanatory variable and be a simple means to clean the data set from noise and thus increase the model accuracy.

Acknowledgements We want to express our gratitude to Prof. H. Heinimann (Chair of Land Use Engineering, ETH Zurich) for supporting this study. We want to explicitly thank Dr. Johannes Stoffels from the Environmental Sensing and Geoinformatics Group of Trier University for providing the tree species classification map as well as for constructive discussions when it came to interpreting the results. Special gratitude is owed to the State Forest Service of Rhineland-Palatinate, in particular Dr. Joachim Langshausen, Jürgen Dietz and Claus-Andreas Lessander, for collaboration and providing the forest inventory and geodata. We also want to thank Kai Husmann and Christoph Fischer from the Northwest German Forest Research Institution Göttingen for their advice in processing the terrestrial inventory data, and Alexander Massey and Michael Hill for proofreading. We also want to thank two anonymous reviewers for their valuable support to improve the manuscript.

Conflict of Interest The authors declare that they have no conflict of interest.

References

Akaike H (2011) Akaike's Information Criterion, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 25–25. DOI 10.1007/978-3-642-04898-2_110, URL http://dx.doi.org/10.1007/978-3-642-04898-2_110

Bitterlich W (1984) The relascope idea. Relative measurements in forestry. Commonwealth Agricultural Bureaux

Bohlin J, Bohlin I, Jonzén J, Nilsson M (2017) Mapping forest attributes using data from stereophotogrammetry of aerial images and field data from the national forest inventory. SILVA FENNICA 51(2)

Breidenbach J, Astrup R (2012) Small area estimation of forest attributes in the norwegian national forest inventory. European Journal of Forest Research 131(4):1255–1267, DOI 10.1007/s10342-012-0596-7

Breidenbach J, Kublin E, McGaughey R, Andersen HE, Reutebuch SE (2008) Mixed-effects models for estimating stand volume by means of small footprint airborne laser scanner data. Photogrammetric Journal of Finland 21(1):4–15

Breiman L (2001) Random forests. Machine learning 45(1):5–32, DOI 10.1023/A:1010933404324

Brosofske KD, Froese RE, Falkowski MJ, Banskota A (2014) A review of methods for mapping and prediction of inventory attributes for operational forest management. Forest Science 60(4):733–756

Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM (2006) Measurement error in nonlinear models: a modern perspective. CRC press, DOI 10.1201/9781420010138, URL https://doi.org/10.1201/9781420010138

Congalton RG, Green K (2008) Assessing the accuracy of remotely sensed data: principles and practices. CRC press, DOI 10.1201/9781420055139, URL https://doi.org/10.1201/9781420055139

Deo RK, Froese RE, Falkowski MJ, Hudak AT (2016) Optimizing variable radius plot size and lidar resolution to model standing volume in conifer forests. Canadian Journal of Remote Sensing 42(5):428–442, DOI 10.1080/07038992.2016.1220826, URL http://dx.doi.org/10.1080/07038992.2016.1220826, http://dx.doi.org/10.1080/07038992.2016.1220826

Draper NR, Smith H (2014) Applied regression analysis. John Wiley & Sons

Bundesministerium für Ernährung LuV (2011) Aufnahmeanweisung für die dritte Bundeswaldinventur BWI3 (2011 - 2012). URL https://www.bundeswaldinventur.de/index.php?id=421

Fahrmeir L, Kneib T, Lang S, Marx B (2013) Regression: models, methods and applications. Springer Science & Business Media, DOI 10.1007/978-3-642-34333-9, URL http://www.springer.com/us/book/

9783642343322

Gauer J, Aldinger E (2005) Waldökologische Naturräume Deutschlands-Wuchsgebiete. Mitteilungen des Vereins für Forstliche Standortskunde und Forstpflanzenzüchtung 43:281–288

Gregoire TG, Valentine HT (2007) Sampling strategies for natural resources and the environment. CRC Press

Gustafson P (2003) Measurement error and misclassification in statistics and epidemiology: impacts and Bayesian adjustments. CRC Press, DOI 10.1201/9780203502761, URL https://doi.org/10.1201/9780203502761

Hill A, Breschan J, Mandallaz D (2014) Accuracy assessment of timber volume maps using forest inventory data and lidar canopy height models. Forests 5(9):2253–2275, DOI 10.3390/f5092253, URL http://www.mdpi.com/1999-4907/5/9/2253

Hollaus M, Wagner W, Maier B, Schadauer K (2007) Airborne laser scanning of forest stem volume in a mountainous environment. Sensors 7(8):1559–1577, DOI 10.3390/s7081559, URL http://www.mdpi.com/1424-8220/7/8/1559

Husmann K, Rumpf S, Nagel J (2017) Biomass functions and nutrient contents of european beech, oak, sycamore maple and ash and their meaning for the biomass supply chain. Journal of Cleaner Production DOI 10.1016/j.jclepro.2017.03.019, URL https://doi.org/10.1016/j.jclepro.2017.03.019

Jakubowski MK, Guo Q, Kelly M (2013) Tradeoffs between lidar pulse density and forest measurement accuracy. Remote Sensing of Environment 130:245–253, DOI 10.1016/j.rse.2012.11.024, URL https://doi.org/10.1016/j.rse.2012.11.024

Kirchhoefer M, Schumacher J, Adler P, Kändler G (2017) Considerations towards a novel approach for integrating angle-count sampling data in remote sensing based forest inventories. Forests 8(7):239

Koch B (2010) Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment. ISPRS Journal of Photogrammetry and Remote Sensing 65(6):581–590, DOI 10.1016/j.isprsjprs.2010.09.001, URL https://doi.org/10.1016/j.isprsjprs.2010.09.001

Köhl M, Magnussen SS, Marchetti M (2006) Sampling methods, remote sensing and GIS multiresource forest inventory. Springer Science & Business Media

Kublin E (2003) Einheitliche beschreibung der schaftform–methoden und programme–bdatpro. Forstwissenschaftliches Centralblatt 122(3):183–200

Kublin E, Breidenbach J, Kändler G (2013) A flexible stem taper and volume prediction method based on mixed-effects b-spline regression. European journal of forest research 132(5-6):983–997, DOI 10.1007/s10342-013-0715-0

Lamprecht S, Hill A, Stoffels J, Udelhoven T (2017) A machine learning method for co-registration and individual tree matching of forest inventory and airborne laser scanning data. Remote Sensing 9(5), DOI 10.3390/rs9050505, URL http://www.mdpi.com/2072-4292/9/5/505

Latifi H, Nothdurft A, Koch B (2010) Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/lidar-derived predictors. Forestry 83(4):395–407, DOI 10.1093/forestry/cpq022, URL https://doi.org/10.1093/forestry/cpq022

Latifi H, Nothdurft A, Straub C, Koch B (2012) Modelling stratified forest attributes using optical/lidar features in a central european landscape. International Journal of Digital Earth 5(2):106–132, DOI 10.1080/17538947.2011.583992, URL http://dx.doi.org/10.1080/17538947.2011.583992

Liaw A, Wiener M (2002) Classification and regression by randomforest. R News 2(3):18–22, URL http://CRAN.R-project.org/doc/Rnews/

von Lüpke N, Saborowski J (2014) Combining double sampling for stratification and cluster sampling to a three-level sampling design for continuous forest inventories. European journal of forest research 133(1):89–100, DOI 10.1007/s10342-013-0743-9, URL http://dx.doi.org/10.1007/s10342-013-0743-9

Maack J, Lingenfelder M, Weinacker H, Koch B (2016) Modelling the standing timber volume of baden-württemberg—a large-scale approach using a fusion of landsat, airborne lidar and national forest inventory data. International Journal of Applied Earth Observation and Geoinformation 49:107–116, DOI 10.1016/j.jag.2016.02.004, URL https://doi.org/10.1016/j.jag.2016.02.004

Magnussen S, Næsset E, Gobakken T (2010) Reliability of lidar derived predictors of forest inventory attributes: A case study with norway spruce. Remote Sensing of Environment 114(4):700–712, DOI 10.1016/j.rse.2009.11.007, URL https://doi.org/10.1016/j.rse.2009.11.007

Magnussen S, Mandallaz D, Breidenbach J, Lanz A, Ginzler C (2014) National forest inventories in the service of small area estimation of stem volume. Canadian Journal of Forest Research 44(9):1079–1090, DOI 10.1139/cjfr-2013-0448, URL https://doi.org/10.1139/cjfr-2013-0448

Mandallaz D (2008) Sampling techniques for forest inventories. CRC Press, DOI 10.1201/9781584889779, URL https://doi.org/10.1201/9781584889779

Mandallaz D (2013a) Design-based properties of some small-area estimators in forest inventory with two-phase sampling. Canadian Journal of Forest Research 43(5):441–449, DOI 10.1139/cjfr-2012-0381, URL https://doi.org/10.1139/cjfr-2012-0381

Mandallaz D (2013b) A three-phase sampling extension of the generalized regression estimator with partially exhaustive information. Canadian Journal of Forest Research 44(4):383–388, DOI 10.1139/cjfr-2013-0449, URL https://doi.org/10.1139/cjfr-2013-0449

Mandallaz D, Breschan J, Hill A (2013) New regression estimators in forest inventories with two-phase sampling and partially exhaustive information: a design-based monte carlo approach with applications to small-area estimation. Canadian Journal of Forest Research 43(11):1023–1031, DOI 10.1139/cjfr-2013-0181, URL https://doi.org/10.1139/cjfr-2013-0181

Massey A, Mandallaz D (2015) Design-based regression estimation of net change for forest inventories. Canadian Journal of Forest Research 45(12):1775–1784, DOI 10.1139/cjfr-2015-0266, URL https://doi.org/10.1139/cjfr-2015-0266, https://doi.org/10.1139/cjfr-2015-0266

Massey A, Mandallaz D, Lanz A (2014) Integrating remote sensing and past inventory data under the new annual design of the swiss national forest inventory using three-phase design-based regression estimation. Canadian Journal of Forest Research 44(10):1177–1186, DOI 10.1139/cjfr-2014-0152, URL https://doi.org/10.1139/cjfr-2014-0152

Mathworks (2017) Matlab version 9.2.0.538062 (r2017a)

McRoberts RE, Tomppo EO, Næsset E (2010) Advances and emerging issues in national forest inventories. Scandinavian Journal of Forest Research 25(4):368–381, DOI 10.1080/02827581.2010.496739, URL http://dx.doi.org/10.1080/02827581.2010.496739

McRoberts RE, Næsset E, Gobakken T, Bollandsås OM (2015) Indirect and direct estimation of forest biomass change using forest inventory and airborne laser scanning data. Remote Sensing of Environment 164:36–42, DOI 10.1016/j.rse.2015.02.018, URL https://doi.org/10.1016/j.rse.2015.02.018

Naesset E (2014) Area-based inventory in norway – from innovation to an operational reality. In: Forest Applications of Airborne Laser Scanning - Concepts and Case Studies, Springer, chap 11, pp 216–240, DOI 10.1007/978-94-017-8663-8

Nink S, Hill J, Buddenbaum H, Stoffels J, Sachtleber T, Langshausen J (2015) Assessing the suitability of future multi-and hyperspectral satellite systems for mapping the spatial distribution of norway spruce timber volume. Remote Sensing 7(9):12,009–12,040, DOI 10.3390/rs70912009, URL http://www.mdpi.com/2072-4292/7/9/12009

Næsset E (1997) Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment 61(2):246 – 253, DOI https://doi.org/10.1016/S0034-4257(97)00041-2, URL http://www.sciencedirect.com/science/article/pii/S0034425797000412

Pinheiro J, Bates D (2000) Mixed-Effects Models in S and S-PLUS. Springer, DOI 10.1007/b98882, URL http://link.springer.com/book/10.1007%2Fb98882

R Core Team (2016) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org/

Saborowski J, Marx A, Nagel J, Böckmann T (2010) Double sampling for stratification in periodic inventories—infinite population approach. Forest ecology and management 260(10):1886–1895, DOI 10.1016/j.foreco.2010.08.035, URL https://doi.org/10.1016/j.foreco.2010.08.035

Schreuder HT, Gregoire TG, Wood GB (1993) Sampling methods for multiresource forest inventory. John Wiley & Sons

Stoffels J, Hill J, Sachtleber T, Mader S, Buddenbaum H, Stern O, Langshausen J, Dietz J, Ontrup G (2015) Satellite-based derivation of high-resolution forest information layers for operational forest management. Forests 6(6):1982–2013, DOI 10.3390/f6061982, URL http://www.mdpi.com/1999-4907/6/6/1982

Straub C, Dees M, Weinacker H, Koch B (2009) Using airborne laser scanner data and cir orthophotos to estimate the stem volume of forest stands. Photogrammetrie-Fernerkundung-Geoinformation 2009(3):277–287, DOI 10.1127/0935-1221/2009/0022, URL https://doi.org/10.1127/0935-1221/2009/0022

von Thünen-Institut (2014) Dritte Bundeswaldinventur 2012. URL https://bwi.info, accessed: 2017-02-03

Tonolli S, Dalponte M, Vescovo L, Rodeghiero M, Bruzzone L, Gianelle D (2011) Mapping and modeling forest tree volume using forest inventory and airborne laser scanning. European Journal of Forest Research 130(4):569–577, DOI 10.1007/s10342-010-0445-5, URL http://dx.doi.org/10.1007/s10342-010-0445-5

White JC, Coops NC, Wulder MA, Vastaranta M, Hilker T, Tompalski P (2016) Remote sensing technologies for enhancing forest inventories: A review. Canadian Journal of Remote Sensing 42(5):619–641, DOI 10.1080/07038992.2016.1207484, URL http://dx.doi.org/10.1080/07038992.2016.1207484

Zianis D, Muukkonen P, Mäkipää R, Mencuccini M, et al (2005) Biomass and stem volume equations for tree species in Europe. Silva Fennica Monographs 4

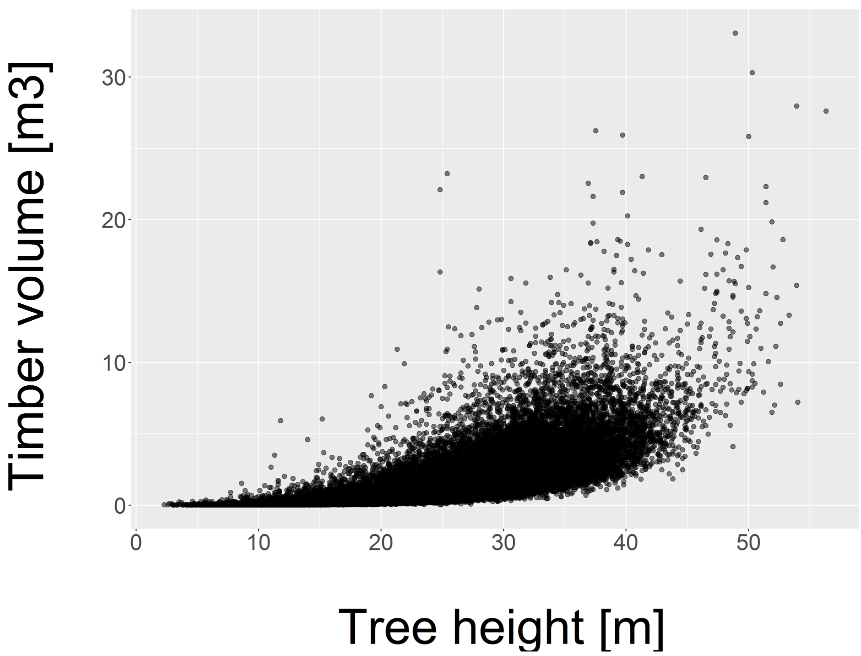
European Journal of Forest Research
Supplementary Material

Combining canopy height and tree species information for large scale timber volume estimations under strong heterogeneity of auxiliary data and variable sample plot sizes

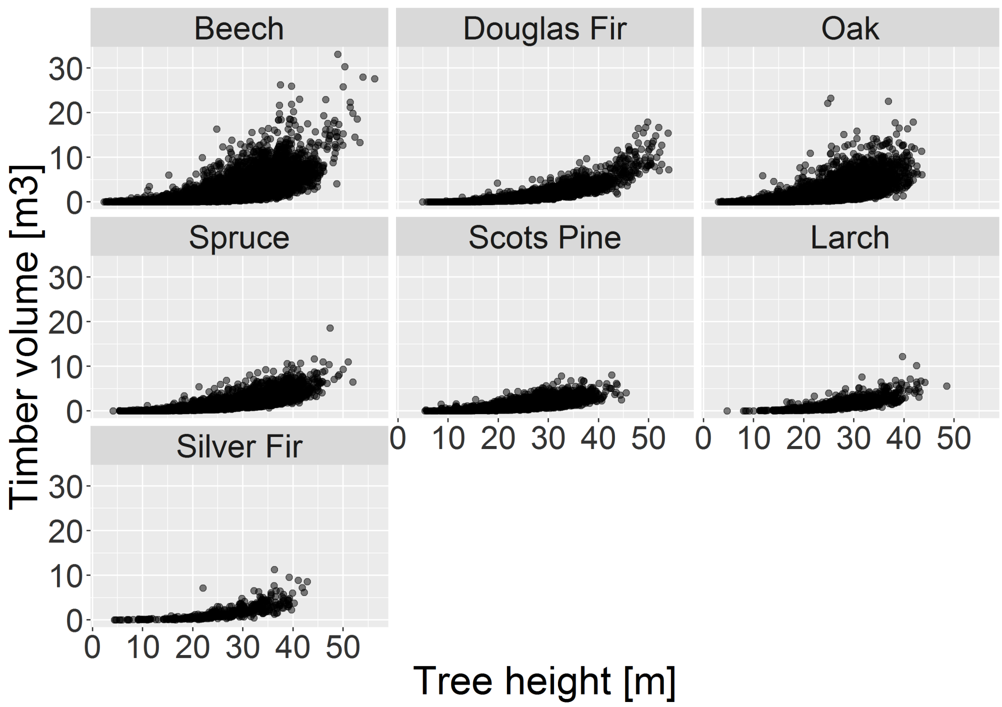
Andreas Hill*, Henning Buddenbaum , Daniel Mandallaz

*Corresponding author: andreas.hill@usys.ethz.ch

1: Timber volume - height relationship on single tree level



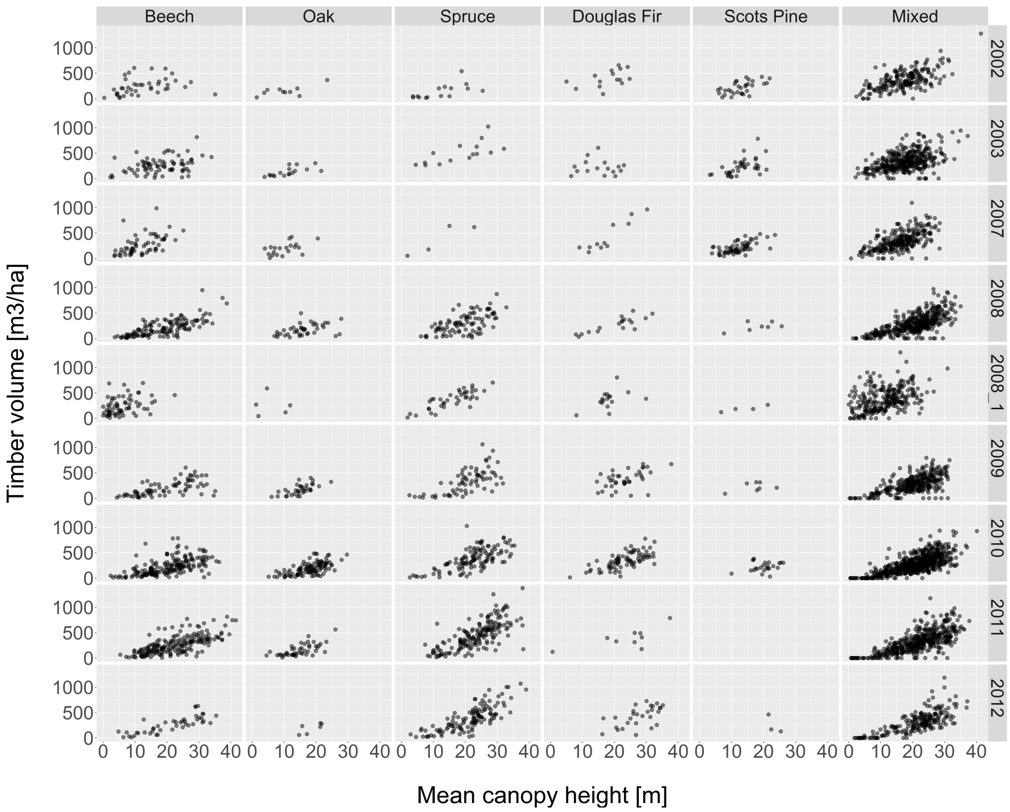
(a) Timber volume vs. tree height



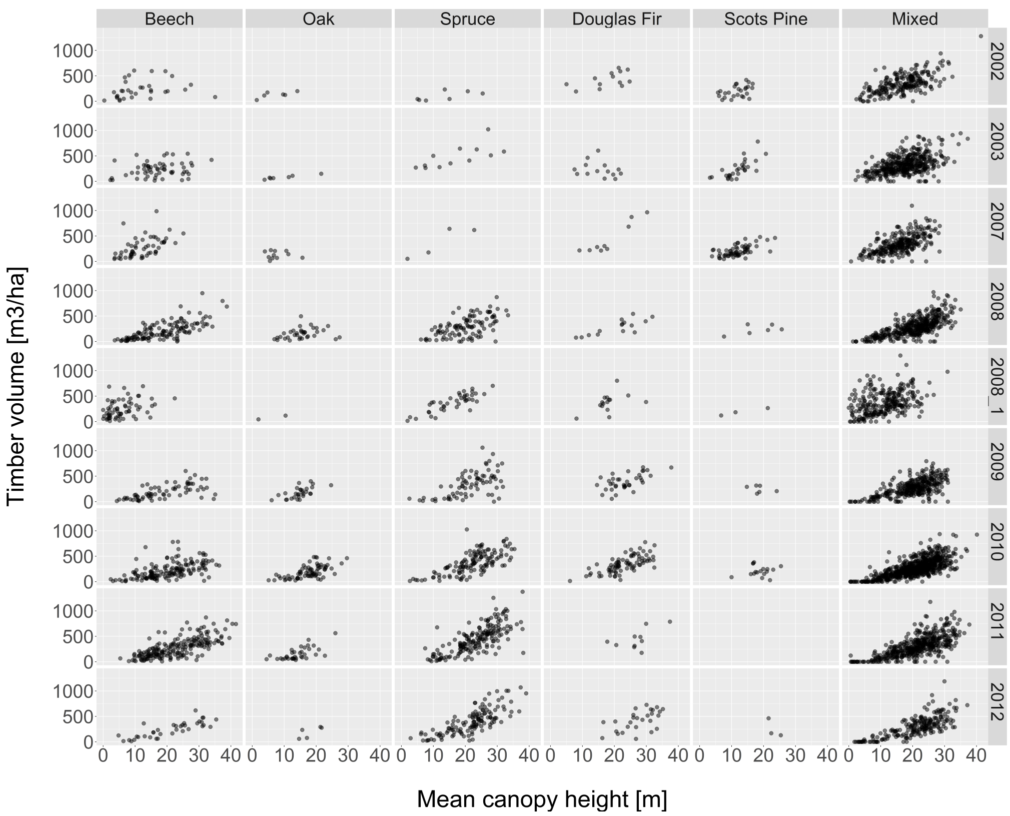
(b) Timber volume of sample trees plotted against tree height and tree species used in the BWI3 taper functions

Figure 1: Timber volume relationships on single tree level of all BWI3 sample trees within RLP

2: Timber volume on plot level vs. predictor variables



(a) Timber volume on plot level vs. LiDAR *meanheight* stratified by the error-free *treespecies* variable



(b) Timber volume on plot level vs. LiDAR *meanheight* stratified by calibrated *treespecies* variable

Figure 2: Timber volume on sample plot level stratified by the *lidaryears* and *treespecies*

3: Classification accuracies of *treespecies* variable

Table 1: User's accuracies realized under various support choices for deriving the major tree species of a sample location. *class*: major tree species class of sample plot, *prod.acc*: producer's accuarcy, *use.acc*: user's accuracy, *oaa*: overall accuracy, $prod.acc_{cal}$: producer's accuarcy after calibration (use.acc and oaa respectively), $n.ref$: number of validation data per tree species.

class	support	threshold	prod.acc	$prod.acc_{cal}$	use.acc	$use.acc_{cal}$	oaa	oaa_{cal}	n.ref
Beech	ind	0%	51.00	69.95	66.36	58.02	52.20	56.67	1857
Douglas Fir	ind	0%	54.66	44.33	43.75	50.00	52.20	56.67	397
Oak	ind	0%	65.53	39.20	41.67	47.57	52.20	56.67	824
Spruce	ind	0%	58.92	68.66	64.38	62.35	52.20	56.67	1037
Scots pine	ind	0%	67.45	62.06	45.51	59.93	52.20	56.67	593
Mixed	ind	0%	2.87	16.48	8.20	42.16	52.20	56.67	522
Beech	ind	50%	51.60	69.30	64.14	55.93	50.31	54.07	1723
Douglas Fir	ind	50%	56.45	45.97	44.21	49.71	50.31	54.07	372
Oak	ind	50%	66.24	39.95	40.70	46.90	50.31	54.07	776
Spruce	ind	50%	58.87	67.64	63.48	60.83	50.31	54.07	992
Scots pine	ind	50%	67.95	60.44	44.32	59.67	50.31	54.07	546
Mixed	ind	50%	7.67	18.51	18.05	35.02	50.31	54.07	821
Beech	ind	60%	48.56	68.79	62.12	54.52	47.55	52.33	1631
Douglas Fir	ind	60%	53.74	43.49	44.29	47.29	47.55	52.33	361
Oak	ind	60%	64.20	39.51	39.56	46.15	47.55	52.33	729
Spruce	ind	60%	57.90	68.19	63.95	61.54	47.55	52.33	962
Scots pine	ind	60%	66.53	55.92	41.69	58.17	47.55	52.33	490
Mixed	ind	60%	14.19	22.71	22.03	35.35	47.55	52.33	1057
Beech	ind	80%	44.16	37.47	50.15	50.24	45.45	53.77	1121
Douglas Fir	ind	80%	48.32	37.25	42.99	48.47	45.45	53.77	298
Oak	ind	80%	63.61	21.45	28.06	40.45	45.45	53.77	415
Spruce	ind	80%	54.70	61.17	60.70	64.78	45.45	53.77	788
Scots pine	ind	80%	65.70	35.74	30.33	55.00	45.45	53.77	277
Mixed	ind	80%	36.94	69.11	51.96	53.33	45.45	53.77	2331
Beech	ind	100%	38.58	19.90	41.42	48.80	50.82	63.86	819
Douglas Fir	ind	100%	46.70	23.35	38.13	49.53	50.82	63.86	227
Oak	ind	100%	61.69	10.73	21.73	40.00	50.82	63.86	261
Spruce	ind	100%	50.86	49.45	55.38	65.22	50.82	63.86	637
Scots pine	ind	100%	60.11	22.47	22.96	50.00	50.82	63.86	178
Mixed	ind	100%	52.90	88.19	68.59	65.95	50.82	63.86	3108
Beech	q25	0%	50.03	72.02	66.99	58.83	52.14	56.88	1919
Douglas Fir	q25	0%	55.36	43.39	42.37	50.00	52.14	56.88	401
Oak	q25	0%	65.69	40.49	41.77	49.22	52.14	56.88	857
Spruce	q25	0%	59.16	67.75	63.85	61.42	52.14	56.88	1048
Scots pine	q25	0%	68.35	61.11	45.52	59.51	52.14	56.88	594
Mixed	q25	0%	3.75	12.92	10.81	37.30	52.14	56.88	534
Beech	q25	50%	51.07	73.54	65.07	56.94	50.36	55.56	1784
Douglas Fir	q25	50%	56.65	46.54	42.43	49.72	50.36	55.56	376
Oak	q25	50%	66.38	40.17	40.53	50.00	50.36	55.56	809
Spruce	q25	50%	59.80	69.45	63.40	61.44	50.36	55.56	1005
Scots pine	q25	50%	68.74	60.88	43.87	60.22	50.36	55.56	547
Mixed	q25	50%	6.97	15.75	18.07	36.59	50.36	55.56	832
Beech	q25	60%	49.62	72.71	62.27	54.04	47.97	52.55	1693
Douglas Fir	q25	60%	55.07	41.92	42.49	47.96	47.97	52.55	365
Oak	q25	60%	65.49	36.35	38.77	47.59	47.97	52.55	762
Spruce	q25	60%	59.49	68.51	63.46	59.86	47.97	52.55	975
Scots pine	q25	60%	68.02	56.62	40.78	59.53	47.97	52.55	491
Mixed	q25	60%	10.68	19.31	22.31	34.86	47.97	52.55	1067
Beech	q25	80%	43.39	36.95	50.05	47.86	43.98	51.93	1180
Douglas Fir	q25	80%	48.01	36.75	42.27	54.41	43.98	51.93	302
Oak	q25	80%	63.12	22.40	27.79	40.74	43.98	51.93	442
Spruce	q25	80%	55.12	62.25	60.66	62.09	43.98	51.93	800
Scots pine	q25	80%	65.11	35.97	29.05	55.25	43.98	51.93	278
Mixed	q25	80%	33.86	65.33	48.74	51.00	43.98	51.93	2351
Beech	q25	100%	36.72	16.42	41.55	48.81	48.65	62.58	877
Douglas Fir	q25	100%	42.17	21.74	40.25	48.54	48.65	62.58	230
Oak	q25	100%	57.14	8.01	20.53	46.94	48.65	62.58	287
Spruce	q25	100%	52.16	54.17	56.05	62.68	48.65	62.58	648
Scots pine	q25	100%	60.34	20.67	22.69	60.66	48.65	62.58	179
Mixed	q25	100%	50.29	87.64	64.05	64.06	48.65	62.58	3132
Beech	q50	0%	50.75	73.17	67.31	59.42	52.89	58.36	1931
Douglas Fir	q50	0%	54.98	47.01	42.02	56.08	52.89	58.36	402

Continued on next page

class	support	threshold	prod.acc	$prod.acc_{cal}$	use.acc	$use.acc_{cal}$	oaa	oaa_{cal}	n.ref
Oak	q50	0%	67.44	43.49	41.76	52.31	52.89	58.36	860
Spruce	q50	0%	59.83	68.38	64.95	63.21	52.89	58.36	1053
Scots pine	q50	0%	70.25	62.86	45.39	59.37	52.89	58.36	595
Mixed	q50	0%	2.80	12.69	13.04	38.20	52.89	58.36	536
Beech	q50	50%	51.28	73.39	65.60	57.96	50.85	55.81	1796
Douglas Fir	q50	50%	55.17	47.48	42.80	53.92	50.85	55.81	377
Oak	q50	50%	67.36	44.83	40.85	52.53	50.85	55.81	812
Spruce	q50	50%	60.12	69.15	64.81	62.12	50.85	55.81	1008
Scots pine	q50	50%	70.44	58.58	44.99	57.42	50.85	55.81	548
Mixed	q50	50%	7.89	14.59	18.59	30.73	50.85	55.81	836
Beech	q50	60%	48.47	69.84	64.63	55.79	48.58	53.45	1704
Douglas Fir	q50	60%	51.91	44.26	45.56	51.59	48.58	53.45	366
Oak	q50	60%	64.31	45.75	40.13	50.58	48.58	53.45	765
Spruce	q50	60%	58.18	69.33	66.32	63.25	48.58	53.45	978
Scots pine	q50	60%	67.48	55.89	43.12	56.47	48.58	53.45	492
Mixed	q50	60%	18.94	20.43	24.52	32.25	48.58	53.45	1072
Beech	q50	80%	40.91	43.69	53.00	52.11	46.48	54.34	1188
Douglas Fir	q50	80%	45.21	39.27	48.07	60.10	46.48	54.34	303
Oak	q50	80%	60.67	24.04	29.61	39.78	46.48	54.34	445
Spruce	q50	80%	52.18	62.39	63.20	63.66	46.48	54.34	803
Scots pine	q50	80%	62.95	41.37	31.59	62.50	46.48	54.34	278
Mixed	q50	80%	42.88	66.14	49.46	53.04	46.48	54.34	2360
Beech	q50	100%	30.39	20.75	45.58	46.10	53.58	61.82	882
Douglas Fir	q50	100%	35.22	21.74	46.29	44.25	53.58	61.82	230
Oak	q50	100%	49.65	11.46	23.25	44.00	53.58	61.82	288
Spruce	q50	100%	43.47	53.61	58.11	60.70	53.58	61.82	651
Scots pine	q50	100%	58.10	26.26	28.42	52.22	53.58	61.82	179
Mixed	q50	100%	63.62	84.59	63.64	64.50	53.58	61.82	3147
Beech	q80	0%	51.16	73.41	67.05	59.67	53.33	58.52	1937
Douglas Fir	q80	0%	55.45	47.52	42.83	56.80	53.33	58.52	404
Oak	q80	0%	67.82	43.75	41.74	52.07	53.33	58.52	864
Spruce	q80	0%	60.95	68.63	65.75	63.34	53.33	58.52	1055
Scots pine	q80	0%	71.19	60.97	45.95	58.43	53.33	58.52	597
Mixed	q80	0%	2.03	14.73	11.96	42.78	53.33	58.52	543
Beech	q80	50%	51.17	72.42	66.05	57.90	51.70	56.06	1802
Douglas Fir	q80	50%	55.94	46.44	45.11	53.99	51.70	56.06	379
Oak	q80	50%	67.52	45.10	41.43	51.90	51.70	56.06	816
Spruce	q80	50%	60.99	69.50	66.17	62.46	51.70	56.06	1010
Scots pine	q80	50%	70.67	59.20	46.75	59.74	51.70	56.06	549
Mixed	q80	50%	12.20	17.89	23.25	34.09	51.70	56.06	844
Beech	q80	60%	47.25	70.23	65.69	56.78	48.76	54.59	1710
Douglas Fir	q80	60%	51.90	46.47	49.61	52.29	48.76	54.59	368
Oak	q80	60%	63.98	47.33	41.80	49.66	48.76	54.59	769
Spruce	q80	60%	56.94	69.59	67.55	63.44	48.76	54.59	980
Scots pine	q80	60%	66.53	55.58	44.99	60.22	48.76	54.59	493
Mixed	q80	60%	23.70	23.70	24.31	36.83	48.76	54.59	1080
Beech	q80	80%	37.80	39.90	55.68	51.63	47.91	53.67	1193
Douglas Fir	q80	80%	40.33	37.70	51.46	60.21	47.91	53.67	305
Oak	q80	80%	57.49	25.73	30.85	41.52	47.91	53.67	447
Spruce	q80	80%	49.00	60.57	64.27	63.99	47.91	53.67	804
Scots pine	q80	80%	62.01	40.86	35.97	66.28	47.91	53.67	279
Mixed	q80	80%	50.13	67.07	49.05	51.71	47.91	53.67	2372
Beech	q80	100%	25.54	22.94	50.45	49.88	57.41	62.61	885
Douglas Fir	q80	100%	29.13	23.91	53.17	52.88	57.41	62.61	230
Oak	q80	100%	40.83	13.49	25.76	52.70	57.41	62.61	289
Spruce	q80	100%	36.96	48.16	61.95	60.15	57.41	62.61	652
Scots pine	q80	100%	51.67	30.00	34.19	60.00	57.41	62.61	180
Mixed	q80	100%	74.43	85.84	63.53	64.62	57.41	62.61	3164
Beech	q100	0%	53.73	72.96	67.46	59.55	55.24	58.53	1945
Douglas Fir	q100	0%	55.88	47.30	45.97	54.21	55.24	58.53	408
Oak	q100	0%	71.35	47.18	43.27	53.25	55.24	58.53	869
Spruce	q100	0%	62.70	66.29	66.14	63.02	55.24	58.53	1059
Scots pine	q100	0%	73.83	65.67	46.98	63.14	55.24	58.53	600
Mixed	q100	0%	0.18	11.23	12.50	33.33	55.24	58.53	552
Beech	q100	50%	48.45	71.66	68.95	57.49	51.87	56.27	1810
Douglas Fir	q100	50%	49.74	48.69	55.39	56.02	51.87	56.27	382
Oak	q100	50%	68.21	49.45	46.40	53.63	51.87	56.27	821
Spruce	q100	50%	57.69	67.85	68.74	62.77	51.87	56.27	1014
Scots pine	q100	50%	68.78	63.70	54.85	61.69	51.87	56.27	551
Mixed	q100	50%	26.55	15.09	21.23	30.50	51.87	56.27	855
Beech	q100	60%	40.98	68.45	69.02	55.84	48.56	52.81	1718

Continued on next page

class	support	threshold	prod.acc	$prod.acc_{cal}$	use.acc	$use.acc_{cal}$	oaa	oaa_{cal}	n.ref
Douglas Fir	q100	60%	40.70	42.86	57.85	51.79	48.56	52.81	371
Oak	q100	60%	61.24	46.64	48.12	50.35	48.56	52.81	774
Spruce	q100	60%	51.12	67.78	70.75	61.93	48.56	52.81	984
Scots pine	q100	60%	63.03	58.18	57.78	60.00	48.56	52.81	495
Mixed	q100	60%	45.28	19.98	25.78	29.22	48.56	52.81	1091
Beech	q100	80%	27.55	40.15	63.22	51.06	51.15	53.49	1198
Douglas Fir	q100	80%	26.30	28.57	63.78	58.67	51.15	53.49	308
Oak	q100	80%	47.12	26.55	36.47	44.12	51.15	53.49	452
Spruce	q100	80%	36.80	60.35	69.72	62.92	51.15	53.49	807
Scots pine	q100	80%	57.35	44.09	52.29	64.06	51.15	53.49	279
Mixed	q100	80%	71.08	67.27	48.96	51.79	51.15	53.49	2389
Beech	q100	100%	5.29	19.69	55.29	40.98	59.36	59.51	889
Douglas Fir	q100	100%	7.79	10.39	69.23	40.68	59.36	59.51	231
Oak	q100	100%	11.26	7.85	27.73	30.26	59.36	59.51	293
Spruce	q100	100%	11.81	43.87	77.78	52.87	59.36	59.51	652
Scots pine	q100	100%	20.56	25.00	50.68	52.94	59.36	59.51	180
Mixed	q100	100%	94.51	84.07	59.89	63.13	59.36	59.51	3188

4: Model accuracies

Table 2: Model accuracies realized under various support choices for the CHM- and **uncalibrated** *treespecies* explanatory variables

$support_{chm}$	$support_{tspec}$	$threshold$	R^2_{adj}	$rmse_{cv}$	$RMSE_{cv}[\%]$	AIC	$R^2_{adj,ref}$	$RMSE_{cv,ref}$	$RMSE_{cv,ref}[\%]$	AIC_{ref}
ind	ind	0%	0.44	142.03	44.20	64408.57	0.45	139.71	43.48	64256.93
ind	ind	50%	0.44	141.74	44.11	64401.58	0.45	139.85	43.52	64268.78
ind	ind	60%	0.44	141.95	44.18	64405.91	0.45	139.48	43.41	64255.23
ind	ind	80%	0.44	141.68	44.09	64384.81	0.46	138.92	43.23	64225.80
ind	ind	100%	0.44	141.97	44.18	64385.38	0.46	138.85	43.21	64204.65
ind	q25	0%	0.44	140.67	44.21	64927.75	0.46	139.25	43.77	64780.23
ind	q25	50%	0.44	140.56	44.18	64912.21	0.46	139.37	43.81	64795.25
ind	q25	60%	0.44	140.63	44.20	64914.25	0.46	139.23	43.76	64783.04
ind	q25	80%	0.45	140.77	44.25	64906.60	0.46	138.57	43.56	64749.77
ind	q25	100%	0.44	140.72	44.23	64909.56	0.46	138.14	43.42	64725.81
ind	q50	0%	0.44	141.62	44.71	65195.36	0.46	139.36	44.00	65057.03
ind	q50	50%	0.44	141.58	44.70	65186.32	0.46	139.60	44.07	65072.06
ind	q50	60%	0.45	141.55	44.69	65184.31	0.46	139.53	44.05	65060.08
ind	q50	80%	0.44	141.61	44.70	65197.21	0.46	138.90	43.85	65026.35
ind	q50	100%	0.44	142.05	44.85	65230.07	0.46	138.60	43.75	65002.79
ind	q80	0%	0.45	140.73	44.58	65392.45	0.46	139.43	44.16	65282.95
ind	q80	50%	0.45	140.75	44.58	65391.12	0.46	139.71	44.25	65297.81
ind	q80	60%	0.44	141.25	44.74	65414.38	0.46	139.58	44.21	65286.72
ind	q80	80%	0.45	141.20	44.72	65397.48	0.46	138.97	44.02	65252.00
ind	q80	100%	0.44	142.03	44.98	65482.59	0.46	138.84	43.98	65227.79
ind	q100	0%	0.45	140.81	44.34	65701.65	0.46	138.83	43.72	65570.10
ind	q100	50%	0.45	140.58	44.27	65688.26	0.46	139.13	43.81	65585.38
ind	q100	60%	0.44	140.87	44.36	65708.90	0.46	139.14	43.81	65575.20
ind	q100	80%	0.44	141.69	44.62	65772.71	0.46	138.39	43.58	65541.28
ind	q100	100%	0.42	143.37	45.14	65885.80	0.46	138.25	43.53	65513.09
q25	ind	0%	0.44	142.57	44.60	64938.10	0.46	140.09	43.82	64776.64
q25	ind	50%	0.44	142.53	44.59	64935.55	0.46	140.28	43.88	64790.59
q25	ind	60%	0.44	142.64	44.62	64936.98	0.46	139.96	43.78	64778.10
q25	ind	80%	0.44	142.39	44.54	64908.14	0.46	139.43	43.62	64756.67
q25	ind	100%	0.44	142.06	44.44	64895.64	0.46	138.92	43.46	64729.99
q25	q25	0%	0.45	141.99	45.44	65965.52	0.47	139.91	44.78	65808.37
q25	q25	50%	0.45	141.87	45.40	65955.14	0.47	140.07	44.82	65822.64
q25	q25	60%	0.45	141.89	45.41	65957.07	0.47	139.77	44.73	65809.39
q25	q25	80%	0.45	141.91	45.42	65944.31	0.47	139.33	44.59	65783.19
q25	q25	100%	0.45	142.05	45.46	65955.03	0.48	138.53	44.33	65747.03
q25	q50	0%	0.45	139.70	44.82	66237.98	0.47	138.15	44.33	66087.78
q25	q50	50%	0.45	139.61	44.80	66229.45	0.47	138.31	44.38	66101.77
q25	q50	60%	0.45	139.71	44.83	66226.29	0.47	138.09	44.31	66088.89
q25	q50	80%	0.45	139.87	44.88	66241.23	0.47	137.55	44.13	66062.63
q25	q50	100%	0.45	140.32	45.02	66280.18	0.47	136.83	43.90	66026.09
q25	q80	0%	0.46	140.42	45.24	66429.54	0.47	139.34	44.89	66311.89
q25	q80	50%	0.46	140.36	45.22	66432.75	0.47	139.72	45.02	66325.46
q25	q80	60%	0.45	140.78	45.36	66453.60	0.47	139.52	44.95	66313.60
q25	q80	80%	0.46	140.53	45.27	66439.86	0.47	139.07	44.80	66286.42
q25	q80	100%	0.45	142.02	45.75	66532.00	0.48	138.48	44.61	66249.45
q25	q100	0%	0.45	142.36	45.55	66767.76	0.47	140.09	44.83	66626.46
q25	q100	50%	0.46	141.92	45.41	66755.71	0.47	140.32	44.90	66640.38
q25	q100	60%	0.45	142.18	45.49	66783.49	0.47	140.42	44.93	66629.48
q25	q100	80%	0.45	143.32	45.86	66854.10	0.47	140.16	44.85	66602.55
q25	q100	100%	0.43	145.09	46.43	66974.26	0.48	139.97	44.79	66563.15
q50	ind	0%	0.45	141.43	44.24	64854.59	0.47	138.82	43.42	64686.53
q50	ind	50%	0.45	141.41	44.24	64850.86	0.47	138.93	43.46	64698.27
q50	ind	60%	0.45	141.50	44.27	64853.84	0.47	138.70	43.39	64689.17
q50	ind	80%	0.45	141.33	44.21	64827.37	0.47	138.27	43.25	64672.63
q50	ind	100%	0.45	141.16	44.16	64827.16	0.47	137.91	43.14	64652.47
q50	q25	0%	0.46	140.61	45.00	65874.22	0.48	138.36	44.28	65710.42
q50	q25	50%	0.46	140.45	44.95	65862.51	0.48	138.43	44.30	65722.03
q50	q25	60%	0.46	140.49	44.96	65864.18	0.48	138.19	44.22	65711.95
q50	q25	80%	0.46	140.59	44.99	65854.77	0.48	137.97	44.15	65691.89
q50	q25	100%	0.46	140.58	44.99	65860.14	0.48	137.41	43.98	65664.81
q50	q50	0%	0.46	138.57	44.46	66145.64	0.48	136.89	43.92	65989.18
q50	q50	50%	0.46	138.52	44.44	66138.78	0.48	137.05	43.97	66001.00
q50	q50	60%	0.46	138.62	44.48	66134.47	0.48	136.89	43.92	65991.20
q50	q50	80%	0.46	138.78	44.53	66146.85	0.48	136.48	43.79	65971.13
q50	q50	100%	0.46	139.30	44.70	66189.99	0.48	135.98	43.63	65943.10
q50	q80	0%	0.47	138.34	44.57	66333.05	0.48	137.22	44.21	66211.70
q50	q80	50%	0.47	138.35	44.57	66341.81	0.48	137.60	44.33	66223.75
q50	q80	60%	0.46	138.77	44.71	66361.90	0.48	137.43	44.28	66214.72
q50	q80	80%	0.46	138.54	44.63	66347.78	0.48	137.11	44.17	66193.69
q50	q80	100%	0.45	140.08	45.13	66444.51	0.48	136.63	44.02	66165.28
q50	q100	0%	0.46	141.39	45.24	66681.45	0.48	138.94	44.46	66528.70
q50	q100	50%	0.47	140.88	45.08	66665.99	0.48	139.14	44.52	66540.96
q50	q100	60%	0.46	141.14	45.16	66693.59	0.48	139.27	44.56	66532.93
q50	q100	80%	0.45	142.17	45.49	66764.60	0.48	139.15	44.52	66511.93
q50	q100	100%	0.44	143.86	46.03	66879.02	0.48	138.81	44.42	66480.34
q80	ind	0%	0.45	141.92	44.40	64868.26	0.47	139.22	43.55	64695.13
q80	ind	50%	0.45	141.87	44.38	64862.40	0.47	139.41	43.61	64711.20
q80	ind	60%	0.45	142.00	44.42	64866.26	0.47	139.20	43.54	64701.77
q80	ind	80%	0.45	141.89	44.39	64847.05	0.47	138.77	43.41	64681.35
q80	ind	100%	0.45	141.79	44.35	64851.41	0.47	138.41	43.30	64658.70

Continued on next page

$support_{chm}$	$support_{tspec}$	$threshold$	R^2_{adj}	$RMSE_{cv}$	$RMSE_{cv}[\%]$	AIC	$R^2_{adj,ref}$	$RMSE_{cv,ref}$	$RMSE_{cv,ref}[\%]$	AIC_{ref}
q80	q25	0%	0.46	140.65	45.01	65888.64	0.48	138.44	44.30	65720.92
q80	q25	50%	0.46	140.50	44.96	65877.65	0.48	138.60	44.35	65737.10
q80	q25	60%	0.46	140.54	44.98	65878.63	0.48	138.46	44.31	65726.72
q80	q25	80%	0.46	140.65	45.01	65869.87	0.48	138.20	44.23	65703.95
q80	q25	100%	0.46	140.70	45.03	65878.28	0.48	137.46	43.99	65674.03
q80	q50	0%	0.46	138.78	44.53	66159.84	0.48	137.07	43.98	66000.20
q80	q50	50%	0.46	138.74	44.51	66154.54	0.48	137.26	44.04	66016.68
q80	q50	60%	0.46	138.88	44.56	66151.52	0.48	136.96	43.94	66005.75
q80	q50	80%	0.46	139.03	44.61	66162.48	0.48	136.59	43.82	65982.82
q80	q50	100%	0.46	139.62	44.80	66209.86	0.48	136.14	43.68	65952.40
q80	q80	0%	0.46	138.03	44.47	66348.58	0.48	136.86	44.09	66222.79
q80	q80	50%	0.46	138.08	44.48	66357.36	0.48	137.28	44.23	66239.71
q80	q80	60%	0.46	138.50	44.62	66378.68	0.48	137.06	44.16	66229.36
q80	q80	80%	0.46	138.41	44.59	66367.72	0.48	136.71	44.04	66205.47
q80	q80	100%	0.45	139.94	45.09	66461.68	0.48	136.34	43.93	66174.95
q80	q100	0%	0.46	141.52	45.28	66698.65	0.48	138.91	44.45	66542.06
q80	q100	50%	0.46	140.95	45.10	66684.27	0.48	139.16	44.53	66558.80
q80	q100	60%	0.46	141.31	45.22	66715.73	0.48	139.20	44.54	66550.13
q80	q100	80%	0.45	142.29	45.53	66783.55	0.48	139.01	44.48	66527.02
q80	q100	100%	0.44	143.91	46.05	66895.38	0.48	138.74	44.39	66492.22
q100	ind	0%	0.39	149.44	46.75	65391.23	0.41	147.18	46.04	65241.42
q100	ind	50%	0.39	149.34	46.72	65379.90	0.40	147.32	46.08	65260.46
q100	ind	60%	0.39	149.43	46.75	65382.54	0.41	147.10	46.01	65249.02
q100	ind	80%	0.39	149.50	46.77	65377.00	0.41	146.57	45.85	65224.48
q100	ind	100%	0.39	149.15	46.65	65366.75	0.41	146.01	45.67	65194.50
q100	q25	0%	0.40	147.61	47.24	66442.67	0.42	145.59	46.59	66291.14
q100	q25	50%	0.40	147.46	47.19	66430.07	0.41	145.73	46.63	66310.71
q100	q25	60%	0.40	147.36	47.16	66428.18	0.42	145.51	46.56	66299.12
q100	q25	80%	0.40	147.39	47.17	66416.52	0.42	145.21	46.47	66271.49
q100	q25	100%	0.40	147.60	47.23	66426.82	0.42	144.46	46.23	66233.99
q100	q50	0%	0.40	146.76	47.09	66720.28	0.42	145.07	46.54	66572.83
q100	q50	50%	0.40	146.63	47.05	66712.95	0.41	145.32	46.62	66592.86
q100	q50	60%	0.40	146.85	47.11	66708.05	0.41	145.14	46.57	66581.36
q100	q50	80%	0.40	147.27	47.25	66722.92	0.42	144.59	46.39	66553.52
q100	q50	100%	0.39	147.64	47.37	66759.50	0.42	144.08	46.23	66516.11
q100	q80	0%	0.40	145.20	46.78	66917.50	0.42	144.07	46.42	66796.64
q100	q80	50%	0.40	145.12	46.76	66917.72	0.41	144.47	46.55	66816.62
q100	q80	60%	0.40	145.58	46.90	66935.84	0.42	144.29	46.49	66805.78
q100	q80	80%	0.40	145.62	46.92	66932.29	0.42	143.97	46.39	66776.81
q100	q80	100%	0.39	147.12	47.40	67015.72	0.42	143.54	46.25	66739.98
q100	q100	0%	0.40	148.90	47.65	67272.02	0.42	146.01	46.72	67121.18
q100	q100	50%	0.40	148.28	47.45	67258.42	0.41	146.33	46.82	67140.71
q100	q100	60%	0.40	148.59	47.55	67283.11	0.42	146.45	46.86	67130.60
q100	q100	80%	0.39	149.75	47.92	67353.11	0.42	146.13	46.76	67102.80
q100	q100	100%	0.38	151.55	48.49	67469.16	0.42	145.80	46.66	67062.13

Table 3: Model accuracies realized under various support choices for the CHM- and **calibrated** *treespecies* explanatory variables

$support_{chm}$	$support_{tspec}$	$threshold$	R^2_{adj}	$RMSE_{cv}$	$RMSE_{cv}[\%]$	AIC	$R^2_{adj,ref}$	$RMSE_{cv,ref}$	$RMSE_{cv,ref}[\%]$	AIC_{ref}
ind	ind	0%	0.45	140.12	43.61	64291.15	0.45	139.71	43.48	64256.93
ind	ind	50%	0.45	140.60	43.75	64317.74	0.45	139.85	43.52	64268.78
ind	ind	60%	0.45	140.13	43.61	64294.65	0.45	139.48	43.41	64255.23
ind	ind	80%	0.46	139.10	43.29	64236.04	0.46	138.92	43.23	64225.80
ind	ind	100%	0.45	139.92	43.54	64280.23	0.46	138.85	43.21	64204.65
ind	q25	0%	0.45	139.90	43.97	64835.11	0.46	139.25	43.77	64780.23
ind	q25	50%	0.45	140.06	44.02	64852.35	0.46	139.37	43.81	64795.25
ind	q25	60%	0.45	139.97	43.99	64846.98	0.46	139.23	43.76	64783.04
ind	q25	80%	0.46	139.17	43.74	64792.34	0.46	138.57	43.56	64749.77
ind	q25	100%	0.46	139.16	43.74	64792.06	0.46	138.14	43.42	64725.81
ind	q50	0%	0.45	139.91	44.17	65093.04	0.46	139.36	44.00	65057.03
ind	q50	50%	0.45	140.25	44.28	65113.44	0.46	139.60	44.07	65072.06
ind	q50	60%	0.45	139.89	44.16	65094.92	0.46	139.53	44.05	65060.08
ind	q50	80%	0.46	139.02	43.89	65033.71	0.46	138.90	43.85	65026.35
ind	q50	100%	0.46	139.19	43.94	65040.22	0.46	138.60	43.75	65002.79
ind	q80	0%	0.46	139.68	44.24	65309.58	0.46	139.43	44.16	65282.95
ind	q80	50%	0.45	139.97	44.33	65336.21	0.46	139.71	44.25	65297.81
ind	q80	60%	0.46	139.97	44.33	65319.59	0.46	139.58	44.21	65286.72
ind	q80	80%	0.46	139.25	44.11	65283.08	0.46	138.97	44.02	65252.00
ind	q80	100%	0.46	139.09	44.05	65251.61	0.46	138.84	43.98	65227.79
ind	q100	0%	0.46	138.89	43.73	65577.54	0.46	138.83	43.72	65570.10
ind	q100	50%	0.46	139.18	43.82	65589.23	0.46	139.13	43.81	65585.38
ind	q100	60%	0.46	139.21	43.83	65580.00	0.46	139.14	43.81	65575.20
ind	q100	80%	0.46	138.43	43.59	65543.72	0.46	138.39	43.58	65541.28
ind	q100	100%	0.46	138.34	43.56	65521.67	0.46	138.25	43.53	65513.09
q25	ind	0%	0.45	140.80	44.04	64815.81	0.46	140.09	43.82	64776.64
q25	ind	50%	0.45	140.79	44.04	64815.82	0.46	140.28	43.88	64790.59
q25	ind	60%	0.45	140.73	44.02	64812.34	0.46	139.96	43.78	64778.10
q25	ind	80%	0.46	139.93	43.77	64780.19	0.46	139.43	43.62	64756.67
q25	ind	100%	0.45	140.61	43.99	64846.27	0.46	138.92	43.46	64729.99
q25	q25	0%	0.46	141.17	45.18	65882.42	0.47	139.91	44.78	65808.37
q25	q25	50%	0.46	140.96	45.11	65878.58	0.47	140.07	44.82	65822.64
q25	q25	60%	0.46	140.82	45.07	65871.86	0.47	139.77	44.73	65809.39
q25	q25	80%	0.47	139.69	44.70	65814.08	0.47	139.33	44.59	65783.19
q25	q25	100%	0.47	139.64	44.69	65826.77	0.48	138.53	44.33	65747.03
q25	q50	0%	0.46	138.51	44.44	66122.18	0.47	138.15	44.33	66087.78
q25	q50	50%	0.46	138.65	44.49	66132.92	0.47	138.31	44.38	66101.77
q25	q50	60%	0.46	138.51	44.44	66140.26	0.47	138.09	44.31	66088.89
q25	q50	80%	0.47	137.57	44.14	66087.41	0.47	137.55	44.13	66062.63
q25	q50	100%	0.47	137.32	44.06	66076.87	0.47	136.83	43.90	66026.09
q25	q80	0%	0.46	139.64	44.99	66345.17	0.47	139.34	44.89	66311.89
q25	q80	50%	0.46	139.99	45.10	66355.29	0.47	139.72	45.02	66325.46
q25	q80	60%	0.46	139.57	44.97	66349.49	0.47	139.52	44.95	66313.60
q25	q80	80%	0.47	139.47	44.93	66318.88	0.47	139.07	44.80	66286.42
q25	q80	100%	0.47	139.12	44.82	66310.46	0.48	138.48	44.61	66249.45
q25	q100	0%	0.47	140.12	44.83	66631.70	0.47	140.09	44.83	66626.46
q25	q100	50%	0.47	140.34	44.90	66642.41	0.47	140.32	44.90	66640.38
q25	q100	60%	0.47	140.46	44.94	66633.43	0.47	140.42	44.93	66629.48
q25	q100	80%	0.47	140.20	44.86	66606.19	0.47	140.16	44.85	66602.55
q25	q100	100%	0.47	140.13	44.84	66576.55	0.48	139.97	44.79	66563.15
q50	ind	0%	0.46	140.05	43.81	64745.65	0.47	138.82	43.42	64686.53
q50	ind	50%	0.46	139.97	43.79	64748.33	0.47	138.93	43.46	64698.27
q50	ind	60%	0.46	139.64	43.68	64731.81	0.47	138.70	43.39	64689.17
q50	ind	80%	0.47	139.05	43.50	64712.28	0.47	138.27	43.25	64672.63
q50	ind	100%	0.46	139.66	43.69	64771.24	0.47	137.91	43.14	64652.47
q50	q25	0%	0.47	139.57	44.67	65788.59	0.48	138.36	44.28	65710.42
q50	q25	50%	0.47	139.75	44.72	65802.74	0.48	138.43	44.30	65722.03
q50	q25	60%	0.47	139.60	44.68	65789.21	0.48	138.19	44.22	65711.95
q50	q25	80%	0.48	138.54	44.34	65728.34	0.48	137.97	44.15	65691.89
q50	q25	100%	0.48	138.42	44.30	65735.17	0.48	137.41	43.98	65664.81
q50	q50	0%	0.47	137.47	44.11	66034.28	0.48	136.89	43.92	65989.18
q50	q50	50%	0.47	137.59	44.15	66037.77	0.48	137.05	43.97	66001.00
q50	q50	60%	0.47	137.33	44.06	66033.77	0.48	136.89	43.92	65991.20
q50	q50	80%	0.48	136.78	43.89	66003.35	0.48	136.48	43.79	65971.13
q50	q50	100%	0.48	136.62	43.84	66002.02	0.48	135.98	43.63	65943.10
q50	q80	0%	0.48	137.80	44.39	66248.74	0.48	137.22	44.21	66211.70
q50	q80	50%	0.47	137.94	44.44	66261.63	0.48	137.60	44.33	66223.75
q50	q80	60%	0.47	137.77	44.39	66256.59	0.48	137.43	44.28	66214.72
q50	q80	80%	0.47	137.69	44.36	66250.53	0.48	137.11	44.17	66193.69
q50	q80	100%	0.48	137.13	44.18	66202.85	0.48	136.63	44.02	66165.28
q50	q100	0%	0.48	138.91	44.45	66527.59	0.48	138.94	44.46	66528.70
q50	q100	50%	0.48	139.18	44.53	66542.04	0.48	139.14	44.52	66540.96
q50	q100	60%	0.48	139.30	44.57	66535.91	0.48	139.27	44.56	66532.93
q50	q100	80%	0.48	139.17	44.53	66513.80	0.48	139.15	44.52	66511.93
q50	q100	100%	0.48	138.82	44.42	66479.85	0.48	138.81	44.42	66480.34
q80	ind	0%	0.46	140.31	43.89	64746.39	0.47	139.22	43.55	64695.13
q80	ind	50%	0.46	140.23	43.87	64748.57	0.47	139.41	43.61	64711.20
q80	ind	60%	0.46	140.00	43.80	64737.02	0.47	139.20	43.54	64701.77
q80	ind	80%	0.47	139.16	43.53	64702.03	0.47	138.77	43.41	64681.35
q80	ind	100%	0.46	140.13	43.84	64782.06	0.47	138.41	43.30	64658.70
q80	q25	0%	0.47	139.67	44.70	65812.79	0.48	138.44	44.30	65720.92
q80	q25	50%	0.47	139.78	44.73	65816.46	0.48	138.60	44.35	65737.10
q80	q25	60%	0.47	139.63	44.68	65803.87	0.48	138.46	44.31	65726.72
q80	q25	80%	0.48	138.66	44.37	65740.91	0.48	138.20	44.23	65703.95

Continued on next page

$support_{chm}$	$support_{tspec}$	$threshold$	R^2_{adj}	$RMSE_{cv}$	$RMSE_{cv}[\%]$	AIC	$R^2_{adj,ref}$	$RMSE_{cv,ref}$	$RMSE_{cv,ref}[\%]$	AIC_{ref}
q80	q25	100%	0.48	138.48	44.32	65747.54	0.48	137.46	43.99	65674.03
q80	q50	0%	0.47	137.55	44.13	66041.96	0.48	137.07	43.98	66000.20
q80	q50	50%	0.47	137.67	44.17	66057.53	0.48	137.26	44.04	66016.68
q80	q50	60%	0.47	137.33	44.06	66042.83	0.48	136.96	43.94	66005.75
q80	q50	80%	0.48	136.92	43.93	66016.64	0.48	136.59	43.82	65982.82
q80	q50	100%	0.48	136.83	43.90	66016.69	0.48	136.14	43.68	65952.40
q80	q80	0%	0.47	137.32	44.24	66254.83	0.48	136.86	44.09	66222.79
q80	q80	50%	0.47	137.49	44.30	66267.03	0.48	137.28	44.23	66239.71
q80	q80	60%	0.47	137.41	44.27	66262.63	0.48	137.06	44.16	66229.36
q80	q80	80%	0.47	137.17	44.19	66247.32	0.48	136.71	44.04	66205.47
q80	q80	100%	0.48	137.14	44.18	66237.84	0.48	136.34	43.93	66174.95
q80	q100	0%	0.48	139.02	44.48	66548.99	0.48	138.91	44.45	66542.06
q80	q100	50%	0.48	139.28	44.56	66566.73	0.48	139.16	44.53	66558.80
q80	q100	60%	0.48	139.23	44.55	66553.30	0.48	139.20	44.54	66550.13
q80	q100	80%	0.48	139.00	44.48	66526.64	0.48	139.01	44.48	66527.02
q80	q100	100%	0.48	138.84	44.43	66501.70	0.48	138.74	44.39	66492.22
q100	ind	0%	0.40	147.90	46.26	65284.59	0.41	147.18	46.04	65241.42
q100	ind	50%	0.40	147.96	46.28	65287.17	0.40	147.32	46.08	65260.46
q100	ind	60%	0.40	147.88	46.26	65290.73	0.41	147.10	46.01	65249.02
q100	ind	80%	0.40	147.33	46.09	65256.46	0.41	146.57	45.85	65224.48
q100	ind	100%	0.40	147.94	46.28	65310.65	0.41	146.01	45.67	65194.50
q100	q25	0%	0.41	146.40	46.85	66352.97	0.42	145.59	46.59	66291.14
q100	q25	50%	0.41	146.37	46.84	66364.04	0.41	145.73	46.63	66310.71
q100	q25	60%	0.41	146.44	46.86	66359.76	0.42	145.51	46.56	66299.12
q100	q25	80%	0.41	145.80	46.66	66319.25	0.42	145.21	46.47	66271.49
q100	q25	100%	0.41	145.44	46.54	66311.30	0.42	144.46	46.23	66233.99
q100	q50	0%	0.41	145.73	46.76	66629.70	0.42	145.07	46.54	66572.83
q100	q50	50%	0.41	145.90	46.81	66641.20	0.41	145.32	46.62	66592.86
q100	q50	60%	0.41	145.50	46.68	66621.92	0.41	145.14	46.57	66581.36
q100	q50	80%	0.41	145.08	46.55	66597.85	0.42	144.59	46.39	66553.52
q100	q50	100%	0.42	144.76	46.45	66570.46	0.42	144.08	46.23	66516.11
q100	q80	0%	0.41	144.43	46.53	66840.17	0.42	144.07	46.42	66796.64
q100	q80	50%	0.41	144.61	46.59	66852.78	0.41	144.47	46.55	66816.62
q100	q80	60%	0.41	144.50	46.56	66831.91	0.42	144.29	46.49	66805.78
q100	q80	80%	0.41	144.40	46.52	66819.43	0.42	143.97	46.39	66776.81
q100	q80	100%	0.42	144.23	46.47	66799.38	0.42	143.54	46.25	66739.98
q100	q100	0%	0.42	146.18	46.78	67128.38	0.42	146.01	46.72	67121.18
q100	q100	50%	0.41	146.48	46.87	67147.05	0.41	146.33	46.82	67140.71
q100	q100	60%	0.42	146.49	46.88	67131.35	0.42	146.45	46.86	67130.60
q100	q100	80%	0.42	146.17	46.77	67105.03	0.42	146.13	46.76	67102.80
q100	q100	100%	0.42	145.81	46.66	67064.03	0.42	145.80	46.66	67062.13