

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Bont, Leo Gallus; Hill, Andreas; Waser, Lars T.; Bürgi, Anton; Ginzler, Christian; Blattert, Clemens

Title: Airborne-laser-scanning-derived auxiliary information discriminating between broadleaf and conifer trees improves the accuracy of models for predicting timber volume in mixed and heterogeneously structured forests

Year: 2020

Version: Accepted version (Final draft)

Copyright: © 2020 Elsevier B.V. All rights reserved.

Rights: CC BY-NC-ND 4.0

Rights url: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Please cite the original version:

Bont, L. G., Hill, A., Waser, L. T., Bürgi, A., Ginzler, C., & Blattert, C. (2020). Airborne-laser-scanning-derived auxiliary information discriminating between broadleaf and conifer trees improves the accuracy of models for predicting timber volume in mixed and heterogeneously structured forests. *Forest Ecology and Management*, 459, Article 117856.
<https://doi.org/10.1016/j.foreco.2019.117856>

1 **Airborne-laser-scanning-derived auxiliary information discriminating**
2 **between broadleaf and conifer trees improves the accuracy of models for**
3 **predicting timber volume in mixed and heterogeneously structured forests**

4

5 Leo Gallus Bont ^{a,*}, Andreas Hill ^b, Lars T. Waser ^c, Anton Bürgi ^a, Christian Ginzler ^c, Clemens Blattert ^d

6 **Corresponding Author*

7 *^a Sustainable Forestry Group, Swiss Federal Institute for Forest, Snow and Landscape Research (WSL),*
8 *Zuercherstrasse 111, CH 8903 Birmensdorf, Switzerland,*

9 *^b Forest Service Rhineland-Palatinate, Beethovenstrasse 3, 54516 Wittlich, Germany*

10 *^c Department of Land Change Science, Swiss Federal Institute for Forest, Snow and Landscape*
11 *Research (WSL), CH 8903 Birmensdorf, Switzerland*

12 *^d Department of Biological and Environmental Science, University of Jyväskylä, P.O. Box 35, 40014*
13 *Jyväskylä, Finland*

14

15 **Highlights:**

- 16 • Information on forest type proportions improved the accuracy of volume predictions
17 • Forest type proportions weighted by canopy height outperformed area proportions
18 • The best model performance was obtained using forest type maps from leaf-off
19 LiDAR

20

21 **Abstract**

22 Managing forests for ecosystem services and biodiversity requires accurate and spatially
23 explicit forest inventory data. A major objective of forest management inventories is to
24 estimate the standing timber volume for certain forest areas. In order to improve the
25 efficiency of an inventory, field based sample-plots can be statistically combined with
26 remote sensing data. Such models usually incorporate auxiliary variables derived from
27 canopy height models. The inclusion of forest type variables, which quantify broadleaf and
28 conifer volume proportions, has been shown to further improve model performance.
29 Currently, the most common way of quantifying broadleaf and conifer forest types is by
30 calculating the proportions of the corresponding areas of the canopy cover. This practice
31 works well for single-layer forests with only a few species, but we hypothesized that this is
32 not best practice for heterogeneously structured and mixed forests, where the area
33 proportion does not accurately reflect the timber volume proportion. To better represent
34 the broadleaf and conifer volume proportions, we introduced two new auxiliary variables in
35 which the area proportion is weighted by height information from a canopy height model.

36 The main objectives of this study were: (1) to demonstrate the advantage of including forest
37 type (broadleaf/conifer distinction) information in ordinary least squares regression models
38 for timber volume prediction using widely available data sources, and (2) to investigate the
39 hypothesis that including the broadleaf and conifer proportions, weighted by canopy height
40 information, as additional auxiliary variables is favourable over including simple area
41 proportions. The study was conducted in three areas in Switzerland, all of which have
42 heterogeneously structured and mixed forests. Our main findings were that the best model
43 performance can generally be achieved: (1) by deriving conifer and broadleaf proportions
44 from a high-resolution broadleaf/conifer map derived from leaf-off airborne laser scanning

45 data, and (2) by using broadleaf/conifer proportions weighted by height information from a
46 canopy height model. Incorporating the so-derived conifer and broadleaf proportions
47 increased the model accuracy by up to 9 percentage points in root mean square error
48 (RMSE) compared with models not using any forest type information, and by up to 2
49 percentage points in RMSE compared with models using conifer and broadleaf proportions
50 based solely on the corresponding area proportions, as done in current practice. Our findings
51 are particularly relevant for mixed and heterogeneously structured forests, such as those
52 managed to achieve multiple functions or to adapt effectively to climate change.

53

54 **Keywords:** airborne laser scanning, best fit models, canopy height model, forest type map,
55 high-precision forest inventory, image-based point clouds, mixed and heterogeneously
56 structured forest, ordinary least squares regression models, merchantable timber volume

57

58 **1 Introduction**

59 Forest ecosystems provide multiple benefits for humans and are particularly important for
60 the conservation of biodiversity (Bäck et al., 2017; MEA, 2005). These manifold contributions
61 of forests make their management a complex and challenging task requiring accurate and
62 spatially explicit information. The aim of forest inventories in general is to obtain reliable
63 information on the condition and development of the forest (Barrett et al., 2016). As the
64 census of an entire forest area is usually impossible, because of the high costs involved,
65 sampling concepts are used in practice. With these methods, the local hectare density of
66 timber volume, basal area and many other forest attributes are derived from measurements
67 of the trees in randomly or systematically distributed sample-plots in the forest area. This
68 data is then used to estimate mean values and totals for the entire forest area, for example
69 the mean or the total timber stock.

70 More accurate information on an entire forested area or a small area (e.g. parts of a forest
71 enterprise) can be obtained by complementing field based inventories with remote sensing
72 data. Furthermore, this can be a cost-effective alternative to increasing the number of field
73 based sample units. The principle of such two-phase inventories is to use statistical models
74 to predict response variables, such as basal area or timber volume, for the population units
75 where no field data is available. Many studies have already demonstrated the potential of
76 these methods (e.g. Hill et al., 2018; Magnussen et al., 2014; Mandallaz et al., 2013; Næsset,
77 2004, 2002; Steinmann et al., 2013). Usually, such statistical models are based on auxiliary
78 variables derived from a canopy height model (CHM) (Xu et al., 2019). Other sources of
79 information, such as tree species or forest type (we use the term 'forest type' for the
80 distinction between either conifer and broadleaf trees or evergreen and deciduous trees,

81 see section 2.2) have been added in a few case studies to improve the performance of the
82 models (Gabriel et al., 2018; Hill et al., 2018).

83 For Switzerland, a freely available national forest type map (FTM) based on optical remote
84 sensing data (Waser et al., 2017) exists in the framework of the Swiss National Forest
85 Inventory (NFI). FTMs can also be derived using multi-temporal (i.e. leaf-on, leaf-off) laser
86 scanning data. Such an approach entails collecting two Light Detecting and Ranging (LiDAR)
87 datasets at the same site under leaf-off and leaf-on conditions (Liang et al., 2007).
88 Alternatively, forest types can be derived based on return intensity and ranking distributions
89 of laser scans from a single point in time, preferably under leaf-off conditions for the
90 conifer/broadleaf differentiation (Liang et al., 2007; Ørka et al., 2009; Parkan, 2018).
91 However, deciduous conifers (i.e. conifers which defoliate in autumn, such as larch) cannot
92 be identified as conifers with any of these three approaches (Fassnacht et al., 2016).

93 FTM information has been used in many studies to improve standing timber volume
94 estimations. For example, Breidenbach et al. (2008) used a continuous variable 'conifer
95 proportion' and its interaction term with the average canopy height to include FTM
96 information derived from leaf-off LiDAR data. Latifi et al. (2012) formed a FTM based on
97 colour infrared (CIR) orthoimages and included the forest type as a categorical variable in
98 their model. The forest type of the sample-plots was assigned to either broadleaf or conifer
99 if the proportion of the pixels of one particular type exceeded 70%, and it was assigned to
100 the mixed category in all other cases. Straub et al. (2009) used CIR orthoimages to create a
101 FTM. They derived the conifer and broadleaf proportions from the percentages of the
102 corresponding pixels in each sample-plot, and this information was then included in the
103 model to estimate the stem volume of forest stands. Hill et al. (2018) included information
104 on five tree species as a categorical variable to improve timber volume predictions. Finally, in

105 Nordic countries such as in Finland, including tree species information as auxiliary data to
106 predict timber volume is quite common, as described by Kukkonen et al.(2019, 2018),
107 Packalén and Maltamo (2006) and Rätty et al. (2016).

108 All of the approaches mentioned above have led to a significant improvement in the
109 accuracy of standing timber volume predictions by including forest type or tree species
110 information in the models. However, in all these studies forest type proportion information
111 was derived based on the area covered by the canopy of the corresponding forest type. This
112 current practice works well for even-aged and single-layer forests. However, we hypothesize
113 that this approach is not best practice for heterogeneously structured and mixed forests, in
114 which different age classes and tree species can occur across a small surface (e.g. within a
115 sample-plot). In such forests the area proportion does not adequately reflect the timber
116 volume proportion. There are two possible reasons for this discrepancy. First, the mean tree
117 size might differ depending on the canopy height. To illustrate this possibility we consider a
118 sample-plot on which conifer and broadleaf trees cover about the same amount of area and
119 where the conifers are all mature whereas the broadleaf trees are much younger. In this
120 case the volume on the half of the sample-plot with the young broadleaf trees is clearly
121 smaller than that on the half with the mature coniferous trees. This point has been
122 confirmed in studies about allometric relationships, such as the work by Reineke, (1933) in
123 establishing the self-thinning rule. In this case, using the area proportion of conifers leads to
124 an underestimation of the proportion of conifer timber volume. Second, stand density (the
125 number of stems in a certain area) also differs for the different species. This point was
126 shown by Pretzsch and Biber (2005) and by Rivoire and Le Moguedec (2012), who
127 generalized the self-thinning relationship of Reineke (1933) for multi-species and mixed-size
128 forests.

129 To better represent the volume proportion of different forest types, in this study we
130 introduced two new forest type variables (FTVs), referred to as 'weighted-canopy-height
131 proportions', in which the area proportions are weighted by height information from a
132 canopy height model. We hypothesized that the 'weighted-canopy-height-proportion' FTVs
133 are favourable over simple area proportion FTVs in mixed and heterogeneously structured
134 forests.

135 The overall objectives of the present study were: (1) to determine the advantage of including
136 forest type information in regression models for timber volume prediction using existing and
137 forthcoming data sources, and (2) to investigate the hypothesis that the 'weighted-canopy-
138 height proportion' FTVs are favourable over simple area proportion FTVs. Both aims are of
139 high practical interest because the underlying data sources used to derive the FTVs, such as
140 the national broadleaf/conifer map or the national leaf-off airborne laser scanning (ALS)
141 data, are widely available and can be integrated with a small effort into current inventories.
142 However, neither of these sources have been used operationally for inventory purposes. This
143 work was embedded in the implementation of design-based regression estimators
144 (Mandallaz, 2013) for predicting standing timber volume. As these design-based regression
145 estimators rely on ordinary least squares (OLS) regression, we used OLS regression models in
146 our study.

147 We addressed the following specific research questions:

148 (1) Is there a gain in model performance when the new 'weighted-canopy-height proportion'
149 FTVs are incorporated into OLS regression models for predicting timber volume, compared
150 with models including the simple area proportion FTVs and models with no forest type
151 explanatory variables?

152 (2) Is there a common best practice for integrating FTM information, such as the spatial
153 resolution considered, that is independent of the individual LiDAR and FTM characteristics?

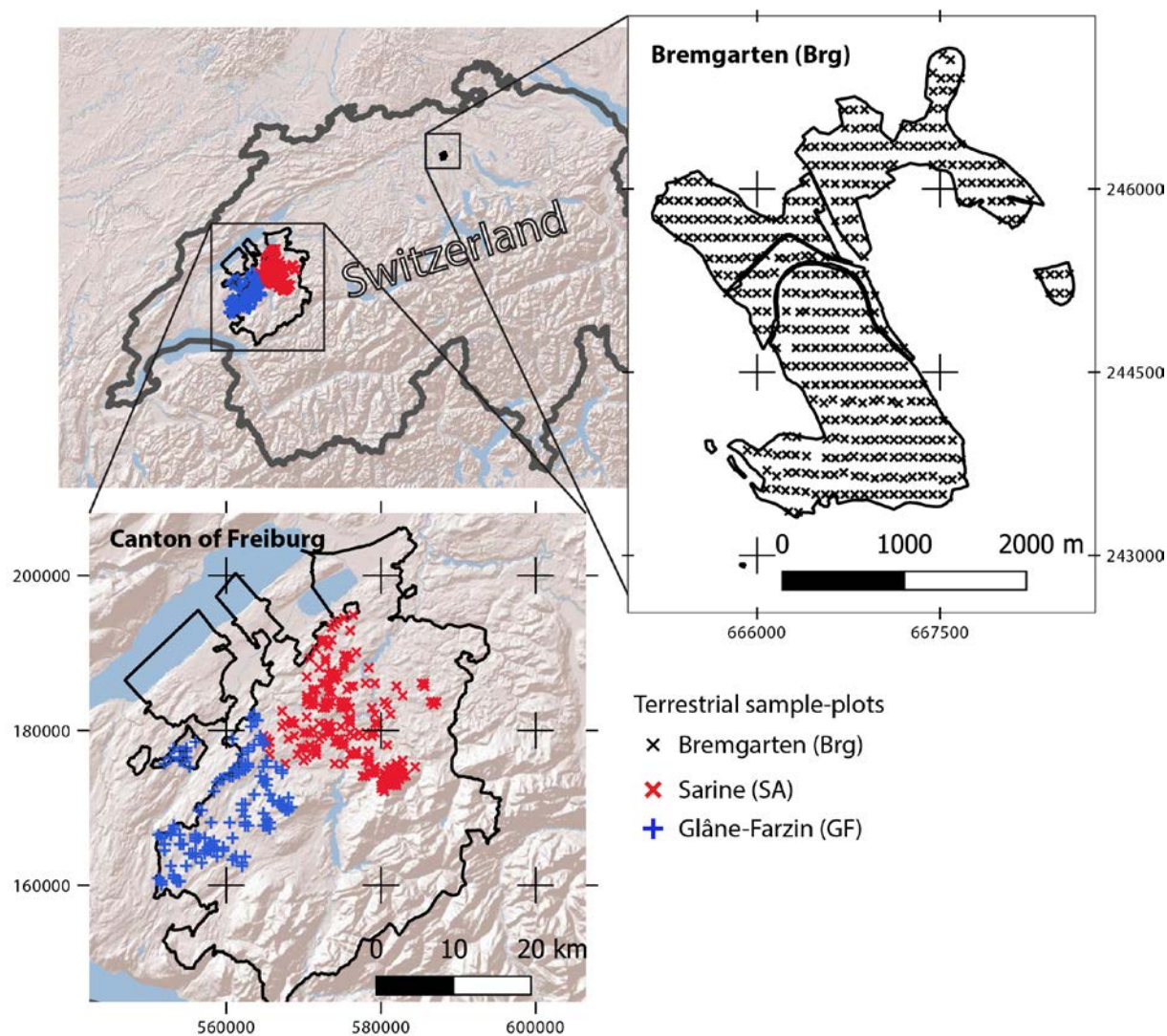
154 **2 Materials and Methods**

155 **2.1 Study areas**

156 We studied the effect of incorporating different FTV alternatives in three independent study
157 areas in Switzerland: Bremgarten (Brg), Glâne-Farzin (GF) and Sarine (SA) (Figure 1). All three
158 study areas are heterogeneously structured and mixed temperate forests of conifer and
159 broadleaf species, but they have different properties in terms of age-class structure.

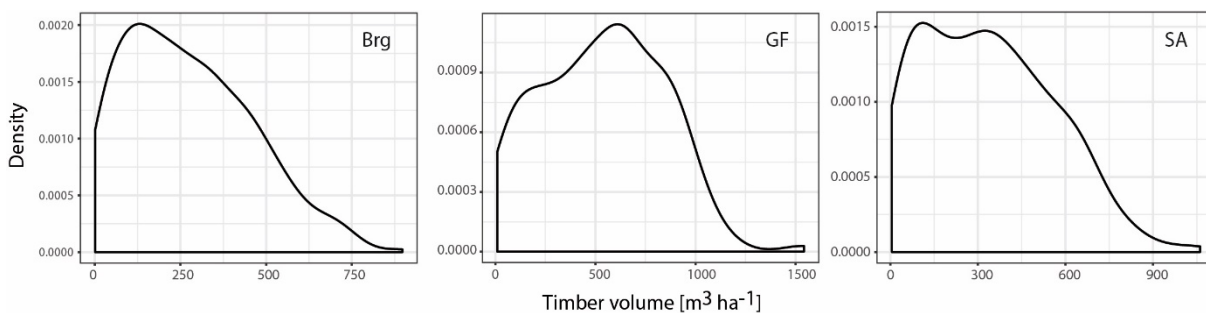
160 **2.1.1 Bremgarten (Brg)**

161 In Brg, the latest forest inventory data from 2011/2012 was used as input. This data is based
162 on permanent sample-plots on an 80 m x 150 m grid, as described by Schmid-Haas (2003).
163 The location of the sample-plots is shown in Figure 1. A total of 363 sample-plots were
164 measured, with an average timber volume of $274 \text{ m}^3 \text{ ha}^{-1}$, 55% of which is conifer wood
165 (Table 1). The distribution of the timber volumes of the sample-plots indicates that stands
166 with large timber volumes are under-represented compared with a normal age-class
167 structure (Salo and Tahvonen, 2002) (Figure 2). This unbalanced age-class structure of the
168 forest was caused by the storm Lothar in December 1999, which primarily damaged mature
169 stands. Figure 3 shows that sample-plots with a large timber volume are dominated by the
170 conifer volume, whereas sample-plots with a small volume are dominated by the broadleaf
171 volume. LiDAR data was acquired in November 2011, providing excellent temporal
172 synchronization between the field based and the remote sensing data. The density of the
173 LiDAR raw data is at least 8 points m^{-2} and is comparable with values in the other study areas
174 (Table 1).



175

176 **Figure 1: Investigated case study areas in Switzerland and the distribution of the permanent sample-plots**
 177 **(coordinate system: EPSG 21781, CH 1903 / LV 03).**



178

179 **Figure 2: Local densities of the measured timber volume distributions on the field based sample-plots for**
 180 **Bremgarten (Brg), Glâne-Farzin (GF) and Sarine (SA). The scales for both axes differ among panels.**

181

	Bremgarten (Brg)	Glâne-Farzin (GF)	Sarine (SA)
Number of field based sample-plots	363	137	202
Number of sample-plots after cleaning (see section 2.6)	341	131	194
Number of sample-plots after cleaning and removing sample-plots containing larch	304	125	184
Theoretical grid size of sample-plots	80 m x 150 m	400 m x 400 m	400 m x 400 m
Exact measurement of the sample-plot centres	Yes, DGPS measurement with precision of 1 m available	No, only theoretical position known	
Recording method	400 m ² circle, min DBH threshold for recording: 12 cm (Schmid-Haas et al., 1993)	3 concentric circles: [I] 200 m ² circle with DBH threshold of 12 cm [II] 300 m ² circle with DBH threshold of 16 cm [III] 500 m ² circle with DBH threshold of 36 cm (Keller, 2013)	
Date of measurements on the sample-plots	Autumn – Winter 2011/2012	Autumn 2016	Autumn 2017
Date of the LiDAR flight	9.11.2011	07.10.2016 until 12.02.2017, mostly leaf-off condition but leaf-on also partially available	
Point density of the LiDAR raw data	≥ 8 points m ⁻²	≥ 5 points m ⁻²	≥ 5 points m ⁻²
GPS receiver and precision	DGPS receiver, 1 m precision	SXBlue II+ GNSS, 2.5 m horizontal precision	SXBlue II+ GNSS, 2.5 m horizontal precision

183 **Table 1: Properties of the study areas Bremgarten, Glâne-Farzin and Sarine (DBH = diameter at breast height,**
184 **1.3 m above the ground).**

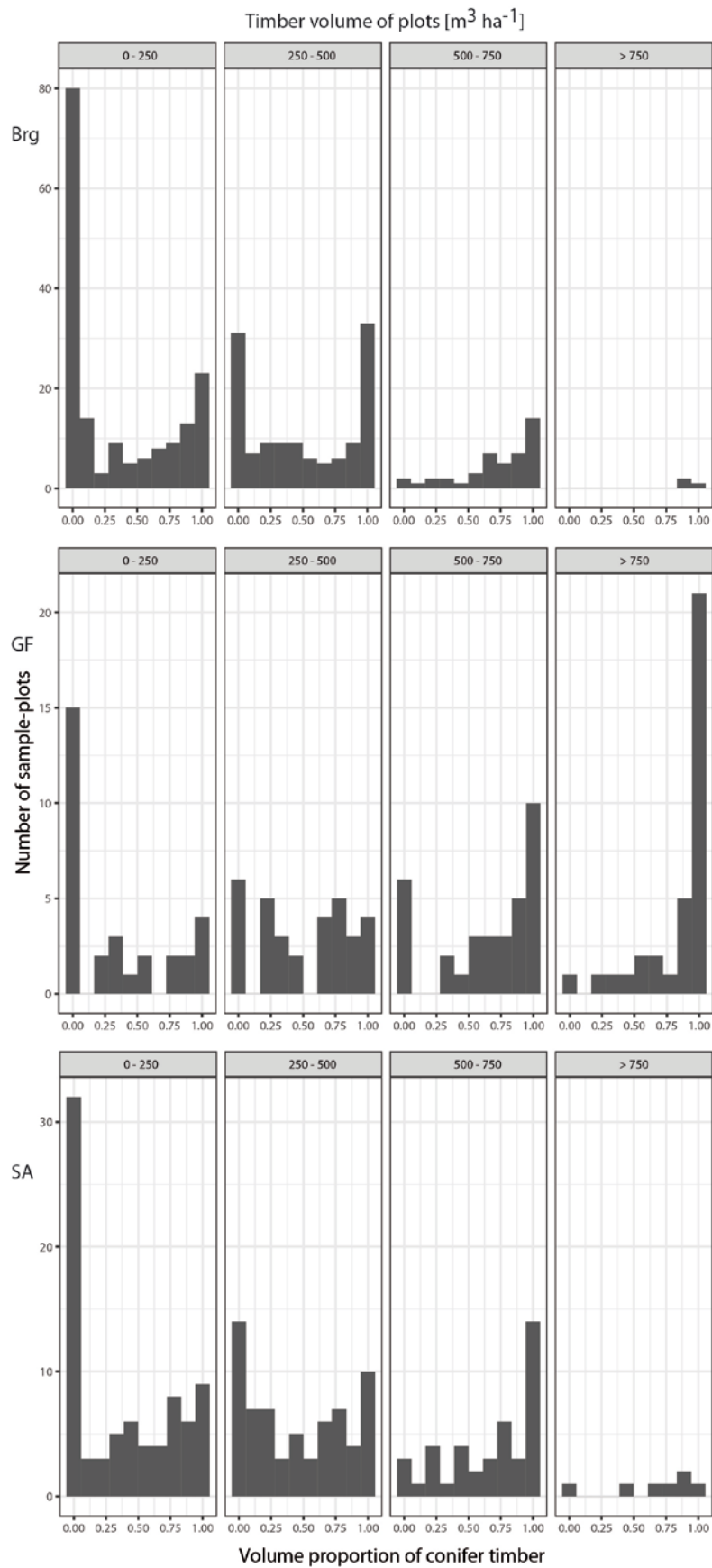
185 **2.1.2 Glâne-Farzin (GF)**

186 In GF, the inventory data from 2016 was used as input, which is based on 137 permanent
187 sample-plots on a 400 m x 400 m grid (Figure 1). With a horizontal accuracy of 2.5 m, the
188 sample-plot centres are less accurate than in Brg, as georeferencing with a differential GPS
189 (DGPS) was not available. However, this precision was provided by the manufacturer of the
190 GPS receiver, which might be too optimistic under a forest cover (Lamprecht et al., 2017).
191 Even with DGPS an accuracy of 2.5 m is not feasible (Lamprecht et al., 2017). Further
192 characteristics and properties of the study area are listed in Table 1. Compared with Brg, the
193 forests in GF are more homogeneous and mainly dominated by conifers. The mean timber
194 volume is 534 m³ ha⁻¹, 70% of which is conifer wood. Sample-plots with a large timber

195 volume are overrepresented compared with a normal age-class structure (Figure 2), and,
196 according to the volume ratio, conifers dominate these sample-plots (Figure 3). The LiDAR
197 flight was carried out between November 2016 and February 2017, and it again provided
198 excellent temporal synchronization between terrestrial and remote sensing data. In the
199 majority of cases, data was collected under leaf-off conditions. The density of the LiDAR raw
200 data is at least 5 points m⁻².

201 **2.1.3 Sarine (SA)**

202 The inventory in SA consists of 202 sample-plots and was last conducted in 2017. SA is, like
203 GF, located in the canton of Freiburg and thus has the same inventory design (Table 1). The
204 average timber volume of the sample-plots is 335 m³ ha⁻¹, 55% of which is conifer wood.
205 Inventory sample-plots indicate that stands with a small timber volume are slightly
206 overrepresented compared with a normal forest age-class structure (Figure 2). Overall,
207 however, forests in this study area are well balanced in terms of age-class structure.
208 Compared with GF, where most sample-plots are conifer dominated, most sample-plots in
209 SA are mixtures of conifer and broadleaf trees (Figure 3). The LiDAR data originates from the
210 same flight as used for GF and has the same characteristics.



211

212 **Figure 3: Proportion of conifer timber volume on the terrestrial sample-plots in Bremgarten (Brg), Glâne-**
 213 **Farzin (GF) and Sarine (SA) for different timber volume densities (< 250 , $250-500$, $500-750$ and $> 750 \text{ m}^3 \text{ha}^{-1}$).**
 214 **The scale on the y-axis differs among panels.**

2.2 Forest type maps (FTM)

We used five types of FTMs. Table 2 gives an overview of them and Figure 4 shows the information they contain and their resolution for a detail in the study area Brg. As some FTMs distinguish between trees that are foliated throughout the year and trees that are only foliated during the vegetation period (evergreen/deciduous) whereas others distinguish between conifers and broadleaf trees, Table 2 additionally indicates the forest type differentiation of each map. This differentiation is particularly relevant if larch (*Larix decidua*, a deciduous conifer) is present. However, to have a proper experimental design that differentiates between broadleaf and conifer trees for all FTMs, sample-plots that include larch were removed, as explained in detail in section 2.6.

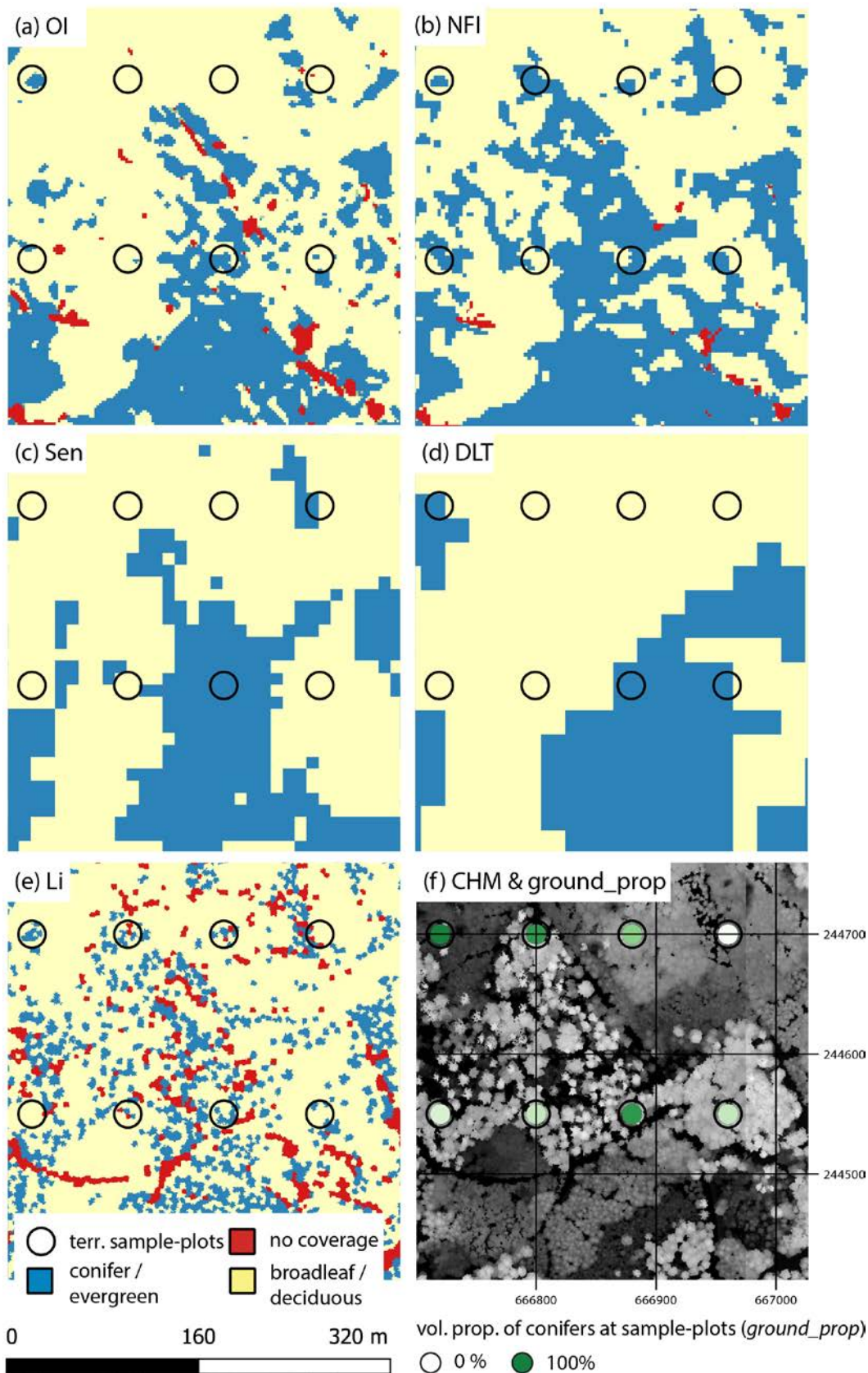
Description	Resolution [m]	Acquisition of raw data [year]	Forest type distinction
Orthoimage (OI)	2 m x 2 m	2014 (only Brg available)	evergreen / deciduous
Swiss NFI Orthoimage (NFI)	3 m x 3 m	2010–2016	conifer / broadleaf
Sentinel NFI (Sen)	10 m x 10 m	2016–2017	conifer / broadleaf
Copernicus Dominant Leaf Type (DLT)	20 m x 20 m	2012 (Brg) / 2015 (GF, SA)	conifer / broadleaf
LiDAR (based on return intensity leaf-off) (LI)	0.5 m x 0.5 m	2014 (Brg) / 2016 (GF, SA)	evergreen / deciduous

Table 2: Overview of the forest type maps (FTMs) used in the study. Brg = Bremgarten, GF = Glâne-Farzin (GF) and SA = Sarine.

2.2.1 Orthoimage (OI)

The FTM orthoimage (OI) is based on leaf-off (winter 2014) and leaf-on (summer 2014) digital aerial stereo imagery (Table 2). A digital surface model was calculated using the stereo-images from summer 2014. It was normalized with the digital terrain model 'swissAlti3D' to derive a canopy height model (CHM). This CHM was used to orthorectify the image datasets from both winter and summer 2014. For both orthoimages the normalized difference vegetation index (NDVI) was calculated. Thus, a winter dataset that represents evergreen vegetation taller than 3 m and a summer dataset that represents both evergreen

235 and broadleaf vegetation taller than 3 m were produced. The two datasets were combined
236 to retrieve the FTM. Therefore, this FTM distinguishes between evergreen and deciduous
237 trees (Figure 4a).



238

239 **Figure 4:** (a) – (e) Forest type maps (FTMs) and (f) canopy height model (CHM) for a sub-area around eight
 240 field based inventory sample-plots in the study area Bremgarten (coordinate system: EPSG 21781, CH 1903 /
 241 LV 03) (Li = LiDAR, Sen = Sentinel NFI, NFI = Swiss NFI Orthoimage, OI = Orthoimage, DLT = Copernicus
 242 Dominant Leaf Type). For (a) and (e) the FTM differentiates between evergreen and deciduous foliage; for
 243 (b), (c) and (d) the FTM differentiates between conifer and broadleaf trees.

244 **2.2.2 Swiss NFI Orthoimage (NFI)**

245 This map distinguishes between broadleaf and conifer trees (Waser et al., 2017) and was
246 produced for the Swiss National Forest Inventory (NFI). It covers all of Switzerland and has a
247 spatial resolution of 3 m x 3 m. The methodology is based on the classification of more than
248 1700 four-band aerial photo strips. Prior to applying it to all of Switzerland, different
249 classification methods, i.e. Support Vector Machine, Logistic Regression and Random Forest
250 (RF) were tested for selected areas in Switzerland regarding accuracy, computing time and
251 minimum number of required training values. The tests revealed that RF performed best
252 regarding accuracy and processing time (Waser et al., 2017; Figure 4b).

253 **2.2.3 Sentinel NFI (Sen)**

254 The FTM sentinel NFI is a national FTM that distinguishes between broadleaf and conifer
255 trees at a spatial resolution of 10 m.

256 This map is based on freely available spaceborne Sentinel-1 / -2 data from the European
257 Space Agency's Copernicus Programme (ESA, 2019). In order to get a cloud-free coverage of
258 all of Switzerland, a total of 50 Sentinel-2 images from June, July and August of 2016–2018
259 were used. The Sentinel-1 SAR data was acquired in summer 2016 and 2017 and processed,
260 i.e. flattening terrain and increasing the spatial resolution, according to Small (2012). This
261 map is a product of the Swiss NFI, and it is a follow-up and improved version of the
262 preliminary dataset (referred to as FTM NFI) and is free from overestimations of conifers
263 caused by topographic and illumination effects. This new approach incorporates a RF
264 classifier. According to Breiman (2001), this is a widely used ensemble classifier that
265 produces multiple decision trees using a randomly selected subset of training samples and
266 variables (Figure 4c).

267 **2.2.4 Copernicus Dominant Leaf Type (DLT)**

268 The FTM Dominant Leaf Type (DLT) is a product of the Copernicus Land Monitoring Service
269 coordinated by the European Environment Agency. The DLT FTM provides information on
270 the dominant leaf type (broadleaf or conifer) at a 20 m x 20 m resolution, and it was derived
271 from multi-temporal satellite image data using Support Vector Machine (Langanke, 2017;
272 Figure 4d).

273 **2.2.5 LiDAR (Li)**

274 The FTM LiDAR differentiates between deciduous and evergreen foliage and was computed
275 by using the corrected return intensity of leaf-off airborne laser scanning (ALS) data, as
276 proposed in the Digital Forestry Toolbox (Parkan, 2018). The FTM LiDAR has a spatial
277 resolution of 0.5 m and was computed for GF and SA from the same leaf-off data as used for
278 the CHM and for Brg from a leaf-off flight from March 2014 with a density of > 10 points m^{-2} .
279 This map has great potential for application because leaf-off ALS data is widely available and
280 calculation with the Digital Forestry Toolbox is straightforward (Figure 4e).

281 **2.3 Canopy height model (CHM)**

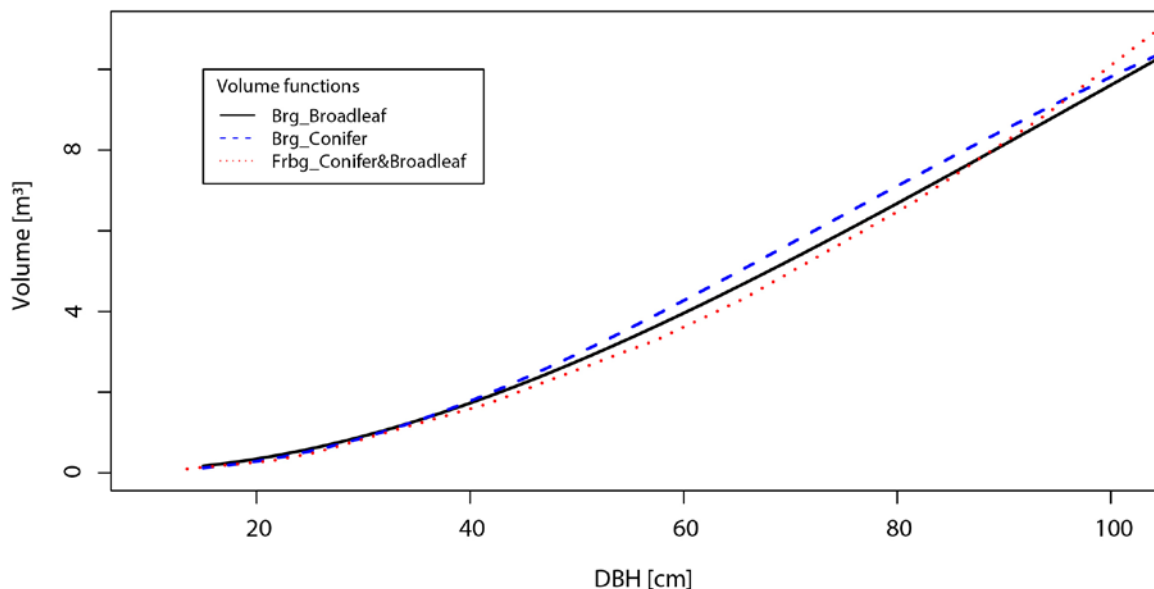
282 We used the LiDAR raw point data (Table 1) to compute pit-free CHMs. This approach for
283 CHM calculation avoids gaps ('pits'), following the methods presented by Isenburg (2014)
284 and Khosravipour et al. (2014). This algorithm was implemented in LAStools (Rapidlasso
285 GmbH, Gilching, Germany), but also using the package 'lidR' (Roussel and Auty 2017) in R
286 version 3.5 (R Core Team, 2018). An example of a pit-free CHM is given in Figure 4. The
287 CHMs produced in this study have a spatial resolution of 0.33 m (Figure 4f).

288 **2.4 Response variable**

289 The response variable is the local density of the merchantable timber volume [$m^3 ha^{-1}$]
290 (volume of the stem and branches with a diameter ≥ 7 cm) and is referred to as *VOL* in the
291 model formulation. The field measurements were carried out according to the sampling

292 protocol of Keller (2013) for GF and SA and of Schmid-Haas et al. (1993) for Brg. In all study
293 sites, only the DBH (diameter at breast height, measured at 130 cm above the level of the
294 terrain) was recorded. For the calculation of the timber volume of a single tree (*VOL_ST*) in
295 Brg, the following one-parameter volume function of Hoffmann (1982) was used:
296 $VOL_ST(DBH) = \exp(a_1 + a_2 \ln(DBH) + a_3 (\ln(DBH))^4)$, with the parameters a_1 , a_2 and
297 a_3 and with *DBH* as the only input variable. In GF and SA a volume function based on volume
298 tables was used (Schweizer, 2012), again with DBH as the only input variable. This volume
299 function was chosen to ensure comparability with earlier inventories. While in SA and GF the
300 same volume function was used for trees of both leaf types, in Brg separate volume
301 functions were used for broadleaf and coniferous trees. However, a comparison of the
302 functions showed that the differences in the predicted tree volumes between the volume
303 functions differentiating by forest type (Brg) and the mixed volume functions (GF and SA)
304 were marginal (Figure 5). Clear differences only emerged at ≥ 100 cm DBH. However, this has
305 no relevance for the present study, as only four sample trees were above this DBH threshold.
306 As volume functions are not valid for large areas, they have to be calibrated locally. The
307 parametrization of the volume function in Brg was done during the inventory of 1986. For
308 this purpose, the height (*h*), DBH and diameter at 7 m height (*d*₇) was measured for all
309 sample trees in one-eighth of each sample-plot area (for trees with a DBH < 20 cm), in one-
310 quarter of each sample-plot area (20 cm \leq DBH < 50 cm) or in the entire sample-plot (DBH \geq
311 50 cm) (Schmid-Haas et al., 1993). With the help of three-parameter volume functions
312 requiring *h*, DBH and *d*₇ as input, a reference volume was then calculated for all sample trees
313 that were measured in detail. Finally, a nonlinear curve fitting method was used to
314 determine the locally adapted parameters a_1 , a_2 and a_3 (Hoffmann, 1982).

315 To ensure comparability with previous inventories, we therefore used the parameters
 316 derived in 1986. A validation of the volume functions used in Brg was done by Kaufmann
 317 (2001). He reported a coefficient of determination (R^2) of 95% and a standard deviation of
 318 the residuals (R_s) of 20% (conifer) and 27% (broadleaf) of the mean. The local density of the
 319 total timber volume on each terrestrial sample-plot was based on the timber volume and
 320 inclusion probabilities of individual trees and was calculated using the Horvitz Thompson
 321 Estimator (Mandallaz, 2007).



322
 323 **Figure 5: Volume functions for Glâne-Farzin and Sarine in canton Freiburg (Frbg) and for Bremgarten (Brg).**

324 **2.5 Auxiliary variables**

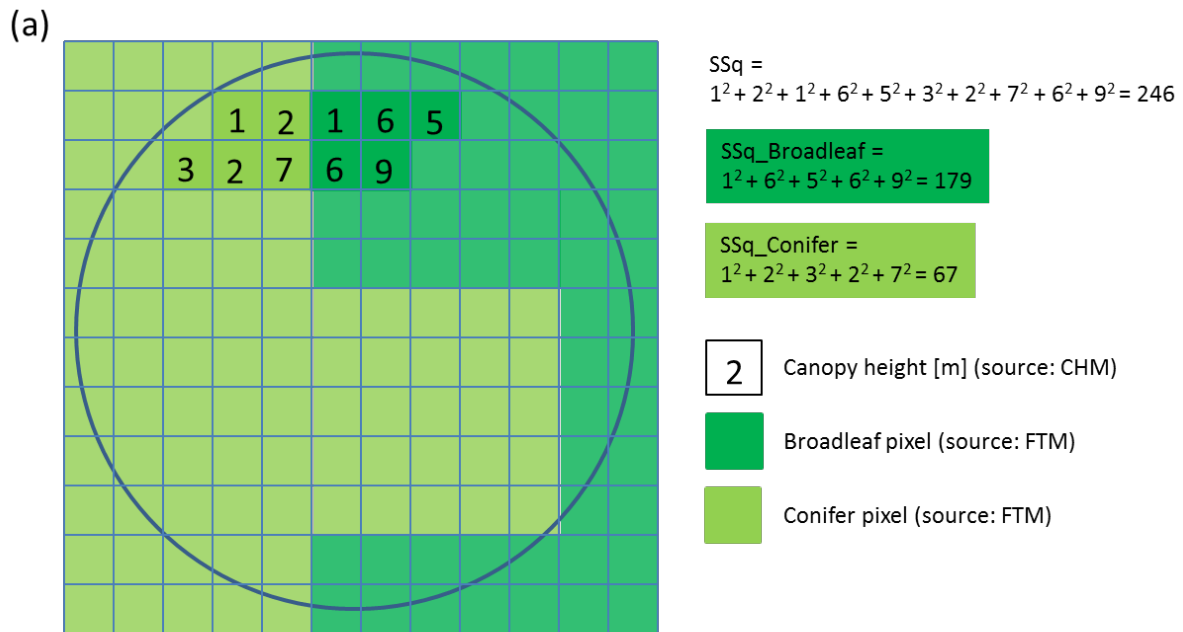
325 **2.5.1 Auxiliary variables based on canopy height models**

326 Auxiliary variables form the basis for model building and were used to derive a model for the
 327 response variable (*VOL*) from the sample data. Auxiliary variables can come from various
 328 sources, such as LiDAR raw data, CHMs and FTMs. An overview of auxiliary variables that are
 329 potentially interesting for forestry can be found in McCallum et al. (2014) and Saarela et al.
 330 (2015). Auxiliaries derived from the CHM, which were used in our models, are listed in Table
 331 3.

332 Most of the auxiliary variables included in this study are well established for forest inventory
 333 purposes. However, to the best of our knowledge, the variable *SSq* has not previously been
 334 described in the literature. It is presented in detail in Figure 6a and represents the sum of
 335 the squares of the height values of the individual CHM-pixels within a sample-plot. The idea
 336 behind introducing this auxiliary variable was to make it possible to map the allometric
 337 relationships more appropriately, as the stand volume is usually not a linear function of the
 338 canopy height (Eichhorn, 2013; Pretzsch, 2001).

	Description	Abbreviation	Unit
Percentile – values	Specifies the height percentiles of the pixels in the sample-plot. Example: 95th percentile: P95	PXX XX = {99, 95, 90, 80, 75, 70, 60, ... , 30, 25, 20, 10, 05, 01}	[m] (same unit as the CHM)
Coverage – values	Indicates the vegetation cover at a certain height. Example: 2 m coverage: C02 = 0.9 means that 90% of the pixels of the vegetation height model in the plot are higher than or equal to 2 m above the ground.	CXX XX = {00, 02, 05, 10, 15, 20, ... , 50}	[0 – 1]
Minimum	Minimum pixel value within a plot	MIN	[m]
Maximum	Maximum pixel value within a plot	MAX	[m]
Standard deviation	Standard deviation of the individual pixels within a plot	STD	[m]
Average	Average value (height) of the individual pixels within a plot	MEAN	[m]
Sum of squares	Sum of the squares of the individual pixels within a plot	SSq	
Average sum of squares	Sum of squares divided by number of pixels in a plot	MSSq	
Coefficient of variation	MEAN / STD	CV	

339 **Table 3: Auxiliary variables derived from the canopy height model (CHM).**



(b)

	①	②	③
	number of pixels	sum of canopy height	SSq of canopy height
conifer	5	15	67
broadleaf	5	27	179
total	10	42	246
↓	area_prop	area_chm_prop	area_chm_sq_prop
proportion of conifer	0.50	0.36	0.27

340

341 **Figure 6: Principle of the combined evaluation of different grids. The two grids forest type map (FTM)**
 342 **(conifer/broadleaf) and canopy height model (CHM) are superimposed and evaluated for each attribute of**
 343 **the FTM. a) Illustration of the derivation of the auxiliary variable SSq (sum of squares), based on a combined**
 344 **evaluation of the CHM and the FTM. The circle represents the boundary of the sample-plot. b) Computation**
 345 **of the different forest type variable alternatives: 1) ratio of the number of pixels (*area_prop*), 2) ratio of the**
 346 **sums of canopy height (*area_chm_prop*), and 3) ratio of the sums of squares of the canopy heights**
 347 **(*area_chm_sq_prop*). Most pixels in (a) are empty to simplify the comprehensibility of the calculations. If the**
 348 **FTM and the CHM have different spatial resolutions, the analysis is performed on the basis of the grid with a**
 349 **higher spatial resolution.**

350 2.5.2 Forest type variables (FTV)

351 In order to include FTM information in the statistical models, auxiliaries (referred to as forest
 352 type variables, FTV) were introduced that describe the proportion of conifer trees. There
 353 were five FTV alternatives: four continuous and one categorical. The continuous FTV
 354 alternatives were: (1) the ratio of the number of pixels of different forest types (*area_prop*),
 355 (2) the ratio of average heights of the forest types (*area_chm_prop*), (3) the ratio of the sum
 356 of squares of the canopy heights of the forest types (*area_chm_sq_prop*) (see Figure 6b),

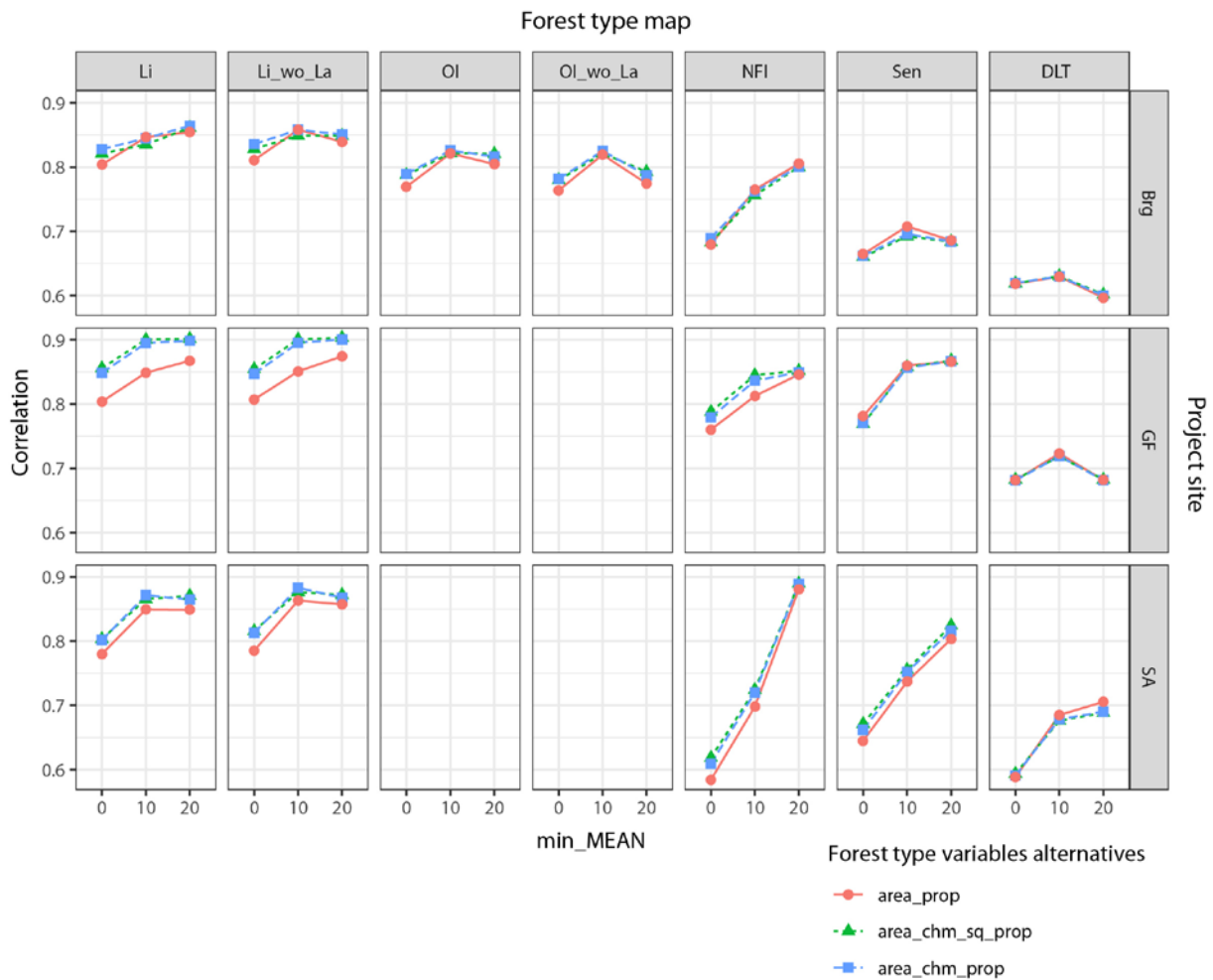
357 and (4) the variable *ground_proportion*, which is introduced in section 2.5.3. *area_prop* is
358 the classical formulation and represents the proportion of the area that is covered by the
359 canopies of conifers. *area_chm_prop* and *area_chm_sq_prop* were our new ‘weighted-
360 canopy-height proportion’ FTVs, for which the area proportion was weighted by information
361 from a CHM. Therefore, the two grids (CHM and FTM) were superimposed and a separate
362 evaluation was performed for each forest type, as presented in Figure 6a and b. For the
363 categorical variable, we chose a setup that assigns the sample-plot to a main forest type
364 (broadleaf or conifer) if the proportion of the pixels of one particular type exceeds 70% and
365 to the category ‘mixed’ otherwise, as done by Breidenbach et al. (2008) and Latifi et al.
366 (2012). Other thresholds were not explored in this study, even though they may also have an
367 impact on model performance (Hill et al. 2018). However, this study was not focused on the
368 investigation of categorical variables, as categorical variables have the disadvantage of
369 requiring an additional degree of freedom for each level. This effect is amplified when
370 interaction terms between the FTV and other auxiliary variables are included in the model.
371 This is particularly problematic when the number of sample-plots is small, as in this study. A
372 rule of thumb says that 10–20 observations per auxiliary variable are required to be able to
373 detect reasonable-sized effects with reasonable power (Harrell, 2017).

374 **2.5.3 Forest type variable *ground_proportion* and classification accuracy**

375 FTVs derived from FTMs (section 2.5.2) are usually not error free, as shown, for example, by
376 Hill et al. (2018) and Straub et al. (2009). We evaluated classification accuracy to understand
377 the precision and the errors of the different FTVs and FTMs. This was done by evaluating the
378 correlation of the continuous FTVs *area_prop*, *area_chm_prop* and *area_chm_sq_prop* with
379 *ground_proportion* (Figure 7). *ground_proportion* represented the conifer timber volume
380 proportion for each sample-plot based on field measurements. This variable was derived

381 from the volumes of the individual trees, which were predicted by means of the volume
382 functions with DBH as an input variable (see section 2.4). *ground_proportion* was not an
383 error-free variable, since the underlying volume functions were not error free, but field
384 measurements are usually more precise than variables derived from remote sensing data.

385 The analyses showed that the LiDAR-derived FTVs had the strongest correlation with
386 *ground_proportion* for all study areas. The FTVs *area_chm_sq_prop* and *area_chm_prop*
387 were more reliable than *area_prop*, except for the FTMs Sen and DLT, which had low
388 resolutions of 10 m and 20 m, respectively. The classification accuracy generally increased
389 with average canopy height in the sample-plot. Therefore, including an interaction term
390 between the FTVs and *MEAN* had the potential to further improve model accuracy, and it
391 was considered in some model formulations. To explore the potential benefit of having a
392 highly precise FTM (i.e. without classification error) based on field measurements, we
393 considered *ground_proportion* as a FTV in our statistical models. Although classification
394 errors of forest type are typically small when they are based on accurate field observations,
395 the *ground_proportion* FTV used in our study was not error free because its derivation was
396 based on volume estimations from volume functions.



397

398 **Figure 7: Correlation of the proportion of conifer timber volume, derived using the different FTMs and FTV**
 399 **alternatives compared with the variable *ground_proportion*. The figure is separated into different average**
 400 **classes of the CHM that were considered. min_MEAN = 20 means that only sample-plots with a minimum**
 401 **average canopy height of 20 m were considered; a value of 0 means that all sample-plots were considered**
 402 **for the analysis of the correlation. Brg = Bremgarten, GF = Glâne-Farzin (GF) and SA = Sarine.**

403 2.6 Statistical methods

404 For model building, we excluded sample-plots with a timber volume value of zero and
 405 sample-plots in which harvests had taken place in the period between the field
 406 measurements and the acquisition of the remote sensing data. This step is referred to as
 407 ‘cleaning’ in Table 1. The FTMs derived from orthoimage (OI) and LiDAR (Li) differentiated
 408 between deciduous and evergreen foliage. To differentiate accurately between conifer and
 409 broadleaf trees for these FTMs, we removed sample-plots that include larch (*Larix decidua*).
 410 We labelled these cases with the suffix ‘wo_La’ (without larch). The number of remaining

411 sample-plots in each study area is listed in Table 1. Each FTM, except OI, was tested in all
 412 study areas. The abbreviations of all study area and FTM combinations are given in Table 4.

	Orthoimage (OI)	Swiss NFI Orthoimage (NFI)	Sentinel NFI (Sen)	LiDAR (return intensity) (Li)	Copernicus Dominant Leaf Type (DLT)
Glâne-Farzin	-	GF_NFI	GF_Sen	GF_Li / GF_Li_wo_La	GF_DLT
Sarine	-	SA_NFI	SA_Sen	SA_Li / SA_Li_wo_La	SA_DLT
Bremgarten	Brg_OI / Brg_OI_wo_La	Brg_NFI	Brg_Sen	Brg_Li / Brg_Li_wo_La	Brg_DLT

413 **Table 4: Combinations of study areas and FTMs considered in the statistical analyses.**

414 Root mean square error (RMSE) from leave-one-out cross-validation (LOOCV) was used to
 415 denote model accuracy (Equation 1). Table 5 lists the multiple linear regression models used
 416 in this study, as basic formulations without FTVs. Each basic formulation was combined with
 417 each FTV alternative. To evaluate the gain in model performance when any forest type
 418 information was included, each basic formulation with no FTVs was analysed. The statistical
 419 software R (version 3.5) was used for model analyses (R Core Team, 2018).

$$420 \quad \mathbf{RMSE} = \sqrt{\frac{\sum_{x \in s} (\hat{Y}(x) - Y(x))^2}{n}} \quad (\text{Equation 1})$$

421 where $Y(x)$ is the observed local density of the timber volume on the sample-plot level at
 422 location $x \in s$ [$\text{m}^3 \text{ha}^{-1}$], $\hat{Y}(x)$ is the predicted volume density on the sample-plot level at
 423 location $x \in s$ [$\text{m}^3 \text{ha}^{-1}$], and s is the modelling dataset composed of n sample-plots.

424 The results are discussed in terms of the relative RMSE, defined as the RMSE relative to the
 425 mean y_{mean} of the observed values (Equation 2).

$$426 \quad \mathbf{RMSE}[\%] = \frac{\mathbf{RMSE}}{y_{mean}} * 100, \text{ with } y_{mean} = \frac{1}{n} \sum_{x \in s} Y(x) \quad (\text{Equation 2})$$

427 The OLS regression model formulation is defined in Equation 3

428 $Y(x) = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_p Z_p + \varepsilon(Z)$ (Equation 3)

429 with error term $\varepsilon(Z)$ independent and identically distributed, $E(\varepsilon(Z)) = 0$, and

430 $Var(\varepsilon(Z)) = \sigma^2$,

431 where $Y(x)$ is the response (see Equation 1), $\beta_0 \dots \beta_p$ are the regression coefficients, $Z_1 \dots Z_p$

432 denote the auxiliary variables, and p is the number of auxiliary variables (explanatory or

433 predictor variables). All models are displayed using the following R-style formatting:

434 $Y \sim Z_1 + Z_2 + \dots + Z_p$ (see Table 5).

Model #	Basic formulation
1	VOL ~ MEAN
2	VOL ~ MEAN + STD
3	VOL ~ MEAN + STD + C02
4	VOL ~ MEAN + STD + P90 + C30
5	VOL ~ MEAN + STD + C30 + C02
6	VOL ~ MEAN + STD + C30 + C02 + MEAN:STD
7	The model variables were selected based on the Akaike Information Criterion (AIC; Akaike, 2011). All variables from Table 3, such as <i>MEAN</i> , <i>STD</i> , <i>CXX</i> and <i>PXX</i> , were available for selection. The selected variables are listed in Table A1 in the Appendix.
8	Same variables as Model #7, but with the following interaction terms: MEAN + MEAN:STD + (MEAN:foresttype) + (MEAN^2:foresttype) After adding the interaction terms, variable selection based on the AIC was again performed (for each forest type alternative)
9	Same variables as #7, but with the following interaction terms: MEAN + MEAN:STD + (MEAN:foresttype) + (MEAN^2:foresttype) VOL ~ MEAN + STD + MEAN:STD
10	For the forest types with the following interaction terms: foresttype + (MEAN:foresttype)+(MEAN^2:foresttype) variable selection based on the AIC was performed (for each forest type alternative) VOL ~ MEAN + STD + MEAN:STD
11	For the forest types with the following interaction terms: foresttype + (MEAN:foresttype)+(MEAN^2:foresttype) (same variables as #10, but without AIC variable selection)

435 **Table 5: Basic model formulations for the multiple linear regression. The interaction terms are indicated**
 436 **by ‘:’. ‘foresttype’ stands for the forest type variable and is either *area_prop*, *area_chm_prop*,**
 437 ***area_chm_prop*, *categorical* or *ground_proportion*.**

438 Models 1–5 were quite simply models without interaction terms (Table 5). Model 6 included

439 the interaction between *MEAN* and *STD*. Model 7 was evaluated by performing variable

440 selection based on the Akaike Information Criterion (AIC; Akaike, 2011). Models 8 and 9

441 included further interaction terms. For Model 8, variable selection (based on AIC) was

442 performed a second time after the interaction terms were added. Models 10 and 11 had

443 only three variables but included interaction terms. Variable selection (based on AIC) was
444 performed to find a satisfactory relationship between the goodness of fit and the simplicity
445 of the model in Model 10 but not in Model 11. The formulations for the best performing
446 model (Model 10) are given in Table 7.

447 We additionally computed the adjusted coefficient of determination (adjusted r-squared) to
448 facilitate comparisons with other related publications. The adjusted r-squared values were of
449 limited use for the evaluation of the models, as over-fitted models, such as Models 9 and 11,
450 also had high adjusted r-squared values. Models with many variables are generally at risk of
451 over-fitting. Therefore, cross-validation with RMSE % was considered better suited for the
452 evaluation of our OLS regression models. We did not evaluate results based on the adjusted
453 r-squared values, but these values can be found in the Appendix (Figures A1 and A2).

454 **3 Results**

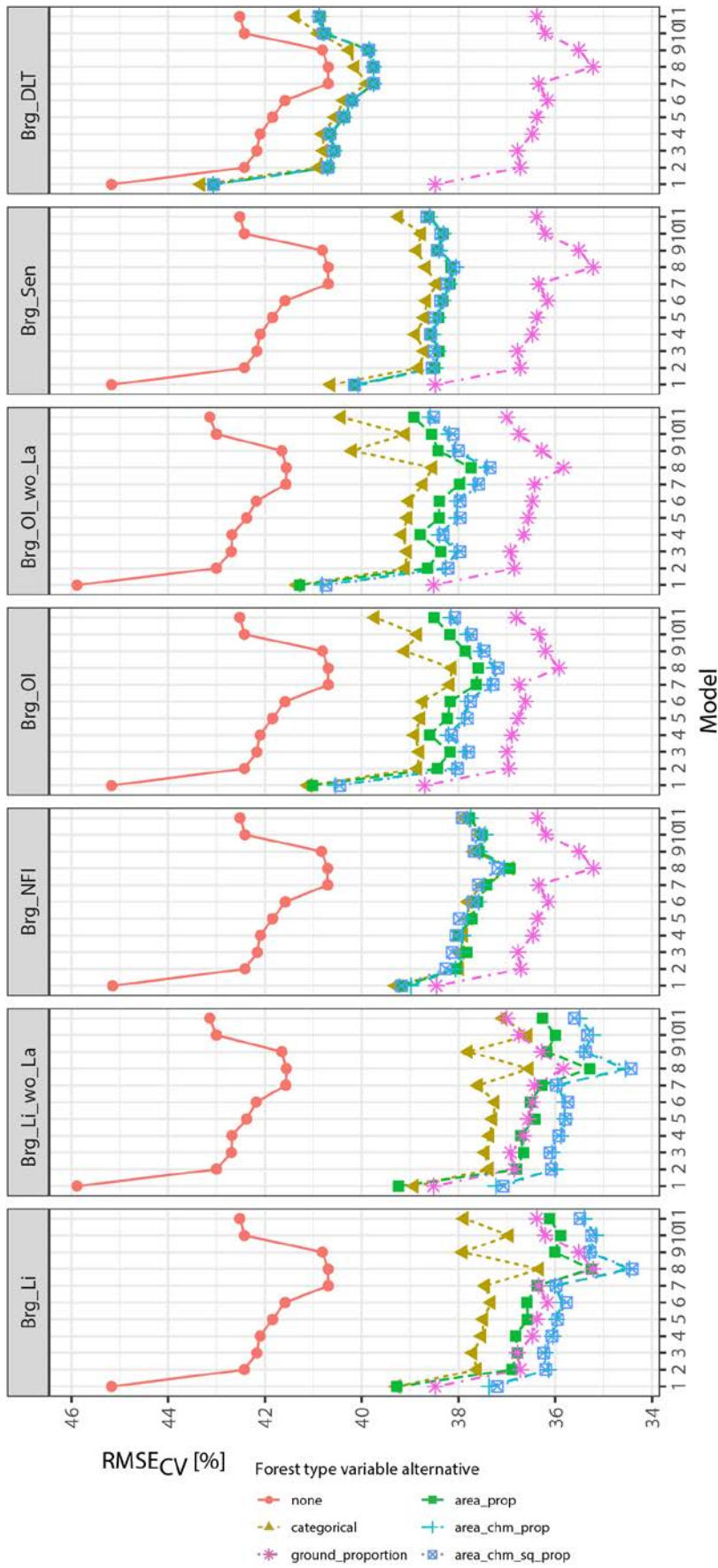
455 **3.1 Bremgarten (Brg)**

456 The results of the leave-one-out cross-validation (LOOCV) (RMSE %) are depicted in Figure 8,
457 which shows the accuracies of the OLS regression models, defined in Table 5, achieved for
458 models including all combinations of the four different FTV alternatives and without FTVs
459 ('None').

460 Including the FTM information substantially reduced the cross-validated RMSE for all
461 models. The best model performance (lowest RMSE) for all FTMs was reached with Model 8
462 (Figure 8), where a double stepwise selection procedure was applied that also allowed
463 interaction terms to remain in the model. The best model performance was achieved by
464 adding FTVs based on the LiDAR dataset (Brg_Li and Brg_Li_wo_La). The reduction in RMSE
465 was up to 9 percentage points in comparison to models without any FTM information (Figure
466 9). Further, the FTV alternative influenced the accuracy of the model. The FTV alternatives

467 *area_chm_sq_prop* and *area_chm_prop* showed similar results and performed better than
468 *area_prop*, which in turn outperformed the categorical variable. It is quite remarkable that
469 models including the FTV alternatives *area_chm_sq_prop* and *area_chm_prop* derived from
470 LiDAR data outperformed the model with *ground_proportion* (conifer timber volume
471 proportion, derived from field measurements). Model performance improved further when
472 inventory sample-plots with larch were excluded (Brg_Li_wo_La; Figure 8).

473 The FTVs derived from the orthoimages (Brg_NFI, Brg_OI and Brg_OI_wo_La) reduced the
474 RMSE by 4–6 percentage points in comparison to models without any FTM information (e.g.
475 from 45% to 39% for Model 1 in Brg_NFI). Whereas almost no difference could be observed
476 among models including the different FTV alternatives for Brg_NFI, slight differences existed
477 among the models depending on the FTV alternative included for Brg_OI and Brg_OI_wo_La,
478 which were based on a 2 m x 2 m orthoimage. Models based on Sentinel data (Brg_Sen)
479 showed slightly higher RMSE values compared with those derived from orthoimages. The
480 lowest model accuracy was achieved by incorporating FTVs derived from the DLT FTM, with
481 a model improvement of only of 1–2 percentage points.



482

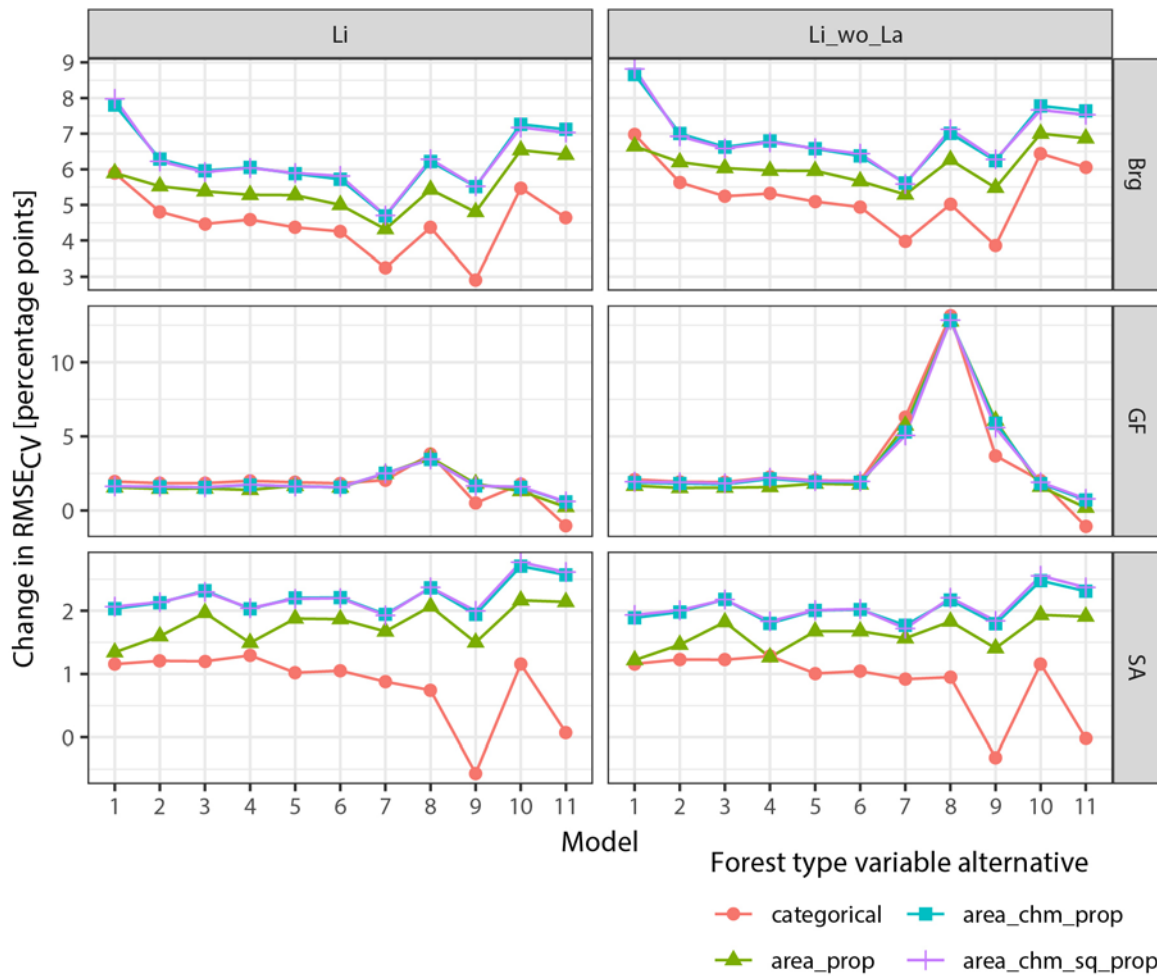
483 **Figure 8: Leave-one-out cross-validated RMSE_{CV} [%] for each regression model, FTM and combination of FTV**
 484 **alternatives in the case study area Bremgarten (Brg).**

485 **3.2 Glâne-Farzin and Sarine**

486 When interpreting the results of GF and SA, it should be kept in mind that the positional
487 accuracy of the sample-plot centres is less precise (Table 1). Thus, we expected smaller
488 improvements in model accuracy by adding FTVs. The RMSE values for GF and SA are
489 presented in Figure 10. Like in Brg, Model 8 achieved the lowest RMSE overall, and for SA_Li
490 Model 10 was equivalent to Model 8. The largest model improvement was reached by
491 adding the LiDAR based FTVs (SA_Li and SA_Li_wo_La) to the regression models and by using
492 the FTV alternatives *area_chm_sq_prop* and *area_chm_prop*. Including these variables led to
493 better model performance than when *ground_proportion* was included. The gain in model
494 accuracy was more than 2.5 percentage points (Model 10; Figure 9). Auxiliary variable
495 combinations of the remaining FTVs (SA_NFI, SA_Sen and SA_DLT) showed similar patterns
496 and reduced the RMSE by about 1–2 percentage points. Comparing the results of GF with
497 the other study areas, the RMSE without FTVs was already lower (RMSE \approx 32%, Figure 10)
498 than the RMSE of the best models containing FTVs in Brg (RMSE \approx 34%) or SA (RMSE \approx 36%).
499 Including forest type information in the regression models led to an improvement in model
500 performance (Figure 10). FTVs derived from the LiDAR FTM showed again the lowest RMSE
501 values. In comparison to the other study areas, including FTVs had a smaller effect on model
502 accuracy for GF. In some cases the categorical variable performed best (GF_Li) by a very
503 small margin. However, including FTM information in the model resulted in a decrease in
504 RMSE in the best case (RMSE_{CV} \approx 28% for Model 8 and GF_Li_wo_La; Figure 10). Contrary to
505 results for SA and Brg, including *ground_proportion* as a predictor resulted in a slightly better
506 model performance than using FTVs from maps. GF was the study area with the smallest
507 number of sample-plots; for this reason, GF had the highest risk of model over-fitting among
508 the three study areas, and over-fitting was also detected for the combination of Models 7–9,

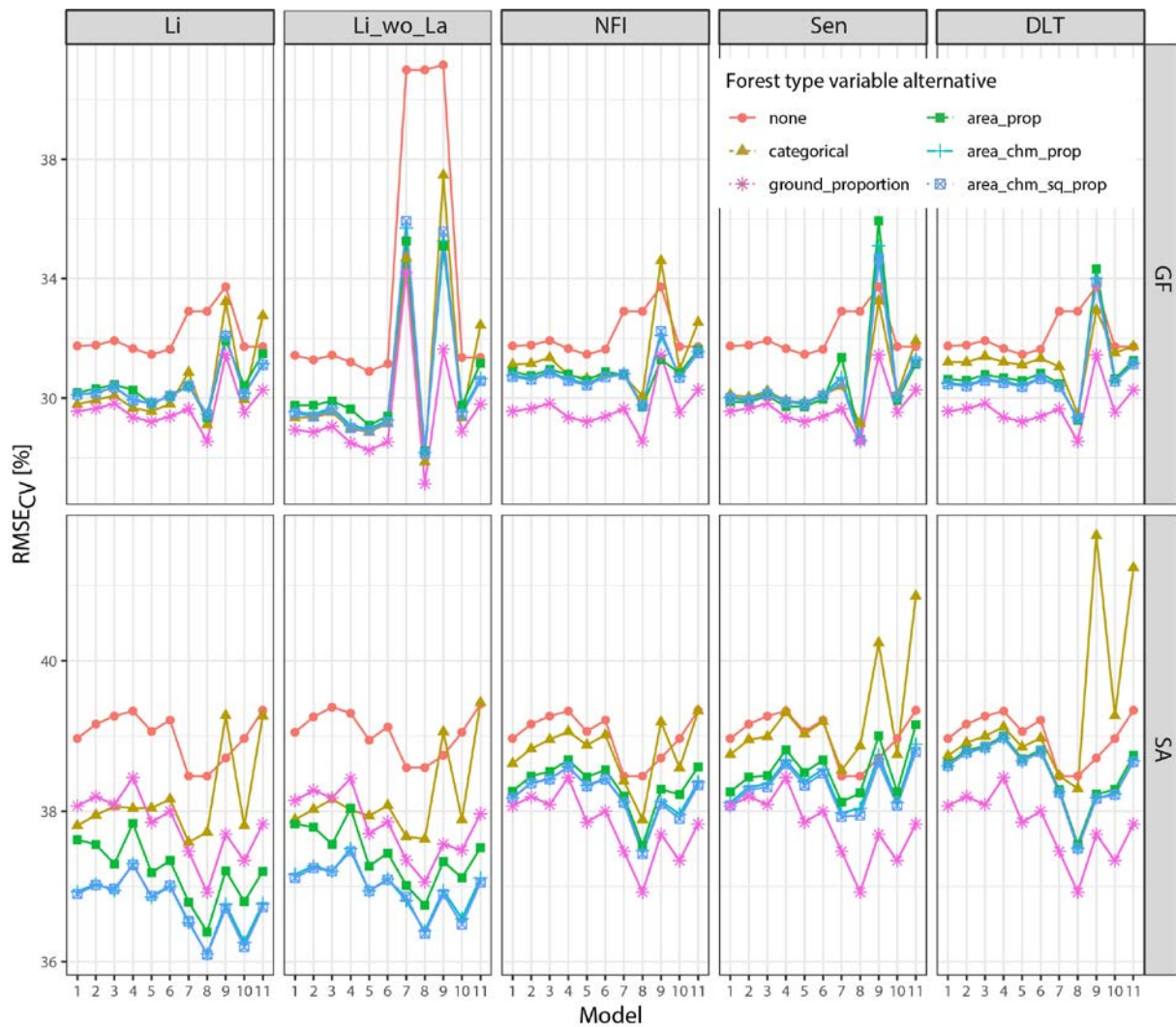
509 FTM Li_wo_La and without FTV. The adjusted r-squared values (Figure A2 in the Appendix)
510 indicate an excellent model performance (high adjusted r-squared value), whereas the high
511 $RMSE_{CV}$ value in Figure 10 indicates the opposite. However, this was the only case of over-
512 fitting. A further comparison of adjusted r-squared values (Figure A2) demonstrates that the
513 remaining models with the highest adjusted r-squared values (Model 9 and GF_Li_wo_La)
514 had rather low RMSE values.

515 Table 6 shows the model formulations for the best performing models. Although some
516 models had many predictors, they were not over-fitted, as shown through the leave-one-out
517 cross-validated RMSE (Figures 8 and 10). For Brg and SA, Model 10 provided the second best
518 model performance according to RMSE, which was only slightly worse than that of Model 8
519 (Table 7) but included far fewer variables.



520

521 **Figure 9: Change in $RMSE_{cv}$ [%] values under the different model types and forest type variables (FTV)**
 522 **alternatives compared with a model without FTVs in the three case study areas (Brg = Bremgarten, GF =**
 523 **Glâne-Farzin, SA = Sarine). Results are only displayed for variables derived from the LiDAR FTM. The peak in**
 524 **GF and Li_wo_La for Models 7–9 was caused by over-fitted models without FTVs.**



525

526 **Figure 10: Leave-one-out cross-validated $RMSE_{CV}$ [%] for each regression model, forest type map and**
 527 **combination of forest type variable alternatives in the case study areas Glâne-Farzin (GF) and Sarine (SA).**

Study area	Model	Forest type variable	Auxiliary variables used
Brg_Li	8	area_chm_sq_prop	P40 + P30 + P25 + P01 + C10 + C20 + foresttype + MEAN + foresttype:MEAN
Brg_Li	8	area_chm_prop	P40 + P30 + P25 + P01 + C10 + C20 + foresttype + MEAN + foresttype:MEAN
Brg_Li_wo_La	8	area_chm_sq_prop	P99 + P60 + P01 + C10 + C20 + CV + foresttype + MEAN + foresttype:MEAN
Brg_Li_wo_La	8	area_chm_prop	P99 + P60 + P50 + P30 + P01 + C10 + C20 + CV + foresttype + MEAN + foresttype:MEAN
GF_Li	8	categorical	P60 + P25 + C02 + C10 + C15 + C25 + C30 + foresttype + MEAN
GF_Li_wo_La	8	categorical	C02 + C10 + C15 + C25 + C30 + foresttype + MEAN + foresttype:MEAN
SA_Li	8	area_chm_sq_prop	P20 + C05 + C15 + C20 + C25 + C35 + C50 + foresttype + MEAN + foresttype:MEAN

SA_Li	8	area_chm_prop	P20 + C05 + C15 + C20 + C25 + C35 + C50 + foresttype + MEAN + foresttype:MEAN
SA_Li_wo_La	8	area_chm_sq_prop	MSSq + P25 + P05 + C02 + C15 + C20 + C25 + C35 + C50 + chm_sq_average + foresttype + MEAN + foresttype:MEAN
SA_Li_wo_La	8	area_chm_prop	MSSq + P25 + P05 + C02 + C15 + C20 + C25 + C35 + C50 + chm_sq_average + foresttype + MEAN + foresttype:MEAN

528 **Table 6: Selected model formulations for the best performing forest type model combinations for Model 8.**
529 **Interaction terms are indicated by ‘:’. ‘foresttype’ stands for the forest type variable and is either *area_prop*,**
530 ***area_chm_prop*, *area_chm_prop*, *categorical* or *ground_proportion*.**

531

Study area	Model	Forest type variable	Auxiliary variables used
Brg_Li	10	area_chm_sq_prop	MEAN + STD + foresttype + MEAN:foresttype
Brg_Li	10	area_chm_prop	MEAN + STD + foresttype + MEAN:foresttype
Brg_Li_wo_La	10	area_chm_sq_prop	MEAN + STD + foresttype + MEAN:foresttype
Brg_Li_wo_La	10	area_chm_prop	MEAN + STD + foresttype + MEAN:foresttype
GF_Li	10	categorical	MEAN + STD + foresttype + MEAN:STD
GF_Li_wo_La	10	categorical	MEAN + foresttype
SA_Li	10	area_chm_sq_prop	MEAN + foresttype + MEAN:foresttype
SA_Li	10	area_chm_prop	MEAN + foresttype + MEAN:foresttype
SA_Li_wo_La	10	area_chm_sq_prop	MEAN + foresttype + MEAN:foresttype
SA_Li_wo_La	10	area_chm_prop	MEAN + foresttype + MEAN:foresttype

532 **Table 7: Selected model formulations for the best performing forest type model combinations for Model 10.**
533 **Interaction terms are indicated by ‘:’. ‘foresttype’ stands for the forest type variable and is either *area_prop*,**
534 ***area_chm_prop*, *area_chm_prop*, *categorical* or *ground_proportion*.**

535 **4 Discussion**

536 **4.1 Influence of different forest type variables and forest type maps on model** 537 **performance**

538 Our results confirm the findings of previous studies (Breidenbach et al., 2008; Latifi et al.,
539 2012) that incorporated FTM information to predict timber volume. Both earlier studies
540 reported an increase in model accuracy (measured as the LOOCV RMSE) of the OLS
541 regression models for their investigated study areas. In doing so, Latifi et al. (2012) found an
542 improvement of 2–4 percentage points in RMSE when stratifying into broadleaf and conifer

543 forest types. We found similar improvements when the categorical FTV was included in
544 models (Figure 9). In our study, the best performing FTM in all study areas was the one
545 derived from leaf-off LiDAR data. This is not particularly surprising, as this map had both the
546 highest resolution and the best correlation with the values derived from field measurements
547 (*ground_proportion*), as visualized in Figure 7. The further order of the best suited FTMs was
548 less clear and varied between the study areas. In general, a map with a higher resolution
549 gave better results than one with a coarser resolution, even if both maps showed the same
550 classification accuracy, as derived in section 2.5.3. This was observed in Brg (Figure 8), where
551 we had the following order of accuracy according to the RMSE: Brg_OI (2 m x 2 m) \approx Brg_NFI
552 (3 m x 3 m) > Brg_Sen (10 m x 10 m) > Brg_DLT (20 m x 20 m). The NFI Orthoimage map
553 (*_NFI) is biased due to topographic and illumination effects, i.e. shadows, of optical remote
554 sensing data (Waser et al., 2017).

555 We found that the selection of a FTV alternative strongly affected model accuracy,
556 particularly when FTMs with a high resolution (≤ 2 m) were used, such as the LiDAR-derived
557 map (0.5 m). The continuous variables thus outperformed the categorical variable. Among
558 the continuous variables, *area_chm_prop* and *area_chm_sq_prop* performed about equal
559 and were associated with higher model accuracies than *area_prop*. For *area_chm_prop* and
560 *area_chm_sq_prop*, the forest type proportion was weighted by canopy height information
561 of a CHM. This representation mapped the volume proportions of conifers and broadleaf
562 trees better, in particular for mixed and heterogeneously structured forests. The difference
563 in RMSE between all the different FTV alternatives was up to 2 percentage points. However,
564 in GF there was no clear difference between the different FTV alternatives, possibly because
565 exactly georeferenced sample-plot positions were not available and the model accuracy was
566 already high without incorporating FTVs (Figure 9). The gain in model accuracy was up to 9

567 RMSE percentage points for Brg, and up to 3 percentage points for GF and SA. The
568 improvement in model accuracy, measured as percentage points of the RMSE, was lower in
569 GF and SA compared with Brg, as the results from GF and SA already showed a small RMSE
570 before any forest type information was added to the OLS regression model.

571 **4.2 Best performing model formulation**

572 Our second research question was how to best include FTM information in a regression
573 model. Model 8 performed best for all study areas and FTV combinations. This model was
574 calibrated using a variable selection procedure based on AIC values, followed by the addition
575 of interaction terms, and finally a second variable selection based on AIC values. Except for
576 GF_Li, all of the best performing models included an interaction term between *foresttype*
577 and *MEAN* (Table 6). Therefore, we highly recommend using interaction terms when
578 including FTM information in OLS regression models. Interaction terms are particularly
579 meaningful because, depending on the average vegetation height, the FTMs have a different
580 classification accuracy (Figure 7).

581 Although Model 8 included several variables, over-fitting was not an issue, as shown by the
582 cross-validation results. However, if we were interested in making inferences about the
583 predictors, Model 8 could have failed. In such cases, Model 10 would have been the
584 preferred model formulation because it included remarkably fewer variables (Table 7). Due
585 to its smaller number of variables, Model 10 was also less vulnerable to over-fitting. Further,
586 variance inflation, which occurs as a consequence of including too many correlated variables
587 and is quantified as the variance inflation factor (VIF), could affect the significance of single
588 variables (Fox and Monette, 1992). This was not the case in our study, however, because
589 Model 10 had the best AIC and highly correlated variables were eliminated. The only issue
590 could be that main effects are difficult to interpret in models with interactions.

591 **4.3 Potential of a high-precision forest type variable**

592 The idea behind introducing *ground_proportion* (conifer volume proportion based on field
593 measurements) as a FTV in the models was to explore the potential benefit of having a high-
594 precision FTM without misclassifications. Results from the two study areas Brg and SA show
595 that when the LiDAR-derived FTM is used in combination with the variables *area_chm_prop*
596 and *area_chm_sq_prop*, regression model performance can be even better than when this
597 FTM is used in combination with *ground_proportion*. This result is remarkable and deserves
598 an in-depth discussion. The following issues are relevant for its understanding and
599 interpretation.

600 First of all, it must be emphasized that both the variable *ground_proportion* and the remote-
601 sensing-based FTV contained errors. In the case of *ground_proportion*, the errors were
602 caused by the volume functions used, which were solely based on the DBH. For example,
603 Kaufmann (2001) explored volume functions for the Swiss NFI. He stated that a tree in a later
604 stage of development has a larger stem volume than a tree with the same DBH in an earlier
605 developmental stage. This behaviour was ignored in the volume functions, but we think that
606 this relationship was mapped in the LiDAR-based 'weighted-canopy-height proportion' FTV,
607 as the canopy height is correlated with the developmental stage of a tree. In the case of the
608 remote-sensing-based FTV, the most relevant sources of error were misclassification
609 (confusion of conifer and broadleaf), inadequate temporal synchronization (a time
610 difference between the field measurements and the recording of remote sensing data), and
611 imperfect spatial matching of the sample-plot location with the detail of the remote sensing
612 data.

613 Second, the variables *ground_proportion* and *area_chm_prop / area_chm_sq_prop* can only
614 be compared directly if they are derived identically. This would be the case if

615 *ground_proportion* was also derived by weighting the broadleaf/conifer sample tree
616 information by the respective tree heights – analogous to calculating *area_chm_prop* /
617 *area_chm_sq_prop*. However, this was not the case because tree heights were not recorded.
618 Third, the response variable, the volume per sample-plot, which was calculated by the same
619 volume functions as used to derive the forest type proportion (*ground_proportion*), was also
620 not error free and contains the same errors as the field observed FTV. The applied volume
621 functions ignore some real-existing influences on the tree volume and they therefore return
622 the same volume for different values of some influential factors. The evaluation and
623 comparison of different models and auxiliary variables is therefore always limited by the
624 accuracy of the reference values, which were in our case generated by one-parameter
625 volume functions. In order to allow a more precise interpretation of the results, one would
626 have to use more precise, e.g. three-parameter volume functions. Therefore, it is difficult to
627 say whether there was actually a statistically significant difference between the results
628 obtained by ground based forest type information and remote sensing based forest type
629 information.

630 Overall, the following conclusions can be drawn: (1) Although the FTM-based variables are
631 error prone, they can generate a more powerful signal for modelling than the field based
632 information *ground_proportion*. This supports the statement that the proposed weighting by
633 canopy height information of the FTV generates a very useful signal.

634 (2) Until now, it has always been assumed that field information is the reference in terms of
635 accuracy. However, this concept is limited by the sampling protocol; in this case, no tree
636 heights were measured during the field observation and therefore the corresponding
637 calculation of ground reference including tree height information was not possible. If tree

638 height had been recorded, the ground reference would possibly have been as good or better
639 than the FTM-based variables.

640 Nevertheless, we recommend including *ground_proportion* as a FTV in regression models,
641 particularly during model evaluation, because the variable might serve as a benchmark to
642 assess the effect of including FTM information. Additionally, it might function as an indicator
643 of whether it is worth investing resources into creating an advanced FTM.

644 **4.4 Generalization of the Results**

645 In order to address our second research question, we studied the effect of incorporating
646 different FTV alternatives in three independent study areas. We considered the results in
647 Brg, with its precisely measured forest inventory sample-plots, as the reference. The main
648 findings were that the best results were generally achieved: (1) by deriving forest type
649 variables from a FTM with a high resolution and (2) by concurrently using forest type
650 variables, such as *area_chm_prop* and *area_chm_sq_prop*, derived by a superimposed
651 evaluation of the FTM and the CHM. These findings apply to Brg and SA. For GF, point (2)
652 could not be confirmed, on the one hand because the terrestrial inventory sample-plot
653 centres were not positioned accurately and on the other hand because GF is less
654 heterogeneously structured than the other forests because it is dominated by numerous
655 pure mature conifer sample-plots (Figure 3). To summarize, precisely measured inventory
656 sample-plots and the presence of mixed and heterogeneously structured forests are
657 required for this approach to be beneficial for current inventories. Further, leaf-off LiDAR
658 data is desirable. We worked with a LiDAR data density of 8 and 10 points m^{-2} , which
659 delivered excellent results. Such data is widely available; for example, by 2023 Switzerland
660 will be covered completely with leaf-off LiDAR data with an average density of 15–20 points
661 m^{-2} .

662 **4.5 Relevance for management**

663 Climate change is expected to considerably influence forest ecosystems and their
664 management (Hanewinkel et al., 2013; Schelhaas et al., 2015). In combination with climate
665 change, disturbances are expected to increase and affect forest ecosystems (Seidl et al.,
666 2017, 2014). To face these future challenges, Messier et al. (2019) recommended using
667 forest management to increase the number of tree species and the structural diversity of
668 forests at the landscape scale to improve the resilience of forests. Managing forests for
669 higher resilience requires, however, more accurate and more spatially explicit information
670 on the current situation of forests, which is usually assessed during forest inventories. Such
671 information is used to guide forest practitioners in their management aims, which further
672 include providing the demanded ecosystem services and the conservation of forest
673 biodiversity in the best possible way (Bäck et al., 2017; MEA, 2005). Assessment of the
674 standing timber volume therefore provides important information to practitioners because
675 management activities are very often addressed primarily in forest stands with the largest
676 timber volumes, as they contain the highest accumulated timber values and require cautious
677 planning of harvesting activities and stand regeneration. Further, the standing timber
678 volume is often used as an important indicator to quantify the provisioning of timber
679 production (Blattert et al., 2017; Bugmann et al., 2017), which is still one of the most
680 important ecosystem services in forestry. Improved inventory fundamentals have further
681 implications. They lead to improved management solutions, such as more efficient
682 harvesting or road network layouts (Bont et al., 2019, 2012), as timber volume is usually a
683 major component used to form such solutions (Bont and Church, 2018). With our presented
684 method for incorporating FTM information, predictions about the standing timber volume

685 can be provided for forest management with higher accuracy compared with classical
686 approaches and on larger scales.

687 **5 Conclusions**

688 We draw the following three major conclusions from our study: (1) Incorporating FTM
689 information into ordinary least squares (OLS) linear regression models for predicting timber
690 volume on the sample-plot level increased model accuracy. An improvement in RMSE of up
691 to 9 percentage points in comparison to models not using any forest type information was
692 observed. The highest explanatory power of regression models was achieved by weighting
693 high-resolution FTM information (leaf-off LiDAR) with superimposed canopy height model
694 (CHM) information. This new approach of deriving FTVs improved the RMSE by up to 2
695 percentage points compared with classical approaches. (2) The OLS regression models had
696 the best fit when they included an interaction term between mean canopy height and forest
697 type. (3) Considering *ground_proportion* (the value derived from field measurements on the
698 inventory sample-plots) in the model evaluation could serve as an important benchmark
699 and/or upper bound for assessing the improvement in model accuracy when FTM
700 information is included in model definitions.

701 Overall, our new method of deriving FTVs better reflects timber volume in heterogeneously
702 structured and mixed forests. Detailed standing timber volume assessments are relevant for
703 guiding practitioners in managing forests for multiple ecosystem services, particularly
704 nowadays when resilient forests with a diverse structure and species mixture are needed to
705 face the challenges of climate change. However, further research is required regarding
706 different statistical model types. For example, it would be interesting to know if our findings
707 can be transferred to other modelling approaches. In addition, a detailed differentiation into

708 single tree species or main tree species, beyond the conifer / broadleaf classification, could
709 be a further development of this new approach.

710

711 **Funding:** This research was partially funded by the Swiss Forest and Wood Research Fund (WHFF)
712 (Project 2015.01).

713 **Acknowledgements:** We thank Pierre Cothureau and Robert Jenni from the Canton of Freiburg for
714 supporting this project. We also thank Dr. Markus Kalisch from the Seminar for Statistics (ETH Zurich)
715 for his practical input concerning statistical analysis, and Melissa Dawes for English editing assistance.
716 We are especially grateful to the WSL, in particular Dr. Oliver Thees and Dr. Renato Lemm, for
717 collaboration and supporting this project. We also thank two anonymous reviewers for their valuable
718 support in improving the manuscript.

719 **Conflicts of Interest:** The authors declare no conflict of interest and the funding sponsors had no role
720 in the design of the study; in the collection, analyses or interpretation of data; in the writing of the
721 manuscript; or in the decision to publish the results.

722

References

- 724 Akaike, H., 2011. Akaike's information criterion. *International encyclopedia of statistical science* 25–
725 25.
- 726 Bäck, J., Aszalós, R., Ceulemans, R.J., Glatzel, G., Hanewinkel, M., Kakaras, E., Kotiaho, J.S., Lindroth,
727 A., Lubica, D., Luysaert, S., 2017. Multi-functionality and sustainability in the European
728 Union's forests. *EASAC policy report*.
- 729 Barrett, F., McRoberts, R.E., Tomppo, E., Cienciala, E., Waser, L.T., 2016. A questionnaire-based
730 review of the operational use of remotely sensed data by national forest inventories. *Remote
731 Sensing of Environment* 174, 279–289. <https://doi.org/10.1016/j.rse.2015.08.029>
- 732 Blattert, C., Lemm, R., Thees, O., Lexer, M.J., Hanewinkel, M., 2017. Management of ecosystem
733 services in mountain forests: Review of indicators and value functions for model based multi-
734 criteria decision analysis. *Ecological Indicators* 79, 391–409.
735 <https://doi.org/10.1016/j.ecolind.2017.04.025>
- 736 Bont, L.G., Church, R.L., 2018. Location set-covering inspired models for designing harvesting and
737 cable road layouts. *Eur J Forest Res* 137, 771–792. [https://doi.org/10.1007/s10342-018-
1139-7](https://doi.org/10.1007/s10342-018-
738 1139-7)
- 739 Bont, L.G., Heinimann, H.R., Church, R.L., 2012. Concurrent optimization of harvesting and road
740 network layouts under steep terrain. *Ann Oper Res*. [https://doi.org/10.1007/s10479-012-
1273-4](https://doi.org/10.1007/s10479-012-
741 1273-4)
- 742 Bont, L.G., Maurer, S., Breschan, J.R., 2019. Automated Cable Road Layout and Harvesting Planning
743 for Multiple Objectives in Steep Terrain. *Forests* 10, 687. <https://doi.org/10.3390/f10080687>
- 744 Breidenbach, J., Kublin, E., McGaughey, R.J., Andersen, H.-E., 2008. Mixed-effects models for
745 estimating stand volume by means of small footprint airborne laser scanner data.
746 *Photogrammetric Journal of Finland* 21, 4–15.
- 747 Breiman, L., 2001. Random forests. *Machine learning* 45, 5–32.
- 748 Bugmann, H., Cordonnier, T., Truhetz, H., Lexer, M.J., 2017. Impacts of business-as-usual
749 management on ecosystem services in European mountain ranges under climate change. *Reg
750 Environ Change* 17, 3–16. <https://doi.org/10.1007/s10113-016-1074-4>
- 751 Eichhorn, F., 2013. *Ertragstabellen für die Weißtanne*. Springer-Verlag, Berlin, Heidelberg.
- 752 ESA, 2019. European Space Agency (ESA), Copernicus Sentinel mission. [WWW Document]. URL
753 <https://sentinel.esa.int/web/sentinel/missions>
- 754 Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh,
755 A., 2016. Review of studies on tree species classification from remotely sensed data. *Remote
756 Sensing of Environment* 186, 64–87. <https://doi.org/10.1016/j.rse.2016.08.013>
- 757 Fox, J., Monette, G., 1992. Generalized Collinearity Diagnostics. *Journal of the American Statistical
758 Association* 87, 178–183. <https://doi.org/10.1080/01621459.1992.10475190>
- 759 Gabriel, A., Hill, A., Breschan, J., 2018. Neue Hilfsmittel zur Anwendung zweiphasiger
760 Stichprobenverfahren in der Waldinventurpraxis. *Schweizerische Zeitschrift für Forstwesen*
761 169, 210–219. <https://doi.org/10.3188/szf.2018.0210>
- 762 Hanewinkel, M., Cullmann, D.A., Schelhaas, M.-J., Nabuurs, G.-J., Zimmermann, N.E., 2013. Climate
763 change may cause severe loss in the economic value of European forest land. *Nature Clim
764 Change* 3, 203–207. <https://doi.org/10.1038/nclimate1687>
- 765 Harrell, F.E., 2017. *Regression modeling strategies*, Springer Series in Statistics. Springer, Heidelberg.
- 766 Hill, A., Buddenbaum, H., Mandallaz, D., 2018. Combining canopy height and tree species map
767 information for large-scale timber volume estimations under strong heterogeneity of
768 auxiliary data and variable sample plot sizes. *Eur J Forest Res* 137, 489–505.
769 <https://doi.org/10.1007/s10342-018-1118-z>
- 770 Hoffmann, C., 1982. Die Berechnung von Tarifen für die Waldinventur. *Forstw Cbl* 101, 24–36.
771 <https://doi.org/10.1007/BF02741168>

772 Isenburg, M., 2014. Rasterizing Perfect Canopy Height Models from LiDAR. rapidlasso GmbH. URL
773 <https://rapidlasso.com/2014/11/04/rasterizing-perfect-canopy-height-models-from-lidar/>
774 (accessed 9.24.19).

775 Kaufmann, E., 2001. Estimation of Standing Timber Growth and Cut, in: Brassel, P., Lischke, H. (Eds.),
776 Swiss National Forest Inventory: Methods and Models of the Second Assessment. WSL.

777 Keller, M., 2013. Schweizerisches Landesforstinventar, Felddaufnahme-Anleitung 2013, Internal
778 Report, . WSL, Birmensdorf.

779 Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., Hussin, Y.A., 2014. Generating Pit-free
780 Canopy Height Models from Airborne Lidar. *photogramm eng remote sensing* 80, 863–872.
781 <https://doi.org/10.14358/PERS.80.9.863>

782 Kukkonen, M., Korhonen, L., Maltamo, M., Suvanto, A., Packalen, P., 2018. How much can airborne
783 laser scanning based forest inventory by tree species benefit from auxiliary optical data?
784 *International Journal of Applied Earth Observation and Geoinformation* 72, 91–98.
785 <https://doi.org/10.1016/j.jag.2018.06.017>

786 Kukkonen, M., Maltamo, M., Korhonen, L., Packalen, P., 2019. Multispectral Airborne LiDAR Data in
787 the Prediction of Boreal Tree Species Composition. *IEEE Transactions on Geoscience and*
788 *Remote Sensing* 57, 3462–3471. <https://doi.org/10.1109/TGRS.2018.2885057>

789 Lamprecht, S., Hill, A., Stoffels, J., Udelhoven, T., 2017. A Machine Learning Method for Co-
790 Registration and Individual Tree Matching of Forest Inventory and Airborne Laser Scanning
791 Data. *Remote Sensing* 9, 505. <https://doi.org/10.3390/rs9050505>

792 Langanke, T., 2017. Copernicus Land Monitoring Service – High Resolution Layer Forest: Product
793 Specifications Document.

794 Latifi, H., Fassnacht, F., Koch, B., 2012. Forest structure modeling with combined airborne
795 hyperspectral and LiDAR data. *Remote Sensing of Environment* 121, 10–25.
796 <https://doi.org/10.1016/j.rse.2012.01.015>

797 Liang, X., Hyyppä, J., Matikainen, L., 2007. Deciduous-coniferous tree classification using difference
798 between first and last pulse laser signatures 5.

799 Magnussen, S., Mandallaz, D., Breidenbach, J., Lanz, A., Ginzler, C., 2014. National forest inventories
800 in the service of small area estimation of stem volume. *Can. J. For. Res.* 44, 1079–1090.
801 <https://doi.org/10.1139/cjfr-2013-0448>

802 Mandallaz, D., 2013. Design-based properties of some small-area estimators in forest inventory with
803 two-phase sampling. *Can. J. For. Res.* 43, 441–449. <https://doi.org/10.1139/cjfr-2012-0381>

804 Mandallaz, D., 2007. Sampling Techniques for Forest Inventories. Chapman and Hall/CRC.
805 <https://doi.org/10.1201/9781584889779>

806 Mandallaz, D., Breschan, J., Hill, A., 2013. New regression estimators in forest inventories with two-
807 phase sampling and partially exhaustive information: a design-based Monte Carlo approach
808 with applications to small-area estimation. *Can. J. For. Res.* 43, 1023–1031.
809 <https://doi.org/10.1139/cjfr-2013-0181>

810 McCallum, K., Beaty, M., Mitchell, B., 2014. First Order LIDAR Metrics: A supporting document for
811 LIDAR deliverables. RedCastle Resources Inc., Remote Sensing Applications Center (RSAC),
812 Salt Lake City, Utah.

813 MEA, 2005. Millennium Ecosystem Assessment - Ecosystems and human well-being. Island press
814 Washington, DC:

815 Messier, C., Bauhus, J., Doyon, F., Maure, F., Sousa-Silva, R., Nolet, P., Mina, M., Aquilué, N., Fortin,
816 M.-J., Puettmann, K., 2019. The functional complex network approach to foster forest
817 resilience to global changes. *For. Ecosyst.* 6, 21. <https://doi.org/10.1186/s40663-019-0166-2>

818 Næsset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne
819 scanning laser. *Scandinavian Journal of Forest Research* 19, 164–179.
820 <https://doi.org/10.1080/02827580310019257>

821 Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical
822 two-stage procedure and field data. *Remote Sensing of Environment* 80, 88–99.
823 [https://doi.org/10.1016/S0034-4257\(01\)00290-5](https://doi.org/10.1016/S0034-4257(01)00290-5)

824 Ørka, H.O., Næsset, E., Bollandsås, O.M., 2009. Classifying species of individual trees by intensity and
825 structure features derived from airborne laser scanner data. *Remote Sensing of Environment*
826 113, 1163–1174. <https://doi.org/10.1016/j.rse.2009.02.002>

827 Packalén, P., Maltamo, M., 2006. Predicting the Plot Volume by Tree Species Using Airborne Laser
828 Scanning and Aerial Photographs. *for sci* 52, 611–622.
829 <https://doi.org/10.1093/forestscience/52.6.611>

830 Parkan, M., 2018. *Digital Forestry Toolbox for Matlab/Octave*.

831 Pretzsch, H., 2001. Modellierung des Waldwachstums. *Parey*.

832 Pretzsch, H., Biber, P., 2005. A Re-Evaluation of Reineke’s Rule and Stand Density Index. *for sci* 51,
833 304–320. <https://doi.org/10.1093/forestscience/51.4.304>

834 R Core Team, 2018. *R: a language and environment for statistical computing*. R Foundation for
835 Statistical Computing, Vienna, Austria.

836 Rätty, J., Vauhkonen, J., Maltamo, M., Tokola, T., 2016. On the potential to predetermine dominant
837 tree species based on sparse-density airborne laser scanning data for improving subsequent
838 predictions of species-specific timber volumes. *Forest Ecosystems* 3, 1.
839 <https://doi.org/10.1186/s40663-016-0060-0>

840 Reineke, L.H., 1933. Perfecting a stand-density index for even-aged forests. *Journal of Agricultural*
841 *Research* 7, 627–638.

842 Rivoire, M., Le Moguedec, G., 2012. A generalized self-thinning relationship for multi-species and
843 mixed-size forests. *Annals of Forest Science* 69, 207–219. [https://doi.org/10.1007/s13595-](https://doi.org/10.1007/s13595-011-0158-z)
844 [011-0158-z](https://doi.org/10.1007/s13595-011-0158-z)

845 Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., Hyyppä,
846 J., 2015. Model-assisted estimation of growing stock volume using different combinations of
847 LiDAR and Landsat data as auxiliary information. *Remote Sensing of Environment* 158, 431–
848 440. <https://doi.org/10.1016/j.rse.2014.11.020>

849 Salo, S., Tahvonen, O., 2002. On the Optimality of a Normal Forest with Multiple Land Classes. *for sci*
850 48, 530–542. <https://doi.org/10.1093/forestscience/48.3.530>

851 Schelhaas, M.-J., Nabuurs, G.-J., Hengeveld, G., Reyer, C., Hanewinkel, M., Zimmermann, N.E.,
852 Cullmann, D., 2015. Alternative forest management strategies to account for climate change-
853 induced productivity and species suitability changes in Europe. *Reg Environ Change* 15,
854 1581–1594. <https://doi.org/10.1007/s10113-015-0788-z>

855 Schmid-Haas, P., 2003. Die Idee der Kontrollstichproben: ihre Entstehung und ihre Zukunft | The
856 Swiss Continuous Forest Inventory: its genesis and its future. *Schweizerische Zeitschrift für*
857 *Forstwesen* 154, 102–111. <https://doi.org/10.3188/szf.2003.0102>

858 Schmid-Haas, P., Baumann, E., Werner, J., forêt, la neige et le paysage (Birmensdorf) I. fédéral de
859 recherches sur la, 1993. *Kontrollstichproben: Aufnahmeinstruktion*. Eidgenössische
860 Forschungsanstalt für Wald Schnee und Landschaft.

861 Schweizer, S., 2012. *Schweizerischer Forstkalender 2013*. Huber, Frauenfeld.

862 Seidl, R., Schelhaas, M.-J., Rammer, W., Verkerk, P.J., 2014. Increasing forest disturbances in Europe
863 and their impact on carbon storage. *Nature Climate Change* 4, 806–810.
864 <https://doi.org/10.1038/nclimate2318>

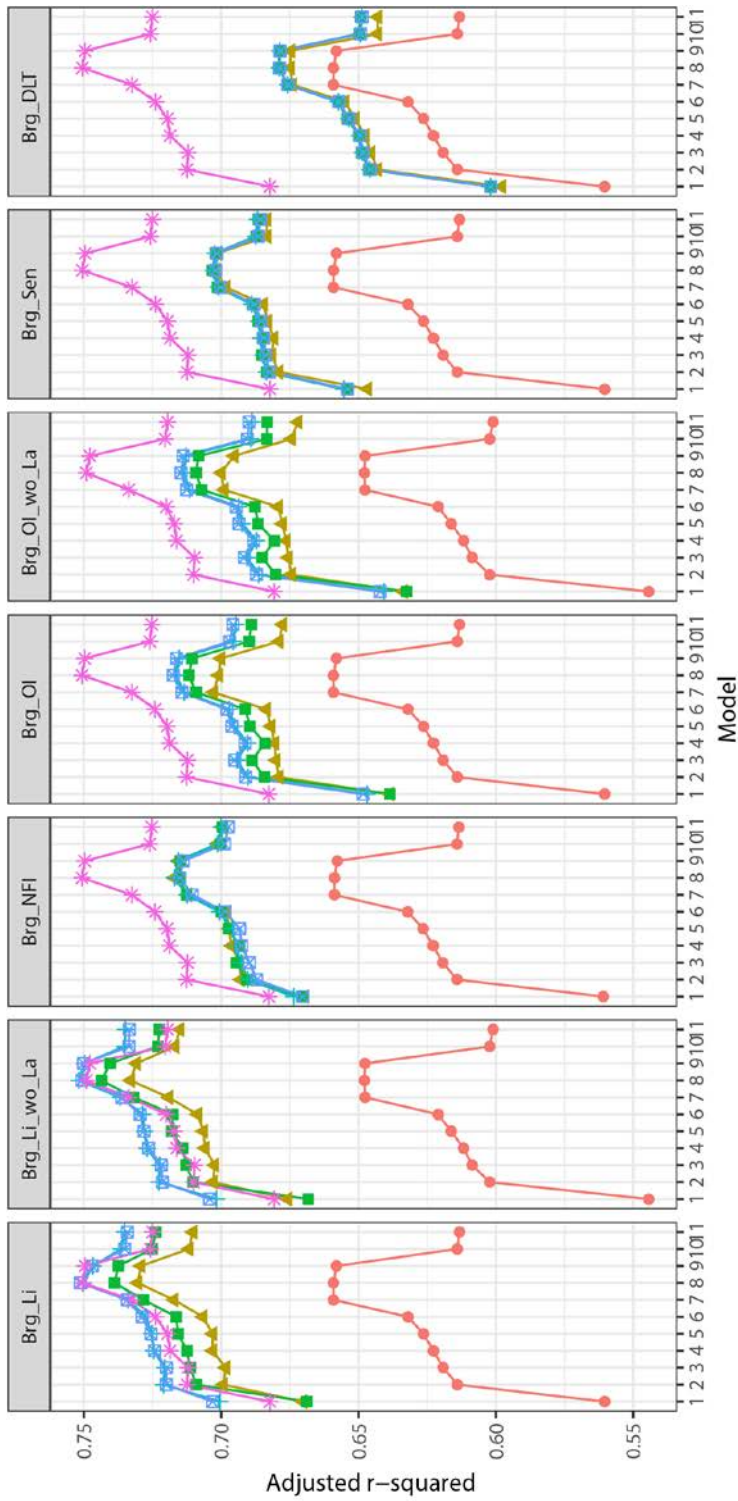
865 Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D.,
866 Petr, M., Honkaniemi, J., Lexer, M.J., Trotsiuk, V., Mairota, P., Svoboda, M., Fabrika, M.,
867 Nagel, T.A., Reyer, C.P.O., 2017. Forest disturbances under climate change. *Nature Climate*
868 *Change* 7, 395–402. <https://doi.org/10.1038/nclimate3303>

869 Small, D., 2012. SAR backscatter multitemporal compositing via local resolution weighting, in: 2012
870 IEEE International Geoscience and Remote Sensing Symposium. Presented at the 2012 IEEE
871 International Geoscience and Remote Sensing Symposium, pp. 4521–4524.
872 <https://doi.org/10.1109/IGARSS.2012.6350465>

873 Steinmann, K., Mandallaz, D., Ginzler, C., Lanz, A., 2013. Small area estimations of proportion of
874 forest and timber volume combining Lidar data and stereo aerial images with terrestrial data.

875 Scandinavian Journal of Forest Research 28, 373–385.
876 <https://doi.org/10.1080/02827581.2012.754936>
877 Straub, C., Dees, M., Weinacker, H., Koch, B., 2009. Using Airborne Laser Scanner Data and CIR
878 Orthophotos to Estimate the Stem Volume of Forest Stands [WWW Document]. URL
879 [https://www.ingentaconnect.com/content/schweiz/pfg/2009/00002009/00000003/art0001](https://www.ingentaconnect.com/content/schweiz/pfg/2009/00002009/00000003/art00010)
880 0 (accessed 9.16.19).
881 Waser, L.T., Ginzler, C., Rehus, N., 2017. Wall-to-Wall Tree Type Mapping from Countrywide
882 Airborne Remote Sensing Surveys. Remote Sensing 9, 766.
883 <https://doi.org/10.3390/rs9080766>
884 Xu, Y., Li, C., Sun, Z., Jiang, L., Fang, J., 2019. Tree height explains stand volume of closed-canopy
885 stands: Evidence from forest inventory data of China. Forest Ecology and Management 438,
886 51–56. <https://doi.org/10.1016/j.foreco.2019.01.054>
887

Appendix



Forest type variable alternative

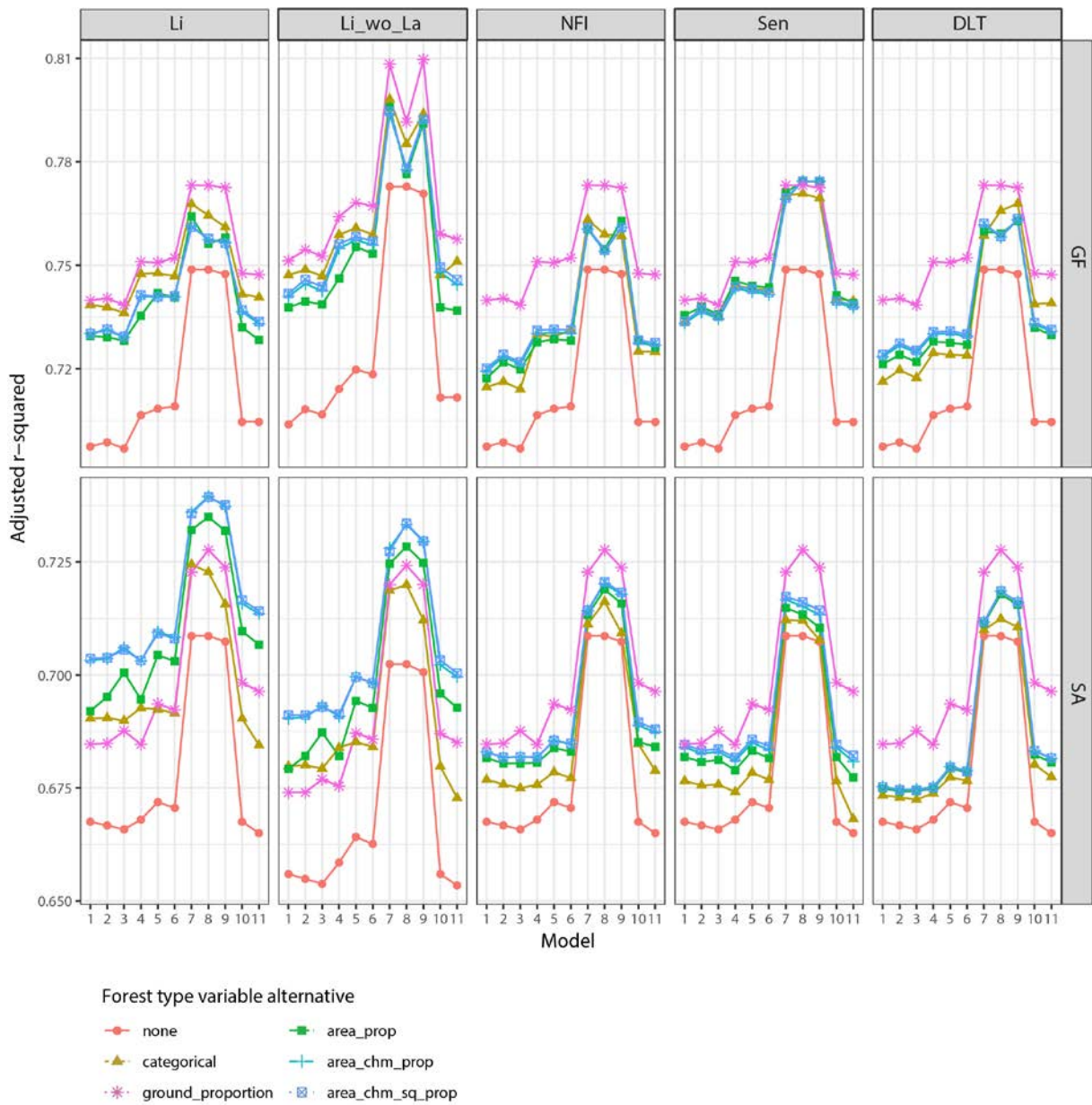
- none
- ▲ categorical
- * ground_proportion
- area_prop
- + area_chm_prop
- area_chm_sq_prop

889

890

891

Figure A1: Adjusted r-squared values for each regression model, forest type map source and combination of forest type variables in the case study area Bremgarten (Brg).



893

894 **Figure A2: Adjusted r-squared values for each regression model, forest type map source and combination of**
 895 **forest type variables in the case study areas Glâne-Farzin (GF) and Sarine (SA).**

896

Study area	Variables
Brg_OI	MSSq + STD + MAX + P40 + P30 + P25 + P05 + P01 + C00 + C02 + C10 + C15 + C20 + MEAN ²
Brg_OI_wo_La	STD + P99 + P60 + P50 + P30 + P25 + P05 + P01 + C00 + C10 + C20 + CV + STD:MEAN
Brg_NFI	MSSq + STD + MAX + P40 + P30 + P25 + P05 + P01 + C00 + C02 + C10 + C15 + C20 + MEAN ²
GF_NFI	P95 + P80 + P60 + P40 + P25 + P10 + P01 + C00 + C02 + C05 + C10 + C15 + C20 + C25 + C30 + C35 + C45 + CV
SA_NFI	P99 + P95 + P20 + C02 + C05 + C15 + C20 + C25 + C35 + C50 + P60
Brg_DLT	MSSq + STD + MAX + P40 + P30 + P25 + P05 + P01 + C00 + C02 + C10 + C15 + C20 + MEAN ²
GF_DLT	P95 + P80 + P60 + P40 + P25 + P10 + P01 + C00 + C02 + C05 + C10 + C15 + C20 + C25 + C30 + C35 + C45 + CV
SA_DLT	P99 + P95 + P20 + C02 + C05 + C15 + C20 + C25 + C35 + C50 + P60
Brg_Li	MSSq + STD + MAX + P40 + P30 + P25 + P05 + P01 + C00 + C02 + C10 + C15 + C20 + MEAN ²
Brg_Li_wo_La	STD + P99 + P60 + P50 + P30 + P25 + P05 + P01 + C00 + C10 + C20 + CV + STD:MEAN
GF_Li	P95 + P80 + P60 + P40 + P25 + P10 + P01 + C00 + C02 + C05 + C10 + C15 + C20 + C25 + C30 + C35 + C45 + CV
GF_Li_wo_La	STD + P95 + P80 + P60 + P40 + P30 + P10 + P01 + C00 + C02 + C05 + C10 + C15 + C20 + C45 + MEAN ²
SA_Li	P99 + P95 + P20 + C02 + C05 + C15 + C20 + C25 + C35 + C50 + P60
SA_Li_wo_La	MSSq + P25 + P05 + C02 + C05 + C15 + C20 + C25 + C35 + C50 + MEAN ²
Brg_Sen	MSSq + STD + MAX + P40 + P30 + P25 + P05 + P01 + C00 + C02 + C10 + C15 + C20 + MEAN ²
GF_Sen	P95 + P80 + P60 + P40 + P25 + P10 + P01 + C00 + C02 + C05 + C10 + C15 + C20 + C25 + C30 + C35 + C45 + CV
SA_Sen	P99 + P95 + P20 + C02 + C05 + C15 + C20 + C25 + C35 + C50 + P60

899 **Table A1: Selected variables for Model 7. In addition to the listed variables, the FTV can be added**
900 **(*area_prop, area_chm_prop, area_chm_prop, categorical or ground_proportion*).**