Canadian Journal of Forest Research

A design-based assessment of an expanded set of auxiliary information for forest growth estimation

Journal:	<i>Canadian Journal of Forest Research</i>
Manuscript ID	cjfr-2020-0529.R1
Manuscript Type:	Article
Date Submitted by the Author:	22-Mar-2021
Complete List of Authors:	Massey, Alexander; Swiss Federal Institute for Forest Snow and Landscape Research Lanz, Adrian; Swiss Federal Institute for Forest Snow and Landscape Research Ferretti, Marco; Swiss Federal Institute for Forest Snow and Landscape Research
Keyword:	model-assisted estimation, elastic net, forest inventory, optical remote sensing, growth estimation
Is the invited manuscript for consideration in a Special Issue? :	Not applicable (regular submission)

SCHOLARONE™
Manuscripts

A design-based assessment of an expanded set of auxiliary information for forest growth estimation

Alexander Massey, Adrian Lanz, and Marco Ferretti

Draft

Alexander Massey. Corresponding author, e-mail: stats.massey@gmail.com
Swiss Federal Research Institute WSL
Zürcherstrasse 111, CH 8903 Birmensdorf, Switzerland
Adrian Lanz. e-mail adrian.lanz@wsl.ch
Swiss Federal Research Institute WSL
Zürcherstrasse 111, CH 8903 Birmensdorf, Switzerland
Marco Ferretti. Tel.: +41 44 739 22 51, email: marco.ferretti@wsl.ch
Swiss Federal Research Institute WSL
Zürcherstrasse 111, CH 8903 Birmensdorf, Switzerland

Abstract:

To improve the design-based precision of gross increment estimates from forest inventories, we propose an assessment of an expanded set of auxiliary information grouped from five sources: (i) a vegetation height model, (ii) satellite imagery, (iii) spatial data, (iv) topography and (v) variables identified by external forest monitoring and research networks. The former two are from optical remote sensing and the latter three are chosen based on interpretable and proved connections with forest growth. We evaluate each source individually as well as collectively for the Swiss National Forest Inventory using two-phase estimation with the elastic net. In terms of relative efficiency, all individual groups demonstrated improvement over one-phase estimation by 7 to 29% with the interpretable sources consistently outperforming those based on optical remote sensing. However, the interpretable sources do not provide significant additional gains when combined together whereas the optical remote sensing consistently and substantially supports other sources of auxiliary data when combined, leading to a 50 to 71% improvement overall. Given the availability of data from e.g. international monitoring programs, expanding the set of auxiliaries to include both interpretable sources and optical remote sensing is a feasible and promising option for National Forest Inventories.

Résumé:

Keywords: model-assisted estimation, elastic net, growth estimation, forest inventory, remote sensing

Introduction

It is commonplace for countries to implement National Forest Inventories (NFIs) that use auxiliary information to improve the design-based estimation of key forest attributes. Most often this auxiliary information consists of vegetation height models (VHM) derived via optical remote sensing in the form of LiDaR or stereo-photogrammetry. There has been great success in using VHMs to derive auxiliary variables to improve the estimation precision of attributes of forest state as well as for forest growth (Esteban et al. 2019; Næsset et al. 2013; McRoberts et al. 2015). This has been demonstrated using model-based techniques (Saarela et al. 2020) and also with model-assisted techniques (Irulappa-Pillai-Vijayakumar et al. 2019; Magnussen et al. 2018) whose validity depends solely on the randomization of the plots regardless whether the model holds or not.

Whereas it seems relatively straightforward why a VHM derived variable would serve as a good predictor of timber volume at a given location, the nature of its correlation with forest growth is far more ambiguous. For example, mean canopy height may be correlated with growth only indirectly via timber volume because a single volume measurement indicates both the result of growth as well as the potential for growth in the future. This issue is well-known in statistics as omitted-variable bias (Wooldridge 2009). While for causal inference omitted variable bias can be a major problem because bias in the estimated regression coefficients can lead to completely false conclusions, for predictive modeling it is only important insofar as the loss of predictive performance through the omission of, or failure to recognize, potentially useful covariates. Model-assisted techniques are design-based (a frequent governmental requirement for NFI reporting) and fall in the latter case because gains of precision depend almost entirely on the predictive performance of the assisting model (Breidt and Opsomer 2017). A more intuitive auxiliary variable for estimating forest growth is the change in a VHM variable between two time points. However, this can be confounded by difficult to detect disturbances as well as the practical limitations of the technique; many aerial remote sensing campaigns are prone to temporal misalignment between aerial remeasurements and the evaluations of the field plots (Avitabile and Camia 2018). Passive sensor data from satellites have more frequent updates but are still less efficient than active sensor data such as LiDaR (Petrrou et al. 2015).

To overcome these challenges, we present a novel approach that expands the set of auxiliary information by incorporating spatial and topographical variables as well as new variables that have been directly identified by existing forest monitoring and research networks as being environmental drivers of forest growth. In Europe such a network is the International Co-operative Program on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) which operates under the auspices of the UNECE Air Convention (Ferretti and Fischer 2013). ICP Forests have two main components: a so-called Level I network, based on systematic sampling, and the Level II network, based on purposive sampling. Level I studies temporal and spatial trends of the forest condition on an annual basis; Level II investigates relationships between environmental drivers and forest response. The co-

occurrence of NFI and ICP Forests schemes is frequent across European countries, including Switzerland where the Swiss NFI and the Level II ICP Forest monitoring network, known as the Long-Term Forest Ecosystem Research programme (LWF), are conducted and managed by the same institute.

Assessments of auxiliary information for forest inventory do exist in the literature (Bergseng et al. 2015; Rätty et al. 2018). However, they typically focus on estimating forest state attributes utilizing optical remote sensing data from satellites or LiDaR. For forest change attributes, Breidenbach et al. (2021) combined NFI data with Landsat-based Forest Cover Loss and one-time wall-to-wall airborne laser scanning (ALS) data in order to improve the estimation precision for C-stock dynamics in greenhouse gas inventories. Here we seek to increase the scope of potential auxiliary information beyond just optical remote sensing. We will use a model-assisted framework to evaluate precision gains for estimating the growth variable – gross increment with ingrowth – under an expanded set of auxiliary information with data from the Swiss NFI. We consider five main sources of auxiliary information: (i) a vegetation height model, (ii) satellite imagery, (iii) spatial data, (iv) topography and (v) variables directly identified as drivers of forest growth by LWF (see below). LWF variables are based on meteorological measurements as well as nitrogen deposition and ozone dose which are mapped using geostatistical techniques. Auxiliary sources and their possible combinations are assessed using two-phase estimation. The relative efficiency of each combination of auxiliary sources is compared to the basic one-phase estimator. Sources are also assessed relative to other combinations of auxiliaries when a given source is removed. We call the latter variant marginal relative efficiency which quantifies how well a particular auxiliary source supports other sources. As the consideration of interaction terms can greatly inflate the dimensionality of the feature space, this assessment utilizes the elastic net framework for penalized linear regression for the assisting model to simultaneously perform variable selection and model fitting.

Materials

Data and target variables

The study spans between the third and the fourth Swiss NFIs (NFI3 and NFI4 respectively). The Swiss NFI, referred to herein as NFI, is based on a systematic sampling design with 20 638 plots, each consisting of two concentric circles of 200 m^2 and 500 m^2 respectively, that are allocated uniformly across Switzerland on a $\sqrt{2}\text{ km}$ by $\sqrt{2}\text{ km}$ grid (Fischer and Traub 2019). In 2009, the NFI changed from a periodic inventory that assessed all permanent terrestrial plots every ten years to an annual design that measures one-ninth of the total sample every year on interpenetrating $3\sqrt{2}\text{ km}$ subgrids referred to as annual panels. As each periodic inventory took three years to complete, NFI3 plots may have been measured anytime between 2004 and 2006 whereas NFI4 plots can be attributed to the specific year where the field measurement occurred. Hence, the NFI3 to NFI4 transition features

remeasurement periods between 3 and 12 years necessitating the consideration of annualized target variables. The layout of the plots locations across Switzerland can be found in Figure 1 and an example of how estimates are reported for each year can be found in Table 1.

The main target variables for forest growth in the Swiss NFI are called growth components. Six growth components were readily available: i) survivor growth, ii) ingrowth, iii) ongrowth, iv) loss, v) gross increment without ingrowth and (vi) gross increment with ingrowth (GI). All of these variables are measured in terms of (annualized) change in $\frac{m^3}{ha}$ in overbark bole volume and are available on the full $\sqrt{2} km$ by $\sqrt{2} km$ terrestrial grid. Ultimately, GI, being defined as the sum of all other change components, was selected as the target variable of interest because it provides key summary information on forest vitality.

An important consideration when interpreting this study's results is the restriction of the study area to the 2017 forest mask defined using aerial stereo-photography. We note that the forest mask definition has a small error rate compared to the ground truth definition; approximately 1.4% of the total plots are misclassified compared to the official definition. We ignore the non-forest mask because including a large number of non-forest plots, whose GI is zero, can lead to overoptimistic results concerning the model's performance in predicting growth in the forest. For the purposes of publishing official NFI estimates one could use the forest mask, which is available wall-to-wall with a 1 m by 1 m resolution, as a true stratification. However, notwithstanding the effect of the stratification weights, the contribution to the overall variance of the non-forest stratum is clearly close to zero for global estimation which is why it is dropped from the scope of this study.

Auxiliary information

A total of 42 auxiliary variables from multiple sources were considered for the analysis. As the design-based performance of the assisting model is predominantly a prediction problem, we opt to take a machine learning perspective (Bzdok et al. 2018) where one includes all potential sources of auxiliary information and allow the prediction algorithm to determine the optimal predictors. Admittedly, the auxiliary sources considered here may not be feasible in all countries due to practical concerns such as budgetary constraints or computational effort which is why all combinations of categories are also evaluated. Table 2 summarizes the 42 variables that were selected in the final full model of the case study.

Optical remote sensing

The optical remote sensing sources are based on an updated VHM and publicly accessible satellite data collected mostly by the European Environment Agency's Copernicus Land Monitoring Service (Langanke 2018). The satellite variables of interest are remote estimates of tree cover density, a broadleaf/conifer classification variable as well as a Normalized Density Vegetation Index (NDVI) that is commonly accessible via satellite sources. The

VHM was derived by stereo-image matching of a digital surface model (DSM) generated by aerial stereo imagery gathered between 2011 and 2016 and a previously existing digital terrain model. The DSM uses ADS40/ADS80 provided by the Swiss Federal Office of Topology (swisstopo) during the leaves-on season and a digital terrain model to calculate vegetation heights on a 1 m by 1 m resolution over 500 m^2 interpretation areas centered over the nominal centers of the terrestrial plots. Comprehensive details about the automated workflow and the stereo-image matching strategy is described in Ginzler and Hobi (2015). The VHM variables of interest are plot-level summary statistics of the canopy characteristics, such as mean, median, standard error and quantiles.

The distribution of the temporal alignment between aerial and field measurements can be found in Figure 2. The majority of the annual panels tend to consist of NFI plots whose terrestrial measurements occurred before their respective VHM measurements. The annual panels occurring after 2012 exhibit a gradual increase in the inclusion of plots whose VHM year occurred before the NFI measurement and in 2017 all terrestrial measurements either occurred in the same year or after their associated VHM measurements. This temporal mismatch between VHM and NFI measurements is unavoidable because it was not practical for the flight campaign to provide full uniform coverage for all of Switzerland for a given year.

Expanded sources with direct interpretation for growth prediction

We consider an expanded set of auxiliaries that have more direct explanatory roles in predicting forest growth.

Spatial (Sp) and topographic (Topo) data represent site attributes that intuitively influence growth but do not change much over time. Sp includes latitude, longitude and a categorical variable for each of the five production regions: Jura, Plateau, Prealps, Alps and South Alps. Topo includes geographic orientation, slope azimuth and gradient, topographical indexes, elevation and topsoil pH. The latter was obtained from a model based on 1500 soil profiles and included here as it was obtained from the same database and was not directly identified by the LWF.

The source denoted LWF was included on the basis of previous results obtained by the LWF in Switzerland. The motivation was to identify growth predictors from long-term forest research carried out on nineteen sites purposely distributed across important Swiss forest ecosystems (Thimonier et al. 2010). Since 1994, LWF has been collecting data on several attributes related to biological, chemical and physical characteristics of atmosphere, vegetation, and soil with different time resolutions ranging from continuous measurements to 10 year remeasurement cycles. On the basis of these plots, Etzold et al. (2014) focused on net primary productivity (NPP) that is expected to be highly correlated with GI. The two auxiliary variables that showed the strongest correlation to NPP were nitrogen (N) deposition and mean summer temperature. N deposition was also identified as the most important environmental driver of growth after a larger, European-scale study (Etzold et al. 2020). On the other hand, since negative effects of tropospheric ozone on growth were reported for Switzerland (Braun et al. 2014), we also considered the phytotoxic

ozone dose for beech (*Fagus sylvatica* L.) and Norway spruce (*Picea abies* (L.) Karst.). It is worth noting that the role of environmental factors, and in particular N deposition and climate, on different growth-related metrics (NPP, volume, basal area increment) has been pointed out by studies carried out in several European countries (Bontemps et al. 2011; de Vries et al. 2009; Ferretti et al. 2014; Magnani et al. 2007; Trotsiuk et al. 2020). Thus, modeled wall-to-wall analogues of these variables (owner: Federal Office for the Environment FOEN; see Rihm and Künzle (2019); Thimonier et al. (2019) for N deposition and see Braun and Rihm (2012) for ozone dose) were selected as auxiliary variables grouped in LWF. Meteorological data were provided by MeteoSwiss.

Methods

The estimators described in this study are presented in the model-assisted design-based framework in the context of two-phase estimation (also known as double sampling). Although the ideas borrow heavily from Särndal et al. (2003), which uses the finite population approach where the sampling units are plots, the approach here is rooted in the infinite population approach (Mandallaz 2008; Hill et al. 2021) where the fundamental sampling units are points rather than plots. This avoids the need for circular plots to be approximated by hexagons in order to be tessellated over the forest area. It also implies that the population size is no longer dependent on an individual plot's surface area. Results between the infinite and finite population approaches tend to differ at most negligibly. Here the infinite population notation has been adapted somewhat so that the formulas match more closely with their counterparts in the finite population literature.

One-phase sampling

In the infinite population approach we consider a well-defined population of infinite points, ω , whose union spans the entire forested area F . The surface area of F is denoted as $S(F)$ and is assumed to be known or easily estimable. Every point has an associated local density of the response variable, denoted $y(\omega)$, that is assumed to be measurement error-free and defined as a Horvitz-Thompson estimator of sample size 1 for the parameter of interest, i.e. the spatial mean of all $y(\omega)$ across F , which we call $\bar{y}$. Note that in the infinite case the inclusion probabilities are defined as the probability that a point lies in an inclusion ring drawn around each tree and occur behind the scenes in the technical mathematical definition of $y(\omega)$ (see Massey (2015) p. 5 for full definition and explanation of boundary effect treatment). For the context of this case study, $y(\omega)$ can be simply understood as the density measured in $\frac{m^3}{ha}$ of GI associated with the point located at the center of a field plot.

The sampling design is chosen so that it links to the spatial mean using the following identity

$$[1] \quad \mathbb{E}_{\omega}(y(\omega)) = \frac{1}{S(F)} \int_F y(\omega) d\omega = \bar{y}$$

It is straightforward to see that we can estimate [1] from a sample of ω and multiply by $S(F)$ to obtain the population total. The design-based inference is assumed to be based on a uniform and independent (uid) selection of n_2 points that form the terrestrial sample s_2 . In a slight abuse of notation regarding the uncountably infinite $\omega \in F$, we will switch to the more familiar $y(\omega) := y_i$ to match the finite population literature. This leads to the one-phase estimator

$$[2] \quad \hat{y}_{1p} = \frac{1}{n_2} \sum_{i \in s_2} y_i$$

whose variance can be estimated using

$$[3] \quad \hat{V}(\hat{y}_{1p}) = \frac{1}{n_2} \hat{V}_{s_2}(y_i)$$

where $\hat{V}_{s_2}$ is the estimated sample variance in s_2 . Table 1 shows the one-phase estimates by year that will serve as a baseline for this study. Both $\hat{y}_{1p}$ and $\hat{V}(\hat{y}_{1p})$ are exactly design-unbiased and design-consistent because y_i are assumed to be uid. For mathematical purposes we assume a uid selection of points when in reality they are systematically distributed across the country. This is generally considered to be acceptable as it tends to yield conservative variance estimates as long as the periodicity of the sample grid does not coincide with the spatial correlation structure of the forest.

Two-phase sampling

The motivation of this study is to improve the design-based precision by exploiting available auxiliary data. The most common approach is to use so-called double sampling where a large sample s_1 of n_1 points are selected uid to collect vectors of auxiliary data, denoted $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ip})^T$. Then a subsample $s_2 \subset s_1$ is selected using simple random sampling without replacement (SRSwoR) into order to collect the ground truth, i.e. y_i . In the infinite population approach it is easy to show that this is mathematically equivalent to selecting s_2 uid from F and then augmenting that sample with another $n_1 - n_2$ points selected uid from F to construct s_1 because the probability of selecting the same point twice is zero, i.e. they are also selected without replacement (Sunter 1977). This is useful for this study where a few of the auxiliary variables were only immediately available on a $\sqrt{2} \text{ km}$ by $\sqrt{2} \text{ km}$ first-phase grid, but could be densified if need be in the future.

Model-assisted estimation

The difference estimator

The relationship between $\mathbf{x}_i$ and y_i is typically defined by an arbitrary predictive model of y_i , denoted $m(\mathbf{x}_i)$. For simplicity we will restrict ourselves to the consideration of a linear assisting model of the form

$$[4] \quad y_i = \mathbf{x}_i^T \boldsymbol{\beta} + r_i$$

with $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$. In this framework, the assisting model does not need to hold in the model-dependent sense to maintain its design-based properties, but is used to provide structure for the prediction, $\hat{m}(\mathbf{x}_i) := \mathbf{x}_i^T \hat{\boldsymbol{\beta}}$. This leads to the two-phase difference estimator defined as the mean of the model predictions observed over s_1 corrected by the mean of the residuals, denoted $\hat{r}_i := y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}$, observed over s_2 :

$$[5] \quad \hat{y}_{2p} = \frac{1}{n_1} \sum_{i \in s_1} \mathbf{x}_i^T \hat{\boldsymbol{\beta}} + \frac{1}{n_2} \sum_{i \in s_2} \hat{r}_i$$

We briefly consider the special case where $\hat{\boldsymbol{\beta}}$ is assumed to be fit independently of s_1 and denote the estimator as $\hat{Y}_{2p, diff}$. It is easily checked that this is exactly design-unbiased using the identity $\mathbb{E}(\cdot) = \mathbb{E}_{s_1}(\mathbb{E}_{s_2|s_1}(\cdot))$ regardless whether the underlying model assumptions hold or not. In this context, $\hat{m}(\mathbf{x}_i)$ can be viewed merely as dimension reduction for $\mathbf{x}_i$. Furthermore, since $(\mathbf{x}_i, y_i)$ are iid and assuming $\hat{\boldsymbol{\beta}}$ is fit independently of s_1 , we can explicitly write the exact theoretical variance using the decomposition

$$[6] \quad \begin{aligned} \mathbb{V}(\hat{Y}_{2p, diff}) &= \mathbb{V}_{s_1} \mathbb{E}_{s_2|s_1}(\hat{Y}_{2p, diff}) + \mathbb{E}_{s_1} \mathbb{V}_{s_2|s_1}(\hat{Y}_{2p, diff}) \\ &= \frac{1}{n_1} \mathbb{V}(y_i) + \left(1 - \frac{n_2}{n_1}\right) \frac{1}{n_2} \mathbb{V}(r_i) \end{aligned}$$

It is intuitively clear that the precision of $\hat{y}_{2p}$ depends on choosing $m(\mathbf{x}_i)$ to minimize $\mathbb{V}(r_i)$ because n_2 is constrained by the high cost of acquiring more ground sample and n_1 can be made arbitrarily large.

In practice it is not always feasible to find an independent sample to fit the assisting model. As a result, it is common to define $\hat{m}(\mathbf{x}_i)$ as some model fit using s_2 as a training set and plugging it directly into [5]. This plug-in estimator is sometimes referred to as the "model-assisted" estimator, although the precise definition tends to vary somewhat in the literature. The classical plug-in variance estimator for [6] is

$$[7] \quad \hat{\mathbb{V}}(\hat{Y}_{2p}) \approx \frac{1}{n_1} \hat{\mathbb{V}}_{s_2}(y_i) + \left(1 - \frac{n_2}{n_1}\right) \frac{1}{n_2} \hat{\mathbb{V}}_{s_2}(\hat{r}_i)$$

It is important to point out that this plug-in variance completely ignores the effect of fitting the assisting model in s_2 . In practice this will lead to a loss of independence among $\hat{m}(\mathbf{x}_i)$ in s_2 because the model fit depends on all

the selected $i \in s_2$. This likely leads to an underestimation of the variance that tends to increase as the number of parameters increases relative to the sample size. Fortunately, it has been demonstrated that for many common model families [5] remains asymptotically design-unbiased in the plug-in framework and [7] is design-consistent (see Breidt and Opsomer (2017) for a thorough overview in the finite population context). This can usually be expected to hold in the infinite population case as well (see Mandallaz (1991), pp. 173-176 for general proof concerning a large class of root-n consistent linear models) where the asymptotics are relatively simplified. We forgo further technical justifications here, but emphasize that properties of estimators in the infinite approach intuitively follow from results obtained in the finite population case, which pertain to $\mathbb{E}_{s_2|s_1}(\hat{Y}_{2p,diff})$ and $\mathbb{V}_{s_2|s_1}(\hat{Y}_{2p,diff})$ in [6].

The classical regression estimator and its drawbacks

We obtain the well-known classical regression (REG) estimator when we define $\hat{\beta}$ as the ordinary least squares (OLS) solution

$$[8] \quad \hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i \in s_2} (y_i - \mathbf{x}_i^T \beta)^2$$

For many usage cases, the REG is quite attractive for its familiarity and its flexibility. It is asymptotically design-unbiased, the plug-in variance estimator is design-consistent and an improved variance estimator – the g-weight variance estimator – is available via first-order Taylor approximation to account for the effect of fitting the model internally on the variance (see Mandallaz (2008), pp. 115-117). It also possesses other useful properties such as being a generalization of post-stratification and calibration of the survey weights to the auxiliary means of the first-phase sample (Massey et al. (2014), p. 1180 for short overview). However, given the high-dimensionality of the auxiliary data considered in this study, the REG is problematic for following reasons:

- There is no guaranteed unique solution to [8] when the number covariates is greater than the number of observations (i.e. $p > n_2$) which is the case when all five sources of auxiliary information are included with their respective interaction terms;
- The inclusion of irrelevant auxiliary variables leads to overfitting that can increase the design-based variance to the extent where it is bigger than the one-phase variance estimate;
- Popular variable selection methods like best subset and stepwise selection by AIC are either computational infeasible or lead to undesirable local minima with respect to all possible subsets of selected covariates.

As a result, this study requires a more general framework for optimizing $\hat{\beta}$ that is better suited to a high-dimensional feature space.

Elastic Net

A common approach to overcome the computational drawbacks of high-dimensional auxiliary data is to include a penalty term in the optimization program for the model's coefficient vector. We consider the elastic net criterion proposed by Zou and Hastie (2005)

$$[9] \quad \hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i \in s_2} (y_i - \mathbf{x}_i^T \beta)^2 + \lambda \left(\alpha \sum_{j=1}^p |\beta_j| + (1 - \alpha) \sum_{j=1}^p \beta_j^2 \right)$$

where $\lambda > 0$ is a tuning parameter that controls the degree of penalization and $\alpha \in [0, 1]$ controls the amount of mixing between the penalty under the L1 norm, i.e. the sum of the absolute values of the coefficients, and the L2 norm, i.e. the sum of the squared coefficients. This convex optimization system yields unique predictions when $p > n_2$ and is capable of selecting more than p auxiliary variables for $\alpha \in [0, 1)$. A non-negative λ will intentionally add bias to the coefficient estimates, which are unbiased under OLS (i.e. $\lambda = 0$) in the model-dependent sense. This is intended to avoid overfitting by shrinking less relevant variable coefficients towards zero under the constraint of the penalty term. The degree of penalization is frequently chosen using cross-validation in order to minimize the prediction error and by extension the resulting variance under the bias-variance tradeoff.

The elastic net formulation of $\hat{\beta}$ includes notable special cases. When $\alpha = 1$ we get the LASSO criterion proposed by Tibshirani (1996), which acts as an aggressive automatic variable selector that removes unnecessary covariates by shrinking their coefficients directly to zero with larger choices of λ . When $\alpha = 0$, we get the ridge regression solution proposed by Hoerl and Kennard (1970) that shrinks all coefficients proportionally but does not completely remove any of the auxiliary variables from the selected model. For this case study we want to include variable selection in order to identify efficient predictors, which would lead us to prefer $\alpha > 0$. However, the LASSO solution has a somewhat undesirable property in the presence of multicollinearity in that it tends to select either one variable or the other when they are highly correlated. As a result, there is a tendency to select individual categories from categorical variables rather than the entire variable. In an effort to mitigate this behavior, the elastic net allows us to mix the ridge and lasso penalties by selecting $\alpha \in (0, 1)$. McConville et al. (2020) provide a useful guide for the design-based implementation of these special cases. Technical derivations of the design-based asymptotics for the model-assisted LASSO and ridge regression estimators are provided in McConville et al. (2017) and Park and Yang (2008) respectively.

Implementation

Analyses using the elastic net are conducted in R 4.0.2 (R Core Team 2020) using the package glmnet (Friedman et al. 2010) where 10-fold cross validation is used to select λ for ridge regression ($\alpha=0$), LASSO ($\alpha=1$) and the elastic net with α fixed at 0.5. The same cross-validation folds are used for each variant. The training set for

$\hat{\beta}$ pools all annual panels of NFI4 together because it provides nearly uniformly improved precision compared to refitting the models separately for each year. Pooling the annual panels is essentially equivalent to the strategy of borrowing strength that is commonly used in small area estimation where each annual panel can be thought of as a domain (e.g. the pseudo-synthetic estimator in Mandallaz (2013) or the modified GREG estimator in Rao and Molina (2015)). This also has the theoretical advantage that eight-ninths of the training set can be considered independent of s_1 , making the estimator for any single year very similar to $\hat{Y}_{2p,diff}$ and, by consequence, the variance estimate based on [7] less prone to underestimating the true variance.

Each combination of auxiliary sources is evaluated using one-phase estimation as a reference method. Since all variants of estimators considered here are at least asymptotically design-unbiased, it suffices to compare the ratio of their respective variance estimates in order to define the relative efficiency as

$$[10] \quad RE(\cdot) = \frac{\hat{V}(\hat{Y}_{1p})}{\hat{V}(\hat{Y}_{2p}(\cdot))}$$

The numerator is based on [3] and the denominator is the plug-in sample variance [7]. $RE(\cdot) = 1$ implies no improvement in precision over one-phase estimation while values greater than one imply improved precision.

We also want to evaluate the contribution of a particular source of auxiliary information when it is included with a combination of other auxiliaries. To do this we will use as a reference the two-phase variance where a given auxiliary source is removed from the input combination and call this the marginal relative efficiency

$$[11] \quad MRE(source, comb) = \frac{\hat{V}(\hat{Y}_{2p}(comb))}{\hat{V}(\hat{Y}_{2p}(source \cup comb))}$$

where *comb* is the set of auxiliary sources in a considered combination, *source* is the auxiliary source to be assessed and $source \cap comb = \emptyset$. Thus, $MRE(LWF, \cdot)$ denotes the relative efficiency when LWF is included with some other combination of auxiliaries. This describes how well a given auxiliary source supports other combinations of auxiliary groupings. Values greater than one imply greater predictive potential.

Results

The results are presented for global estimation for each year in NFI4 spanning from 2009 to 2017 where the assisting model was fit using all annual panels pooled together. Efficiency metrics are based on the extrapolation $n_1 = \infty$ because the first-phase sample size in the context of producing official NFI estimates would practically be chosen to be very large¹. This was convenient for this study where the LWF auxiliaries were not readily available beyond the standard $\sqrt{2} km$ by $\sqrt{2} km$ grid because the relative efficiency (RE) and the marginal relative efficiency

¹ Point estimates where each year was refit separately and $n_1 = 6117$ can be found in the Supplementary Materials but provide no compelling evidence of being statistically different than their respective one-phase estimates.

(MRE) are calculable knowing only s_2 and the sample sizes n_1 and n_2 . Recall that for the implementation here, when we say a particular auxiliary source is included it means that all auxiliary variables from that source are considered in the model (with potential 2-way interactions), but the model fit may result in zero coefficients for some of those variables.

To assess the asymptotic design unbiasedness of the variance estimator, we compare the main protocol of the case study (i.e. $n_1 = \infty$ and pooling all annual panels to create a single training set) to true difference estimators where the target annual panel is excluded from the training set. The true difference estimators have exactly design-unbiased variance estimators but their respective training sets are approximately eight-ninths the size of the pooled training set, which has 6117 units. Figure 3 compares the pooled approach to the true difference estimator in terms of RE^2 . We observe potential evidence for overestimation in RE due to the modest underestimation of the asymptotic variance estimator which becomes more pronounced as the number of selected variables increases with respect to n_2 . In an operational setting, it is recommended to address this variance underestimation either by applying a variant of the g-weight variance estimator or by bootstrapping (see McConville et al. (2020) for full discussion in the finite population case). However, none of these options was feasible for this case study due to the lack of availability of certain auxiliary variables on the full first-phase grid. We will proceed regardless as the potential variance inflation does not appear to be influential with regards to comparing competing sources of auxiliary information.

Figure 4 assesses the choice of α by comparing the RE of each estimator for each year and combination of auxiliary sources. Both the LASSO and the elastic net ($\alpha = 0.5$) outperformed ridge regression but were only negligibly different from each other in overall predictive performance. We continue here with only the $\alpha = 0.5$ variant because it should provide a fuller picture of which variables are helpful given the high multicollinearity of the data.

Table 3 shows auxiliary variable membership for each individual auxiliary source as well as the best performing combinations by number of input sources. Note that membership could be in the form of interaction terms which are not explicitly displayed in the table. The consideration of all possible two-way interaction terms was necessary to improve performance but made the selected variables quite difficult to interpret. The most common interaction terms are associated with PRODREG which is intuitive as the underlying growth model is expected to vary region to region. However, as many of the selected interaction terms have no clear interpretation, it is advisable to consider entire groups of auxiliary sources as a whole and assess them purely in terms of the predictive performance of the estimator (e.g. relative efficiency).

Tables 4 and 5 shows the RE and average MRE for every combination of auxiliaries using the elastic net. The

²See Supplementary Materials for direct comparison of the variance estimates.

sort order is based on the mean RE across all years. Mean RE for estimates of GI show an 11 to 21% improvement over one-phase estimation when the auxiliary sources are considered individually and further substantial increases when more than one source is taken into account. Surprisingly, the optical remote sensing (RS) categories, i.e. VHM and Sa, which are arguably the most widely used types of auxiliary information in NFIs, are at the bottom of the list when included individually. Sp, Topo and LWF all lead to a higher relative efficiency as compared to VHM and Sa when they are the only source of auxiliary information considered. This is not the case when optical RS is mixed with the expanded set of auxiliary sources. Table 5 demonstrates that the optical RS categories made consistently meaningful contributions to the predictive performance in terms of MRE for nearly all combinations of sources. Thus, the optical RS auxiliaries are clearly desirable to have, but have untapped potential for growth prediction when used by themselves. This is in stark contrast to Sp, Topo and LWF that appear only make meaningful contributions individually or when mixed with optical RS, but not when mixed with each other.

Figure 5 plots RE and MRE for all possible combinations pertaining to each of the five sources. Trend lines are added by the number of sources included in the combination. It is evident that LWF, Topo and Sp all perform better than VHM individually (i.e. when the number of sources is 1). However, there are diminishing returns in the MRE as the number of sources in the input combination increases, particularly for LWF and Sp. This is apparent as the darker trend lines approach $RE = 1$ for Sp, Topo and LWF.

There appears to be some evidence for later years that Topo and LWF have increased performance whereas VHM has declining performance. A potential explanation for this is that the expanded auxiliary set, being more or less static over time, performs better for longer remeasurements periods. This is in contrast to the VHM that appears to decline as a higher proportion of the plots have field measurement years before the VHM year. This decline in performance may possibly be due to random yearly fluctuations in the data but is noteworthy as it was also observed when refitting the VHM separately for each year³.

Conclusion

This case study indicates that design-based estimation precision of forest growth (e.g. gross increment with ingrowth) can be significantly improved by expanding sources of auxiliary information beyond just optical remote sensing. Spatial data, topography and the variables identified from long-term monitoring programs (LWF) all outperformed the optical remote sensing sources (VHM and satellite) when each auxiliary source is applied individually. However, they tended to produce diminished benefits when combined together. In contrast, the VHM and satellite data play an important role in supporting other sources of auxiliary information. Thus, in the context of growth estimation it can be advisable for NFIs that use optical remote sensing to expand their auxiliary infor-

³ See Supplementary Materials.

mation, and include those with interpretable connections to growth. As for the latter, several studies demonstrated the relationship between growth and e.g. N deposition and climate, within and beyond Switzerland. Incorporating this expanded set of auxiliaries can therefore help in capturing important, large-scale correlates of growth and is technically feasible, at least as far as the sources of data considered in this study: spatial and topographical data are widely available from direct on-site measurements and terrain models, and variables as those obtained from LWF are available at continental scale and beyond from existing monitoring programs (e.g. ICP Forests) and models (e.g. EMEP, the European Monitoring and Evaluation Programme). Finally, as an assisting model, the elastic net proved to be an effective strategy for automating variable selection across many potential predictors and their interactions.

Acknowledgements

This project was financed entirely by WSL. We are grateful to Simpal Kumar and Käthi Liechti their assistance in providing technical support for accessing data from LWF and MeteoSwiss. We also would like to thank Andri Baltensweiler for providing topographical data, Reto Meier for providing variables for ozone dose and nitrogen deposition, and Minna Pulkkinen for allowing us to access her dataset with variables related to the VHM and satellite imagery. Also, thanks to Rita Ghosh and the rest of the Statistics Lab group at WSL for their support in developing the mathematical framework of the case study. We are also grateful to Urs-Beat Brändli, Lorenzo Fattorini, Arthur Gessler and Esther Thürig for discussion and support in the preparation of the project proposal and for their comments on an earlier version of the manuscript. Finally, many thanks to Daniel Mandallaz for his suggestions and thoughtful feedback throughout the course of this project.

References

- Avitabile, V. and Camia, A. 2018. An assessment of forest biomass maps in Europe using harmonized national statistics and inventory plots. *Forest Ecology and Management*, **409**: 489–498.
- Bergseng, E., et al. 2015. Assessing forest inventory information obtained from different inventory approaches and remote sensing data sources. *Annals of Forest Science*, **72**(1): 33–45.
- Bontemps, J.-D., et al. 2011. Nitrogen footprint in a long-term observation of forest growth over the twentieth century. *Trees*, **25**(2): 237–251.
- Braun, S. and Rhim, B. 2012. Ozonbelastung von Waldbäumen in der Schweiz und damit verbundene Wachstumseinbußen. *Schweizerische Zeitschrift für Forstwesen*, **163**: 383–388.
- Braun, S., Schindler, C., and Rhim, B. 2014. Growth losses in Swiss forests caused by ozone: epidemiological data analysis of stem increment of *Fagus sylvatica* L. and *Picea abies* Karst. *Environ. Pollut.*, **192**: 129–138.
- Breidenbach, J., et al. 2021. Improving living biomass C-stock loss estimates by combining optical satellite, airborne laser scanning, and NFI data. *Canadian Journal of Forest Research*.
- Breidt, F. J. and Opsomer, J. D. 2017. Model-Assisted Survey Estimation with Modern Prediction Techniques. *Statistical Science*, **32**(2): 190–205.
- Bzdok, D., Altman, N., and Krzywinski, M. 2018. Statistics versus machine learning. *Nature Methods*, **15**(4): 233–234.
- de Vries, W., et al. 2009. The impact of nitrogen deposition on carbon sequestration by European forests and heathlands. *Forest Ecology and Management*, **258**(8): 1814–1823.
- Esteban, J., et al. 2019. Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data. *Remote Sensing*, **11**(16): 1944.
- Etzold, S., et al. 2014. Tree growth in Swiss forests between 1995 and 2010 in relation to climate and stand conditions: Recent disturbances matter. *For. Ecol. Manag.*, **311**: 41–55.
- Etzold, S., et al. 2020. Nitrogen deposition is the most important environmental driver of growth of pure, even-aged and managed European forests. *Forest Ecology and Management*, **458**: 117762.
- Ferretti, M. and Fischer, R., editors 2013. *Forest Monitoring*, volume 12. Elsevier, first edition.

- 383 Ferretti, M., et al. 2014. On the tracks of Nitrogen deposition effects on temperate forests at their southern European
384 range - an observational study from Italy. *Global Change Biology*, **20**(11): 3423–3438.
- 385 Fischer, C. and Traub, B. 2019. *Swiss National Forest Inventory Methods and Models of the Fourth Assessment*.
386 Springer.
- 387 Friedman, J., Hastie, T., and Tibshirani, R. 2010. Regularization Paths for Generalized Linear Models via Coordi-
388 nate Descent. *J. Stat. Softw.*, **33**(1): 1–22.
- 389 Ginzler, C. and Hobi, M. 2015. Countrywide Stereo-Image Matching for Updating Digital Surface Models in the
390 Framework of the Swiss National Forest Inventory. *Remote Sensing*, **7**: 4343–4370.
- 391 Hill, A., Massey, A., and Mandallaz, D. 2021. The R Package forestinventory : Design-Based Global and Small
392 Area Estimations for Multiphase Forest Inventories. *Journal of Statistical Software*, **97**(4).
- 393 Hoerl, A. E. and Kennard, R. W. 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Tech-*
394 *nometrics*, **12**(1): 55–67.
- 395 Irulappa-Pillai-Vijayakumar, D. B., et al. 2019. Increasing Precision for French Forest Inventory Estimates using
396 the k-NN Technique with Optical and Photogrammetric Data and Model-Assisted Estimators. *Remote Sensing*,
397 **11**(8): 991.
- 398 Langanke, T. 2018. *Copernicus Land Monitoring Service High Resolution Layer Forest: Product Specifications*
399 *Document*. European Environment Agency and HRL Forest service providers.
- 400 Magnani, F., et al. 2007. The human footprint in the carbon cycle of temperate and boreal forests. *Nature*,
401 **447**(7146): 849–851.
- 402 Magnussen, S., Nord-Larsen, T., and Riis-Nielsen, T. 2018. Lidar supported estimators of wood volume and
403 aboveground biomass from the Danish national forest inventory (2012–2016). *Remote Sensing of Environment*,
404 **211**: 146–153.
- 405 Mandallaz, D. 1991. *A unified approach to sampling theory for forest inventory based on infinite population and*
406 *superpopulation models*. Ph.D. thesis, ETH Zürich, Zürich. Diss. Techn. Wiss. ETH Zürich, Nr. 9378.
- 407 Mandallaz, D. 2008. *Sampling Techniques for Forest Inventories*. Chapman and Hall, Boca Raton FL.
- 408 Mandallaz, D. 2013. Design-based properties of some small-area estimators in forest inventory with two-phase
409 sampling. *Can. J. For. Res.*, **43**: 441–449.

- 410 Massey, A. 2015. *Multiphase estimation procedures for forest inventories under the design-based Monte Carlo*
411 *approach*. Ph.D. thesis, ETH Zürich, Zürich. Diss. Techn. Wiss. ETH Zürich, Nr. 23025.
- 412 Massey, A., Mandallaz, D., and Lanz, A. 2014. Integrating remote sensing and past inventory data under the new
413 annual design of the Swiss National Forest Inventory using three-phase design-based regression estimation. *Can.*
414 *J. For. Res.*, **44**(10): 1177–1186.
- 415 McConville, K. S., Moisen, G. G., and Frescino, T. S. 2020. A Tutorial on Model-Assisted Estimation with
416 Application to Forest Inventory. *Forests*, **11**(2): 244.
- 417 McConville, K. S., et al. 2017. Model-Assisted Survey Regression Estimation with the Lasso. *Journal of Survey*
418 *Statistics and Methodology*, **5**(2): 131–158.
- 419 McRoberts, R. E., et al. 2015. Indirect and direct estimation of forest biomass change using forest inventory and
420 airborne laser scanning data. *Remote Sensing of Environment*, **164**: 36–42.
- 421 Næsset, E., et al. 2013. Model-assisted estimation of change in forest biomass over an 11 year period in a sample
422 survey supported by airborne LiDAR: A case study with post-stratification to provide "activity data". *Remote*
423 *Sensing of Environment*, **128**: 299–314.
- 424 Park, M. and Yang, M. 2008. Ridge Regression Estimation for Survey Samples. *Communications in Statistics -*
425 *Theory and Methods*, **37**(4): 532–543.
- 426 Petrou, Z., et al. 2015. Discrimination of Vegetation Height Categories With Passive Satellite Sensor Imagery
427 Using Texture Analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*,
428 **8**(4): 1442–1455.
- 429 R Core Team 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical
430 Computing, Vienna, Austria.
- 431 Rao, J. and Molina, I. 2015. *Small Area Estimation*. John Wiley & Sons, Inc, Hoboken, NJ, USA.
- 432 Rätty, M., Heikkinen, J., and Kangas, A. 2018. Assessment of sampling strategies utilizing auxiliary information
433 in large-scale forest inventory. *Canadian Journal of Forest Research*, **48**(7): 749–757.
- 434 Rihm, B. and Künzle, T. 2019. Mapping Nitrogen Deposition 2015 for Switzerland. Technical Report on the
435 Update of Critical Loads and Exceedance, including the years 1990, 2000, 2005 and 2010. Technical report,
436 Meteotest, Bern, commissioned by the Federal Office for the Environment (FOEN).

- 437 Saarela, S., et al. 2020. Mapping aboveground biomass and its prediction uncertainty using LiDAR and field data,
438 accounting for tree-level allometric and LiDAR model errors. *Forest Ecosystems*, **7**(1): 43.
- 439 Särndal, C., Swenson, B., and Wretman, J. 2003. *Model Assisted Survey Sampling*. Springer Series in Statistics,
440 New York.
- 441 Sunter, A. B. 1977. List Sequential Sampling with Equal or Unequal Probabilities Without Replacement. *Journal*
442 *of the Royal Statistical Society Series C*, **26**(3): 261–268.
- 443 Thimonier, A., et al. 2010. Does exceeding the critical loads for nitrogen alter nitrate leaching, the nutrient status of
444 trees and their crown condition at Swiss Long-term Forest Ecosystem Research (LWF) sites? *European Journal*
445 *of Forest Research*, **129**(3): 443–461.
- 446 Thimonier, A., et al. 2019. Total deposition of nitrogen in Swiss forests: Comparison of assessment methods and
447 evaluation of changes over two decades. *Atmos. Environ.*, **198**(1): 335–350.
- 448 Tibshirani, R. 1996. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society:*
449 *Series B (Methodological)*, **58**(1): 267–288.
- 450 Trotsiuk, V., et al. 2020. Assessing the response of forest productivity to climate extremes in Switzerland using
451 modeldata fusion. *Global Change Biology*, **26**(4): 2463–2476.
- 452 Wooldridge, J. M. 2009. *Introductory Econometrics: A Modern Approach*. Cengage Learning, Mason, OH, fifth
453 edition.
- 454 Zou, H. and Hastie, T. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal*
455 *Statistical Society: Series B (Statistical Methodology)*, **67**(2): 301–320.

Table 1. One-phase estimates for gross increment with ingrowth by year.

Year	n_2	$\hat{y}_{1p}$	$\hat{\nabla}(\hat{y}_{1p})$
2009	687	7.66	0.077
2010	685	8.04	0.067
2011	691	7.77	0.064
2012	687	8.18	0.072
2013	704	7.17	0.055
2014	693	7.51	0.065
2015	687	7.21	0.058
2016	635	7.06	0.053
2017	648	7.97	0.063

Note: Study area is restricted to the 2017 forest mask definition which classifies 1238043.0 hectares ($\approx 30.0\%$) of the Switzerland as forest.

Fig. 1. Map of study area in Switzerland. All plots are measured in both NFI3 and NFI4 after removing all plots not identified by the 2017 forest mask. An example of the distribution of a single annual panel is highlighted for 2017.

Source: Swiss Federal Institute for Forest, Snow and Landscape Research WSL (Produced in R with ggplot2)

Fig. 2. Distribution of the differences between the aerial measurements for the VHM and the terrestrial plot evaluation in the NFI.

Fig. 3. Comparison to the relative efficiency of the true difference estimator when $n_1 = \infty$. Assisting models for the difference estimator are fit excluding the target year in order to ensure that the variance estimator is exactly unbiased. Assisting models for the rest of this case study use a single training set pooling all nine years of data causing the variance estimator to only be asymptotically unbiased.

Fig. 4. Comparison of relative efficiency for ridge regression, LASSO and elastic net ($\alpha = 0.5$) when various numbers of auxiliary sources are included in the full model; $n_1 = \infty$.

Note: The axes are drawn where the relative efficiency equals one indicating that the auxiliary information offered no improvement over one-phase estimation. The dotted line indicates equality in the relative efficiencies.

Fig. 5. Relative efficiency and marginal relative efficiency by year using the elastic net ($\alpha = 0.5$) and $n_1 = \infty$. Trend lines are in grayscale according to the number of auxiliary sources included in the full model. Relative efficiency uses estimation with no auxiliary information as a reference where the column indicates inclusion of an auxiliary source. Marginal relative efficiency uses a reference model where the source indicated by the column is removed from a given combination.

Table 2. Summary of auxiliary variable names and descriptions by source.

Variable	Source	Description
WTEMP_SE	LWF	Std. Err. Winter Temperature in $^{\circ}C$ (December to February, 2005-2015)
WTEMP_MEAN	LWF	Mean Winter Temperature in $^{\circ}C$ (December to February, 2005-2015)
STEMP_SE	LWF	Std. Err. Summer Temperature in $^{\circ}C$ (June to August, 2005-2015)
STEMP_MEAN	LWF	Mean Summer Temperature in $^{\circ}C$ (June to August, 2005-2015)
SPRECIP_SE	LWF	Std. Err. Summer Precipitation in mm (June to August, 2005-2015)
SPRECIP_TOT	LWF	Total Precipitation in mm (June to August, 2005-2015)
ND_CONIF	LWF	Nitrogen Deposition (conifer) in $kgN/(ha * Year)$ averaged over 2005, 2010 and 2015
ND_BROAD	LWF	Nitrogen Deposition (broadleaf) in $kgN/(ha * Year)$ averaged over 2005, 2010 and 2015
BD1_MEAN	LWF	Ozone Dose (beech) in $mmol/(m^2 * year)$ averaged from 2004 to 2014
FD1_MEAN	LWF	Ozone Dose (N. Spruce) in $mmol/(m^2 * year)$ averaged from 2004 to 2014
PGRAD	Topo	Potential Global Radiation May to October
SOIL_pH	Topo	Topsoil pH model
SCURV	Topo	Surface curvature smoothed with Gaussian filter
SLOPE	Topo	Slope in degrees
EASTNESS	Topo	West-east gradient of aspect obtained by the cosine function
NORTHNESS	Topo	North-south gradient of aspect obtained by the sine function
MRVBF	Topo	Multiresolution Index of Valley Bottom Flatness
TPI	Topo	Topographic position index (hill, valley)
TWL.SAGA	Topo	Topographic Wetness Index
MASL	Topo	Meters above sea level
PRODREG	Sp	Production region
X	Sp	Latitude
Y	Sp	Longitude
H_MIN	VHM	Minimum raster point height in VHM
H_MAX	VHM	Maximum raster point height in VHM
H_MEAN	VHM	Mean raster point height in VHM
H_SE	VHM	Standard error of raster point heights in VHM
H_MED	VHM	Median raster point height in VHM
H_Q10	VHM	10% quantile of raster point heights in VHM
H_Q20	VHM	20% quantile of raster point heights in VHM
H_Q30	VHM	30% quantile of raster point heights in VHM
H_Q40	VHM	40% quantile of raster point heights in VHM
H_Q60	VHM	60% quantile of raster point heights in VHM
H_Q70	VHM	70% quantile of raster point heights in VHM
H_Q80	VHM	80% quantile of raster point heights in VHM
H_Q85	VHM	85% quantile of raster point heights in VHM
H_Q90	VHM	90% quantile of raster point heights in VHM
H_Q95	VHM	95% quantile of raster point heights in VHM
H_Q99	VHM	99% quantile of raster point heights in VHM
TCD	Sa	Tree Cover Density (2015)
FTYPE	Sa	Conifer/broadleaf/unclassified forest classification (2015)
NDVI	Sa	Median Normalized Density Vegetation Index without clouds (June to August, 2015-2018)

Note: All variables are continuous with the exception of PRODREG and FTYPE which are categorical.

Table 3. Selected variables for best performing auxiliary input combinations.

	Variable	LWF	Topo	Sp	VHM	Sa	Sa Topo	Sa VHM Topo	Sa VHM Topo LWF	Sp Sa VHM Topo LWF
LWF	WTEMP_SE	X							X	X
	WTEMP_MEAN	X							X	X
	STEMP_SE	X							X	X
	STEMP_MEAN	X							X	X
	SPRECIP_SE	X							X	X
	SPRECIP_TOT	X							X	X
	ND_CONIF	X							X	X
	ND_BROAD	X							X	X
	BD1_MEAN	X							X	X
	FD1_MEAN	X								X
Topography	PGRAD		X				X	X	X	X
	SOIL_pH		X				X	X	X	X
	SCURV		X				X	X	X	X
	SLOPE		X				X	X	X	X
	EASTNESS		X				X	X	X	X
	NORTHNESS		X				X	X	X	X
	MRVBF		X				X	X	X	X
	TPI		X				X	X	X	X
	TWL_SAGA		X				X	X	X	X
	MASL		X				X	X	X	X
Spatial	PRODREG			X						X
	X			X						X
	Y			X						X
VHM	H_MIN							X	X	X
	H_MAX							X	X	X
	H_MEAN				X			X	X	X
	H_SE							X	X	X
	H_MED				X			X	X	X
	H_Q10				X			X	X	X
	H_Q20				X			X	X	X
	H_Q30				X			X	X	X
	H_Q40				X			X	X	X
	H_Q60							X	X	X
	H_Q70							X	X	X
	H_Q80							X		X
	H_Q85									X
	H_Q90									
	H_Q95							X		
	H_Q99									
Satellite	TCD					X	X	X	X	X
	FTYPE					X	X	X	X	X
	NDVI					X	X	X	X	X
	# categories	1	1	1	1	1	2	3	4	5
	# parameters	18	24	11	6	5	35	86	94	136
	Mean RE	1.19	1.21	1.19	1.14	1.11	1.39	1.53	1.58	1.61

Note: All two-way interactions are implicitly considered using the elastic net with $\alpha = 0.5$ for all marked auxiliary variables. The inclusion of a variable could be either as a main effect or as part of an interaction term. # parameters is the number of covariates (both interactions and main effects) with non-zero coefficients in the model fit. The mean RE uses one-phase estimation as the reference and may be slightly overoptimistic for when more parameters are included in the model.

Table 4. Relative efficiency by year using the elastic net ($\alpha = 0.5$) and $n_1 = \infty$ for all combinations of auxiliary sources.

	Year									Mean
	2009	2010	2011	2012	2013	2014	2015	2016	2017	
Sp, Sa, VHM, Topo, LWF	1.59	1.50	1.67	1.57	1.63	1.50	1.71	1.63	1.70	1.61
Sa, VHM, Topo, LWF	1.54	1.47	1.66	1.53	1.57	1.47	1.68	1.59	1.68	1.58
Sp, Sa, VHM, Topo	1.56	1.47	1.61	1.54	1.60	1.47	1.71	1.61	1.62	1.58
Sp, Sa, VHM, LWF	1.56	1.45	1.58	1.53	1.53	1.48	1.60	1.60	1.52	1.54
Sa, VHM, Topo	1.49	1.44	1.56	1.50	1.55	1.39	1.70	1.60	1.58	1.53
Sp, Sa, VHM	1.53	1.43	1.50	1.49	1.51	1.44	1.56	1.58	1.44	1.50
Sp, VHM, Topo, LWF	1.45	1.44	1.58	1.48	1.53	1.43	1.49	1.51	1.56	1.49
Sp, VHM, Topo	1.45	1.42	1.54	1.47	1.52	1.42	1.51	1.52	1.52	1.49
Sa, VHM, LWF	1.47	1.39	1.56	1.45	1.46	1.42	1.54	1.51	1.48	1.48
Sa, Topo, LWF	1.45	1.35	1.50	1.45	1.43	1.34	1.57	1.44	1.64	1.46
VHM, Topo, LWF	1.40	1.40	1.56	1.42	1.48	1.39	1.46	1.46	1.52	1.45
Sp, Sa, Topo, LWF	1.43	1.33	1.49	1.45	1.42	1.35	1.52	1.44	1.58	1.45
Sp, Sa, Topo	1.42	1.32	1.44	1.44	1.40	1.33	1.53	1.44	1.52	1.43
Sa, Topo	1.37	1.30	1.38	1.40	1.36	1.26	1.54	1.42	1.50	1.39
VHM, Topo	1.34	1.36	1.44	1.37	1.42	1.32	1.44	1.43	1.40	1.39
Sp, VHM, LWF	1.39	1.37	1.44	1.39	1.41	1.38	1.35	1.42	1.37	1.39
Sp, Sa, LWF	1.39	1.28	1.39	1.41	1.34	1.32	1.42	1.39	1.43	1.37
Sp, Sa	1.39	1.27	1.34	1.38	1.31	1.29	1.41	1.38	1.36	1.35
Sp, VHM	1.33	1.33	1.37	1.33	1.36	1.34	1.31	1.38	1.30	1.34
Sa, LWF	1.34	1.25	1.39	1.35	1.30	1.28	1.39	1.33	1.41	1.34
VHM, LWF	1.31	1.31	1.39	1.29	1.34	1.32	1.29	1.31	1.31	1.32
Sp, Topo, LWF	1.26	1.22	1.36	1.32	1.30	1.25	1.28	1.28	1.44	1.30
Sa, VHM	1.28	1.25	1.32	1.26	1.30	1.23	1.29	1.31	1.20	1.27
Topo, LWF	1.23	1.19	1.32	1.26	1.27	1.21	1.25	1.24	1.40	1.26
Sp, Topo	1.21	1.19	1.27	1.25	1.23	1.21	1.25	1.25	1.32	1.24
Sp, LWF	1.21	1.18	1.26	1.24	1.22	1.21	1.19	1.21	1.28	1.22
Topo	1.17	1.17	1.21	1.22	1.20	1.16	1.23	1.23	1.29	1.21
Sp	1.18	1.16	1.20	1.21	1.18	1.18	1.18	1.19	1.22	1.19
LWF	1.19	1.14	1.24	1.18	1.18	1.18	1.16	1.16	1.25	1.19
VHM	1.14	1.16	1.16	1.12	1.17	1.14	1.13	1.14	1.09	1.14
Sa	1.12	1.07	1.11	1.13	1.09	1.07	1.14	1.15	1.10	1.11

Note: The relative efficiency is the multiplicative factor of n_2 needed to achieve the same precision under one-phase estimation. Auxiliary sources are spatial (Sp), satellite (Sa), vegetation height model (VHM), topological (topo) and variables derived from the Long-Term Forest Ecosystem Research programme (LWF).

Table 5. Average marginal relative efficiency using the elastic net ($\alpha = 0.5$) and $n_1 = \infty$ for all combinations of auxiliary sources. The reference is two-phase estimation using the combination of sources indicated in the row when the source in the column is removed.

	Expanded aux.			Optical RS	
	LWF	Topo	Sp	VHM	Sa
Sp, Sa, VHM, Topo, LWF	1.02	1.05	1.02	1.12	1.08
Sa, VHM, Topo, LWF	1.03	1.07		1.08	1.08
Sp, Sa, VHM, Topo		1.05	1.03	1.10	1.06
Sp, Sa, VHM, LWF	1.03		1.04	1.12	1.11
Sa, VHM, Topo		1.21		1.10	1.10
Sp, Sa, VHM			1.18	1.11	1.12
Sp, VHM, Topo, LWF	1.01	1.08	1.03	1.15	
Sp, VHM, Topo		1.11	1.07	1.20	
Sa, VHM, LWF	1.16			1.10	1.12
Sa, Topo, LWF	1.05	1.09			1.16
VHM, Topo, LWF	1.05	1.10		1.15	
Sp, Sa, Topo, LWF	1.01	1.05	0.99		1.11
Sp, Sa, Topo		1.06	1.02		1.15
Sa, Topo		1.26			1.15
VHM, Topo		1.22		1.15	
Sp, VHM, LWF	1.04		1.05	1.14	
Sp, Sa, LWF	1.02		1.03		1.12
Sp, Sa			1.21		1.13
Sp, VHM			1.18	1.13	
Sa, LWF	1.21				1.13
VHM, LWF	1.16			1.11	
Sp, Topo, LWF	1.05	1.06	1.03		
Sa, VHM				1.15	1.12
Topo, LWF	1.04	1.06			
Sp, Topo		1.04	1.02		
Sp, LWF	1.03		1.03		
Topo		1.21			
Sp			1.19		
LWF	1.19				
VHM				1.14	
Sa					1.11

Note: The marginal relative efficiency can be interpreted as the multiplicative factor of n_2 needed to achieve the same precision under two-phase estimation while excluding the auxiliary source indicated in each column. The sort order is based on the average relative efficiency over all years (the same as Table 4).