Estimating Timber Volume using Airborne Laser Scanning Data based on Bayesian Methods

J. Breidenbach¹ and E. Kublin²

¹ Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, P.O. Box 5003, NO-1432 Ås; Fon: +47 6496 5737; Johannes.Breidenbach@umb.no
² Forest Research Institute of Baden-Württemberg, Department of Biometrics, Wonnhaldestr. 4, 79199 Freiburg, Germany; Edgar.Kublin@forst.bwl.de

Introduction

Extending pointwise measurements from sample plot inventories over space (i.e., upscaling) is an active research area in the interdisciplinary field of biometrics and remote sensing. For this task, airborne laser scanning (ALS) has become an operational tool in the recent years, particularly in the Nordic countries (Næsset et al. 2004). Examples include the use of parametric regression methods for the estimation of biomass (Næsset & Gobakken 2008) or diameter distributions (Bollandsås & Næsset 2007). The primary use of nonparametric methods has been to estimate multivariate responses such as timber volume by tree species (Packalén & Maltamo 2006; Hudak et al. 2008).

While comparing different estimation methods, we found that (non-parametric) kNN methods usually outperform linear regression models in terms of precision. We ascribe this to possible nonlinear relations between covariates and the response. Due to well-known drawbacks of kNN methods, in this paper we shortly describe a semiparametric regression method that allows a flexible modeling of covariates.

Material and Methods

The study area is located in the Forbach state forest, approximately 10 km south-east of Baden-Baden, Germany (Breidenbach et al. 2008). As ground reference we used 1310 circular plots from a systematic sample plot inventory on a 100x200 m grid. The inventory is part of the standard monitoring procedure of the Baden-Württemberg forest service. A sample plot consists of 4 concentric circle plots. Thus, the inclusion probability of a tree depends on the diameter class to which it belongs. For example, only trees with a diameter at breast height (DBH) greater than 30 cm will be sampled on the largest circle plot, with a radius of 12 m (Breidenbach, Gläser & Schmidt 2008). Plot total volume (m³ ha⁻¹) was derived from single tree observations. Additionally, a digital stand map allowed the consideration of stands topology and the clustering of sample plots within stands.

Low resolution (0.5-1 m²) first- and last return airborne laserscanning data (ALS) are available for the study area. Consequently, the area based method was applied: A digital terrain model (DTM) was derived from last pulse data. This DTM was used to derive vegetation heights from the first pulse data. Various height metrics as well as canopy cover (cc) and the coniferous proportion (cp) (Breidenbach et al. 2008) were derived from the ALS data for every plot.

A Geodaadditive Model (Kammann & Wand 2003), a special case of Structured Additive Regression Models (STAR) (Fahrmeir, Kneib & Lang 2004), was fitted to the data. The response is assumed to be (conditional) Gaussian distributed

\[ y|\zeta, \sigma^2 \sim N(\eta, \sigma^2 I) \]

with \( \zeta = \) coefficients of the linear predictor \( \eta \), \( \sigma^2 = \) model variance and \( I = \) identity matrix. The general form of the linear predictor is given by

IUFRO Division 4 meeting: Extending Forest Inventory and Monitoring over Space and Time. 19-22 May 2009 Quebec City, Canada
\[ \eta = f_1(x_1) + \ldots + f_j(x_j) + \ldots + f_p(x_p) + f_{\text{spatial}}(s) + u' \gamma \]

where \( f_j \) denote Bayesian P-Splines with second order random walk penalties and \( u' \gamma \) are the linear effects of covariates. The Spline functions are used to depict the non-linear effect of metric covariates. Nonlinear interactions of metric covariates are modeled based on two-dimensional tensor product P-Splines. The spatial effect is composed of an unstructured (i.e., “random” effect) and a structured part (i.e., a spatial trend):

\[ f_{\text{spatial}}(s) = f_{\text{str}}(s) + f_{\text{unstr}}(s) \]

For the structured part, the mean effect of every region \( s \) is assumed to be an unweighted average of the effects of the neighboring regions. This results in a smooth regional effect and is known as a Markov random field. The model parameters are estimated based on full Bayesian inference using MCMC simulation techniques as implemented in BayesX (Brezger, Kneib & Lang 2005). Let 

\[ f_{\text{str}}(s) = f_j(s) = \beta_{js} \]

then the prior for the Markov random field is given by

\[ \beta_{js} | \beta_{s'}, s \neq s', \tau_j^2 \sim N \left( \frac{1}{N_s} \sum_{s' \in \partial_s} \beta_{js'}, \frac{\tau_j^2}{N_s} \right) \]

where \( s' \in \partial_s \) denotes \( s' \) a neighbor of \( s \) and \( N_s = \) number of neighbors.

**Results and discussion**

We use the same predictor variables as in the study of Breidenbach et al. (2008) (i.e., no variable selection was carried out). However, instead of simple linear predictors we use some smooth functions to model possibly nonlinear effects of the predictor variables. The deduced model has the form

\[ y_{lm} = f_1(l, \text{mean}_{lm}) + f_2(cc_{lm}) + f_3(cp_{lm}) + f_4(l, \text{mean}_{lm}, cc_{lm}) + f_5(l, \text{mean}_{lm}, cp_{lm}) \]

\[ + f_{\text{MRF}}(ID_l) + z' \delta + \varepsilon_{lm} \]

where \( l, \text{mean} \) is the mean ALS vegetation height in stand \( l \) and plot \( m \), \( cc = \) canopy cover, \( cp = \) coniferous proportion, \( z' \delta = \) random effect, \( f_j = \) (tensor product) P-Splines, \( f_{\text{MRF}} = \) Markov random field and \( ID = \) stand polygon-ID (i.e., region \( s \)). The root mean squared error relative to the mean observed volume (RMSE\%) of this model is 31.3%.

As can be seen from the plot of the bivariate interaction effect of the covariates in Figure 1 and Figure 2, a higher estimation of timber volume is associated with increasing mean vegetation height, canopy cover and coniferous proportion. However, given low \( cc \) or \( cp \), the increase of timber volume with mean vegetation height is significantly lower than for higher values of \( cc \) or \( cp \). For \( cp \), this may be attributed to the fact that stem parts within the crown area of coniferous trees are considered as usable timber, whereas this is usually not the case for broadleaved trees. Thus, timber volume will be significantly higher for conifers than for broadleaved trees of the same vegetation height.
The influence of the unstructured spatial effect (“random effect”) is significantly less than the structured spatial effect and does, in fact, not differ significantly from zero for any stand. The structured spatial effect reveals that the regions roughly located on a diagonal line through the study area from south-west to north-east have a positive effect on the estimation of timber volume. The effect of the stands located in the south-east and north-west has tendency (Figure 3). This may be caused by an underlying variable not considered in the model, such as site index. Since such a variable would (usually) not be expected to change alter abruptly at a stand border, the Markov random field seems well suited for modeling such an effect. However, as for the unstructured effect, the effect does not differ significantly from zero for most stands (Figure 4).

**Conclusion**

STAR models allow for a very flexible modeling of the influence of covariates on the response, also in the case of bivariate interactions. In addition, the consideration of spatial effects shades some light on potentially missing or unobservable variables. If requested, appropriate (parametric) transformations of covariates could now be applied based on these results. In our case, however, a parametric transformation would be difficult due to the irregular curve progression of the functions within the linear predictor.
Figure 3: Effect of the Markov random field on the timber volume estimate (m³ ha⁻¹).

Figure 4: Significance of the effect of the Markov random field on the timber volume estimate (-1 = significantly negative, 1 = significantly positive).

Literature Cited


scanning and aerial photographs., Forest Science 52 : 611-622.