GENERALIZED SPATIAL MODELS OF FOREST STRUCTURE USING AIRBORNE MULTISPECTRAL AND LASER SCANNER DATA

HOOMAN LATIFI[*], BARBARA KOCH

Dept. of Remote sensing and Landscape information systems, University of Freiburg, Tennenbacherstraße. 4, D-79106 Freiburg-Germany - (hooman.latifi, Barbara.koch)@felis.uni-freiburg.de

KEY WORDS: LiDAR, Colour Infrared, Nonparametric spatial models, Genetic Algorithm, Stratification

ABSTRACT:

Forest structure was modelled to explore the nonparametric approaches for inventory of small forests. Models were built using information from airborne optical and LiDAR data across two adjacent forest sites using the remote sensing data collected by a similar instrument. Off-site samples from the inventoried reference stands were combined with drawn samples from the target area for modelling responses. The applied combinations included two imputation methods, three sampling designs, and two predictor subset sizes, where a Genetic Algorithm (GA) was used to prune the predictors. Diagnostic tools including Root Mean Square Error, bias and Standard Error of Imputation were employed to evaluate the results. Results showed that Random Forests produced more accurate results than Most Similar Neighbour inference, yet was slightly more biased than MSN. Appending systematically stratified samples from the target dataset yielded more accurate results, yet it was most influential up to a low-to-medium intensity. The use of pruned predictors resulted in reduced bias. By bootstrapping the RMSE, the majority of the simulations lied within the confidence intervals. The applied methods show positive potential towards producing spatial models of forest structure.

1. INTRODUCTION

Remote sensing data and approaches have been widely used to model and map biophysical forest parameters such as biomass, being often deemed as a prerequisite for studying processes such as carbon sequestration. Light Detection and Ranging (LiDAR) data has been employed to provide useful information through modelling of structural forest attributes (Hudak et al. 2008; Yu et al. 2011). Koch (2010) provides a review on the developments in the integrated application of optical/LiDAR remote sensing for forest biomass assessment. The term "small area" is commonly used to describe a small domain, i.e. a small subpopulation in a large geographical area. Sample survey data of a small area or subpopulation can be used to derive reliable estimates of totals and means for large areas. However, the usual direct survey estimators based on the sampled data are likely to yield large standard errors due to an improper sample size (Ghosh and Rao, 1994). This makes sense in e.g. regional or state-wide forest inventories where the sample size is typically small because e.g. the overall sample size in a survey is commonly determined to provide specific accuracy at a higher level of aggregation than that of small areas.

In regional forest inventory, the aim is often to derive the desired forest structural attributes in an aggregated scale which requires the use of spatial statistics. Due to the high cost associated with sample-based forest inventory, some technical problems often arise with the inventory of some forested areas, specifically within small privately-owned forest patches where conducting reasonably intensive periodical inventory is often a challenge. Thus, a trade-off should be found to retrieve explicit estimates of mean forest attributes in those areas. The regional-scale remote sensing data in 2D and 3D forms will be applicable for retrieving forest inventory attributes. Provided that the inventoried stands are within the span of stands for which forest attributes are of interest (i.e. comparable structure and variation), off-site samples from the inventoried stands can be used for prediction of target variables in target stands of interest (Lappi, 2001, Nothdurft et al. 2009). Those off-site samples can facilitate regionalization of response variables using spatial models e.g. nonparametric approaches (Mandallaz, 1993).

We used aerial LiDAR and multispectral metrics. Two nonparametric methods of Random Forests (RF) (Breiman 2001) and Most Similar Neighbour (MSN) inference (Moeur and Stage 1995) were employed, which differ primarily in the way they handle the distance calculations and NN imputations.

To explore the effect of combining off-site sample data, three sampling designs were utilized to draw samples from target units. This enabled us to pursue trends of selected diagnostics over the varying number of auxiliary samples from target area. Combinations of inventory data from the previously-inventoried forests (off-site data) with sampled data from the neighbouring area (as private target area) were applied to predict the forest structure in the target data. Further, varying sampling intensities were tested. It also was attempted to address the issue of how the auxiliary data could be drawn from the target area by using three sampling designs.

Standard Error of Imputation (SEI) was used as a diagnostic tool together with the conventional error terms of relative RMSE and Bias as diagnostic tools. Moreover, nonparametric confidence intervals (CIs) were built based on bootstrap resampling of RMSE yielded from NN inferences.

To address how MSN and RF methods perform in presence of correlated and uncorrelated predictors from the remotely sensed data, two sets of pruned and unpruned auxiliary variable sets were used, where a GA was used to prune the auxiliary variables for both imputation methods. Additionally, an "off the shelf" case was carried out to impute the values using all predictor variables. In this study, the functionality of RF and MSN methods in a specific case of imputation using off-site samples was considered to explore the "mobility" of the

---

[*] Corresponding author

methods related to a set of affecting factors, including the size of predictor variable set.

2. MATERIALS

2.1 Study area

Two sites, Karlsruhe (hereafter called KA) (8° 24' 09" E and 49° 03' 37" N to 8° 25' 49" E and 49° 01' 15" N) and Ettlingen (hereafter called ET) (8° 24' 52" E and 48° 57' 47" N to 8° 27' 08" E and 48° 56' 11" N), are located near to the city of Karlsruhe in BW (Fig. 1). KA is located north of Karlsruhe and covers over 900 ha of managed forests. ET is approximately 8 km to the south and covers 370 ha. These temperate forests are dominated by Scots Pine (*Pinus sylvestris* L.), European Beech (*Fagus sylvatica* L.), Sessile Oak (*Quercus petraea* Liebl.) and Pedunculate Oak (*Q. robur* L.). Other species play a minor role. The mean standing timber volume for each location is 264 m^3 ha^{-1} (KA) and 320 m^3 ha^{-1} (ET), which are both less than the average rate in BW (362 m3 ha-1).
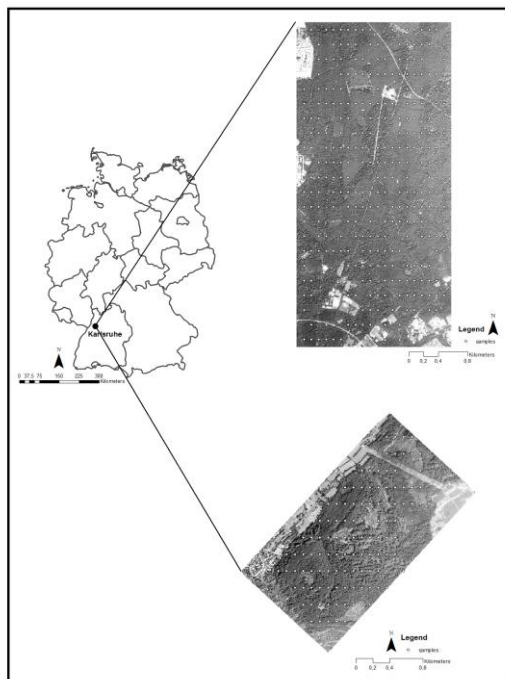

Fig.1. Geographical location of the study area in Germany, overlaid by the ground samples

2.2 Field Data

The field data was collected in the summer of 2006 during a forest inventory based on permanent circular sample plots. The collected data contains measurements from 300 plots in KA and 122 plots in ET. Both datasets were collected using a regular 200 m × 100 m grid with various radii, on which different tree characteristics including diameter at breast height (DBH) were measured. The height of two dominant trees of each main species and one dominant from the other mixed species are measured, and the heights of remaining trees are predicted by uniform stand height curves. Single tree timber volume was calculated using the taper functions of Kublin (2003), and the total timber volume in m^3 ha^{-1} was derived by summing the single tree volumes weighted by the inverse of the corresponding sample plot area. The above-ground biomass of single trees (including bark, branches and needles) was

estimated with Zell (2008) parameters for an allometric equation, and the total biomass on each sample plot was derived by summing all single tree biomass estimates. Total standing timber volume [m^3 ha^{-1}], stem count and total biomass [tons ha^{-1}] were used as response variables in the simulations. As for the stem count, only trees with DBH> 25 cm were considered as the first-pulse LiDAR metrics primarily characterize the older stems in the overstory of the stands.

2.3. Remote sensing data

Full wave LiDAR data was acquired for both sites in August 2007 by Toposys GmbH using the Harrier56 LiDAR system on a helicopter using a Riegl LMS-Q560 laser scanner. The scanner featured a beam divergence of 0.5 mrad, and the data were acquired at a flight altitude of 700 m, with a flight speed of 30 m/s, average footprint size of 0.35 m and density of 16 points m^{-2}. RiANALYZE560 software was used to extract the total number of echoes within each laser beam, out of which the first pulse data were applied for the study. Various laser metrics were derived from the data at sample plot level. The height metrics consisted of quantities including measures of central tendency, measures of dispersion, L-moments, height percentiles and height proportions which were computed directly from the normalized point cloud. The intensity values were also used to derive metrics (Boyd and Hill, 2007).

Optical data in 4 spectral bands was recorded on July 2008 by Toposys GmbH using a RGB/NIR line scanner installed on a Falcon II system. The data covers an approximate range of 450 nm in the electromagnetic spectrum. In addition to the plot-scale mean values of four original spectral bands, the metrics were derived as mean values of variance of the bands, NDVI, two components of principal component analysis (PCA), and texture indices of all four original channels. Texture analysis was carried out using a mean Euclidean distance algorithm developed by Iron and Petersen (1981).

To remove the possible effects of substantial differences in the values of predictor variables, the two study sites were selected in a way that the remote sensing data could be provided from the similar optical and LiDAR sources recorded by the same instruments in the same flight.

3. METHODS

3.1 Variable screening

Reducing the dimensionality of large predictor datasets, especially in those associated with an abundance of correlated predictors, is a preliminary aim in modelling procedures (Cadima et al. 2004). The local search heuristics such as GA have been successfully applied to select approximately uncorrelated predictors and alternatives for NN models by e.g. Bakken and Jurs (2000) and Latifi et al (2010a). GA is a variable search procedure based on the principle of evolution by natural selection. The procedure works by evolving sets of variables (called chromosomes) which fit certain criteria from an initial random population via cycles of differential replication, recombination and mutation of the fittest chromosomes. Although RF has been reported to come with its own variable importance measure which performs roughly the same as GA (Svetnik et al. 2003), we performed GA on both cases of RF and MSN inferences to conduct consistent variable screening for both modelling procedures. The GA variable search was based on maximising the Tau squared index $\tau^2 = 1 - \Lambda^{\frac{1}{r}}$, where the maximization of τ^2 is

equivalent to the minimization of the Wilks' Lambda criterion Λ, and r is the rank of hypothesis (effect description) matrix. A thorough description of the method is provided by Cadima et al. (2004). The initial population was set to 100, the probability of mutation was 0.01, maximum number of identical replicates (clones) was 1, and the algorithm was run for 200 generations and 1000 solutions. Accordingly, 12 covariates were selected for consequent modelling, as reported to be an optimum amount of predictors by e.g. Maltamo et al. (2009) and Latifi et al. (2010b).

3.2. Predictions

The advantages of non-parametric methods are that they retain the full range of variation of the data as well as the covariance structure of the population (Moeur and Stage, 1995). Amongst NN methods, MSN and RF have been shown to provide plausible imputation results for forest structure (Hudak et al. 2008; Eskelson et al. 2009). The MSN inference uses canonical correlation analysis to produce a weighting matrix used to select NNs from reference units. It searches for linear transformations U_k and V_k for the set of predictors X and response variables Y which maximizes their correlations:

$$U_k = \alpha_k Y$$
and
$$V_k = \gamma_k X$$

which will be used as weighting matrix for the MSN distance function:

$$D_{uj}^2 = (X_u - X_j)\Gamma\Lambda^2\Gamma(X_u - X_j)' \quad (1)$$

Where X_u is the vector of known descriptor variables from the target observation, X_j is the vector of descriptor variables from the reference observations, Γ is the matrix of canonical coefficients of predictor variables, and Λ is the diagonal matrix of the squared canonical correlations.

The regression tree method of RF (Breiman, 2001) was used as a second method. The RF has been reported to be very efficient, especially when the number of descriptors is very large (Svetnik et al. 2003). The algorithm works as follows (Liaw and Wiener, 2002):

   (1)   Bootstrap samples are drawn from the original data.

   (2)   For each bootstrap sample, an *unpruned* regression tree is grown. The best splits are chosen from the randomly-sampled variables at each node.

   (3)   New predictions are made by aggregating the predictions of the total number of trees. That is, the mode votes from the total trees will be the predicted value of the respective variable.

Here, the RF was set to generate 300 classification trees per response variable to ensure stable results in repeated runs (Hudak et al. 2008).

The number of NNs k considered for predictions of the target unit can be set to any number smaller than the entire number of samples. As using multiple k could make biased predictions by shifting the predictions towards the sample mean, a single NN was used here. Using more NNs would presumably return more accurate results, as averaging would provide a wider variety of responses (See LeMay and Temesgen, 2005). Yet, this should be carried out cautiously, particularly in cases where small sample sizes are used.

The models were implemented on pruned dataset, as well as on all-predictor dataset. There were two "extreme" cases, where 1) NONE of auxiliary sample was selected from KA, i.e. all the KA units were predicted by means of ET data as the reference dataset and 2) ALL of the KA units were modelled using KA samples as references (without including any ET units). For all other models, all samples from the ET were combined with varying intensities of sampled units from KA to outline the reference dataset, where KA units served as target units to be predicted. The amount of sampled data ranged from 10 to 200 samples (in 10-sample increments), drawn by means of three sampling designs of random, stratified random and stratified systematic methods.

For the stratification (in which unequal probabilities between strata should be used) vector of first order inclusion probabilities of KA samples $\pi_i = P(I_i = 1)$ were computed, in which I_i denotes sample inclusion indicators, $\pi_i > 0, i = 1, 2, \cdots, N$ and selection probabilities are equal in each stratum (Tillé and Matei, 2009). As for the random stratified method, the units were randomly sampled within each stratum, whereas in the last method they were systematically drawn within each stratum. The stratum sample sizes were determined in proportional to their size.

3.3. Diagnostics

Model performance was assessed using leave-one-out cross validation. For each of the combinations, the models were run in a looping process of 100 repetitions from which the mean values of diagnostics were summarized and reported (Maltamo et al. 2009). The previously-sampled KA units were excluded prior to calculation of statistics to avoid reporting optimistic results. A range of diagnostic tools were used, including:

- Statistics based on comparison of observed and predicted values in the simulated target data set. They comprise of RMSE

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad Bias = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n}$$

and the bias. Both statistics were reported as relative forms (absolute RMSE and bias normalized to the mean of the observed values).

- Standard Error of Imputation (SEI): According to Stage and Crookston (2007), error properties of imputation errors differ from those of regression estimates, as the two methods include a different mix of error components. In an imputation, sources of errors are divided to different sources including Measurement error and Pure error. If Pure and Measurement errors are assumed as the stochastic components drawn from distributions having zero mean and zero covariance, the estimate of twice the sum of variances of pure error and measurement error can be obtained by averaging the squared differences for some fraction of the units with short Mahalanobis distances (MMSD0), where Mahalanobis distances are calculated in the space of normalized, uncorrelated X (predictor) variables. This distance is selected because other distance functions may transform the X variables such that the dimension of the space spanned by the transformed X variables is lower than that of the original X space. The SEI will be yielded using:

$$SEI = \sqrt{RMSD^2 - RMMSD0^2/2} \quad (2)$$

where $RMSD^2$ denotes the squared RMSE of the underlying model, and $RMMSD0^2$ indicates the squared RMSE of the Mahalanobis distance model using the similar predictor feature space. A proof can be found in Stage and Crookston (2007).

As one might consider obtaining more accurate information about distribution of the error terms in a regression analysis, the

RMSEs of predictions in each of those combinations were simulated using bootstrap resamples for both imputation methods. Nonparametric bootstrap allows estimation of the sampling distribution of a statistic empirically without making assumptions about the form of the population, and without explicit derivation of the sampling distribution. Here, the RMSE (as prominent error term of the underlying predictions) was bootstrapped 500 times to obtain information on its distribution. The resampling method was balanced bootstrap (Davison et al. 1986).

The major application of distributions from bootstrap is to calculate the CIs, which help to assess the uncertainty. Here, 95% basic and percentile bootstrap CIs were applied as described by Davison and Hinkley (1997).

4. RESULTS

4.1. RMSE

Fig 2 depicts the relative RMSE% for all combinations. In each graph, the left and right ends show the diagnostics rates when NONE samples (left end) and ALL samples (right end) from KA have been included as auxiliary plots. Both statistics were calculated once, and thus remained unchanged in all three sampling designs. Whereas a tremendously high RMSE was observed when using an all-descriptor set for MSN prediction of the entire KA plots without any auxiliary variables (Fig 2: column a), RF showed to be more robust than "off-the-shelf" mode, where all KA samples were predicted using all predictors at once (Fig 2: column b). Both methods showed a constant reduction in RMSE% as more auxiliary samples were added to the reference set. However, the error was still higher than the "ideal case" of accessing all plot information from KA, even when using 200 supplementary samples. In this case both methods performed relatively similar, though the differences in the right end of the MSN graphs are vaguely masked in the illustration, as a result of plot scaling.

Using a GA-pruned set of descriptors greatly enhanced MSN compared to its previous performance, especially when used without auxiliary plots at the left end (Fig 2: column c). Still, RF showed roughly the same trend as before, except that the errors slightly increased compared to its "off-the-shelf" mode (Fig 2: column d). Again, both imputation schemes showed a constant reduction in RMSE%. In general, RF performed better than MSN, as was expected. The best results were obtained when drawing systematic stratified samples, which performed only slightly better as compared to stratified random (second best performance) and random sampling. The stem count achieved the highest RMSEs compared to standing volume and biomass, where the two latter responses performed roughly similar. In all cases, RMSE began to rapidly increase as the first set of auxiliary samples were added and then gradually abated as additional data were appended.

4.2. Bias

In Fig 3, the relative bias% rates in the absence and presence of all auxiliary plots from KA are depicted at left and right ends of each graph, respectively.

In the absence of auxiliary samples for MSN, following an exceedingly biased prediction at the beginning, the bias was reduced notably as additional plots were added to the reference dataset (Fig 3: column a). As for the similar case in RF, the predictions were still biased at the beginning (though much less

biased than those from MSN), and yielded lower-biased estimates as more supplementary samples were added. It then remained nearly constant until the end, though the most accurate predictions (using all KA samples) were more biased than those obtained from a combination of ET and KA samples (Fig 3: column b).

When pruned predictors were used, the two imputation methods behaved almost similarly. In the absence of KA plots, the more biased estimates of MSN (yet markedly better than those using all predictors) tended to improve notably as more KA samples were added to the reference set (Fig 3: column c). Comparably, RF predictions followed a similar trend (Fig 3: column d). Yet, MSN predictions had surprisingly lower-bias (sometimes negligibly unbiased) in contrast to those of RF. Sampling design did not contributed to a notable difference in Bias.

4.3. SEI and simulated RMSE

The SEI output is presented as absolute values in m3ha-1 (trees ha-1 for stem count) in Fig 4. In an all-predictor case (Fig 4, 1st row), MSN showed to be considerably more sensitive than RF when supportive samples were added, whereas the latter behaved more stable (particularly at the beginning).

However, both methods performed somewhat differently depending on whether random or stratified samples were drawn from KA. Whereas SEI tended to increase slightly when drawing random samples, it showed a reduction (in MSN) or remained fairly unaffected (in RF) when drawing stratified samples. The drastic changes at the beginning were gradually flattened as further supplementary samples were added. In the 12 predictor-case (Fig 4, 2nd row), both imputation methods performed quite similarly. However, SEI increased a small amount when random samples were drawn, whereas it showed a more stable trend for both inferences in either of the stratification designs.

In terms of bootstrapped RMSE, The two applied 95% nonparametric CIs corresponded quite well (Fig.5). Furthermore, the majority of observed values of RMSE lied between low-and upper boundaries of CIs.

5. DISCUSSION AND CONCLUSIONS

The suitability of diverse remote sensing data types for retrieval of forest parameters compels the use of multi-sensorial data for forest modelling (Koch, 2010). This study concerned the application of several error properties of two well-known nonparametric processes, which can serve as means to fill in missing datasets by imputing values of measured observation units to interspersed, less frequently inventoried forest patches.

In an all-predictor case and without any auxiliary samples from KA, MSN entirely failed to project the reference to the target units. This may produce even worse results, especially in cases where there are less-overlapping populations of descriptor variables or when large distances are used. Although we forced MSN to scale both sets of predictors prior to projecting them into canonical space, the method was still unable to produce reliable results unless supplementary plots are added to the reference dataset. In contrast, RF uses ensembles of trees to improve the performance of decision trees, and can be trained more rapidly compared to similar methods (Breiman, 2001). Moreover, RF was rather insensitive to using unreduced predictor variables, as descriptor selection is intrinsic to tree growing process.
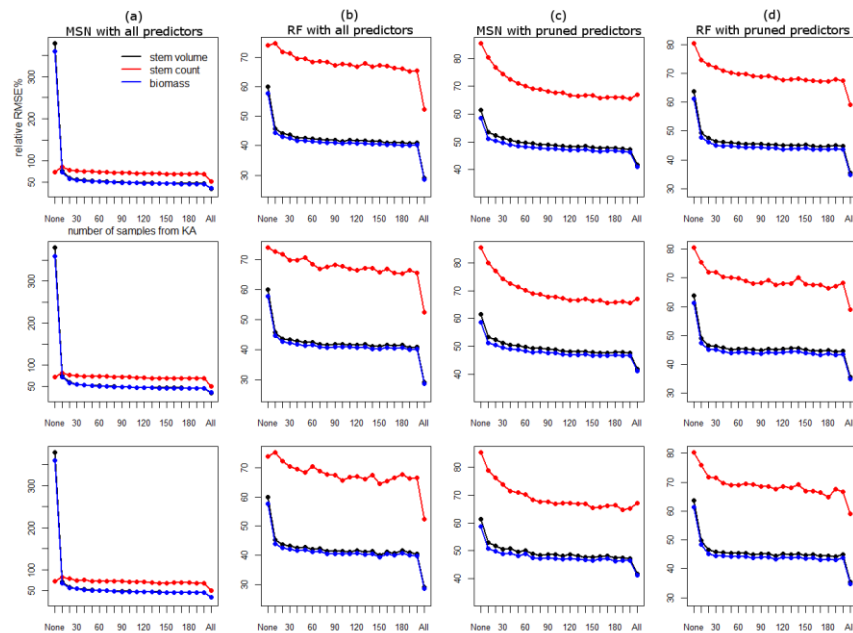
Fig.2. Relative RMSE% of NN predictions. In each row, from left to right: MSN with all predictors (a), RF with all predictors (b), MSN with pruned predictors (c), and RF with pruned predictors (d). Each column shows predictions using random (top), stratified random (middle) and systematic stratified (bottom) samples.
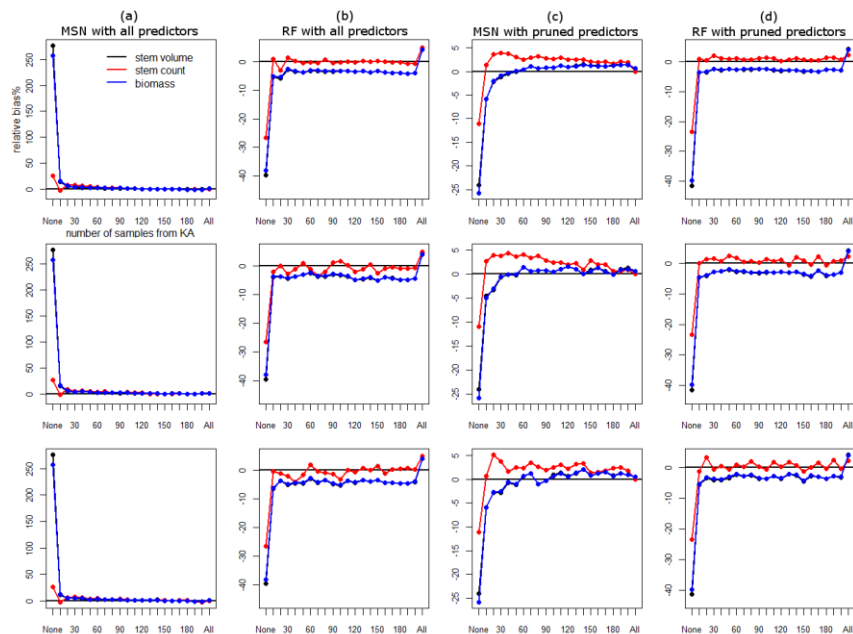

Fig.3. Relative bias% of NN predictions for the entire combinations. Description is similar to Fig. 2

Previous studies by Vauhkonen et al. (2010) and Latifi et al. (2010a) reported the general dominance of RF irrespective of the number of predictors. Here, it was shown that RF models can be even applied to off-site imputations.

Appending auxiliary units showed to be most influential at the early stages, i.e. as 10 to 30 samples (approximately 3 to 10% of intensity) were added to the reference set. We used three sampling procedures to draw auxiliary units, which allow capturing the properties of the individual strata that are unique to each individual stratum. The results indicated the slight enhancement of accuracy following systematic stratified sampling. We used a proportional allocation to define the stratum sample size. However, other methods e.g. Neyman allocation (Tillé and Favre 2005) are promoted to be examined. The result of MSN in all-predictor mode was poorly biased when used to predict without auxiliary plots. The bias was trimmed down to approx. 14% (by adding 10 auxiliary samples) and later to 6% (by adding 20 samples).

The results showed a negligible bias when using 40 supplementary samples onwards. The results were slightly improved by using a pruned predictor set. This supports reports by Latifi et al (2010a) concerning lesser bias of MSN compared to RF predictions.

We used samples as combinations to impute units from another population, which is different from a conventional sampling situation, where a sample is selected to produce estimates for the same population. However, the MSN estimates still produced negligibly low-biased predictions, though one may note that there is no reason for such estimates to be (approximately) unbiased.
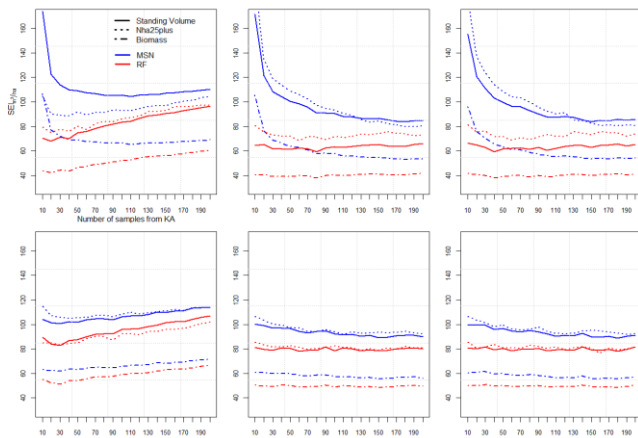
Fig.4. Absolute SEI of NN predictions for all combinations for single NN. Rows show values using all descriptors (a) and pruned descriptors (b). In each row, from left to right: random, stratified random and systematic stratified samples
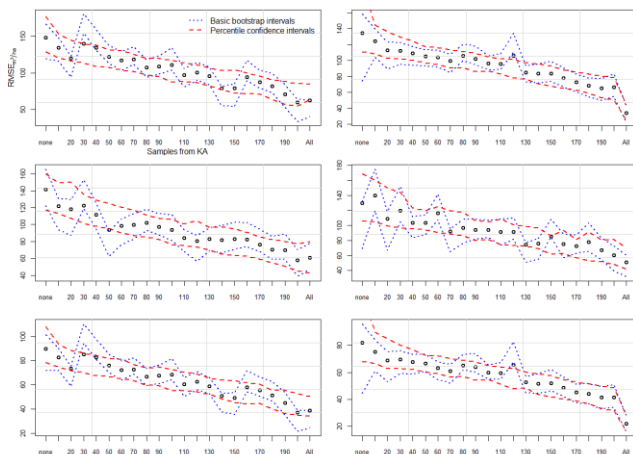

Fig.5. Bootstrap 95% CIs of simulated absolute RMSEs [m3ha-1] resulted from predictions using pruned descriptors, and systematic-stratified samples in a single NN case. Columns show MSN (a) and RF (b). In each column, from top to bottom: standing volume, stem count, and biomass.

This follows Lappi (2001) who stated that off-site plots basically cannot be used if one insists on obtaining completely unbiased estimates for an area.

In terms of SEI, RF models showed lower SEI values than those from MSN. Whereas the value of the SEI showed a sudden reduction when adding up to 30 sampled units for MSN inferences, it showed to be rather steady in case of RF. when using reduced predictors, both imputation methods changed steadily as more samples were added. One remarkable case was the sensitivity to how the additional samples were drawn. As a stratified sample can provide greater precision than a simple random sample of the same size, it seems that simple random sampling adds up the variance of the reference set slightly more than in a stratification process. Thus, statistics such as SEI are conditional on distribution of descriptor values in the reference set and could be used to guide decisions on whether to augment that reference set (Stage and Crookston, 2007).

The majority of our observed RMSEs for single NN, pruned models lie within range of both bootstrap CIs. The RMSE was selected as the statistic to be bootstrapped here, as it is addressed in a variety of relevant literature as a baseline measure of predictive accuracy. We assessed the bootstrap

simulations using basic and percentile CIs. These methods are considered appropriate to obtain information about the bootstrap distribution, as long as the distribution reveals a symmetric shape (Davison and Hinkley, 1997).

This study aimed at making use of regional remote sensing, to find potential implications for retrieval of forest attributes of small areas. Concerning increasing trend of the availability of airborne remote sensing information in BW, the applied procedures could be of operational use to produce mean values of forest structural attributes. It is concluded that the combined use of remote sensing and forest inventory could be an interesting topic to be considered for describing forest structure. In this manner, the use of RF approach is suggested as a more robust method to deal with problems such as overfitting and large numbers of descriptors.

ACKNOWLEDGEMENT

REFERENCES

Bakken, G. A., Jurs, P. C., 2000, Classification of multidrug-resistance reversal agents using structure-based descriptors and linear discriminant analysis. *Journal of Medicinal Chemistry,* 43, pp. 4534-4541.

Boyd, D. S., Hill, R. A., 2007, Validation of airborne LiDAR intensity values from a forested landscape using HYMAP data: preliminary analysis. Proceedings of the ISPRS Workshop 'Laser Scanning 2007 and SilviLaser 2007' Part 3 / W52, Espoo-Finland, 71-76 pp.

Breiman, L., 2001, Random Forests. *Machine Learning,* 45(1), pp. 5-32.

Cadima, J., Cerdeira, J. Orestes Monhoto, M., 2004, Computational aspects of algorithms for variable selection in the context of principal components. *Computational Statistics and Data Analysis,* 47, pp. 225-236.

Davison, A.C., Hinkley, D. V., Schechtman, E., 1986, Efficient bootstrap simulation. *Biometrika,* 73(3), pp. 555-566.

Davison, A.C., Hinkley, D.V., 1997, Bootstrap methods and their application, pp. 191-220 (Cambridge: Cambridge University Press).

Eskelson, B.N.I., Barrett, T.M., Temesgen, H., 2009, Imputing mean annual change to estimate current forest attributes. *Silva Fennica,* 43(4), pp. 649-658.

Ghosh. M., Rao.J, N, K., 1994, Small Area Estimation: An Appraisal. *Statistical Science,* 9(1), pp. 55-76.

Hudak, A., Crookston, N., Evans, J., Hall, D. and Falkowski, M., 2008, Nearest neighbour Imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment,* 112, pp. 2232-2245.

Iron, J. R., Petersen, G. W., 1981, Texture Transforms of Remote Sensing Data. *Remote Sensing of Environment,* 11, pp. 359-370.

Koch, B., 2010, Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment. *ISPRS Journal of Photogrammetry and Remote Sensing,* 65, pp. 581-590.

Kublin, E., 2003, Einheitliche Beschreibung der Schaftform – Methoden und Programme – BDATPro. *Forstwissenschaftliches Centralblatt,* 123, pp. 183-200 (in German with English summary).

Lappi, J., 2001, Forest inventory of small areas combining the calibration estimator and a spatial model. *Canadian Journal of Forest Research,* 31(9), pp. 1551–1560.

Latifi, H., Nothdurft, A. Koch, B., 2010a, Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR –derived predictors. *Forestry,* 83(4), pp. 395-407.

Latifi, H., Nothdurft, A., Straub, C., Koch, B., 2010b, Optical/LiDAR feature search for nonparametric prediction of stratified forest attributes using an improved genetic procedure. Proceedings of *Silvilaser2010*, 14-17 September 2010, Freiburg-Germany.

LeMay, V., Temesgen, H., 2005, Camparison of nearest neighbour methods for estimating basal area and stems per hectare using aerial auxiliary variables. *Forest Science,* 51(2), pp. 109-119.

Liaw, A. Wiener, M., 2002, Classification and regression by randomForest. *RNews,* 2, pp. 18−22.

Maltamo, M., Næsset, E., Bollandsås, O.M., Gobakken, T., Packalén, P., 2009, Non-parametric prediction of diameter distribution using airborne laser scanner data. *Scandinavian Journal of Forest Research,* 24, pp. 541-553.

Mandallaz, D., 1993, Geostatistical methods for double sampling schemes: application to combined forest inventories. Habilitation thesis. ETH Zurich-Switzerland.

Moeur, M. Stage. A. R., 1995, Most similar neighbor: An improved sampling inference procedure for natural resource planning. *Forest Scienc.,* 41, pp. 337–359.

Nothdurft, A., Saborowski, J. Breidenbach, J. 2009. Spatial prediction of forest stand variables. *European Journal of Forest Research,* 128(3), pp. 241-251.

Stage, A. R. Crookston, N. L., 2007, Partitioning error components for accuracy-assessment of near-neighbour methods of imputation. *Forest Science,* 53(1), pp. 62−72.

Svetnik, V., Liaw, A., Tong, C., Cullberson, C., Sheridan, R. P. Feuston, B. P., 2003, Random Forest: a classification and regression tool for compound classification and QSAR modelling. *Journal of Chemical Information and Computer Sciences,* 43, pp. 1947-1958.

Tillé, Y., Favre, A.C., 2005, Optimal allocation in balanced sampling. *Statistics and Probability Letters,* 74, pp. 31–37.

Tillé, Y., Matei., A, 2009. Sampling: Survey Sampling. R package version 2.2. Available online at http://cran.r-project.org/web/packages/sampling/index.html (accessed 16.02.2011).

Vauhkonen, J., Korpela, I., Maltamo, M. Tokola, T., 2010, Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sensing of Environment,* 114, pp. 1263–1276.

Yu X., Hyyppä, J., Vstarana, M., Holopainen, M., Viitala, R., 2011, Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing,* 66(1), pp. 28-37.

Zll, J., 2008, Methoden für die Ermittlung, Modellierung und Prognose der Kohlenstoffspeicherung in Wäldern auf Grundlage permanenter Großrauminventuren. Ph.D thesis, Faculty of Forest and Environmental Studies, University of Freiburg, Germany. 162p (in German with English summary).