Regionalization of Variables of Sample Based Forest Inventories at the District Level

Dissertation

zur Erlangung des Doktorgrades der Fakultät für
Forstwissenschaften und Waldökologie der
Georg-August-Universität Göttingen

vorgelegt von
Jens Nieschulze
geboren in Uelzen

Göttingen 2003

D7

1. Berichterstatter: Prof. Dr. J. Saborowski

2. Berichterstatter: Prof. Dr. A. Akça

3. Berichterstatter: Prof. G. S. Biging, PhD

Tag der mündlichen Prüfung: 14. Februar 2003

Diese Arbeit wurde durch die Niedersächsische Staats- und Universitätsbibliothek, SUB-Göttingen, unter folgender Internetadresse veröffentlicht:

http://webdoc.sub.gwdg.de/diss/2003/nieschulze/

Worüber wir nicht sprechen können, darüber müssen wir schweigen.

(That we cannot speak about, we must pass over in silence.)

[Ludwig Wittgenstein]

Danksagung

Contents

List of Figures

List of Tables

Abbreviations

B°   Bestockungsgrad (stocking density)

BI   Betriebsinventur (forest inventory at the district level)

CIR   Color-Infrared: color images, where one of the three light sensitive layers records in the near infrared spectrum

d_G Diameter of the "Grundflächenzentralstamm" (median cross sectional area tree): the diameter by which the tree population ordered by basal area is divided into two equal halves

dbh   Diameter at breast height: diameter of a tree at 1.3 m above ground

DSM   Digital surface model: representation of a landscape and its features like trees, buildings etc., by elevation above sea level (asl.)

DTM   Digital terrain model: representation of a landscape by its ground elevation above sea level (asl.)

VIII

FVA   Forstliche Versuchsanstalt (research station of the forest service)

GIS   Geographical Information System: a software, that enables
spatial localization, analysis and visualization of data

ISODATA   An algorithm used in cluster analysis

MSN   Most similar neighbor

REML   Restricted maximum likelihood

RMSE   Root mean square error

SI   Site Index

STR_R   Standraum$_{Reduziert}$ (reduced growing space)

UK   Universal kriging

1 Introduction

The concept of sustainability has been becoming widely popular among people of all walks of life after the earth summit 1992 in Rio de Janeiro (United Nations, 1992), where public awareness was raised to new heights. Sustainability has a spatial and a time component. Assessment of this concept hinges on comparing conditions over time.

The concept of sustainability has a long tradition in German forestry. The term was first mentioned by Carlowitz (1713). Ever since it has been the guiding management principle and today is anchored in the federal state's (Länder) forest law (Niedersächsische Landesregierung, 1992). Its interpretation has been developed over time. Originally, sustainable forest use was concerned with sufficient supply of timber. Nowadays the concept is applied to the complex and dynamical ecosystem forest with its multi-functional products. Criteria and quantitative indicators of sustainable use have been proposed for temperate wood-lands (Schneider, 1995) and are applied in the certification process of forest enterprises (Brahms and Graulich, 2000). Information that is needed by management activities can be categorized threefold according to Davis (1966) as follows:

1. Information that is external to the forest, e.g. property rights or weather.

2. Direct information that is derived from the forest and its condition, e.g. standing timber volume or diversity.

3. Operational information like budget of management objectives.

The direct information is derived from field measurements. The forest management planning is based on such direct information among other sources and surveys are usually conducted on a 10 year cycle. In the classical survey approach, each stand is visited by an inventory forester and, based on an ocular assessment and some measurements, stand characteristics like volume per hectare are derived from a yield table. Information at the enterprise scale is then obtained by assembling the information at the stand level.

The introduction of silvicultural programs aiming at transforming typical single

species rotation forest management stands into mixed structured continuous forestry system stands (Otto, 1989; Otto, 1991) renders the described approach impractical. The yield table is no longer applicable owing to different growth patterns in temporally and spatially mixed forests, yielding imprecise estimates (Tzschupke, 1991).These changed conditions led to the introduction of sample based inventories at the enterprise level by the state forest department and by private forest owners.

Sample based inventories have advantages over classical approaches (Böckmann et al., 1998), namely feasibility of extensive measurements because the actual surveyed area is reduced, statistical inference of the uncertainty of the measurements, and improved increment estimation and management control. Especially the last point has become very important in the course of forest decline research (Beck, 1999).

However, the major drawback of sample based inventories is that they cannot replace the stand wise assessment of classical forest management planning surveys (Bitter and Merrem, 1997; Tzschupke, 1991). Confining target populations entails lower sampling densities which is problematic because the variation within the target populations will not decrease accordingly. Economically viable sample densities for forest enterprises of around 6-7 ha per sampling unit mean that few if any stands will comprise enough sample points to allow sound statistical inference. Some stands will even lack any sample at all.

A global assessment of sustainability, e.g. at the country level as an extreme example, does not conform to its definition. The normal-forest model (Hundeshagen, 1826) implies per definitionem sustainability at the global level although it entails clear cutting large areas. It should be clear that sustainability should be evaluated at the spatial scale at which management or land use activities are carried out, that is, at the stand level in case of forest enterprises.

Conventional inventories have operated at the stand level but will be gradually outphased due to cost constraints. The need of stand wise information demands the development of techniques that enable a reliable and evaluable prediction of stand characteristics based on the high quality point information of field sample plots.

2 Objective

The objective of this study is to explore and compare different methods that allow an estimation of characteristics at the stand level based on information surveyed at the district level. From a theoretical viewpoint the objective is equivalent to the generalization of point information (the sample plots) into the spatial domain (the forest stands). In a statistical sense, the problem can either be regarded as a regionalization (Journel and Huijbregts, 1978; Saborowski and Stock, 1994) or as a small area estimation (Särndal et al., 1992; Dees, 1996).

The basic idea underlying all approaches is to consider only similar observations or a weighted average thereof. The approaches differ with respect to how similarity is defined. In the spatial domain, similarity between a sampling unit and an arbitrary location refers to their spatial distance. An appropriate measure then is the spatial correlation structure among the sampling units. Such an approach comprises not only the values of the variables surveyed at the sample locations but also their alignment in space. The resulting spatial interpolation is commonly called kriging.

The "feature space similarity" between a sample location and an arbitrary location refers to how alike the surveyed auxiliary characteristics are. Distance can be derived by multivariate techniques like canonical correlation analysis, can be heuristically modeled, or can be regarded constant as done with simple mean estimators.

All introduced methods depend on auxiliary information, that is, additional information needs to be available at the sample location as well as at all prediction locations. The use of auxiliary information in forest inventories can be regarded as established with analog CIR imagery being the principal source (Akça et al., 1993; Bitter and Merrem, 1997). Another source of auxiliary information is provided by the field assessment ("Bestandesbegang"). The field assessment is implemented by forest administrations as part of the new sample based forest management planning.

The comparison of the investigated methods comprises their complexity, ease of implementation, accuracy, and demands for auxiliary information. Practitioners can thus choose from the repertoire to tailor their implementation according to their data at hand and needs.

3   Data sources

Three data sets of three different forest districts are investigated. The first, provided by the "Forstplanungsamt" (forest planning office) of the forest administration of Lower Saxony, and the second, provided by the Forstliche Versuchsanstalt, FVA (forest research station), of the forest administration of Baden-Württemberg serve the model selection. The third data set, again provided by the "Forstplanungsamt" from the forest service of Lower Saxony, is used for model evaluation and refinement. The first and third data sets are characterized by a rather low diversity compared to the second data set.

Model selection is be guided by cross-validation or leave-one-out prediction (cf. section 4.7). It should be noted that errors obtained from cross-validating point measurements, that is, the BI plots, will be much larger than prediction errors of entire stands. This is due to the relative small size of the BI plots and the inhomogeneity of stands.

3.1   Lower Saxony: Grünenplan

The forest management planning of the forest service of Lower Saxony is based on a 2 phase sampling design. Its optimization leads to an average density of around 6 ha per terrestrial phase 2 unit (Saborowski and Dahm, 1996).

In phase 1, plots are located on analog CIR imagery with scale 1:12500. The plot alignment follows a 100 m × 100 m raster and plots are classified according to their species group and their age. The arrangement of species groups is binary with "predominantly deciduous" and "predominantly coniferous" . Age classes are "< 40 y.", "≥ 40 & < 80 y.", "≥ 80 & < 120 y.", and "≥ 120 y.". The allocation of phase 1 plots to one out of the possible eight strata is determined by the conditions at the grid-nodes, that is, the phase 1 plots are not representative for the forest stands they are located in but only for a small circle around the node roughly covering a terrestrial phase 2 plot.

In phase 2, a subset of the phase 1 plots are surveyed on circular plots on the ground.

The tree population is divided into three sub-populations which differ with respect to selection probabilities. Trees with a dbh < 7 cm are counted as regeneration by height levels on a circle with radius 3 m. Trees with a dbh ≥ 7 cm and < 30 cm are surveyed on a circle with radius 6 m whereas trees with a dbh ≥ 30 cm are surveyed on circles with radius 13 m. All three circles have the same center. For each phase 2 plot the compartment, sub-compartment, and sub-sub-compartment are recorded according to the present forest delineation. No auxiliary management units ("Hilfsflächen") are considered.

For each surveyed tree the species, polar-coordinates, dbh, and age are recorded. Heights are measured only for one tree per species group as introduced above and per layer. The height tree is chosen according to its dbh which should represent the d_G of its respective layer. The arrangement of the layers is guided by economical criteria as follows:

- Layer 1 is the main stand with economical focus.

- Layer 2 is regeneration under the main stand that should become the next layer 1.

- Layer 3 is under-storey, trees and shrubs underneath the main stand with mainly tree shading or soil cover functions or to reduce hazards.

- Layer 4 are remnant trees of the former main stand with stocking density of ≤ 0.3, towering the main stand.

The missing heights are completed according to Nagel (1999).

All investigated methods depend on auxiliary information, that is, information that is easy to obtain, that has a functional relationship with the variables to be predicted, that should be cheap, and that must be available at the sampling locations as well as at all prediction locations.

The CIR imagery of the forest management planning are taken one to two years in advance of the terrestrial surveys of the phase 2 plots. In addition to the stratification of phase 1 plots the CIR imagery form the basis of the new forest delineation. This

delineation also comprises auxiliary management units and is reviewed in the field as part of a field assessment ("Bestandesbegang"). The field assessment provides quantitative estimates of site index, species mixture proportions, and stocking densities. All estimates rely on expert knowledge only, no measurements are carried out.

The field assessment and the phase 1 plots provide auxiliary variables as part of the already implemented forest management planning procedure. Additional auxiliary information stipulate additional input.

On a large scale, remotely sensed data are the only economically viable source for such additional auxiliary information. Using remote sensing data other than the already available CIR imagery is not considered operational owing to cost constraints. Analog processing of the imagery is also not considered operational because of the large areas inventoried every year and the thus resulting high costs. Digitizing the CIR imagery is cheap (around €25 per image) and enables automatic processing. Therefore, the CIR images are scanned, with a resolution resulting in a pixel size of ca. 0.3 m.

There are two principal approaches to extract auxiliary information from the digital images at hand. The first approach employs the spectral information contained in the three channels (red, green, and blue).

The images are ortho-rectified with the aid of a DTM provided by the cadastre service to account for relief distortions (Hoffmann, 2001). Changes in illumination conditions are accounted for by interactively mosaicing the images by flight path. The selection of which part of an image should be retained in the mosaic is done on-screen. The mosaics are then classified into seven classes using the unsupervised cluster algorithm ISODATA of the software PCI (Tou and Gonzalez, 1974). Clusters are formed on the basis of three variables, the spectral bands.

Super- and unsupervised classification are classical procedures of digital image processing. They represent a mapping aiming at a reduction of dimensions thus facilitating information extraction. In the present case the mapping is from a 3 dimensional space with 256^3 values onto a 1 dimensional one with 7 values.

The selected number of clusters is based on an educated guess. For the different features present in a CIR image, one class is reserved for non-vegetated area. The

vegetated area is divided into crown matter and grass or meadows. Crown matter should be further separated into deciduous and coniferous species. Within species, one class should represent bright crown matter and and another shaded crown matter. The seventh cluster is reserved for gaps occurring in stands. Of course, in general the seven clusters will not correspond to the listed classes, rather, classes will comprise areas of several clusters. More than seven clusters means less reduction in information content not necessarily implying better information extraction. Also, more classes means more data, usually in the Giga-bytes. The influence of the number of clusters on the quality of prediction is studied during model evaluation.

The clustering algorithm selects a set of centroids evenly spaced in the space spanned by three vectors representing the three spectral bands. Each pixel is visited at a time and its Euclidian distances to all centroids computed. The pixel is assigned to the closest centroid. The mean and variance of the controids are re-calculated and the algorithm proceeds with the next pixel. Once all pixels are assigned to the centroids mean and variances of the latter are evaluated. Clusters with close centroids may be merged and clusters with large variances may be split dependent on thresholds set by the user. Then the algorithm repeats the assignment of pixels to clusters. The classification terminates once the number of re-assigned pixels falls below a threshold or the number of iterations exceeds a threshold, both times set by the user.

Any auxiliary information must be attached to the sample plots. This is accomplished using a GIS to produce an intersection file. This file contains circles with radius 13 m aligned on a 100 m × 100 m grid. The size of the circles corresponds to the phase 2 plots and their alignment to the phase 1 raster. This data layer is intersected with the classified images. Intersection means that only the geometries of the intersected layer (the images) are retained which fall within the geometries (the circles) of the intersection layer. In addition, the attributes of the intersection layer are attached to the newly created geometries. One such attribute is the ID of the phase 1 plots by which auxiliary information can be joined to the phase 2 plots.

After intersection the original raster data are converted to vector format. Adjacent raster cells with same value are lumped into one polygon. In a classified image tree crowns are composed of polygons of different classes. Size, shape, and frequency of polygons differ for different species. Allocation of observed polygons to tree species is

tried by own programs that compute per class total area, polygon frequency, maximum and average polygon size, and average ratio of area to circumference.

Tree tops in digital images are often characterized by large spectral values and represent bright spots. Such bright spots get larger class values assigned by the ISODATA algorithm than gaps or shaded crown matter. An approximation of tree frequency per plot is tried by the number of polygons that have as neighbors polygons with lower class value only.

The second approach for the extraction of auxiliary information from the CIR images is based on height information. The different position of objects (parallax) on two adjacent images with sufficient overlap enable their height derivation.

The resulting output is a digital surface model, DSM. In principal, its data format is the same as for the spectral images except that grey values are not confined to $[0, 255]$ (8 bit unsigned) but have a 16 bit signed range,$[-32767, 32768]$. A subset of such a DSM is displayed in Fig. 1. Calculated elevation is in meter above sea level. The derivation of stand height requires a normalization, that is, the ground elevation must be subtracted by means of a digital terrain model, DTM.

The normalized DSMs are imported into a GIS and intersected with the intersection file as described above for spectral images. Further processing of the resulting polygons is done by own programs. Derived characteristics are number of polygons per plot, their average size, height difference of the 0.9 quantile to the 0.1 quantile, the 0.85 quantile as plot height, summed area of gaps, and summed area of failure pixel. Polygons are classified as gap if their height is lower than $\frac{7}{8}\times$ 0.9-quantile. The 0.9-quantile should model the top plot height. Maximum height values should be avoided because of outliers. The reduction factor $\frac{7}{8}\times$ is chosen heuristically. As an example, on plots with an estimated 0.9-quantile height of 40 m all polygons of less than 35 m height are considered 'gap'.

The derivation of the DSM revealed problems of the algorithms used by the employed software Orthoengine (Brostuen et al., 2001) with pixels with large height gradient. A pixel situated between an understorey of 7 m height and an overstorey tree of 40 m height will most often not get an elevation but will be classified as failure. Reasons

Digital–Surface–Modell
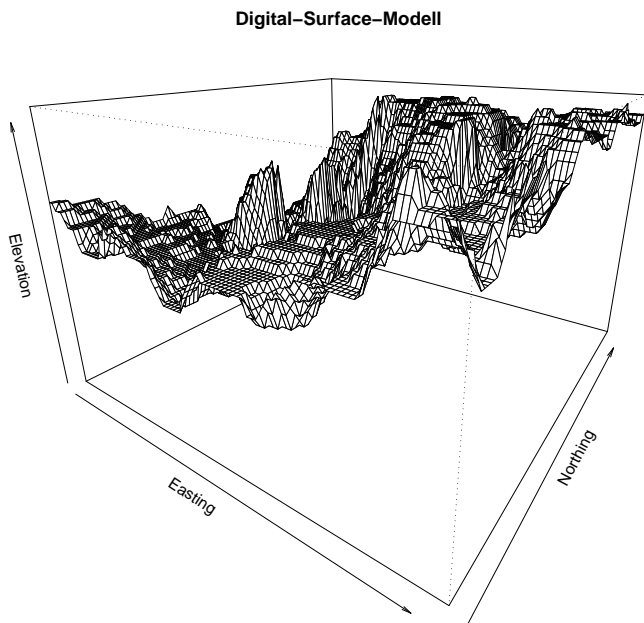
Figure 1.   30 m × 30 m subset of a digital surface model

are the different viewing angles and the search pattern of the matching algorithm.
For example, one image shows the view from above on the tree crown whereas the
other provides the angle from underneath.

The algorithm derives heights of a location from the radial displacement of the cor-
responding homologue pixels. A square sub-window of grey-level values with one
homologue pixel in its center is calculated from one image. Within a given search
radius this window is overlayed with the other image and correlations based on the
grey-level values are calculated. The center-pixel of the window with the highest cor-
relation is taken as the corresponding homologue pixel if the correlation exceeds a
threshold; otherwise a failure value is assigned. Number and area of failure-polygons
are a further source for the derivation of auxiliary variables.

Owing to budget constraints only part of the images covering the forest district
Grünenplan are scanned resulting in an area of approximately 2500 hectares with
441 terrestial plots. Of these terrestrial plots 368 plots are considered here, namely
the plots that comprise at least one spruce (Picea abies, (L.) KARST.) (256 plots)

or one beech (Fagus sylvatica, LINNÉ) (177 plots) in layer 1. Elevation of this subset varies from 175 m asl. to 420 m asl. The distribution of the selected plots is shown in Fig. 2. The two species make up around 80 % of all inventoried trees. As stand characteristic volume per hectare serves as an example.

Data processing revealed that there is no clear cut separation between layers 1, 2, and 4. The diameter distributions of plots with layers 1 and 2 can be quite similar to the ones of plots with layers 1 and 4. Layer 3 contributes little to volume, cannot be assessed in the imagery, and is not present in the investigated area. Against this background it is decided that volume per species is calculated over the layers 1, 2, and 4 but only for plots where spruce or beech occurs in layer 1.
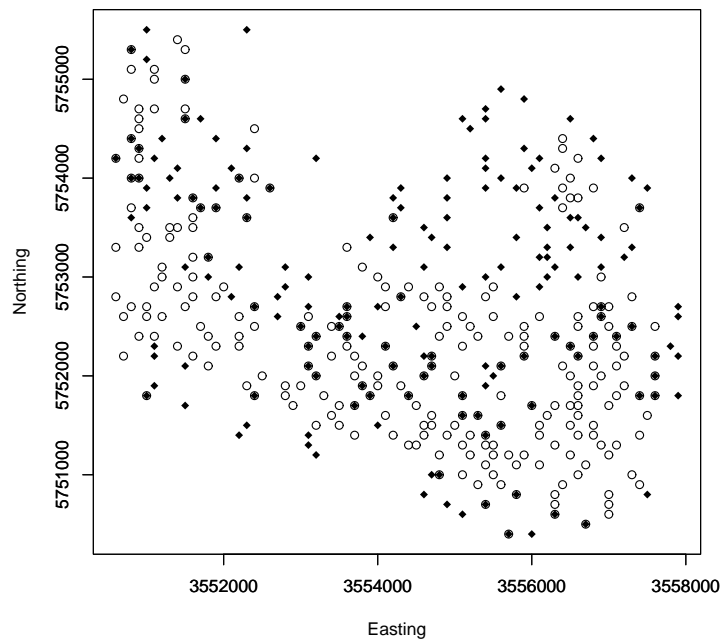
Figure 2.   Lower Saxony, Grünenplan:

Distribution of BI plots; Spruce: Circles; Beech: Diamonds; Gauß-Krüger Coordinates

3.2   Baden-Württemberg: Black Forest

The structure of the forest district as well as the sampling design differ from the ones introduced above for Lower Saxony. Samples are laid out according to a systematic 1 phase design. The distribution of this regular grid is shown in Fig. 3. The forest district is divided into roughly three disconnected sub-regions. Elevation of the field plots varies from 190 m to 770 m above sea level.
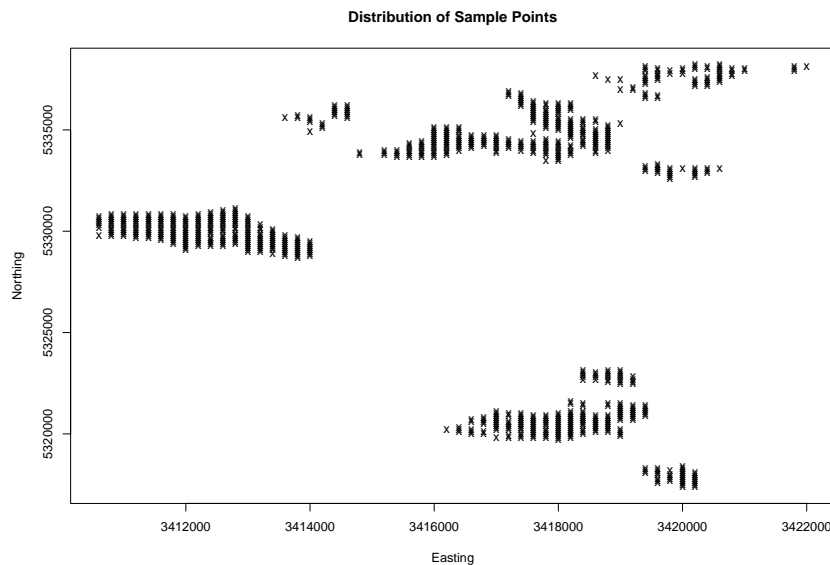
Figure 3.   Baden-Württemberg:

Distribution of all sample plots; Gauß-Krüger Coordinates

According to the methodological description and the surveying instruction of the forest service sample densities are individually derived for forest districts. The guiding principle is a tolerable error of a feature of a given target population. Populations are allocated mainly by age class and stand type and target feature is most often standing volume.

Apart from a tolerable error per population the mean quadratic deviation and the mean of the target variable influence the sample density. An additional rule stipulates at least 500 sample units per forest district. However, cost constraints confine the maximal sample density to 2 hectares per sample. The investigated forest district has been sampled with maximal density and an alignment of the sample units on a

11

regular grid with spacing 200 m×100 m. A total of 791 plots have been surveyed.

The sampling design also divides the tree population into five sub-populations which differ with respect to their selection probabilities. Surveying is also done on concentric circles. Trees with a dbh < 10 cm are inventoried as regeneration by height class. Trees with a height ≤ 1.3 m are counted on a circle with radius 1.5 m whereas taller trees are counted on a circle with radius 2 m. Trees with a dbh ≥ 10 cm and < 15 cm are surveyed on a circle with radius 3 m. Trees with a dbh ≥ 15 cm and < 30 cm are inventoried on a circle with radius 6 m and yet larger trees are surveyed on a circle with radius 12 m. For each tree with dbh ≥ 10 cm species, dbh, and polar-coordinates are recorded.

In the main stand heights ("Oberhöhen") are measured for two trees per species. For the side stands, one height is taken per species. Selection of height-trees is based on expert knowledge aiming at "characteristic" individuals. Age is recorded for height trees only. All other trees are allocated to age classes. The age class is determined by the forest management planning for the entire stand, that is, local deviations from the stand wide age are disregarded. The age classes can be numeric with 10 year interval length, or qualitative like "NV" for natural regeneration or "US" for understorey. Stand layers like used at Grünenplan are not applied.

Almost every stand is visited similar to the "Bestandesbegang" in Lower Saxony. The inventory forester examines delineation of new geometries. Strata are formed from stand types of a forest district. No measurements are taken, all information is based on statistics of the systematic sample and expert knowledge. Volume estimates of tree species are derived from the strata means; the estimates do not necessarily cover all species of a stand. Even species with 20% mixture proportion occur that have no volume assigned.

Owing to the 1 phase sampling design there are no imagery taken in the course of the forest management planning. The only source of large scale hight resolution information are panchromatic airborne images of the cadastre service of Baden-Württemberg. Administrative districts are regularly mapped by means of such imagery and the investigated forest district has been covered in 1998, the year the inventory took place.

Unfortunately, the nominal scale of the images is 1:18500 and they can be obtained on paper print only. The images were scanned at 12 μm, yielding a ground resolution

of around 22 cm. A DTM is not provided by the cadastre service nor does the forest service provide one. An airborne laser scanning is currently carried out but will take some more years before it covers the entire federal state or the forest district under investigation.

3.3   Lower Saxony: Solling

In order to get a realistic assessment of regionalization, knowledge of the true characteristics predicted need be available. No forest stand or compartment will comprise enough phase 2 samples to enable sound statistical inference. The sought information can only be obtained by exhaustive inventory or an extensive sample. Because forest stands vary in size from 2 to 20 ha and different stand-types should be taken into account in an evaluation, an exhaustive inventory is not feasible as part of this study. Instead 12 stands, 4 types by 3 age-classes, were extensively sampled.

The stand types are pure beech (type-id 20), pure spruce (type-id 50), mixed beech-spruce (type-id 25), and mixed spruce-beech (type-id 52). The age classes for the beech dominated stands are "≥ 60 & < 80y.", "≥ 80 & < 120y." and "≥ 120y.". For the spruce dominated stands the age classes are "≥ 50 & < 70y.", "≥ 70 & < 90y." and "≥ 90y.". The selection of the stand types and their age classes reflect their economical importance to the forest service. Pure spruce stands still represent the majority of the Solling stands and mixed stands comprising beech and spruce are the objective of the long-term silvicultural planning. Stands younger than the ones chosen have too low a volume to be of interest here.

The minimum size of a stand to qualify for selection is set to 2 ha and 15 samples are chosen as initial density. The alignment of the samples adheres to the 100 m$\times$ 100 m raster of the phase 1 plots but varies according to stand variability. In homogeneous stands the grid can be pruned to 200 m$\times$ 100 m or can be condensed to 50 m$\times$ 100 m in heterogeneous stands. This flexible alignment ensures that the entire surveyed stand will be represented in the sample. The accuracy aimed at is set to a half confidence interval length of 10 % of the estimated mean, that is,
$(\frac{\hat{\sigma}}{\sqrt{n}} * t_{(n-1,1-\frac{\alpha}{2})}) \leq (0.1 * \hat{\mu}), \quad \alpha = 0.05$, should hold.

All trees with a dbh ≥ 7 cm are inventoried on only one circle. Concentric circles are avoided to reduce variation caused by extrapolation. In young and dense stands the radius of the circle is set to 10 m (3 stands), in old stands to 15 m (6 stands), and in medium-dense stands it is set to 13 m.

Trees are surveyed in accordance with the inventory instructions of the forest service of Lower Saxony. Species, polar-coordinates, and dbh are recorded for every tree. Heights are measured for one approximate d_G-tree per stand-layer and species group. For ease of implementation the evaluation plots are selected from a sub-region of the Solling of approximately 8000 ha, covered by 24 images. The resulting distribution of chosen plots are displayed in Fig. 4 together with 1013 BI-plots located in the sub-region.
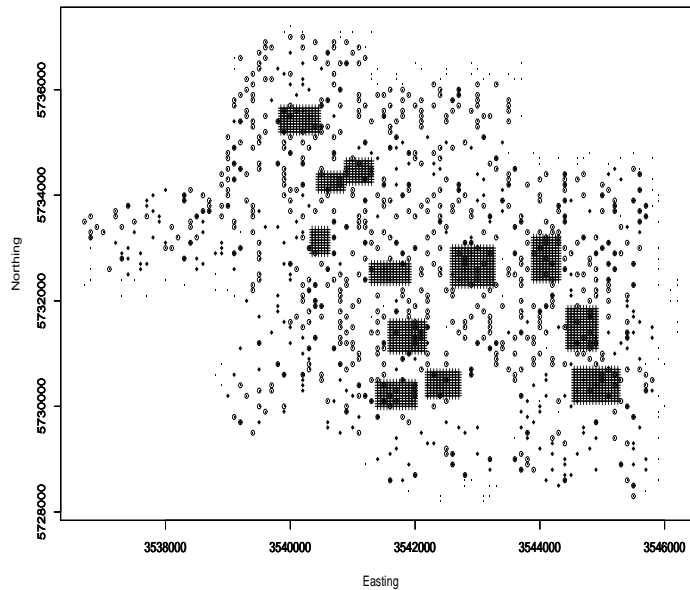
Figure 4. Solling: Distribution of sample plots

BI: points; Evaluated Stands: Crosses; Spruce: Circles, Beech: Diamonds;
Gauß-Krüger Coordinates

Elevation of the BI plots varies from 125 m to 530 m asl. Spruce is inventoried on 690 plots, 373 plots comprise at least one beech, and there are 898 plots with at least one spruce or beech. The data material of the BI and available auxiliary information is as described for Grünenplan except that the CIR images are scanned at 35 μm, resulting into approximately 44 cm ground resolution.

4   Methods

The basic idea underlying all investigated approaches is to consider similar observations only or a weighted average thereof. The methods differ with respect to how similarity is defined and can be classified by

- similarity according to feature distance

- similarity according to spatial distance.

In the feature space similarity between a sample location and an arbitrary location refers to how alike surveyed auxiliary characteristics are. Distance can be inferred by multivariate techniques like the canonical correlation analysis, can be heuristically modeled, or can be regarded constant as done with simple mean estimators.

In the spatial domain, similarity between a sampling unit and an arbitrary location refers to their spatial distance. An appropriate measure is the correlation structure among the sampling units. Such an approach is not confined to the values of the variates surveyed at the sample locations, but also includes their alignment in space. The resulting spatial interpolation is commonly called kriging.

All introduced methods share the dependence on auxiliary information.

4.1   Synthetic Estimator: Allocation of target stand to an appropriate stratum

The formal description is as follows:

$$z_{hi} = \mu_h + e_{hi} \quad (h = 1, ..., L; i = 1, ..., n_h) \tag{1}$$

where z_{hi} is the volume per hectare of the ith sample plot in stratum h, L is the number of strata, n_h the number of samples within stratum h, and e_{hi} are the residuals with zero mean and variance σ_h^2.

The predictor per stratum is the empirical mean

$$\hat{\mu}_h = \lambda_h \sum_{i=1}^{\lambda_h^{-1}} z_{hi} \tag{2}$$

obviously $\lambda_h = \frac{1}{n_h}$. Lambda is chosen in the notation to emphasize its meaning as a weight. The predictor (2) is based on similarity within a feature space. All observations within a stratum have zero distance to themselves and the same constant distance to other stratum members whereas all other observations have infinite distance.

Accordingly, the weights assigned are constant within each stratum and zero else. Rewriting (2) in this context yields

$$\hat{\mu}_h = \sum_{i=1}^{n} \lambda_h z_i \quad with \ \ \lambda_h = 0 \ if \ \ z_i \notin stratum \ h, \ \lambda_h = n_h^{-1} \ \ else \tag{2'}$$

where the summation goes over all n sample plots.

The method is easy to implement. Stands are allocated to homogenous strata based on available auxiliary information. The mean and variance of the variables of interest are calculated and assigned to all stands of the stratum. Each stratum should roughly comprise at least 30 observations to allow reliable variance estimation.

Bitter (1997; 1998) introduced a "typenorientierte Kontrollstichprobe" (stand-type sample based forest inventory) to German forestry based on synthetic estimators. One advantage of his approach is the integration of sample based point information and stand wise information to provide spatially explicit information at the stand level. So far, there hasn't been a comparison of the synthetic estimator with other predictors and Bitter hasn't yet provided an evaluation in terms of accuracy nor precision.

4.2   Ordinary Kriging

Kriging is a collection of generalized regression techniques and has been developed in geostatistics. It is explained in detail in appendix A.1 (p. 109). For further reading Webster et al. (2001) is recommended who provide a good overview of applications in environmental sciences whereas thorough methodological treatment is found in Cressie (1991) or Olea (1999).

The potential geostatistical techniques offer have found their applications in forestry. To name only a few examples: Biondi et al. (1994) and Samra et al. (1989) model height growth or stem size and increment within a stand, that is, with spatial structures with a range below 50 m. Jost (1993) evaluates sampling errors of systematic samples by means of geostatistical analysis. Ramirez-Maldonado (1988) analyzes characteristics of clusters applied by the US forest service. Fouqet and Mandallaz (1993) use kriging techniques in medium scale forest inventories (range below 1000 m) whereas Köhl and Gertner (1992) apply kriging to large scale forest damage inventories (range around 8000 m).

Most sampling schemes in German forestry are design based. Kriging is model based and thus equivalent to model based prediction approaches in survey sampling like ratio or regression sampling (Thompson, 1992), it is even exactly equivalent in case the covariance function of the variable under investigation is known. Kriging provides the BLUP (best linear unbiased predictor) if the model assumptions are met.

It is worthwhile to repeat some model assumptions in more detail:
the random function R honors the intrinsic hypothesis over the sampling domain, that is, the following relationships hold:

$$E[R(s)] = m \tag{3}$$

$$Var(R(s) - R(s + h)) = 2\gamma(h) = E[R(s) - R(s + h)]^2 \tag{4}$$

where $E[\cdot]$ denotes the expectation operator, $R(s)$ the random variable under investigation, $\gamma(\cdot)$ the semivariogram of $R(s)$, s a location vector, and h a distance vector.

The mean m is assumed constant over the sampling domain but is not known. If it were known it could be subtracted from the observations and simple kriging would be applied. The semivariogram is a measure of spatial self-similarity or spatial autocorrelation and is a function of distance only. The predictor $\hat{R}(s_0)$ at location s_0 is given by

$$\hat{R}(s_0) = \sum_{i=1}^{k} \lambda_i R(s_i) \tag{5}$$

subject to the constraint

$$\sum_{i=1}^{k} \lambda_i = 1 \tag{6}$$

Prediction is done as a weighted average of spatial neighbors. The structure of the semivariogram determines the weights λ_i and usually yields larger weights for nearby observations. The number k of neighbors considered can be set by the modeler. In case of no autocorrelation the predictor (5) is equivalent to (2′) with L=1, that is, there is only one stratum.

Here, ordinary Kriging cannot be applied to the raw data, $Z(s)$. The volume estimates don't honor the intrinsic hypothesis of a constant mean (cf. Eq. (12), p. 109). The semivariogram will display a structure but the data will show a trend, as volume is heavily influenced by age. A remedy is to use an external trendfunction $T(s)$ and to define a random function $R(s)$ as the difference of the original random function $Z(s)$, representing volume, to the trendfunction, that is

$$R(s) = Z(s) - T(s). \tag{7}$$

The trendfunction T need not be correct as long as the residuals honor the intrinsic hypothesis.

4.2.1 Ordinary Kriging: External Trendfunction

This simple method uses the yield table (Schober, 1995) as external trendfunction. The yield table must be linked to the sample data. Potential link-items are age, d_G, or h_G, the height of d_G, as provided by a field assessment. In case no site index information is available a mean yield class as class 2 or class 3 can be assumed, otherwise the site index at every sample and prediction location determines the yield table entry.

Yield tables represent pure, fully stocked stands. The observed plot volume must be extrapolated to a potential one to represent such conditions and to account for fluctuations in density and mixture proportions. The potential volume is the volume the species could have achieved had it covered the entire plot. It is obtained by multiplying the volume with the ratio of plot area to crown cover area of the species.

The latter crown area is not measured but must be modeled for example as suggested by Nagel (1999) or Dahm (1995). The random function $Z(s)$ in Eq. (7) is thus given by the potential volume.

4.2.2   Ordinary Kriging: Varying Means

Assume one has auxiliary information at all sample locations that relates to a categorical attribute α with L non-overlapping states α_h, $h = 1, \ldots, L$. The 8 strata classified in phase 1 on the CIR in the BI of Lower Saxony would be such a categorical attribute. For each state or stratum the mean $\hat{\mu}_{\alpha_h}$ is computed and serves as a crude trend function (Goovaerts, 1997). The random function $Z(s)$ in Eq. (7) is then represented by the measured volumes. The predicted value at any location is the sum of the kriged residual and the prevailing mean at that location. It must be emphasized that the kriging predictor of the residuals is based on the residuals of all sample plots, not only on those belonging to the same stratum as the target plot.

This technique can be regarded as a generalization of (2) for it can also accommodate spatial structure and leads over to universal kriging, introduced below.

4.3   Universal Kriging

A more formal introduction is given in appendix A.2. This method differs from ordinary kriging in that the assumption of Eq. (3) is replaced by the assumption that the mean itself is a function of the location vector s. It uses the residuals of the original values to a trend function estimated from the data. The flexibility of the trend function allows the incorporation of any auxiliary information (Gotway and Hartford, 1996) as long as the latter is available at all observed locations and locations to be predicted, and the trend can be modeled as a linear combination. Universal kriging is well suited for double sampling schemes (Mandallaz, 1991; Mandallaz, 1996).

The weights λ_i of the predictor and the coefficients of the trend function must be estimated simultaneously. This poses a circular problem because the semivariogram refers to the residuals which are computed after the trend has been modeled whereas

the trend model depends on the covariance structure of the residuals. All solutions are of iterative nature and restricted maximum likelihood (REML) is given preference based on a study of Nieschulze and Saborowski (2001).

In the REML approach a linear transformation of the observed data is sought that has expectation zero and thus filters out the mean of the observations, facilitating the estimation of the parameter vector of the likelihood function (Kitanidis, 1983). This approach is analogous to intrinsic random function kriging (Christensen, 1990). If the residuals show no spatial structure then the trend estimation is the predictor. Otherwise the prediction is improved by incorporating the residuals (cf. Eq. (33), p. 113).

Universal kriging is very flexible but imposes the highest demand on the quality of auxiliary information. In general it is superior to ordinary kriging but can produce worse results if the trend function is ill specified (Zimmerman et al., 1999). In general universal kriging does not aim at a precise or accurate model but at a precise or accurate model with strongly auto-correlated residuals.

For example, assume one model has a lower process variance (explained at p. 35) than a competitor but only a nugget is present. Further assume that the competing model has no nugget but an appropriately sized range. Then, unless the difference between the process variances is too large, one would prefer the competing model although the overall variation in the residuals is larger. Of course, *too large* and *appropriately* depend on the data modeled.

4.4   Variography

Spatial structure is analyzed by means of empirical semivariograms. The same data set is analyzed with same binning and maximum lag distance by the classical estimator in Fig. 5 (a) and by the robust estimator in Fig. 5 (b). The classical estimator is sensitive to outliers in the data. It doesn't capture the spatial autocorrelation as well as the robust estimator (cf. appendix A.1).

The semivariogram-models in Fig. 5 were fitted bye eye based on a first approxima-

(a) Classical semivariogram estimator   (b) Robust semivariogram estimator
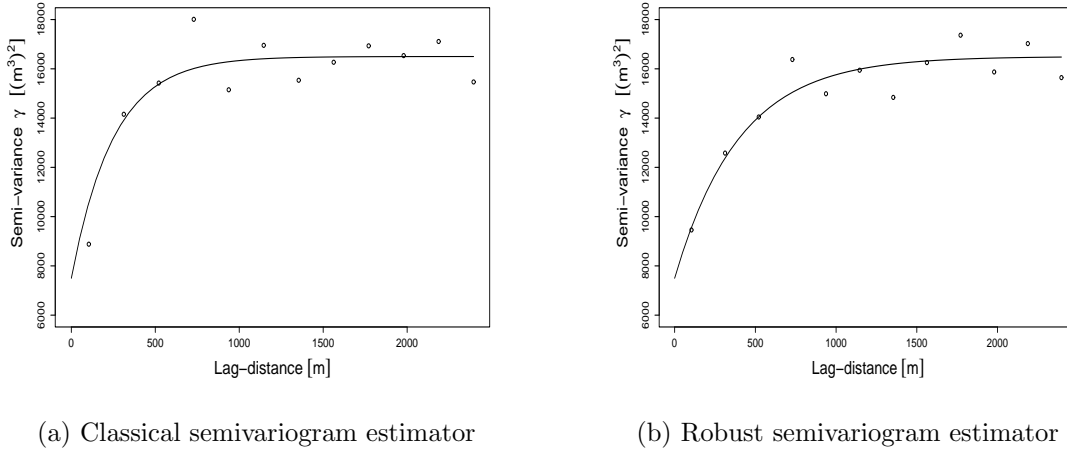
Figure 5.   Spatial Structure Analysis: Variogram estimators

tion by a Gauss-Newton algorithm. Zimmerman and Zimmerman (1991) found that estimators of the semivariogram perform best when the spatial dependence is weak. Also, all estimators investigated by them gave rise to an unbiased kriging predictor, so the kriging predictor seems rather robust with respect to misspecification. This is corroborated by Cressie and Zimmerman (1992) who derive that a prediction based on a misspecified but compatible variogram is asymptotically efficient. The definition of compatibility is based on probability measures and sets of supports, and examples can be found in Stein and Handcock (1989).

Diamond and Armstrong (1984) oppose this view and derive analytically that the robustness of the kriging predictor depends on the condition number of the kriging matrix of the original semivariogram, which is usually not known. The condition number of a matrix is calculated as the ratio of the largest to the smallest absolute eigenvalue. The plot layout they used in their model evaluations is usually not encountered in forestry practice. They used three points where two points were very close to each other. They admit, that in practice such a layout would be altered and the two points combined.

The major differences in the kriging weights were found when a Gaussian semivariogram was applied in the kriging system. Problems with the parameter estimation of such semivariograms are well known. Gaussian variograms are not applied in this

study. Furthermore, Diamond and Armstrong admit that even under their extreme sampling configuration the spherical model is relatively robust.

With the encountered spatial structures fitting by eye has the advantage of precisely capturing the behavior of the semivariogram near the origin (Olea, 1999) and is therefore chosen.

4.5 Nearest Neighbors

4.5.1 Most Similar Neighbor (MSN)

Moeur et al. (1995) introduced this method as a sampling inference tool for natural resource planning. Predictions are based on most similar neighbors where similarity is derived by canonical correlation analysis. An introduction to the latter is provided in appendix B.

With the set up as in appendix B, p. 116, let $\vec{X}$ represent *design* attributes that can be measured on field plots and $\vec{Y}$ be *indicator* attributes that represent auxiliary information, derived from airborne imagery or local databases and that can be measured not only at the same locations as the corresponding *design* attributes but also at all locations to be predicted. The prediction of any field site is then carried out by searching the closest field plot and assigning its *design* attributes to it. The distance measure between a location u and field plot j is similar to the mahalanobis-distance (Mahalanobis, 1936) and derives as follows:

$$D_{uj}^2 = (\vec{Y_u} - \vec{Y_j})_{1 \times q} \; \underset{q \times p'}{\Gamma} \; \underset{p' \times p'}{\Lambda^2} \; \underset{p' \times q}{\Gamma'} \; (\vec{Y_u} - \vec{Y_j})'_{q \times 1} \tag{8}$$

with Γ a matrix of dimension $q \times p'$ where p', $p' \leq p$, denotes the number of non-zero Eigenvalues of Q and is composed of the vectors $\vec{\hat{\beta}}$, the canonical coefficients of $\vec{Y}$, and Λ^2 is a diagonal matrix of squared canonical coefficients. Solving Eq. (8) one sees that the distance measure is a weighted sum of squared canonical variables. The latter are represented by the differences of the *indicator* attributes, and the weights by their squared canonical correlations.

The demand on the auxiliary information is comparable to universal kriging. Canonical correlation analysis can be regarded as a (non-parametric) regression as applied

in universal kriging. Its prediction can be biased, is confined to the range of original values in the sample, and in general is less efficient than a regression. Its main advantages are that is preserves the covariance structure among the *design* attributes and does not require the specification of a parametric regression model. These features make it particularly suitable for multivariate prediction.

4.5.2 Weighted Means

Another feature distance based approach is a generalization of Eq. (2) and resembles the k-nearest neighbor regression proposed by Maltamo and Kangas (1998). Available strata used in the BI are rather coarse in resolution in comparison to the small field plots and for example do not take density into account. In theory, prediction can be improved by applying a weighted mean within strata where the weights are inversely proportional to a feature based distance. Within a given stratum prediction is always based on all members of that stratum. The weights of the members applied in the mean calculation vary for every location to be predicted according to the feature distance to that location.

Distances are a function of available auxiliary variables and can be modeled via regression analysis (Holmström, 2002) or heuristics via normalized euclidian distances (Maltamo and Kangas, 1998), the approach used in this study. The approach by Holmström is much more complicated. For example, it requires the specification of a parametric regression model and the application of generalized least squares, rendering it less suitable. It is therefore not supported.

Predicted values of the weighted means estimator are confined to the range of observed sample values and the resulting predictors need not be unbiased.

4.6 Null-Variant

The forest service of Lower Saxony currently employs a regionalization method based on stratum means similar to the synthetic estimator introduced in 4.1 but modified by mixture proportions and stocking densities.

The crown areas are modeled according to Nagel (1999) and summed per layer for every sampling plot. Let $A_{11}, A_{12}, \ldots, A_{1l}, \ldots, A_{k1}, A_{k2}, \ldots, A_{kl}$ be such summed crown areas for k tree species in l layers within a given stratum. All crown areas are calculated with respect to the larger circle with radius 13 m, that is, crown areas are extrapolated for trees with a dbh < 30 cm. The mixture proportion p_{ij} of species i in layer j is derived by $p_{ij} = \frac{A_{ij}}{\sum_{m=1}^{k} A_{mj}}$.

If $v_{ij}, \quad i = 1, \ldots, k; \quad j = 1, \ldots, l$ denotes volume per hectare of the species i in layer j then the potential volume of species i is represented by $V_{ij.pot} = \frac{v_{ij}}{p_{ij}}$ within layer j at the sample plot.

Prediction is now done by the stratum mean of the potential volume of all inventory plots per species and layer, multiplied by the mixture proportion of the species in the respective layer in the target stand ($p_{ij.target}$). The estimator for any given stratum then becomes

$$\hat{V}_{ij} = \frac{1}{n} \sum_{\theta=1}^{n} V_{ij.pot\theta} * p_{ij.target} = \bar{V}_{ij.pot} * p_{ij.target} \tag{9}$$

where n is the number of sample plots of the respective stratum. The crown closure of layer j of the target stand serves as a proxy of the stand density $B^{\circ}_{j.target}$. Let $\bar{B}^{\circ}_{j}$ be the mean stocking density within a stratum and layer j, then the predicted volume at the target location is

$$\hat{V}_{ij.kor} = \hat{V}_{ij} * \frac{B^{\circ}_{j.target}}{\bar{B}^{\circ}_{j}} \tag{10}$$

Equation (9) employs a "mean of ratios" estimator, appropriate when the variance of the dependent variable is a quadratic function of the independent variable. Equation (10) is based on a "ratio of means" estimator, appropriate when the variance of the dependent variable is directly proportional to the independent variable (Wensel, 1996). It can be assumed that those criteria did not influence the development of the procedure, which was based on a heuristic approach.

Mixture proportions p_{ij} and stocking densities $\bar{B}^{\circ}_{j}$ are calculated from the terrestrial phase 2 plots. The prediction relies on the respective estimation as part of the "Bestandesbegang". This approach employs rather high resolution data, estimation is done per layer. Because the latter are defined by economical criteria this method cannot be modeled with data derived from airborne imagery.

4.7   Cross-validation

Model selection is guided by cross-validation or leave-one-out predictions. As the name suggests, one observation is left out at a time and its characteristics are predicted by the remaining observations. Parameters of the different models are not recalculated during cross-validation. Different methods can be judged relative to each other by the variances of their prediction errors $var\left(Z(s) - \hat{Z}(s)\right)$ with $s = (s_1, \ldots, s_n)'$, and $\hat{Z}(s_i) = \sum_{j \neq i} \lambda_j Z(s_j)$. This variance is hereafter referred to as $\underline{error\ variance}$.

It should be noted that cross-validation is not a hypothesis-testing method (Davis, 1987), nor does it provide useful information about the closeness of the geostatistical model covariance function to the true covariance function (Solow, 1990). The performance of any given model can only be judged against competitors given the particular data set.

5   Applications and Results

5.1   Lower Saxony: Grünenplan

The BI field plots are rather small with 531 m^2 for trees with dbh $\geq$ 30 cm and 113 m^2 for trees with a dbh in the $[7, 30)$ interval. Sample plots even within a stand will show a large variability because they will be situated in dense and sparse parts of a stand.

The variability of per hectare volume of BI plots comprising beech or spruce are summarized by their 0.025, 0.5 and 0.975 quantiles and by age-class in Table 1. Deviations of the chosen quantiles from the median of up to 580 m^3 occur for spruce age-class 4, and up to 339 m^3 for beech age-class 3.

Table 1.   Variation of Volume of BI plots per age class

	Beech			Spruce		
Age-class	2.5% Quantile	Median	97.5% Quantile	2.5% Quantile	Median	97.5% Quantile
1	5	59	231	25	178	392
2	18	104	423	28	287	542
3	30	300	63	97	482	850
4	116	351	676	30	156	738

All methods are compared with respect to a cross-validation based evaluation (cf. section 4.7). The large variability in the BI field plots and their size render a precise prediction difficult. However, on a stand level it is expected that over- and under-prediction of the field plots will partly cancel each other out. Thus, predictions of entire stands will have smaller errors than the ones reported here for the point-predictions during model selection.

Results for the different methods are summarized in Table 2 for spruce and Table 3 for beech.

Table 2.   Grünenplan: Results for Spruce

Method	Auxiliary Variable	Origin	Error Variance
Synthetic Estimator	Age class Stand type	Phase 1	24490
OK: External	d_G	"Bestandesbegang"	25590
Trendfunction	d_G; SI	"Bestandesbegang"	22710
	Age; SI	"Bestandesbegang"	13000
OK: Varying Means	Age class; Stand type	Phase 1	15600
Universal Kriging	Age class Species Proportion	Phase 1 Imagery	16818
	Age; SI Species Proportion	"Bestandesbegang" Imagery	15818
	Age-class Species Proportion; Gap 0.85 Height Quantile	(Phase 1) Imagery	11820
	Species Proportion; Gap 0.85 Height Quantile	Imagery	12910
NN: Canonical Correlation	Age class Spectral values	Phase 1 Imagery	26457
	Spectral values; DSM	Imagery	17988
Euclidian Distance	Age class Spectral values	Phase 1 Imagery	24617
Weighted Means	Age; SI Species Proportion	"Bestandesbegang" Imagery	16100

OK: Ordinary Kriging; NN: Nearest Neighbors; SI: Site Index;

Error-variance from cross-validation

5.1.1   Synthetic Estimator

Age-class and species group are the auxiliary information the synthetic estimator depends upon. These two variables are surveyed as part of the already implemented inventory, no new information needs to be obtained. The analysis at Grünenplan is confined to sample plots where spruce or beech is observed. The synthetic estimator will therefore systematically underestimate their volume.

Nevertheless, the variance of the prediction errors or error variance is comparable to those of other methods. The error variance is rather large with 24490 for spruce and 15200 for beech.

Table 3. Grünenplan: Results for Beech

Method	Auxiliary Variable	Origin	Error Variance
Synthetic Estimator	Age class Stand type	Phase 1	15200
OK: External Trendfunction	d_G	"Bestandesbegang"	13000
	d_G; SI	"Bestandesbegang"	16000
	Age; SI	"Bestandesbegang"	13000
OK: Varying Means	Age class; Stand type	Phase 1	NA
Universal Kriging	Age class; Species group Species Proportion	Phase 1 Imagery	12795
	Age; SI Species Proportion	"Bestandesbegang" Imagery	11120
	Age-class Species Proportion; Gap 0.85 Height Quantile	(Phase 1) Imagery	11335
	Species Proportion; Gap 0.85 Height Quantile	Imagery	9765
NN: Canonical Correlation	Age class Spectral values	Phase 1 Imagery	16100
	Spectral values; DSM	Imagery	10200
Euclidian Distance	Age class Spectral values	Phase 1 Imagery	18200
Weighted Means	Age; SI Species Proportion	"Bestandesbegang" Imagery	11000

OK: Ordinary Kriging; NN: Nearest Neighbors; SI: Site Index

Error-variance from cross-validation

5.1.2   Ordinary Kriging

External trendfunction

Using the external trend-function "yield-table" necessitates the extrapolation of the observed volume to a potential volume. The extrapolation should compensate for stocking densities and mixtures not accounted for in the yield table. The concept of potential volume follows the rationale that trees try to occupy as much growing space as possible (Smith, 1986). Growing space is approximated by crown area, and with crown cover has an equivalent that can readily be assessed in the field. Modeling crown volume would be more appropriate, but requires modeling input usually not provided by management planning inventories.

Crown areas are modeled according to Nagel (1999), who found a good linear relationship between crown width and dbh. His model rests upon field trials where several crown radii are taken per tree. The model does not consider competition or spatial alignment of neighboring trees. To assess the actual crown cover every tree crown is modeled as an ideal circle as displayed in Fig. 6. Wildenhein (1999) summed these crown areas per species and computed mixture proportions as the fraction of total crown area. He calcu-

Figure 6.   Crown cover modeling

lated potential volume as the ratio of observed value to mixture proportion. Stocking density is not accounted for; if a plot comprises only one tree then observed and potential volume will coincide.

With the present data the approach of Wildenhein is not yielding any spatial structure of the residuals. Extrapolation should be based on the ratio of covered area to total plot area. Just summing the modeled crown areas will not yield realistic results owing to overlapping model-crowns (cf. Fig. 6) but can produce extrapolated areas of more than 4 times the plot area. Using such an approach the beech displayed in Fig. 6 would have got a total crown area of 1085 m² and larch 52 m² on a plot of approximately 531 m².

A more realistic crown area similar to Voronoi polygons is computed by own programs.

Each tree crown circle is gridded into quadratic cells with 5 cm side length. If a cell-centroid is covered by more than one tree its area is allocated according to dbh, distance, and height. Only trees with a height difference of less than 5 m to the tallest tree covering the cell are retained. From these remaining trees a dbh weighted distance is calculated and the cell is allocated to the tree with smallest distance. Such an approach yields a total crown area of 642 m^2 for beech and 50 m^2 for larch, the two intersecting trees at the lower left side in Fig. 6.

The overestimation of total crown cover by 130% is due to the concentric layout of the field plots. Plot characteristics are collated from two tree populations. All measurements taken on trees with dbh < 30 cm need to be expanded by approximately 4.7 to be representative for the outer circle. Such an overestimation occurs for beech but is not observed for spruce. In a final step, too large a crown area is normed to match the actual size of the field plot.

The residuals as indicated in Eq. (7) are obtained as the difference of the extrapolated volume to the corresponding yield table volume. Which yield table volume is corresponding depends on how a field plot is linked to the table. There are three link items age, d_G, and h_G and site index which will determine the appropriate yield table volume. The first three auxiliary variables are derived from the surveyed trees for each field plot. Site index is provided by the field assessment. The field assessment also provides estimates of age, d_G, h_G, stocking density, and mixture proportion needed at target stands to calculate standing volume from kriged residuals.

In a first trial the d_G was chosen as link item because an inspection of the data revealed that age estimates frequently come with an error. The d_G is no estimate but calculated from measured data and hence accurate. At this early modeling stage no field assessment had been available. Linking potential to yield table volume by assuming an average SI of yield class 2 for spruce and an average yield class 3 for beech yields the best results.

With data taken from a yield table volume is a sigmoid function of the d_G as shown in Fig. 7 where spruce volume of yield class 2 is plotted over d_G. Volume tapers off with increasing d_G, asymptotically approaching its maximum value of 640 m^3.

The maximum d_G for spruce yield class 2 is 39.3 cm; on field plots values of up to 60 cm are observed. This discrepancy, which also holds for beech, and the sigmoid relationship explain the clustering of residuals at the upper end of the yield table volumes displayed for both species in the first row of Fig. 8. Spatial structure analysis yields a pronounced structure for spruce and a not so clearly discernible structure for beech, cf. Table 4. The error variance is quite high for spruce with 25585 but surprisingly low for beech with 13000 (cf. Tables 2 and 3). The intrinsic hypothesis is not supported for spruce by the plot in the second row of Fig. 8; rather the residuals display a linear relationship with age.
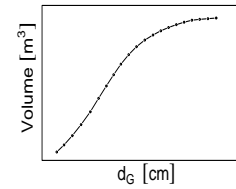
Figure 7.
Volume over d_G

The yield table with average yield class 2 is not a good trend function for spruce. Modeling the yield class on a plot level by the SI from the field assessment yields more evenly distributed residuals at the upper end of yield table volumes. However, the residuals still show a pronounced linear relationship with age (cf. Fig. 8 left column, third and fourth row). Nugget and sill of the empirical variogram are similar to the fixed d_G approach but the range is much smaller (cf. Table 4). The error variance of 22710 is still quite large.

For beech, the residuals are almost exclusively negative beyond 525 m^3 yield table volume (cf. Fig. 8 right column, third row). There is no clear relationship of the residuals with age although beyond age 120 potential volume tends to be larger than the corresponding yield table volume. Spatial structure analysis yields quite a different semivariogram model with a nugget almost double the one of the d_G approach with fixed SI. This increase corresponds to an increase of the error variance to 16000 (cf. Table 3).

Linking potential volume to the yield table via the d_G does not result in residuals supporting the intrinsic hypothesis, at least not for spruce. Using the mean plot age of that species as link item and an individual yield class per plot derived from the field assessment produced more promising results for spruce.

As seen in Fig. 9, left column first and second row, the residuals are evenly distributed over yield table volume and over age. The resulting empirical semivariogram has a slightly smaller nugget of 5000 compared to the other d_G approaches. The sill is less than $\frac{1}{3}$ of the other ones and the range is of around 590 m. The error variance of
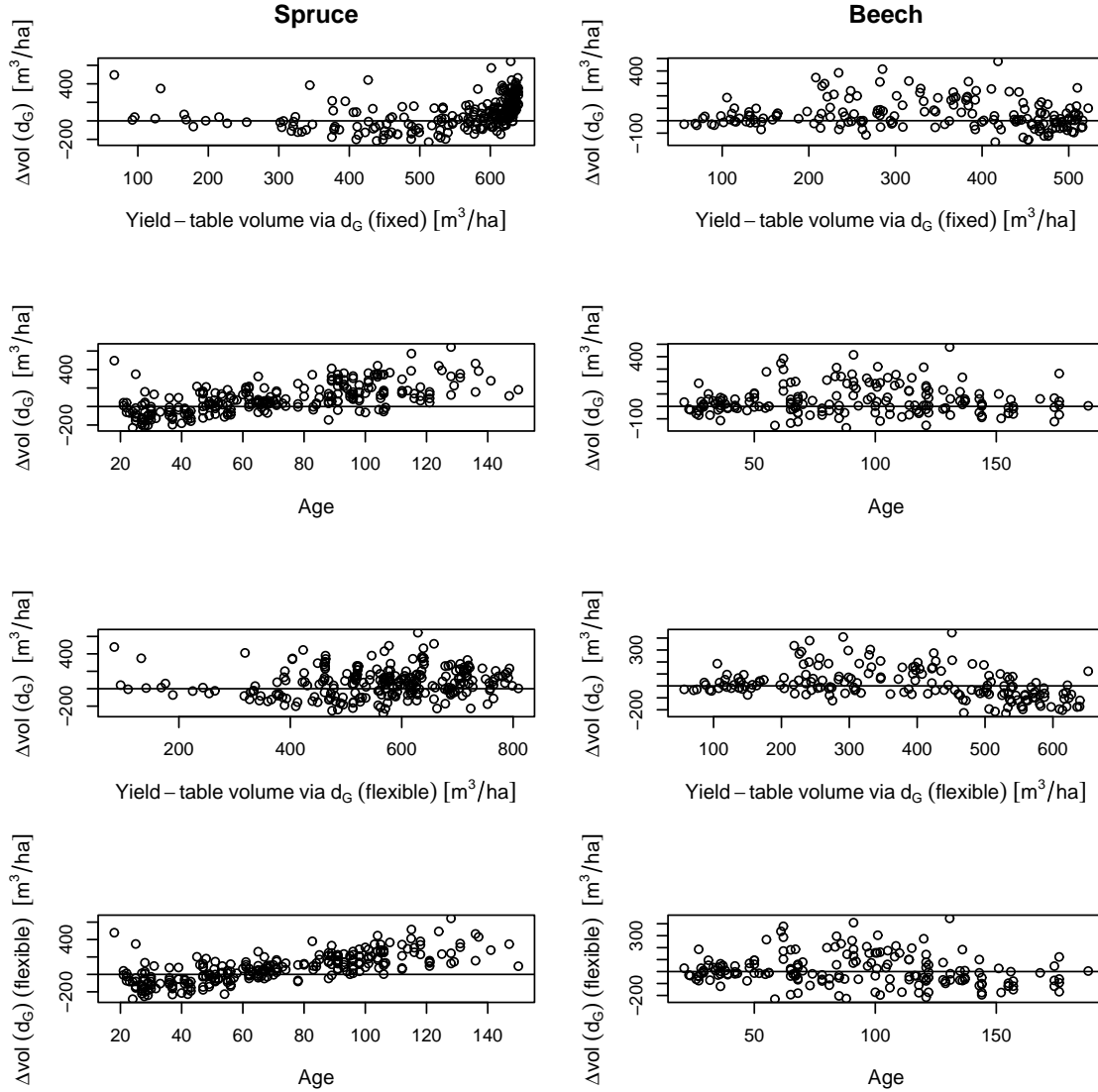
Figure 8.   Residual Analysis: Left: spruce; Right: beech; $\Delta vol = Vol_{pot} - Vol_{Yield-Table}$

13000 is almost half the one of the fixed d_G approach and less than $\frac{3}{5}$ when the d_G is a function of SI (cf. Table 2).

Results are less promising for beech. As seen in Fig. 9, right column first and second row, residuals display a decreasing trend with increasing yield table volume or mean plot age of beech. The empirical semivariogram represents a less pronounced structure than for the fixed d_G approach. Nugget and range are almost twice as large whereas the sill is comparable (cf. Table 4). However, the error variance of 13000 is identical.
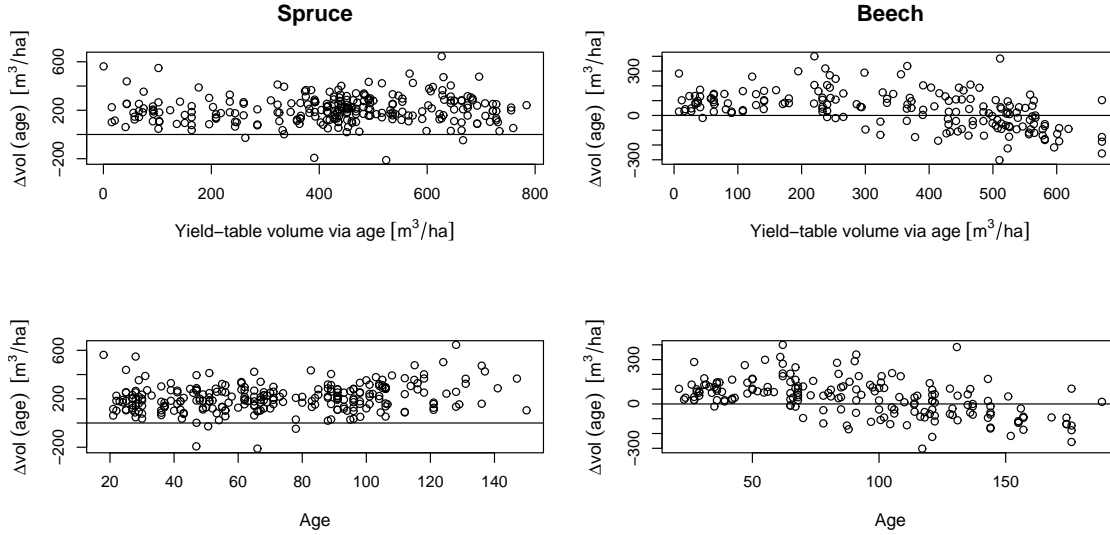
Figure 9. Residual Analysis: Left: spruce; Right: beech; $\Delta vol = Vol_{pot} - Vol_{Yield-Table}$

Varying means

Using local means to de-trend the data is a simple approach which circumvents the problems associated with modeling potential volume. However, a structural analysis of the residuals of measured plot volume to their stratum means failed for beech and spruce. The empirical semivariograms showed no discernible spatial correlation. Varying means are a rather coarse trend; they do not take varying densities into account.

Treating the potential volume per plot as the random function yielded an empirical semivariogram showing spatial correlation for spruce only. The spatial structure seems less pronounced compared to the yield table and d_G based on fixed SI as link item (cf. Table 4) but the resulting error variance of 15600 is less than $\frac{2}{3}$ of that approach (cf. Table 2). A plot of residuals over age shows no systematic deviation. Residuals are evenly distributed around null with values lying within [-320,485].

5.1.3   Universal Kriging

Usually, tree volume of terrestrial plots is calculated as a function of basal area, height, and form factor. Of these only height qualifies as an auxiliary variable because it can

Table 4. Model parameters for ordinary kriging

Method	Beech				Spruce			
	Nugget	Sill	Range	Type	Nugget	Sill	Range	Type
OK_1	4500	8000	225	Exp	8600	19900	1975	Sph
OK_2	8500	7650	265	Exp	7000	20050	300	Exp
OK_3	7500	9000	400	Exp	5000	5800	590	Sph
OK_4	—	—	—	—	10913	5586	720	Exp

OK_1: Ordinary kriging via d_G (fixed SI);

OK_2: Ordinary kriging via d_G (flexible SI);

OK_3: Ordinary Kriging via age (flexible SI);

OK_4: Ordinary kriging: Varying means (potential volume);

Exp: Exponential semivariogram model; Sph: Spherical semivariogram model

be obtained on a large scale from air photos. If no height information is available, like in the first phase of this study, then volume can also be modeled as a function of age, stand density, and site index. Age and site index are regularly recorded during the field assessment ("Bestandesbegang"). At first, the field assessment records haven't been available and age has been approximated by the age-classes of the phase 1 classification.

Stand density can be approximated by number of trees and by crown cover area. Automated tree counts from digital images is not easy to implement, and even on images with a nominal scale as large as 1:5000 results are mixed (Brandtberg and Walter, 1998; Uuttera et al., 1998; Dralle and Rudemo, 1996). Emphasis in this study is put on methods that are cost effective and easy to implement and apply. This may include sophisticated techniques but only when they can be automatized even when covering large areas. This rationale lies behind the decision to employ an un-supervised classification approach although supervised approaches usually yield better results (Lillesand and Kiefer, 2000; Miguel-Ayanz and Biging, 1996).

In the following, all described models are based on linear combinations of variables. Products of auxiliary variables have also been included in the many modeling approaches but have been found less efficient in terms of error variance than the described linear combinations.

Crown cover area modeled according to Nagel (1999) can be estimated based on the auxiliary information derived from the digital images. The ratio of modeled crown cover areas per species to plot area minus area of class 1 are divided into 6 groups with equal frequency. Class 1 refers to the first out of the seven classes classified by ISODATA (cf. section 3.1). It is found that this class typically occurs in gaps; it follows that crown cover area is modeled per species as a proportion of species crown area to total crown area.

A mixture analysis is applied individually to beech and spruce to classify the phase 2 plots into one out of these 6 groups based on the auxiliary information described in section 3.1. Final group membership is assigned as a weighted sum of the group labels, the weights being membership probabilities. It holds for this new auxiliary variable $x_{species}$, $species \in \{beech, spruce\}$, that $x_{species} \in \mathbb{R} \cap [1, 6]$. In the following this new variable is referred to as species proportion.

The mixture analysis is based on the assumption that observed data derive from multiple homogenous populations and that it is unknown which population led to what observation. The method is similar to discriminant analysis except that group memberships are missing. Instead, group membership is regarded as a random variable. The observed distribution function is modeled as a sum, that is, a mixture of distribution functions of the homogenous populations. The aim is the derivation of parameters of the latter enabling a classification (Flury, 1997).

An additional auxiliary variable is the indicator variable "species group" of phase 1. Including this variable improved the error variance for beech to 12795 (cf. Table 3). No improvement is achieved for spruce and therefore it is included in the beech model only (cf. Table 5. The error variance of 16818 is based on the species proportion measure and the age-classes as regressors. The residuals show weak spatial correlation for beech ($\frac{sill}{nugget} = 0.21$) and a more pronounced autocorrelation for spruce ($\frac{sill}{nugget} = 0.37$) (cf. Table 5, first row). However, the overall variation in the residuals is smaller for beech; the process variance (sum of sill and nugget) is only $\frac{2}{3}$ of the one for spruce.

Substituting the mean age per plot for the age-classes, excluding the binary variable species group and including site index improved the models in terms of lower nugget, lower process variance, and resulting lower error variance (cf. Table 2, 3; Age, SI, Species Proportion).

The REML approach does not allow interactive modeling of the semivariogram. Initial estimates of nugget, sill, range, and the type of semivariogram-model of the residuals have to be provided. Such estimates can be obtained from the analysis of an ordinary least square fit. However, the semivariogram of such a fit can be misleading (Nieschulze and Saborowski, 2001). It is recommended to run the model with different initial parameters because of the possibility of a multi-modal profile likelihood (Warnes and Ripley, 1987). For both species a REML fit based on a spherical and on an exponential semivariogram model is obtained. Results are assembled in Table 5, rows 2 and 3.

The two approaches yield quite different parameter estimates; the nugget and sill are lower when using an exponential semivariogram. However, the process variances are quite similar for each species. Comparing the prediction of the two models shows very good agreement with differences of predicted values being in the interval [-28.4,18.2] (50% within [-6,6]) for spruce and in the interval [-19.1,18.3] (50% within [-3.5,3]) for beech, that is, both models resulted into an almost identical prediction. The error variances based on the spherical semivariogram model are 11120 for beech and 15818 for spruce.

In a third approach a substitution for SI and mean age is sought. The plot height is a proxy for the h_G and modeled as the 0.85 quantile of the observed height values, cf. section 3.1. In Fig. 10 the h_G is plotted over the 0.85 quantile. The plot reveals systematic deviations. Terrain elevation is consistently underestimated by ca. 20 m. Negative plot heights are avoided by adding an offset of 100 to the nDSM, the difference of DSM to DTM. The plots in Fig. 10 show a good linear relationship.
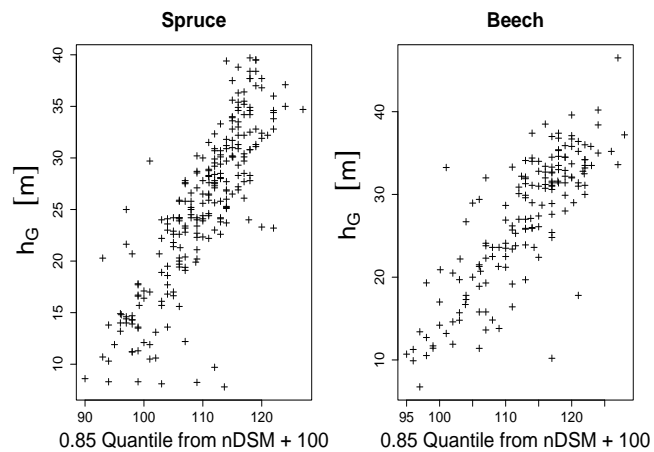
Figure 10. Grünenplan: h_G over 0.85 height quantile
Left: spruce; Right: beech

The h_G always refers to a tree of layer 1. Observed outliers like the one for spruce

with h_G below 10 m and 0.85 quantile of 115 are mostly due to the layer concept. In the mentioned case shading trees of layer 4 with height 35 m tower spruces in layer 1 with height 7.8 m.

Underestimation of tree heights is also reported by Kätsch and Stöcker (2000) although at much smaller magnitude. Kraus (1994) lays down a rule that tree heights derived from analogue aerial images come with an average error of around 0.70 m. In both cases it is stipulated that ground elevation is also measured precisely from the images. In the present case ground elevation is modeled from the DTM obtained from the cadastre service. The DTM is based on measurements taken on a regular grid a with side length of 12.5 m; a continuous DTM is computed by linear interpolation of the grid-node elevations. The cadastre service reports mean elevation errors of ±1.5 m and field assessment showed that such DTM has elevation errors of up to 3.75 m (Döring, personal communication). Against this background the obtained height measure is judged satisfactory.

A linear combination of area of class 1 of the ISODATA classification as gap, the 0.85 quantile, and the above introduced species proportion measure, fitted individually for the 4 age-classes, yields an error variance of 11335 for beech and of 11820 for spruce (cf. Table 3, and 2). The improvement for spruce can be explained by a reduced process variance of 13297 with a small sill of 4135 (cf. Table 5). With this approach in 70 out of 100 cases the differences of predicted to observed value fall within a [-100 m^3, 100 m^3] interval but differences of more than 400 m^3 occur. Apart from imprecise auxiliary variables this can be attributed to the large variation within the data which is hard to model (cf. Table 1). For beech, the process variance of 11426 is comparable to the second introduced approach.

Confining the model to auxiliary variables derived from the digital images, that is, fitting the before-mentioned linear combination over all age-classes improved the error variance for beech down to 9765. The improvement can be attributed to the smaller nugget of 7421 compared to 8736 when fitting for each class individually.

For spruce, the error variance increases to 12910; nugget, sill, and range are similar but a little less favorable, that is, the first two are larger and the range is smaller. The variogram of the residuals for spruce are shown in Fig. 11 together with the fitted model based on a spherical semivariogram. There is a clearly discernible correlation in the residuals with a range of around 750 m. The range reaches beyond stand boundaries as the average stand size is well below 56 ha = 750 m×750 m.
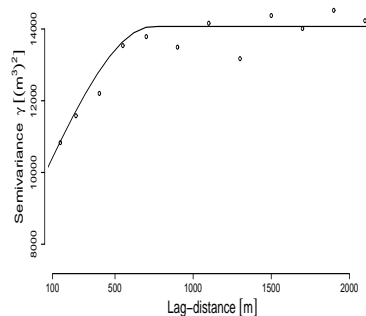
Figure 11. UK_4 (cf. Table 5) structural residual analysis

Table 5. Model parameters for universal kriging

	Beech				Spruce			
Method	Nugget	Sill	Range	Type	Nugget	Sill	Range	Type
UK_1	10545	2218	767	Sph	13577	5890	700	Exp
UK_2	9575	1762	751	Sph	9318	8017	519	Sph
UK_2	7178	4135	127	Exp	7449	9962	188	Exp
UK_3	8736	2690	839	Sph	9162	4135	800	Sph
UK_4	7421	2600	824	Sph	9496	4574	742	Sph

UK_1: universal kriging via age-class and species proportion; species group included for beech
UK_2: universal kriging via age, species proportion, and SI;
UK_3: universal kriging via 0.85 height quantile, age-class, gap and species proportion;
UK_4: universal kriging via 0.85 height quantile, gap, and species proportion: all derived from imagery; Exp: Exponential semivariogram model; Sph: Spherical semivariogram model

5.1.4 Nearest Neighbors

Most similar neighbor (MSN)
In the first project phase a variation of the ”most similar neighbor“ method based on euclidian distances has been incorporated. Such an approach is equivalent to the method introduced in section 4.5.2 except that the weights are constant for all auxiliary variables. Auxiliary information used are the phase 1 variables age-class and species group and the variables derived from the spectral digital image (cf. section

3.1). All variables are standardized. Nearest neighbors are searched within the eight strata determined by the phase 1 information. The error variances of 18200 for beech and 24617 for spruce are rather large.

Using the same auxiliary information as design attributes and as indicator attributes modeled crown area, basal area, volume, and tree frequency the MSN based on canonical correlations yield an error variance of 16100 for beech and 26457 for spruce. Again, nearest neighbors are assigned within the eight strata.

Substituting the phase 1 information by variables derived from the normalized digital surface model improved the prediction. Newly included variables are the 0.85 height quantile as described above, maximum observed height difference on a plot, and the area of pixels below a threshold. Such an approach can be applied to all stands. This is not the case when phase 1 information is used because roughly 5% of all stands are not covered. The error variances decrease substantially down to 10200 for beech and down to 18000 for spruce. For the latter species in 62 out of 100 cases the differences of predicted to observed value fall within the [-100 m^3, 100 m^3] interval.

Weighted Means

As like the second investigated universal kriging approach weighted means are based on the auxiliary variables species proportion derived from imagery, site index SI and mean age of the species, taken from the field assessment and the phase 2 plots. The weight function depends on distance which is a function of three auxiliary variables. The variables are normalized to avoid an influence of scale. Distance is measured as the weighted sum of the absolute differences. The contribution of the variables to the weight function is heuristically modeled. The best results are obtained with a relation of 3 to $\frac{1}{2}$ to 1, that is, the species proportion distance gets multiplied by 3, the site index distance gets halved, and the contribution of the mean age remains unaltered. The error variance of for spruce is in the middle range with 16100 but quite low for beech with 10200.

5.1.5   Null-Variant

The data set enables only a limited evaluation of the null-variant. The delineation of a forest district is updated at every management planning. In particular auxiliary management units ("Hilfsflächen") are created or merged with the compartment they are situated in. These units are not represented by geometries like compartments; the only location information is qualitative like *"in the north-west of sub-compartment **b**"*.

The phase 2 plots comprise, among other information, Gauß-Krüger coordinates, compartment, sub-compartment, and sub-sub-compartment. Auxiliary management units are not recorded because of their short-lived existence. An unequivocal assignment of phase 2 plots to the field-assessment data is possible for stands only that do not comprise such units. This selection reduces the number of suitable spruce phase 2 plots down to 132.

If one assumes that mixture proportions and stocking densities derived from phase 2 plots are good proxies of their counterparts of the field assessment then all spruce plots can be used in the evaluation. Based on this assumption the error variance for spruce volume of layer 1 is 5590. This variance is very low and slightly increases to 6480 if layers 1, 2, and 4 are pooled. However, universal kriging with comparable high resolution data, that is, stocking density, h_G, and mixture proportion, all derived from the phase 2 plots, yield a still lower error variance of 1800 only. Even employing the 0.85 quantile instead of h_G yields a lower error variance of 4095.

Using the phase 2 proxies just serves as a model comparison. Conditions on phase 2 plots will differ from the average condition of the stand they lie in. The error variance will increase when the stand information is used in lieu of the proxies. In case of the above mentioned 132 spruce plots this increase is more than fourfold from 6375 up to 29000.

5.2   Baden-Württemberg: Black Forest

The sampled forest stands are characterized by a large heterogeneity. Confining the investigation to just two species like in Grünenplan where spruce and beech represent over 80% of the population is not tenable. Beech (Fagus sylvatica, LINNÉ) is the most frequent species with 23% of all trees, followed by douglas fir (Pseudotsuga menziesii, MIRBEL) with 13%, spruce (Picea abies, (Linné) KARSTEN) with 12%, and fir (Abies alba, MILLER), with 11%. Oak (Quercus petraea, (Mattfeld) LIBLEIN & Quercus robur, LINNÉ) is the second most frequent deciduous species with 9%. Frequency with respect to the field plots, that is, percentage of plots on which a species is observed, results in 52% for beech. Oak is the second most frequent species with 33% followed by firs with 26%. Spruce is observed on 22 and douglas fir on 30 plots out of 100.

Accordingly, only 14% of all sample plots cover only one species. Plots with two species are the most frequent with 37% of the population; 29% of all plots comprise three species and four and more species are surveyed on still 19 out of 100 plots.

Results for the five species and applicable methods are summarized in Table 6. The error variance, $var(Z(s) - \hat{Z}(s))$, is again based on predictions obtained from cross-validation.

Table 6.   Black Forest: Error-variance from cross-validation

	Method				
Species	Synthetic Estimator	Ordinary Kriging	Universal Kriging[1†]	Universal Kriging[2†]	Weighted Means[1†]
Spruce	27060	35595	22744	24226	25966
Fir	29725	31040	25757	28210	34992
Douglas Fir	49065	15240	28575	32910	46260
Beech	23015	17410	20065	20164	21556
Oak	10795	8703	9177	10206	9829

[1]: parameterization using field assessment data;

[2]: parameterization using field-plot data;

[†]: reduced data set; see section 5.2.3 p. 49 for explanation

5.2.1 Synthetic Estimator

Auxiliary information is not provided by the forest inventory owing to the applied one-phase sampling design. The only already available source for such information is the field assessment, which is still part of the management planning process. During the field survey plots are allocated to one stand type. The allocation is based on the conditions of the stand the plot is situated in and need not reflect plot characteristics. It follows that species can occur in many different stand types and sampling plots of a stand type can be quite heterogeneous. Douglas fir for example is represented in 25 stand types, among them the type "pure beech".

Stand type is the only auxiliary variable suitable for strata allocation. However, using types is impractical because most types do not comprise enough sampling units. Instead, following the concept implemented in Lower Saxony, eight strata are formed based on four age intervals and two species group states. Interval length and definition of species group are as introduced in section 3.1. These strata are rather coarse but still for all species except beech some strata have less than 30 observations. The obtained error variances are comparable but larger for spruce and beech as the respective ones at Grünenplan (cf. Tables 2, 3, and 6). The error variance for douglas fir is rather large with over 49000 but surprisingly low for oak with 10795.

5.2.2 Ordinary Kriging

External trendfunction

Using the yield table as external trendfunction requires the extrapolation of observed volumes to potential volumes as laid out in section 4.2.1. Data processing routines of the forest research station of Baden-Württemberg compute crown cover areas, (STR_I: Standraum ideal), for every tree as a linear function of dbh (Dahm, 1995). These models are parameterized using the yield table entries stem number per hectare and d_G. The modeled crown area is obtained as the ratio of 1 ha to corresponding stem number, that is, $crown\ area = \frac{10,000\ m^2}{stem\ number}$. Total crown area per hectare per plot is normed to match 10,000 m^2.

Plots with low stocking densities get extrapolated crown areas whereas on over-stocked

stands (cf. Fig. 6) crown areas are reduced. Reduction factors differ with respect to age-class; crown area of "regeneration" gets more reduced and "shading trees" or "under-storey" do not get any crown area at all (FVA Baden-Württemberg, 1999). This reduced crown area is referred to as STR_R, (Standraum reduziert).

Individual age is recorded for trees only which have their heights measured. The remaining trees are assigned to an age-class representative of the stand the sampling plot is situated in. Variability of individual tree volume per age-class is larger than for individual age as seen from Fig. 12, where volume of douglas firs is plotted over age-class (left) and age (right). The many trees with age zero in the righthand side of Fig. 12 are trees without height measurements.
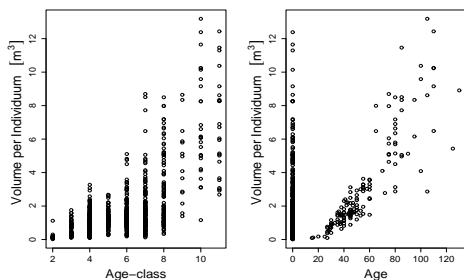
Figure 12.

Volume variability of douglas fir per age-class (left) and age (right)

As mentioned before, potential plot volume can be linked to the yield table by three table items, namely d_G, h_G, and age. The investigated stands are quite heterogenous rendering the d_G unsuitable. The d_G is based on squared diameters so large trees will have a pronounced effect and distort the derivation. However, the corresponding h_G will be less affected. Heights are completed for all trees by data processing routines of the FVA of Baden-Württemberg. Furthermore, height can be estimated rather easily at target stands rendering it a suitable link item. Age-class as a proxy for age is also considered a link item. Using plots where only age is recorded is not appropriate. First, the number of available plots would be significantly reduced and second, species occur with different ages on the same plot making a reliable allocation impossible.

There are four possible approaches to model ordinary kriging with external trend: extrapolation based on "Voronoi" polygons or based on STR_R, and linking based on age or height. The resulting spatial structures are displayed for douglas fir in Fig. 13.

Ordinary kriging is confined to plots having site index information allowing the external trend to be more flexible. This seems appropriate because there are only a few plots without SI entries.
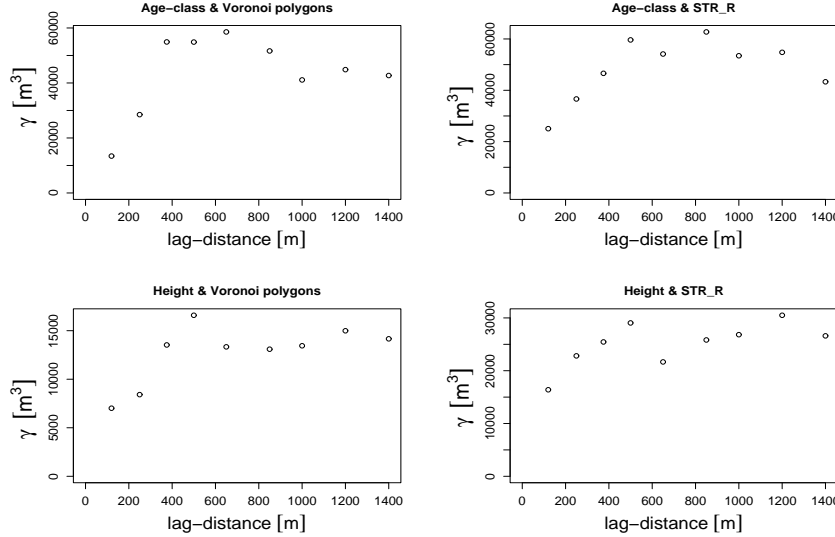
Figure 13. Spatial structure analysis: Douglas Fir

Left: Extrapolation via "Voronoi" polygons; Upper row: link via age-class;
Right: Extrapolation via STR_R; Lower row: link via height; γ: Semi-variance

The observed influences on the spatial structures of the residuals of the different input parameters hold for all species unless specifically mentioned otherwise.

As seen in the upper right part of Fig. 13 using age-class and STR_R leads to a pronounced spatial structure of the residuals. However, the large nugget of around 18000 and sill of around 38000 at a range of 500 m do not seem to justify kriging. Using height in lieu of age-class halves the process variance down to around 28000 which is mostly due to a reduced sill because the nugget dropped only to around 13000 (second row, righthand side).

Using age-class and "Voronoi" polygons (first row, lefthand side) the empirical variogram has the same large process variance as in the first approach but the nugget drops down to null. The range remains almost unaltered. This is also the case in the last trial based on height and "Voronoi" polygons seen in the lower left part of Fig. 13. The process variance of around 15000 is only $\frac{1}{4}$ compared to approaches using age-class and the ratio of nugget to sill is of around $\frac{1}{2}$, rendering a kriging interpolation quite suitable.

Results of the spatial structure analysis are summarized for all species in Table 7. It

holds for all species that approaches using STR_R are outperformed by the respective approach based on "Voronoi" polygons. Spatial structures based on height as link item are more favorable than when based on age-class in all but one case. For spruce modeling based on age-classes almost halved the process variance down to 46000. There is quite a pronounced spatial structure with a ratio of sill to nugget around 1 and a range of 755 m.

Table 7.   Model parameters for ordinary kriging

	Height				Age-Class			
Species	Nugget	Sill	Range	Type	Nugget	Sill	Range	Type
Spruce	44000	34500	300	Exp	21000	25000	755	Sph
Fir	30490	—	—	Nug	44175	—	—	Nug
Doug. Fir	4800	8750	500	Sph	0	56000	550	Sph
Beech	15000	3500	1100	Sph	14000	9000	655	Sph
Oak	3528	—	—	Nug	6315	2905	978	Sph

Exp: Exponential semivariogram model; Sph: Spherical semivariogram model

Nug: Nugget semivariogram model; residuals to yield-table based on "Voronoi" polygons;

There is no discernible spatial auto-correlation within the modeled residuals for fir and oak resulting into pure nugget semivariograms. The nugget of 30490 is quite large for fir whereas for oak the nugget of 3528 is almost $\frac{1}{9}$ of the former. The spatial auto-correlation of the beech residuals is modest with a large nugget of 15000 and a small sill of 3500 reached at 1100 m.

Looking at the process variances three groups can be distinguished which are also reflected by their corresponding estimation variances shown in Table 6. Results in Table 6 are obtained using the most favorable semivariogram model.

Prediction is quite precise for oaks with an error variance of 8703 (pure external trend prediction). The next group comprises beech and douglas fir. Their error variances are roughly two times the one for oak with 17410 and 15240. The error variance of spruce and fir are yet again roughly double the magnitude of the former with 35595 and 31040.

Residuals of douglas fir and beech are displayed in Fig. 14. Residuals based on age-class as link item show a modest non-linear relationship with age for douglas fir (upper left part of Fig. 14). For beech, variation in the residuals first increases until around 75 years and then decreases again (upper right part). No obvious trend with yield table volume can be discerned for douglas fir (2nd row, lefthand side) whereas for beech there seems to be a modest correlation with volume (2nd row, righthand side).
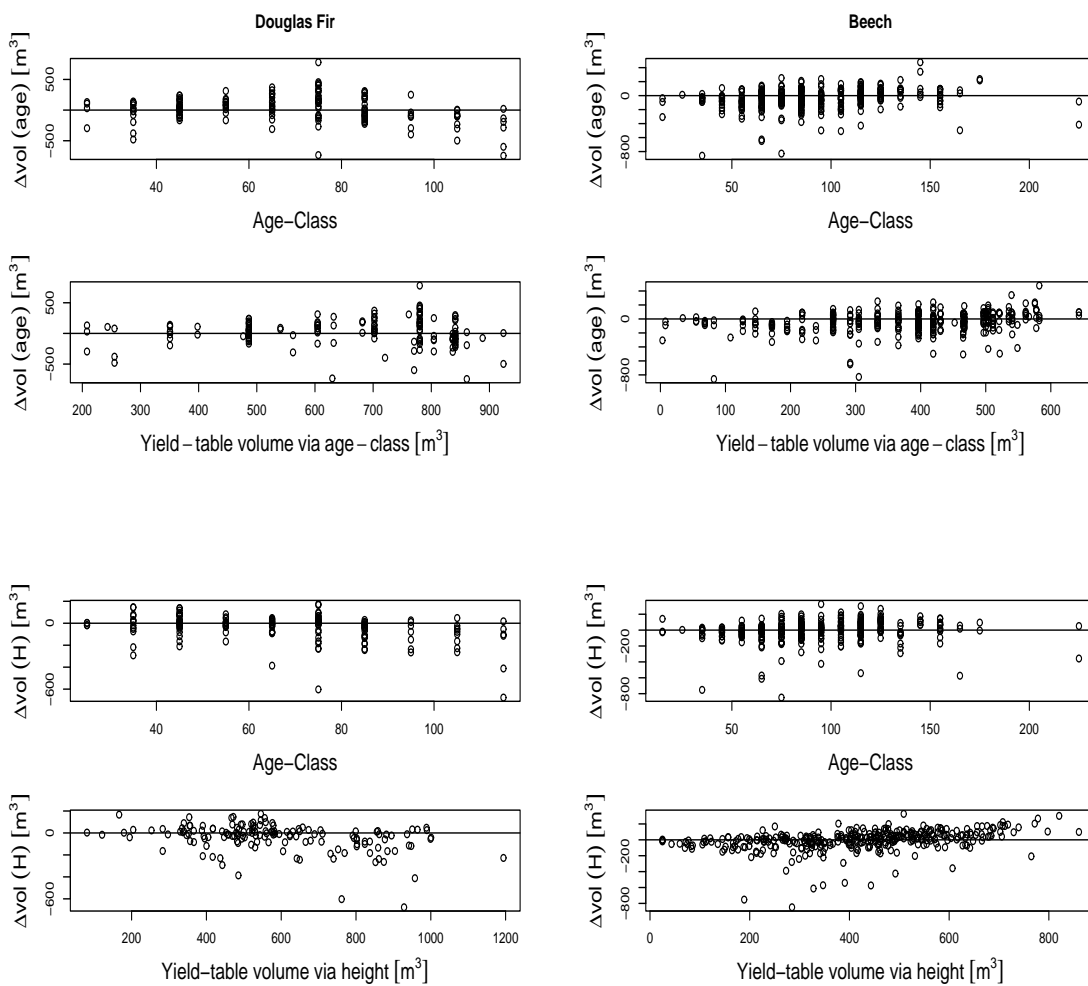
Figure 14.

Residual Analysis: Douglas Fir (left) and Beech (right); $\Delta vol = Vol_{pot} - Vol_{Yield-Table}$

Using height as link item the residuals over age-class (3rd row) show a similar dispersion to their counterparts although overall variation is slightly smaller. Plotting residuals over yield table volume displays a modest decreasing linear relationship with volume for douglas fir and a modest increasing relationship for beech (last row). The observed deviations do not seem to justify a transformation, kriging can be applied directly to the modeled residuals.

Comparing the dispersions in Fig. 14 to the dispersions of the residuals of spruce and beech shown in Fig. 8, p. 32, indicates problems in the extrapolation. There are far less outliers for spruce and beech at Grünenplan than for douglas fir and beech at Baden-Württemberg.

Varying means

This approach failed for all species but oak. The advantage of varying means is that no extrapolation is needed. Using potential volume for calculating the mean-trend as described in section 5.1.2 would counter this advantage and hence is not tried here.

The empirical semivariogram for oak is fit with a spherical model. The nugget is low with 3744. The sill is estimated as 8589, the range as 504 m resulting into a cross-validated error variance of 11176. The prediction is worse than the approach based on "Voronoi" polygons and height (cf. Table 6, Ordinary Kriging) but outperforms the approach based on age-class and "Voronoi" polygons which yields an error variance of 12960.

5.2.3   Universal Kriging

The coincidence of the field survey of the forest inventory and mapping by the cadastre service was by chance. The cadastre service maps the entire federal state on a regular basis but time lags of up to three years to a field survey may occur. Unless field surveys are coordinated with the mapping the panchromatic images cannot be regarded a regular source for auxiliary information. Considering the images in the present study is based on the experience gained at Grünenplan that in order to draw on the strengths of universal kriging variates of high spatial resolution must be used in the prediction process. The variates of the field assessment reflect stand conditions and hence have lower spatial resolution than field plots.

The cadastre service of Baden-Württemberg does not provide a digital terrain model. Such a DTM is needed for ortho-rectifying the images and for the normalization of a digital surface model derived from the images. The derivation of an elevation model from the panchromatic images as a substitute for the DTM failed. Large areas do not get an elevation assigned. Still worse is that there are jumps in elevations of more than 100 m which cannot be attributed to a change in elevation of the landscape. Such jumps occur in large patches of pixels which cannot be filtered out by algorithms provided by the employed software PCI.

An explanation for the failures is the bad quality of the paper prints. The resulting contrast in the digital images is low and the matching algorithm cannot find corresponding points with high enough a correlation within the search environment. The matching algorithm also requires values for minimum and maximum occurring elevation. The large difference of around 500 m of these limits found in most image pairs implies a larger search environment. The latter in conjunction with the low contrast is assumed to be the cause of the before mentioned jumps; spurious matches are found way off sensible search distances.

No automatic derivation of auxiliary information can be obtained from the available panchromatic images. Smoothing of the jumps would require algorithms implemented in a higher programming language like *C* or *C++* that work directly on the image files. The extra expenses of such an implementation would only seem justified if the images were a regular component of the forest management planning like in Lower Saxony. Digital orthophotos of low geometric quality can be computed using these specially filtered images but elevation data would still be unavailable. Of course, another possibility would be using a sophisticated and expensive softcopy photogrammetry station, access to which has not been available in this study.

Instead universal kriging based on variates surveyed during the field assessment has been tried. Results are not expected to be very precise because the variates are obtained by expert knowledge, measurements are not taken, and represent the entire stand. Suitable variates employed are SI, age-class, and mixture proportion. Stocking density has not been available.

Field assessment information is surveyed per spatial units ("Teilflächen"). Each unit

has a unique ID and entries of district, compartment, sub-compartment, sub-sub-compartment, and auxiliary management units. Field plots have description of district, compartment, sub-compartment, sub-sub-compartment. No information on auxiliary management units is recorded. Most often such units do not exist at this early stage of the management planning but are allocated in the course of the field assessment.

The two datasets are joined on a key concatenated of the four location records of the field plots. The join failed for seven field plots; no respective key of the field data is found in the assessment data. Twelve plots are matched to three spatial units, 223 plots to two units, and 549, that is 70%, have a unique match. Further analysis is confined to field plots with unique match because mixture proportion can vary substantially between auxiliary management units even within a sub-sub-compartment.

Species are affected differently by this reduction: there are 77 out of 149 field plots left with douglas fir, 214 out of 340 for beech, 132 out of 172 for fir, 101 out of 142 for spruce, and the number of oak plots is reduced down to 151 from 219.

Species volume is modeled by a linear combination of SI, age-class, and mixture proportion. All variates are provided by the field assessment but the last can also be computed from modeled crown projection areas. However, in the prediction process mixture proportions are provided by the field assessment only. Model parameters of both approaches are summarized in Table 8. As expected spatial structures of the residuals are more pronounced when mixture proportions are computed from plot data. Process variances are half as large except for spruce and all models have a sill although for beech it is less than 10% of the nugget. However, the gain in spatial structure precision does not translate into a gain in prediction precision. As seen in Table 6, p. 41, error variances are always smaller when universal kriging is based on mixture proportions taken from the field assessment.

Predicted values employed in the error variance of Table 6 are computed as the sum of the deterministic trend and cross-validated residuals. Both approaches have in common that the mixture proportion of the field assessment is employed at the prediction location. They differ with respect to trendfunction and semivariogram parameters

and residuals used in cross-validation; each approach uses its residuals, semivariogram parameters, and trendfunction parameters.

Table 8. Model parameters for universal kriging

Species	Field Assessment				Plot Data			
	Nugget	Sill	Range	Type	Nugget	Sill	Range	Type
Spruce	22985	—	—	Nug	6609	6534	232	Sph
Fir	19459	8767	647	Sph	9980	2762	582	Sph
Doug. Fir	22037	13228	2087	Sph	9870	5734	603	Sph
Beech	20162	—	—	Nug	10317	884	822	Sph
Oak	7002	2266	195	Nug	4151	930	399	Sph

Exp: Exponential semivariogram model; Sph: Spherical semivariogram model

Nug: Nugget semivariogram model;

Field Assessment: mixture proportion taken from field assessment

Plot data: mixture proportion derived from plot values

The nugget of the semivariogram indicates short-range components of the random function $Z(s)$ below the shortest sampling distance or measurement errors (Venables and Ripley, 2002). In general, the larger the nugget the larger the variation between adjacent neighbors is. The weighting in the prediction process is determined by the parameters sill and range of the spatial structure. The nugget does not influence the weights as can be seen in Eq. (24), p. 112. The semivariogram values can be expressed as the sum of nugget and spherical or exponential semivariogram. Because $\sum_{i=1}^{n} \lambda_i = 1$, the nugget can be factored out on the lefthand side of Eq. (24) and then cancels with the nugget on the righthand side of the equation.

Correlations between the two mixture proportions can be as low as 0.29 (fir) and do not exceed 0.56 (douglas fir). Hence, plot mixtures are not good proxies of the field assessment mixtures and the cross-validated residuals will not add up sensibly with the trend estimates. Cross-validated residuals based on field assessment mixtures match better with the trend estimates. A reason why this approach always yields lower error variances.

Given the coarse resolution of the employed auxiliary variables, results for universal

kriging are quite satisfactory compared to ordinary kriging. error variances of the latter is only 53% for douglas fir but for beech and oak out-performance is lower with 87% and 94%; in case of spruce and fir universal kriging performs better than ordinary kriging with error variances being 64% and 83% of the latter.

5.2.4 Nearest Neighbors

Most Similar Neighbor

MSN using canonical correlation analysis is not an appropriate inference tool with the given data. Auxiliary information is derived from the field assessment and hence is quite coarse in resolution. Prediction of stands must be based on plots reflecting its variation.

However, in the present case a prediction at any location within a given stand would be based always on the same set of neighbors, where in the worst case the set would comprise only one observation. Using just one observation as a predictor is unreliable owing to the small plot size not being representative of a stand. A set of neighbors can comprise more than one observation because of the discrete nature of data (eg. age-classes, or SI). In such a case MSN would be equivalent to a synthetic estimator with strata counts hardly above 30; prediction would be still unreliable.

Weighted Mean

The weighted mean prediction is based on the same auxiliary variables used by universal kriging: age-class, SI, and mixture proportion. The latter is also taken from the field assessment. The variables are standardized and their contribution to the weight function heuristically modeled. The same weighting is applied to all species. Age-class gets multiplied by 1.5, SI is halved, and mixture proportion is multiplied by 3.5. Results are summarized in Table 6. There is a modest gain in precision compared to the synthetic estimator for all species except for fir where the estimation variance increases by 18% to 34992.

5.3 Lower Saxony: Solling

The first two forest enterprises were chosen for model selection. Given the particular characteristics of the data cross-validation used in the evaluation allows only a comparison of the different methods. Owing to the large variation within the sample plot data (cf. Table 1, p. 26) these points are hard to predict. Stands do not grow homogeneously and the relatively small field plots will fall into dense and sparse parts with resulting over- and under-prediction. Hence the estimation variance is but a rough guide to the prediction quality of entire stands. On the stand level over- and under-prediction will partly cancel out each other. This conjecture can only be validated with fully enumerated or intensively sampled stands.

Full enumeration of a number of stands is too expensive and goes beyond the scope of this study. Instead, 12 stands have been intensively sampled. Their characteristics are summarized in Table 9. The chosen stand-types and their ages reflect their economical importance expressed by the forest service. Pure beech (type 20) and spruce stands (type 50) and different mixtures thereof (type 25: beech prevailing and 52: spruce prevailing) are selected. The three age classes differ according to the prevailing species, see Table 9 for details. Young stands up to 50 years age are not investigated because of their low volume.

There is a large variation within the surveyed stands. Even for a classical pure spruce stand with age around 55 years a minimum of 21 sample plots is needed to meet the initial confidence interval width. In the mentioned stand the applied plot radius was 15 m so plot size was almost seven times the size of BI plots for trees with $dbh < 30$ cm. The initially aimed at half confidence interval length of 10% of the mean estimate at the $\alpha = 0.05$ level proved unrealistic.

For the pure beech stands such an interval length would require from the youngest to the oldest stand 48, 33, and 39 field plots. In mixed stands with beech prevailing the respective number of plots are 797 (!), 368, and 107. This implies that volume cannot be estimated from a sample given the required precision because the summed area of the 797 plots exceeds the actual stand size. Even increasing the half confidence interval length to 20% would require 200, 92, and 27 plots for the mixed stands. The mixed stands with spruce prevailing are less heterogenous except for the oldest one.

This can be attributed to the large percentage of spruce which is 80% in the first two cases but 60% in the last case.

This large encountered variability has led to a redefinition of the initially proposed half confidence interval length up to 20% of the estimated mean at the $\alpha = 0.05$ level. The confidence interval length can be regarded as a function of the number of samples n, ceteris paribus. The function is not linear but a function of the square root of n. This means that there is a decreasing gain in precision with increasing n. The maximal affordable number of sampling plots per stand is set to 30. The costs of an increase of number of samples beyond 30 outweighs the benefits of a better precision.

Table 9.   Characteristics of intensively sampled stands

Stand-Type/ Age [years]	n	n'	$\hat{\mu}_B$	$\hat{\sigma}_B$	CI/2	$mix\%_B$	$\hat{\mu}_S$	$\hat{\sigma}_S$	CI/2	$mix\%_S$
20/ 60-80	17	1	267	90	46	95	10	39	20	2
20/ 80-120	10	3	263	73	52	90	17	53	38	8
20/ >120	14	3	352	107	62	98	16	40	23	2
50/ 50-70	16	5	—	—	—	—	352	78	42	100
50/ 70-90	17	3	6	26	13	2	655	254	131	98
50/ >90	18	4	—	—	—	—	543	129	64	100
25/ 60-80	24	2	50	72	30	48	295	146	62	50
25/ 80-120	20	4	201	157	73	65	95	156	73	35
25/ >120	27	3	205	108	43	45	183	159	63	55
52/ 50-70	19	–	29	55	26	20	409	141	68	80
52/ 70-90	22	4	5	12	5	5	535	105	46	95
52/ >90	15	4	186	174	97	40	226	233	129	60

20: pure beech; 50: pure spruce; 25: predominantly beech; 52: predominantly spruce;

$_B$: Beech; $_S$: Spruce; n: number of field plots; n': number of phase 2 BI-plots;

$\hat{\mu}$: mean volume [m^3]; $\hat{\sigma}$: standard deviation of volume estimates;

CI/2: half confidence interval length; $mix\%$: mixture percentage from field assessment;

The new confidence interval length has been achieved in all pure stands. In mixed stands the only stand with precise enough beech mean estimate is the beech prevailing

type older than 120 years. The other mixed types are much more heterogenous with coefficients of variation based on 22 observations up to 51% (cf. Table 10). There is less variation found in spruce stands. The confidence interval length has also been achieved in the youngest beech prevailing stand type and in the spruce prevailing types except for the oldest stand. In general, the younger the stand and the higher the proportion of spruce the lower the variance of the surveyed spruce volume is.

Only a selection of the models applied at Grünenplan and at Baden-Württemberg are investigated here. The weighted means predictor performed rather well at Grünenplan. Its major drawback is the heuristic derivation of the weights. A general recommendation on which variables should be taken into the model seems feasible, but no reasonable recommendation on the weight function can be made. The weight function must be fitted separately for every forest district and every species, demanding modeling input. In addition, the weighted means predictor is based on auxiliary information derived from both the field assessment and the digital images. Using information just from the former proved inefficient. Employing information only from the digital images is not supported because with this data source the more objective "most similar neighbor" predictor is preferred. Mixing of data sources is not supported either because is does not allow cost cutting.

Based on the same auxiliary information universal kriging performed equally well. It is a more sophisticated and computationally demanding technique but requires less user interaction. Several starting parameters should be tried but a grid search can be implemented in an application software and thus the parameter selection can be shielded from the user. Such an encapsulation is much harder to program for the weighted mean predictor. For these reasons this predictor is not further studied.

The other predictor dropped is ordinary kriging. Using varying means as a coarse trend was only applicable in two out of seven cases and is therfore not further supported. The problems with the application of a yield table as an external trend are twofold: extrapolation and linking.

The only suitable approach to extrapolation without field measurements is via crown cover data. The equivalent on a stand level canopy cover can be assessed quite reliably by experienced inventory foresters. However, modeling crown cover on a plot level is

problematic because of the concentric circles. For example, total crown cover areas greater than the actual plot area are unavoidable. This problem has also been reported by Gill et al. (2000). They tackled this problem by using models with an upper asymptote.

However, the large variation introduced by the extrapolation has a much more severe influence on the ordinary kriging performance. This variation is due to the imprecision introduced by crown cover models. Modeling single tree crown area as a linear function of the *dbh* produced good fits for Nagel (1999). Similar models studied by Gill et al. produced much lower coefficients of determination (around 0.5). The goodness of fit of the resulting canopy cover model at the plot level was worse.

Extrapolation is achieved by multiplying the measured volume by the ratio of plot size to modeled canopy cover. Underestimation of the latter can result into unrealistic high potential volumes and are one explanation of the large deviations observed in Fig. 14. Furthermore, large variation leads to large nuggets lowering the suitability of a kriging approach.

Linking plot data to the yield table is straightforward in pure stands. Age is difficult to define in mixed stands and the d_G of the yield tables refers to pure stands. The h_G seems more appropriate as link item but with the investigated species failed to produce residuals in accordance with the intrinsic hypothesis (cf. Fig. 14). Furthermore, universal kriging produced comparable error variances based on the same source of auxiliary variables in Baden-Württemberg.

The different approaches are compared by their root mean square error (RMSE) and their BIAS. The RMSE is computed as RMSE$= \sqrt{\frac{\sum_{i=1}^{n}(V_i - \hat{V}_i)^2}{n}}$, and the BIAS as BIAS$= \frac{\sum_{i=1}^{n}(\hat{V}_i - V_i)}{n}$, where $\hat{V}_i$ and V_i are predicted and surveyed volume of stand $i = 1, \ldots, n$.

5.3.1 Synthetic Estimator and Null-Variant

The synthetic estimator is covered by the null-variant of the forest service. The Null uses the same strata means but adjusts for stocking densities and mixture proportions.

Densities and mixtures are calculated at the plot from modeled crown cover and at the target stands are replaced by estimates obtained from field assessment. In theory it should be superior to the synthetic estimator because it uses more detailed information. Results for both techniques are summarized in Table 10, row one to three for beech, and row five to seven for spruce.

Stratum membership of a stand is determined by the phase 1 plots. Allocation of a phase 1 plot into one stratum is influenced only by the conditions found on the plot; characteristics of the surrounding stand are ignored. It follows that a mixed stand can comprise phase 1 plots belonging to different strata. This is indicated in Table 10 by superscripts in Greek letters; derivation of corresponding volume estimates are listed below the table.

Setting the mean volume of the field survey as the reference the Null always under-predicts beech stands. Even in pure stands under-prediction can be as large as 58% of the surveyed volume. In mixed stands under-prediction can be even as large as 73% in the oldest stand with spruce prevailing. The resulting RMSE is 84 m^3, and the BIAS 75 m^3.

The synthetic estimator has also a tendency to under-predict beech stand volume although for the stand-type 25 aged above 120 years volume is over-predicted by 8% as 222 m^3. Compared to the Null in five out of ten cases the synthetic estimator produces predictions closer to the surveyed volume. Accordingly, the RMSE with 73 m^3 and the BIAS with 53 m^3 are lower than for the synthetic estimator (cf. Table 10).

No systematic deviation in the prediction by the Null or the synthetic estimator is observed for spruce, although the bias is -15 m^3 for both approaches. The Null predictor performs quite well except for the oldest mixed stand with spruce prevailing, where it over-predicts the surveyed volume by 58%. Relative deviations of volume prediction are large for the "*pure*" beech stands but these stands have low spruce volume, hence absolute deviations are small. The RMSE is rather low with 65 m^3.

The synthetic estimator is inferior to the Null-variant in all but three stands. Its RMSE is almost two times larger with 128 m^3, but the BIAS is identical for both techniques with -15 m^3. Predictions are closer to the surveyed volumes in the oldest

stands of type 20, 25, and 50. The gain in accuracy of the synthetic estimator over the Null-variant for these three stands is lower than the loss in accuracy in the other cases. This is especially true for stands with low spruce percentage where prediction by the synthetic estimator is way off the surveyed value.

Overall, apart from one exception, better performance of the synthetic estimator is found in mixed stands with beech prevailing and in the oldest variant of the pure stands. The inferiority of the Null-variant implies that the employed estimates of mixture proportions and stocking densities are not representative for the respective plot values.

Table 10.   Evaluation of prediction based on "existing" information

	Stand-Type / Age-class [years]											
Method	20/ 60-80	20/ 80-120	20/ >120	25/ 60-80	25/ 80-120	25/ >120	50/ 50-70	50/ 70-90	50/ >90	52/ 50-70	52/ 70-90	52/ >90
Survey ($\hat{\mu}_B$)	267	263	352	50	201	205	—	6	—	29	5	186
$\frac{\hat{\sigma}_{\hat{\mu}_B}}{\hat{\mu}_B}$	8	9	9	29	17	10	—	105	—	44	51	24
Null ($\hat{\mu}_B$)	180	109^{α}	235	15	109^{β}	103^{γ}	—	**1**	—	6	**2**	50
SYN ($\hat{\mu}_B$)	147	102^{ξ}	301	12	142^{ν}	$\mathbf{222}^{\kappa}$	—	15	—	12	12	65
UK ($\hat{\mu}_B$)	**196**	**201**	**337**	**77**	**193**	137	—	51	—	**43**	83	**185**
Survey ($\hat{\mu}_S$)	10	17	16	295	95	183	352	655	543	409	535	226
$\frac{\hat{\sigma}_{\hat{\mu}_S}}{\hat{\mu}_S}$	23	31	67	10	37	17	6	9	6	8	4	26
Null ($\hat{\mu}_S$)	**5**	$\mathbf{24}^{\tau}$	2	255	135^{ϵ}	145^{ζ}	**339**	800	590	**408**	**459**	**357**
SYN ($\hat{\mu}_S$)	47	141^{ψ}	**27**	498	228^{φ}	$\mathbf{162}^{\phi}$	328	498	**498**	328	328	431
UK ($\hat{\mu}_S$)	189	197	201	**269**	220	383	**365**	**550**	488	337	453	393

20: pure beech; 25: predominantly beech; 50: pure spruce; 52: predominantly spruce;

Beech: RMSE_{Null}: 84 m^3; RMSE_{Syn}: 73 m^3; RMSE_{UK}: 44 m^3;

Spruce: RMSE_{Null}: 65 m^3; RMSE_{Syn}: 127 m^3; RMSE_{UK}: 132 m^3;

$\hat{\mu}$: mean volume [m^3]; $\frac{\hat{\sigma}_{\hat{\mu}}}{\hat{\mu}}$: coefficient of variation [%]; $\hat{\sigma}^2$: variance of volume predictions;

$_B$: beech; $_S$: spruce; α: 109=(25+2*151)/3; β: 109=(20+30+2*193)/4; γ: 103=(57+2*126)/3

ξ: 110=(12+2*147)/3; ν: 142=(12+15+2*271)/4; κ: 222=(65+2*301)/3

τ: 24=(18*2+36)/3; ϵ: 135=(179+257+2*52)/4; ζ: 145=(327+2*54)/3

ψ: 141=(47*2+328)/3; φ: 228=(328+498+2*43)/4; ϕ: 162=(431+2*27)/3

Of all field plots that comprise at least one spruce in layer 1 51% are pure and 32% comprise two species. For beech, only 33% of the respective plots are pure but 45% comprise two species. Older stands are more heterogeneous with respect to canopy

cover. The more variation there is in stocking density and mixture, the harder these characteristics are to predict and the less precise these estimates become.

5.3.2 Universal Kriging

Performance of the universal kriging predictor depends, among other things like alignment of the sample plots, on the relationship of the regressors with the variable to be predicted. With high resolution auxiliary data it performed rather well at Grünenplan and produced similar results compared to ordinary kriging at Baden-Württemberg when the same source of auxiliary data was used.

As a substitute to ordinary kriging with external trend function universal kriging employing variates from the field assessment ("Bestandesbegang") is applied. Such variates can only be employed if the phase 2 BI plots can be linked uniquely to the new geometries delineated in the course of the field assessment.

Four different trend functions are investigated:

- a linear combination of SI, age, stocking density, and mixture proportion
- a linear combination of SI, age, and stocking density, each multiplied by mixture proprotion
- linear combinations of the product of age, stocking density, and mixture proportion and SI, where the linear combinations are fitted separately for two age classes (below and above 60 years for spruce, below and above 80 years for beech)
- the same linear combinations as before but the REML fit is based on stocking density and mixture proportions calculated from the phase 2 plots and predictions employ the corresponding variables from the field assessment

The second approach yielded the lowest RMSEs for the two species, followed by the fourth. Results reported refer to the second model unless mentioned otherwise.

A GIS coverage with the new geometries is already available and the new description of compartment, sub-compartment, and sub-sub-compartment is attached to the field

plots by an intersection as outlined in section 3.1 and a subsequent join in a database management system.

The join is based on a key concatenated from the forest district, compartment, sub-compartment, and sub-sub-compartment. The joined data set is confined to entries of layer 1 and where the index of the auxiliary management unit is zero. In general, management units can have characteristics different from the surrounding stand but do not have geometries. If a sub-sub-compartment comprises an auxiliary management unit then the index of the latter is greater than zero.

A unique join of the two datasets is still impossible. In the course of the new stand alignment additional structural elements are classified that have associated geometries. Sometimes these elements coincide with the auxiliary management units but there is no 1:1 mapping between the two characteristics.

So, a sub-sub-compartment can comprise up to five structural elements, translating into five different ages, site indexes, stocking densities, and mixture proportions. The last always adds up to 100 over all species of a sub-sub-compartment, regardless of occurring structural elements. Unique values of the field assessment are obtained for the sub-sub-compartment as a weighted sum, the weights being mixture proportions.

Predictions are summarized in Table 10, and model parameters in Table 11. There is pronounced spatial structure in the beech residuals with a sill of 5563 at a range of 860 m and a ratio of sill to nugget of 0.45. The spatial structure of the spruce residuals is much weaker with a nugget of 22409 and a sill of 3863 reached at 984 m.

Compared to the Null-variant for beech, universal kriging yields always predictions closer to the field survey estimates except where the beech volume is extremely low. Apart from the oldest beech-spruce stand the same holds for the synthetic estimator. This better performance is reflected by the low RMSE of 44 m^3. The BIAS of 6 m^3 is also small.

Table 11.   Universal kriging parameters

Species	Nugget	Sill	Range	Type
Beech	12228	5563	860	Sph
Spruce	22409	3863	984	Sph

Trend variables taken from field assessment
Sph: Spherical semivariogram model

For spruce, universal kriging also over-predicts the very low volume estimates, found in pure beech stands, but deviations are much larger. Besides, prediction is satisfactory for the pure spruce stands. In mixed stands deviations up to 132% of the surveyed volume occur, leading to the large RMSE of 132 m^3.

Compared to Grünenplan the derivation of auxiliary information has been modified. Mosaicing images can be done automatically but usually needs post-processing. Furthermore, for larger areas flight paths will comprise many images with resulting large mosaics but it is the file size of the mosaics that matters. Of course, there can be multiple mosaics per flight path but only at the costs of more interactive user input. The image data needs to be intersected with field data but with large coverages containing more than a million polygons the GIS used (ESRI's ^{TM} Arc/Info, Version 8.0.1 on Unix) reached internal software limits.

Instead, transformed image data are used for the Solling area. The first channel and the ratio of first to second and first to third channel are used in the ISODATA unsupervised classification. Band ratioing helps to compensate for changing illumination conditions, surface slope, aspect, and other sources of variation (Lillesand and Kiefer, 2000), and has been applied in other applications of digitized aerial photographs (Haara and Nevalainen, 2002; Holopainen and Wang, 1998). Thus, it serves the same purpose as the interactive mosaicing at Grünenplan, that is, reducing radiometric distortions. The first channel is retained because modeling experience has shown that it carries more information than the other two channels; this is especially true for the DSM derivation.

The ratio-channels are linearly stretched to brightness values within [0,255] to increase contrast. After the classification is run, points with coordinates representing BI or evaluation plots are imported per image. The coordinates of the BI plots and the evaluation plots are stored in only two separate files. In the evaluation stands there is a point every 50 m. These points are burned into a separate image channel and buffered to actually cover an area of a BI phase 2 plot.

The raster data covered by these buffered regions are converted to vector data and exported to files in Arc/Info readable formats. The entire processing described so far is handled by one script written in EASI, GEOMATICA's scripting language.

The next steps are automatized by an AML script, interacting with self-developed *C* and *Perl* programs. AML is Arc/Info's macro language. In principle, any higher programming language can be chosen but *C* and *Perl* are natural choices in Unix/Linux environments and are also available for all major platforms. *C* does CPU intensive calculations and the second language *Perl* is chosen owing to its unsurpassed text processing capabilities, facilitating data re-formatting.

The vector data are read into GIS coverages, one image at a time. The coverages are intersected with the BI-data or evaluation-data held in separate coverages, thus attaching IDs. The GIS provides a native function that joins coverages into one coverage. With the overlap in the data, usually around 60% in flight direction and 30% perpendicular to it, this native function produced unpredictable results; geometries and even worse polygon attribute values changed. Attribute values not existing at a plot were introduced by the join.

The alternative uses system calls of the GIS, passing the coordinates of a bounding box of each image-coverage to system programs. The programs individually reduce these bounding boxes so that no overlap occurs and pass the coordinates of these boxes via a file back to the GIS. The original image-coverages are clipped by their respective *reduced bounding box coverage* and then joined using the native GIS function. The polygon attribute data of the joined coverages are unloaded from the GIS in plain ASCII format and further processed as described in section 3.1, p. 7.

The resulting processed data files are fed into the relational database management system MySQL. Further joins of spectral, height, field assessment, and forest inventory data are handled by the DBMS. The DBMS is accessed by R (Ihaka and Gentleman, 1996), the software employed for all statistical analysis in this study, locally and over the intra-net of the sponsoring institute. R stores its objects in one binary file located here in an environment accessible by many computers via NFS (network file system) and Samba (acronym for server message block).

The described architecture has several advantages. The data are centrally stored, no versioning is possible, and can be accessed from multiple computers enabling "distributed" computing. Even with a computer with 2.2 GHz processor speed some runs can last more than 12 hours, so having multiple machines at one's disposal is an asset.

R allows batch processing, further facilitating the analysis. Once an R script is developed, it can be applied to the different data sets described below with only minor changes.

During model selection in Grünenplan auxiliary information from the imagery was based on seven clusters representing spectral information and a nDSM with spatial resolution of 0.6×0.6 m, that is, heights got derived for every other pixel. For model evaluation and refinement these categories have been modified. As outlined in section 3.1, seven clusters are thought of as a rather large reduction. Instead, in a first trial, the images were classified into 25 clusters. The resulting amount of data proved too large to be further processed in the GIS as outlined above.

In a next step 20 clusters are selected and successfully prepared for the prediction analysis. Additional reductions are 15 and the already applied 7 clusters. Normalized DSMs are derived with a resolution of 0.44 m (every pixel), and a resolution of 0.88 m (every other pixel).

The higher the resolution of the nDSM the better the visual appearance in terms of crown shapes becomes (Nuske, 2002). However, the resulting nDSM has more variation in the height values and also contains more outliers. The latter require extensive editing of the nDSM before the data can be further processed. A direct consequence of the outliers are more failure pixels because large variation implies low correlation in the image matching. Unfortunately, an entire evaluation stand is not covered by the fully resolved nDSM, and therefore it is decided to consider only the nDSM with 0.88 m resolution here. Of course, failure pixels occur also in this data set but apart from meadows in the forest these areas are small (usually below a crown diameter of a mature beech). The main difference to the fully resolved DSM is that there is enough overlap of elevation data between images.

The height information and the spectral information is used to derive a species proportion measure as described in section 5.1.3, p. 35. Additional measures describing mixtures based on the approach by Wildenhein (cf. section 5.1) and stocking densities are also derived by mixture analysis.

The variates used in universal kriging are the 0.85 height quantile, measures for species

proportion, mixtures, and stocking densities, gap, and growing space. Growing space is calculated by integrating the nDSM height values over the sample plot domain. Only the derived auxiliary variables species proportion, mixture, and stocking density are influenced by the number of ISODATA clusters. A hidden auxiliary information is the occurrence of a species in a given stand; it is assumed that one knows which stands comprise beech or spruce.

The models investigated are:

  - a linear combination of gap, 0.85 height quantile, and species proportion
  - a linear combination of gap and the product of 0.85 height quantile and species proportion
  - a linear combination of stocking density, 0.85 height quantile, and mixture
  - the product of stocking density, 0.85 height quantile, and mixture
  - a linear combination of stocking density and product of 0.85 height quantile and species proportion
  - a linear combination of stocking density, product of 0.85 height quantile and species proportion
  - a linear combination of growing space and species proportion
  - a linear combination of growing space and mixture.

Of these models the first with the linear combination already applied in Grünenplan proved to be the most efficient one. Its parameters are summarized by species and source of auxiliary information in Table 12.

Table 12.   Model parameters for universal kriging

Method	Beech				Spruce			
	Nugget	Sill	Range	Type	Nugget	Sill	Range	Type
UK_7	7079	2195	800	Sph	16046	2511	959	Sph
UK_{15}	6867	1572	677	Sph	14058	3427	425	Sph
UK_{20}	6773	1200	669	Sph	15203	1660	800	Sph

[7,15,20]: Spectral auxiliary information based upon 7, 15, or 20 ISODATA clusters;
Sph: Spherical semivariogram model

Compared to the respective parameters obtained in Grünenplan (cf. Table 5, p. 38), the models for beech have lower nuggets and lower process variances. The decrease is

from 10021 (cf. UK_4 in Table 5) to 7973 (UK_{20}). The sill of 1200 of the latter model is rather low, implying that there is not much spatial autocorrelation in the residuals.

There is more spatial autocorrelation in the residuals of spruce but less than compared to the models fitted in Grünenplan. Also, nuggets and process variances are larger than the ones of the respective models in Grünenplan (cf. Table 12 and Table 5).

Results of the prediction of stand volume are summarized in Table 13 for beech and Table 14 for spruce.

Table 13.   Evaluation of UK prediction: Beech

Stand-Type/ Age [years]	Field Survey				UK_7		UK_{15}		UK_{20}	
	$\hat{\mu}$	CI/2	n'	n	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$
20/ 60-80	267	46	1	31	**268**	7591	222	14269	285	5578
20/ 80-120	264	52	3	48	240	10193	**251**	12663	242	5556
20/ >120	352	62	3	60	307	3765	**351**	4507	313	3283
25/ 60-80	50	30	2	25	110	7796	**50**	2385	65	4273
25/ 80-120	201	73	4	71	169	14273	141	11417	**197**	3079
25/ >120	205	43	3	34	141	5634	220	5884	**198**	7105
52/ 50-70	29	26	–	17	81	6582	**63**	1138	101	3677
52/ 70-90	5	5	4	15	**88**	3175	104	21517	94	11567
52/ > 90	186	97	4	69	**164**	10487	217	13399	239	9201
50/ 70-90	6	13	3	62	61	4619	**56**	6763	124	4529

20: pure beech; 25: predominantly beech; 52: predominantly spruce; 50: pure spruce;

$RMSE_7$: 49 m^3; $RMSE_{15}$: 45 m^3; $RMSE_{20}$: 57 m^3;

$\hat{\mu}$: mean volume [m^3]; CI/2: half confidence interval length; $\hat{\sigma}^2$: variance;

n': number of phase 2 BI-plots; n: number of plots the prediction is based upon

[7,15,20]: Spectral auxiliary information based upon 7, 15, or 20 ISODATA clusters;

For beech, all three approaches perform satisfactorily in the pure beech (type 20) and the predominantly beech (type 25) stands. The approach based on auxiliary information derived from 15 ISODATA clusters (UK_{15} for short) yields predictions always within the confidence interval of the respective surveyed volume of these stands with large beech volume. In mixed stands with low beech volume over-prediction beyond the upper confidence interval occurs, especially in the stands where beech volume is less than 10 m^3 (cf. Table 13).

The UK_{20} beech-model performs similar in terms of predictions within the confidence interval. However, it has the largest RMSE of 57 m^3. It is outperformed by the UK_7 approach with a RMSE of 49 m^3 and a RMSE of 45 m^3 for UK_{15}. The RMSE of the last model is only slightly larger than the 44 m^3 of the universal kriging approach based on auxiliary information of the field assessment and almost $\frac{1}{2}$ of the currently applied Null-variant. It should be noted that all RMSEs obtained for beech by universal kriging are lower than the Null-variant (cf. Table 13 and 10). The BIASes of these universal kriging models are (increasing cluster number) -6 m^3, -11 m^3, and -29 m^3.

Table 14.   Evaluation of UK prediction: Spruce

Stand-Type/ Age [years]	Field Survey				UK_7		UK_{15}		UK_{20}	
	$\hat{\mu}$	CI/2	n'	n	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$
50/ 50-70	352	42	5	45	452	8499	**341**	4920	297	3457
50/ 70-90	656	131	3	62	430	5975	**516**	11913	512	16758
50/ >90	543	64	6	61	315	4228	**462**	6849	455	16603
52/ 50-70	409	68	—	17	**318**	10233	278	6733	290	7309
52/ 70-90	535	46	4	15	**452**	3393	450	18783	394	6270
52/ > 90	226	129	4	69	**295**	13291	332	12844	308	12764
25/ 60-80	295	62	2	25	**312**	13663	239	4057	234	3443
25/ 80-120	95	73	4	71	211	15275	221	8945	**179**	7548
25/ >120	183	63	3	34	**208**	3741	249	4386	225	8059
20/ 60-80	10	20	2	31	178	3168	205	1830	**142**	449
20/ 80-120	17	38	4	48	173	2451	217	6074	**169**	4875
20/ >120	16	23	3	60	**186**	696	224	1477	197	2264

50: pure spruce; 52: predominantly spruce; 25: predominantly beech; 20: pure beech;

RMSE$_7$: 138 m^3; RMSE$_{15}$: 131 m^3; RMSE$_{20}$: 115 m^3;

$\hat{\mu}$: mean volume [m^3]; CI/2: half confidence interval length; $\hat{\sigma}^2$: variance;

n': number of phase 2 BI-plots; n: number of plots the prediction is based upon

$_{7,15,20}$: Spectral auxiliary information based upon 7, 15, or 20 ISODATA clusters;

There is an inverse proportional relationship between the number of clusters used in the derivation of auxiliary information and the RMSE of the prediction of spruce volume they are employed in. The RMSEs are 138 m^3, 131 m^3, and 115 m^3 for UK_7, UK_{15}, and UK_{20}, respectively. The corresponding BIASes are -16 m^3, -33 m^3, and -6 m^3. As for beech, excessive over-prediction occurs where spruce volume is very low,

but also for the stand type 25, age 80-120 years, with 95 m^3. The closest prediction is by UK$_{20}$ with 179 m^3, still beyond the upper confidence limit (cf. Table 14).

In general, predictions of all methods are not within the confidence interval of the surveyed volume. Exceptions are found in four stands: the pure spruce stands younger than 90 years and the oldest representative of both mixed types.

The RMSEs are very large compared to the Null with 65 m^3, but of similar magnitude of the ones of the synthetic estimator and universal kriging based on field assessment information. This holds true except for UK$_{20}$, which comes second to the Null-variant in terms of RMSE. If one excludes the "pure" beech stands with very low spruce volume then the RMSEs are reduced to 128 m^3, 97 m^3, and 97 m^3, where the order is as above. So, even with this exclusion of difficult to predict stands, the smallest RMSE is still one and a half times larger than the one of the Null-variant.

Universal kriging doesn't predict low volumes very well. This holds for both investigated species, although over-prediction is less severe for beech than for spruce. The frequencies of per hectare volume of the BI-field plots used in modeling are displayed by histograms for both species in Fig. 15. Plots with beech volume below 50 m^3 are most frequent whereas for spruce the respective volume is around 275 m^3.

Figure 15. Volume variability within BI plots

The shape of the histograms helps to explain the different performance. The large proportion of small beech volumes mean larger leverage during parameter fit. The result is that the model has a smaller intercept in the universal kriging trend-model and can better predict lower volumes. Per hectare spruce volume of 50 m^3 is less frequent than 500 m^3,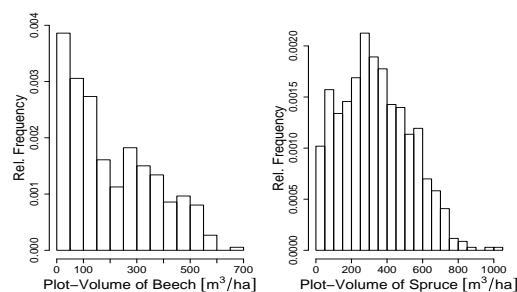 resulting in lower leverage during parameter fit of the trend-function and hence larger intercepts compared to the beech model.

5.3.3   Most Similar Neighbor

Design attributes used in the canonical correlation analysis are number of trees on the plot, and per species h_G, volume per ha, and estimated crown area as described in section 5.1. The derivation of auxiliary variables used as indicator attributes is explained in section 3.1, p. 7. Employed indicator attributes are total and mean area per p-1 ISODATA classes (p=7, 15, 20), mean ratio of circumference to area of the classified polygons, and per class the number of polygons that have larger class value than all adjacent polygons.

Indicator attributes derived from the normalized DSM are area weighted variance of height values, number of polygons representing failure pixels, maximum difference in heights, the number of total polygons and their mean area, the area of polygons classified as gap, and the 0.1 and 0.85 height quantile.

The inclusion of variables is based on expert knowledge guided by an evaluation of the contribution of the variables to the canonical correlations. Inclusion is also influenced by an evaluation of consistencies between correlations among indicator variables and their respective canonical coefficients.

A selection based only on the contribution to the canonical correlations is not appropriate. Increasing the number of indicator attributes does not necessarily increase the canonical correlation as is the case for the coefficient of determination in regression analysis. However, multicollinearity among the indicator variables will be aggravated, leading to unstable predictions. For example, a cause of instability are positively correlated indicator variables that have opposite sign in their canonical coefficients. This was the case when basal area of the investigated species was included. Canonical correlations increased but predictions were very unstable, which was reflected by larger RMSEs than reported below.

Multicollinearity is a matter of degree, not kind (Dougherty, 1992). Here, the required modeling user input is proportional to the number of polygons the spectral information is classified into. With spectral information classified into 15 and 20 clusters some clusters may represent the same feature like shaded crown matter. Including all information of all classes belonging to the same feature type would increase multi-

collinearity. Instead of using information of all classes only a subset is applied based on a heuristic selection .

A first model looked for nearest neighbors within stands that comprised the respective species. As for the universal kriging approach, this model assumes that one knows that a stand comprises the species to be predicted. In such stands the species proportion measure, explained in section 5.1.3, p. 35, can be used as an indicator attribute. In addition, the 0.85 height quantile has been applied to divide the plots of the target stands into two populations: below and above their mean height value. The prediction based on two height populations is inferior to the one based on the entire population. Furthermore, including the species proportion measure does not improve the prediction either.

Results of the most efficient approaches are shown in Table 15 for beech and Table 16 for spruce. Predictions are based on the indicator variables listed above and summarized by the number of spectral classes these indicator variables are derived from.

Prediction is quite good for beech. The RMSE is largest with 77 m^3 for the approach based on 20 ISODATA clusters (MSN$_{20}$ for short) but still smaller than the one of the Null-variant. MSN$_7$ has a smaller RMSE with 61 m^3. The RMSE of the MSN$_{15}$ approach is 48 m^3, only 57% of the Null-variant. The BIASes are (increasing cluster number) 48 m^3, 16 m^3, and 61 m^3. There seems to be a general tendency towards over-prediction, not only for the smaller values as one would assume for this predictor (cf. Table 15). Even the stand with the largest surveyed volume got over-predicted by 29 m^3.

The MSN predictors also seems to over-predict the volume for spruce as the BIASes are (increasing cluster number) 38 m^3, 6 m^3, and 1 m^3. However, in general there is a tendency towards under-prediction, especially for stands with high volume (cf. Table 16). The positive BIASes are due to the excessive over-prediction of the low spruce volume in the "pure" beech stands. The MSN$_7$ predictor has the lowest RMSE with 124 m^3, followed by MSN$_{15}$ with 138 m^3, and MSN$_{20}$ with 141 m^3.

The RMSEs are rather large but similar to the ones obtained by universal kriging

Table 15. Evaluation of MSN prediction[‡] : Beech

Stand-Type/ Age [years]	Field Survey				MSN$_7$		MSN$_{15}$		MSN$_{20}$	
	$\hat{\mu}$	CI/2	n'	n	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$
20/ 60-80	267	46	1	31	337	22628	**244**	26739	396	54725
20/ 80-120	264	52	3	48	325	23625	**253**	42941	252	23789
20/ >120	352	62	3	60	**328**	10377	327	21606	381	7650
25/ 60-80	50	30	2	25	158	17208	**76**	5130	133	11877
25/ 80-120	201	73	4	71	**195**	24542	130	11761	**195**	22604
25/ >120	205	43	3	34	248	19388	**236**	13688	237	17015
52/ 50-70	29	26	–	17	115	19609	102	10315	**87**	5349
52/ 70-90	5	5	4	15	**52**	7376	68	4207	120	9443
52/ > 90	186	97	4	69	220	14306	**210**	22359	270	19288
50/ 70-90	6	13	3	62	**68**	5262	81	5819	105	5506

[‡]: prediction based upon nearest neighbor within BI plots comprising a beech

20: pure beech; 25: predominantly beech; 52: predominantly spruce; 50: pure spruce;

RMSE$_7$: 61 m^3; RMSE$_{15}$: 48 m^3; RMSE$_{20}$: 77 m^3;

$\hat{\mu}$: mean volume [m^3]; CI/2: half confidence interval length; $\hat{\sigma}^2$: variance;

n': number of phase 2 BI-plots; n: number of plots the prediction is based upon

$_{7,15,20}$: Spectral auxiliary information based upon 7, 15, or 20 ISODATA clusters;

(cf. Table 14). If the "pure" beech stands are excluded then the RMSEs decrease to 98 m^3, 111 m^3, and 120 m^3, where the order is by increasing cluster number.

The tendency to over-predict small and under-predict large values is a well known characteristic of nearest neighbor techniques. The differences between the species in the over-prediction of stands with low volume can partly be attributed to the distribution of the BI plot-volume displayed in Fig. 15. For beech, there are more potential neighbors at the lower tail of the distribution of BI plot-volume than for spruce, corresponding to the smaller over-prediction of beech stand volume.

One approach to alleviate over-prediction of low species volume in mixed stands is to search for neighbors within all BI plots, not only the plots that comprise the respective species. Prediction results of such an approach are summarized in Table 17. MSN based on 20 ISODATA clusters is not further investigated because it is inferior to the

Table 16.   Evaluation of MSN prediction[‡]: Spruce

Stand-Type/ Age [years]	Field Survey				MSN_7		MSN_{15}		MSN_{20}	
	$\hat{\mu}$	CI/2	n'	n	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$
50/ 50-70	352	42	5	45	**358**	27981	325	13907	271	19917
50/ 70-90	656	131	3	62	484	35680	472	40879	**572**	76725
50/ >90	543	64	6	61	414	28484	**467**	36895	327	38801
52/ 50-70	409	68	–	17	**389**	19750	186	18186	225	19773
52/ 70-90	535	46	4	15	**521**	20556	461	28160	442	24041
52/ > 90	226	129	4	69	360	29328	**287**	33152	344	26469
25/ 60-80	295	62	2	25	**235**	21530	217	15426	200	17049
25/ 80-120	95	73	4	71	217	25779	164	15617	**152**	11028
25/ >120	183	63	3	34	249	23713	204	33875	**202**	16400
20/ 60-80	10	20	2	31	189	24032	**122**	13950	185	17497
20/ 80-120	17	38	4	48	**141**	12076	249	55820	211	21150
20/ >120	16	23	3	60	238	20308	254	40514	**219**	20500

[‡]: prediction based upon nearest neighbor within BI plots comprising a spruce

50: pure spruce; 52: predominantly spruce; 25: predominantly beech; 20: pure beech;

$RMSE_7$: 124 m³; $RMSE_{15}$: 139 m³; $RMSE_{20}$: 141 m³;

$\hat{\mu}$: mean volume [m³]; CI/2: half confidence interval length; $\hat{\sigma}^2$: variance;

n': number of phase 2 BI-plots; n: number of plots the prediction is based upon

[7,15,20]: Spectral auxiliary information based upon 7, 15, or 20 ISODATA clusters;

other two methods.

Including all BI field plots into the search actually improved the overall prediction for beech in terms of RMSE and BIAS. The RMSE of the MSN_7 predictor is 36 m³, less than $\frac{2}{3}$ of the approach listed in Table 15, and even outperforms the best MSN prediction based on plots comprising beech by 25%. The reduction down to 46 m³ of the RMSE of the MSN_{15} predictor is less pronounced. The BIAS is reduced to -23 m³ for MSN_7 and -7 m³ for MSN_{15}, indicating under-prediciton.

There is systematic under-prediction by MSN_7 of stands where beech occurs. As for the MSN_{15} predictor, the stand with the third largest beech volume, stand type 20 older between 80 and 120 years, gets over-predicted by MSN_{15} by 37 m³. The over-

Table 17. Evaluation of MSN prediction based on all BI plots

Stand-Type/ Age [years]	Field Survey		Beech **MSN$_7$**		Beech MSN$_{15}$		Spruce **MSN$_7$**		Spruce MSN$_{15}$	
	$\hat{\mu}_B$	$\hat{\mu}_S$	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$	$\hat{\mu}$	$\hat{\sigma}^2$
20/ 60-80	267 (46)	10 (20)	**209**	36626	150	22665	**45**	14926	79	26914
20/ 80-120	263 (52)	17 (38)	**259**	43194	300	34107	73	25822	**63**	17158
20/ >120	352 (62)	16 (23)	316	28076	**340**	19132	**25**	4095	38	7431
50/ 50-70	— (—)	352 (42)	13	1613	**0**	0	283	17713	**313**	33164
50/ 70-90	6 (13)	655 (131)	**5**	591	36	5900	**511**	50897	451	46500
50/ >90	— (—)	543 (62)	**14**	1365	53	11643	**426**	50990	386	31662
25/ 60-80	50 (30)	295 (62)	31	6309	**39**	5512	**217**	23873	185	16906
25/ 80-120	201 (73)	95 (73)	**130**	30902	126	22274	146	47174	**115**	24491
25/ >120	205 (43)	183 (63)	158	19163	**218**	21152	113	18440	**122**	19165
52/ 50-70	29 (26)	409 (68)	14	3333	**27**	6750	**348**	45853	229	22491
52/ 70-90	5 (5)	535 (46)	**0**	1	**10**	767	489	26643	**509**	36641
52/ > 90	186 (97)	226 (129)	139	20642	**176**	28472	255	67189	**208**	39675

20: pure beech; 25: predominantly beech; 52: predominantly spruce; 50: pure spruce;
Beech: RMSE$_7$: 36 m^3; RMSE$_{15}$: 46 m^3; Spruce: RMSE$_7$: 73 m^3; RMSE$_{15}$: 101 m^3;
$\hat{\mu}$: mean volume [m^3]; half confidence interval length in parentheses; $\hat{\sigma}^2$: variance; $_B$: beech
$_S$: spruce; $_{7,15}$: Spectral auxiliary information based upon 7 or 15 ISODATA clusters

prediction of stands with low beech volume is greatly reduced by both predictors (cf. Table 17).

There is systematic under-prediction of stands with high volume and over-prediction of stands with low volume for spruce when nearest neighbors are assigned out of all BI field plots (cf. Table 17). Compared to the approach where the search is confined to field plots comprising spruce overall efficiency is improved as the RMSE of MSN$_7$ decreases to 73 m^3, and the one of MSN$_{15}$ to 102 m^3. However, the RMSEs of a prediction excluding the "pure" beech stands with low spruce volume are larger for for MSN$_{15}$ with 114 m^3, but are only 98 m^3 for MSN$_7$.

The over-prediction of stands with low spruce volume is greatly reduced; the predictions for the oldest "pure" beech stands is even within the confidence interval of the surveyed volume (cf. Table 17). This reduction is also reflected by the BIAS, which is now negative for both approaches and of magnitude -33 m^3 for MSN$_7$ and -53 m^3 for MSN$_{15}$.

The mixed beech stand of type 25, age between 80 and 120 years, is one stand for which the under-prediction of beech volume by both MSN$_7$ and MSN$_{15}$ is of similar large magnitude (cf. Table 17). The distribution of the per hectare volume of the evaluation field plots of this stand is displayed in Fig. 16 (a). The stand comprises one plot where surveyed beech volume per hectare is much larger compared to the others. Such a plot, with a per hectare volume of 661 m^3, is hard to predict as there is only 1 (!) BI field plot with a similar high volume (cf Fig. 15). This BI plot has a per hectare volume of 680 m^3.

The second to largest beech BI plot volume is 575 m^3. So, under-prediction of this one evaluation plot is most likely. If this plot is excluded then the mean surveyed stand volume decreases down to 176 m^3 and its under-prediction decreases from 35% to 26% for MSN$_7$, and from 37% down to 28% for MSN$_{15}$.

It should be noted that even though these under-predictions seem large the MSN$_7$ predictions are always within the confidence interval of the surveyed stand volume.

(a) Example for beech
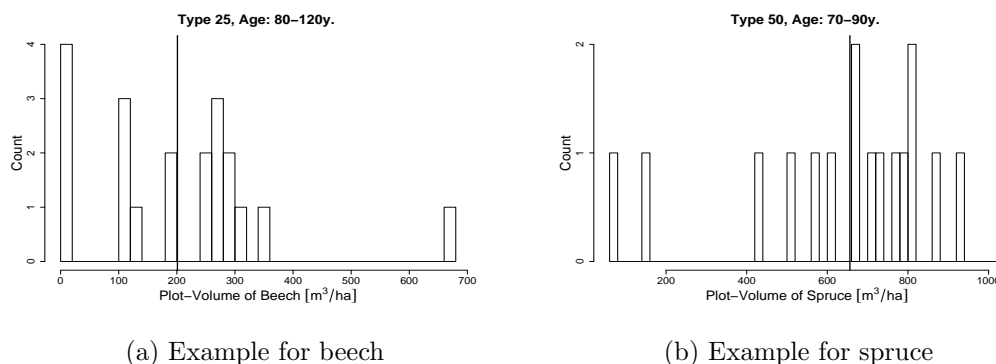
(b) Example for spruce

Figure 16. Volume variability within evaluation stands
vertical line runs at the estimated mean

The distribution of the per hectare volume of the evaluation field plots of the spruce

stand with the largest surveyed volume is displayed in Fig. 16 (b). The very large stand volume of 655 m^3 is under-predicted by 22% (MSN$_7$) and 31% (MSN$_{15}$) (cf. Table 17). The stand volume is hard to predict because only 6.7% (ca. 46 plots) of all BI plots have larger volume per hectare than 655 m^3 and less than 1% (ca. 7 plots) have a volume larger than 790 m^3 (cf. Fig. 15).

When nearest neighbors are assigned out of BI plots comprising beech the MSN predictor employing information from 15 ISODATA classes outperforms the MSN predictor relying on 7 classes. This has been expected since the 7 classes are thought of as rather coarse. If neighbors are assigned out of all BI plots then MSN$_7$ performs better in terms of lower RMSE for beech than MSN$_{15}$.
In case of spruce, the latter is less efficient in both cases, but the RMSE is now 140% of the one of MSN$_7$, compared to 112% when neighbors are confined to BI field plots comprising spruce.

In order to assess the influence of indicator variables derived from the spectral imagery and from the height imagery on the prediction a MSN prediction is run excluding indicator variables based on height information. In addition, instead of just using the nearest neighbor as predictor, further predictors computed as distance-weighted means of k-MSN, $k = 1, \ldots, 7$, are evaluated, where the distance is given by Eq. (8), p. 22.

Anttila (2002) claims that using more than one neighbor decreases the variance of the predictor and the use of weighted means decreases bias without greatly increasing the variance. The design variable h$_G$ is not included in the canonical correlation analysis when the indicator variables derived from the nDSM are excluded. Results are shown in Fig. 17 (a) for MSN$_7$ and in Fig. 17 (b) for MSN$_{15}$.

It has been expected that including height information will result into lower RMSEs, as the estimation variances are lower for the corresponding approaches in Grünenplan (cf. Table 2 and Table 3, pp. 27). However, this is only the case when the spectral information is based on 7 ISODATA clusters. The RMSE of the MSN$_7$ approach using just one neighbor and no height information is more than one and a half times larger than the one including height information (59 m^3 to 36 m^2) (cf. Fig. 17 (a)). For spruce, excluding the height information increased the RMSE of MSN$_7$ by 10% up to 80 m^3.

(a) MSN$_7$ (b) MSN$_{15}$
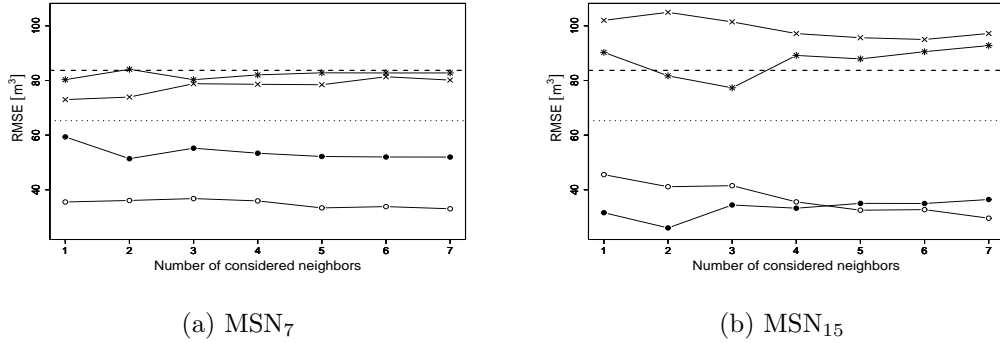
Figure 17.   Root mean square error over k-nearest neighbors

×: spruce (including height information) ; ∗: spruce (excluding height information);
○: beech (including height information) ; •: beech (excluding height information)
$_{7,15}$: Spectral auxiliary information based upon 7 or 15 ISODATA clusters;
dashed line: RMSE of beech by Null-variant; dotted line: RMSE of spruce by Null-variant

For the MSN$_{15}$ approach, the RMSEs of spruce are always lower when the height information is excluded, although canonical correlations are lower compared to the approach where height information is included. Using just one neighbor, the RMSE is 89% of the one including height information. However, the canonical correlations are 0.80, 0.48, and 0.34 compared to 0.82, 0.52, 0.41, and 0.36 when height information is included.

The RMSE for beech, excluding height information, has a minimum for $k=2$ with 26 m^3, the lowest RMSE observed over both species and all methods. There is an overall decreasing tendency for increasing k using MSN$_{15}$ when height information is included (cf. Fig. 17 (a)). In case of spruce, applying k-MSN yielded the lowest RMSE for $k=3$ with 77 m^3 when height information is exlcuded. If the latter is included no clear minimum is discernible (cf. Fig. 17 (b)). Overall, there is more variation in the RMSE of MSN$_{15}$ than in the RMSE of MSN$_7$.

The is no general tendency discernible that with increasing k the BIAS decreases. Excluding height information the BIAS of MSN$_7$ predicting beech decreases from -23 m^3, $k=1$, to -11 m^3 , $k=7$. However, for spruce there is an increase from -33 m^3, $k=1$ up to -38 m^3, $k=7$. For the corresponding MSN$_{15}$, the respective BIASes are (beech) -7 m^3, $k=1$, and -6 m^3, $k=7$, and (spruce) -53 m^3, $k=1$, and -53 m^3, $k=7$.

On average the variances of the MSN predictor decreas with increasing k. For $k=1$ obtained variances can be large but are within the range of the variances of the field survey (cf. Table 9 and 17). Furthermore, even with a variance as large as 50990 the large number of prediction plots, usually $\geq$25, lead to confidence interval lengths relatively small compared to the ones of the field survey.

The final "most similar neighbor" predictor is chosen as MSN_7, $k=1$, including height information. For spruce, this predictor yields the lowest RMSE. The root mean square error of beech is 138% of the minimum obtained by MSN_{15}, $k=3$. However, a general recommendation on k employed by MSN_{15} is difficult as different k lead to a minimum RMSE of spruce and beech. Using different k for different species is considered impractical, also, because the small number of evaluation stands do not warrant a general recommendation. The RMSEs of MSN_7 seem more stable and the average RMSE over the two species is only 0.5 m^3 larger than the "global" minimum achieved by MSN_{15}, $k=2$, excluding height information.

The final predictor yields predictions of spruce stand volume five times inside the confidence interval of the surveyed volume. There is no systematic deviation in that predictions inside the confidence interval occurs for stands only with low, medium, or high volume (cf. Table 17). The Null-variant predicts stand volumes inside the confidence interval of the surveyed volume in 9 out of the 12 stands (cf. Table 10), yet the RMSE of MSN_7 is only 112% of the Null-variant.

For beech, MSN_7 produces predictions 8 times inside the confidence interval of the surveyed volume compared to 3 times of the Null-variant (cf. Table 17 and 10). This different prediction quality is also reflected by the RMSE of the MSN_7 predictor, which is only 43% of the RMSE of the Null-variant.

The number of 7 clusters ISODATA classifies the spectral information into seems well adapted for use by the MSN predictor applied to the given data. Predictions proved to be more efficient than approaches based on 20 or 15 clusters. An additional model run based on 11 clusters produced results similar to MSN_{15} and was thus less efficient than the final predictor employing auxiliary information derived from 7 ISODATA clusters.

6   Discussion

With the analyzed data, **ordinary kriging** is not considered an appropriate prediction tool for reasons outlined in section 5.3, p. 54. It is less flexible than universal kriging or most similar neighbor prediction. The need of an external trend function is its major disadvantage. Linking to the external trend function yield table is problematic. In addition, using the yield table is only appropriate as long as there is a pronounced age structure with only a few age classes. The transition of rotation to continuous cover forest management systems will take years, if not decades, but still, the influence of age structure will vanish gradually over time. So, in the long run, ordinary kriging with external trend function cannot be applied because of the lack of an appropriate external trend function. Furthermore, based on the same source of auxiliary information, universal kriging produced similar results in the Black Forest (cf. Table 6, p. 41).

There are **two further** geostatistical techniques one might think of as appropriate in the context of regionalization of forest inventories: **cokriging** and **double kriging**. Both methods are applied in cases where two sets of variables are measured on a coarse grid and one set of variables has additional measurements on a denser grid. The data sets represent point measurements, that is, even the variables measured on the denser grid do not cover the entire sampling domain, as must be the case in universal kriging.

Cokriging is a multivariate prediction procedure. Suitable data are measurements of two or more variables at the same sampling locations. There is no need of co-occurrence of the sampled variables at every site, on the contrary, cokriging typically performs best when some of the attributes are undersampled (Olea, 1999). The kriging system written in matrix notation, using covariances instead of semi-variances, is almost identical to the kriging system of ordinary kriging. The difference is that elements of the matrices are themselves matrices (Myers, 1982).

One application of cokriging in forest inventories using remotely sensed data can be found in Blodgett et al. (2000). They use Landsat TM images in the derivation of auxiliary variables. The 30 m× 30 m ground resolution of such data renders them unsuitable for application to the problem investigated in this study. Proposed methods

of Blodgett et al. are regression with cokriging of residuals, cokriging, and stochastic simulation. They do not mention if there is spatial structure in the residuals for then a simple ordinary least square regression would not be appropriate. Instead, generalized least squares needs to be applied, but a circularity problem would then arise as outlined in section 4.3. If, on the other hand, there is no spatial structure in the residuals then cokriging would not be appropriate. Blodgett et al. do not discuss these issues, nor do they provide results in terms of precision of prediction.

Cokriging has not been considered because of several problems. The structural analysis required by cokriging is much more demanding than that of ordinary kriging. Besides n covariances required by ordinary kriging, the user must model additional $n(n-1)$ cross-covariances, where n represents the number of regionalized variables. The covariance matrices must be positive definite. Individual selection of permissible models for all covariances and cross-covariances is not a sufficient condition to produce a positive definite covariance matrix (Olea, 1999). Therefore, cokriging demands more user interaction during model fitting. A further disadvantage arises from the encountered variation in the data. The resulting large nuggets do not justify the additional modeling input.

The theoretical foundation of double kriging was laid down by Mandallaz (1996). He developed double kriging as a regionalization tool in the context of double or two phase sampling designs applied in forest inventories, as for example in Lower Saxony.

With the data set-up as described in section 3.1, double kriging works as follows: classify the phase 1 plots into $L=8$ strata; in phase 2, measure a subset of these phase 1 plots on the ground. Compute the L means per stratum of the variable to be regionalized. Use these means at the phase 2 locations to compute the difference of stratum mean to surveyed value of the variable under investigation. So far, the approach is identical to kriging with varying means (cf. section 4.2.2). They differ with respect to the prediction.

In double kriging, the residuals as well as the strata means are considered a realization of a random function. A prediction at an arbitrary location s_0 with unknown stratum information is then carried out as the sum of the prediction of the residuals and the stratum mean for that location s_0. A spatial structure analysis is required for the

trend function given by the strata means, as well as for the residuals to the trend function.

An early application of double kriging is found in Fouquet and Mandallaz (1993). Auxiliary information was derived from analogue airborne images. The phase 2 plots were measured on plots with 10 m radius aligned in a systematic cluster design. Model validation was carried out using a fully inventoried area of 17 hectares.

Variables regionalized were number of stems per unit area and basal area per unit area over all species. For the entire validation area they obtained good results, but if the prediction was confined to 1 ha squares deviations of more than 62% occurred in the prediction of the number of stems.

The data analyzed by Fouquet and Mandallaz are much more detailed than the data provided by the forest service of Lower Saxony. In addition to the four age classes and two species groups surveyed during phase 1 in Lower Saxony, 2 levels for stocking density are recorded. Also, the spatial grids are denser; one phase 1 plot represents 0.9 ha and a phase 2 plot 3.6 ha. The latter are sampled as clusters where one cluster comprises 5 satellites. This set-up reduces the variation in the data, meaning lower process variances and most likely lower nuggets than obtained by a sampling design as applied in Lower Saxony.

The less detailed data provided by the inventory in Lower Saxony explain why a spatial structure analysis of the residuals to the trend function strata means failed (cf. section 5.1 and Table 4, p. 34). The spatial structures of the trend function strata means have large nuggets greater than 13000, process variances around 20000, and a range of around 500 m (not shown). Owing to this variability in the spatial structures double kriging is regarded not appropriate for the given data.

All introduced methods share the need of **auxiliary information**. The two principal sources considered in this study are the field assessment and digital CIR images. The cost of the preparation of auxiliary information based on the latter source are less than € 0.40 ha^{-1}, excluding hard and software. The employed software OrthoEngine is one of the cheapest softwares providing bundle adjustment and automatic digital elevation model extraction, and costs around USD 9,000.- (Brostuen et al., 2001). OrthoEngine

runs on the most common operating systems and is mainly developed for personal computers. Assuming that a forest service is inventorying at least 20,000 hectares per year, software costs would be around € 0.50 ha^{-1} the first year and less in the following years, depending on the chosen support service. With the assumed area inventoried annually, the one time costs for a decent PC are around € 0.15 ha^{-1}. So the total cost of the auxiliary information derived from the digital images employed in this study add up to around € 1.05 ha^{-1}.

The field assessment of the forest service in Lower Saxony costs around € 15 ha^{-1}, although no measurements are taken. It serves multiple purposes, not only the provision of quantitative estimates of stands characteristics. In addition, new geometries delineated on the CIR images are counter-checked on the ground. The field assessment also serves controlling purposes (Beck, 1999).

The difficult economical situation of the forest services of all federal states in conjunction with the introduction of continuous cover forestry systems led to the abandonment of the classical forest management planning surveys. The forest services are keen on reducing inventory related expenses, and initial sampling designs are subject to further cost cutting optimizations (Böckmann et al., 1998). Against this background this study has emphasized on methods that maximize the information extraction based on existing auxiliary data, that is the field assessment, or on methods that, at lower cost, produce similar results. Methods requiring monetary input additional to the expenses of the field assessment are not supported, instead, alternative methods must show a potential to replace or reduce the expensive field assessment.

There is usually a time lag of around one year between the survey of the field plots and the release of the new forest district map. Consequently, predictions relying on auxiliary information of the field assessment will also have a time lag of one year, whereas predictions relying on the digital images can be made once the field survey is completed.

The **Null-variant** is based on data of the field assessment. The auxiliary variables stocking density and mixture proportion are obtained by educated guesses, no measurements are taken. The Null-variant predictions of beech volume are only three times within the confidence interval of the surveyed volume (cf. Table 10, p. 57), always when the surveyed volume is less than 10 m^3 per hectare. The RMSE of 84 m^3

is the largest of all evaluated methods applied to beech.

The opposite is true for spruce. The RMSE of the Null-variant of 44 m³ is the smallest of all evaluated methods. In 9 out of the 12 cases are predictions within the confidence interval of the field survey. The exceptions are the middle aged pure spruce stand with the largest surveyed volume and the two oldest mixed stands where spruce prevails.

Usually, the canopy cover gets less regular with increasing age, rendering stocking density more difficult to estimate. Also, mixture proportions are more difficult to assess when species are more evenly mixed. These two factors help to explain the largest observed over-prediction of the spruce volume by 58% in the stand type 52, older than 90 years (cf. Table 10 and 9, p. 57 and p. 53).

The difficulties in assessing mixture types also help to explain the low performance of the Null-variant for beech. Beech grows mainly in mixed stands, 67% of all inventoried BI-plots comprising beech comprise at least one other species.

The **synthetic estimator** is a simple version of the Null-variant. It does not depend on the field assessment but employs information of the phase 1 plots. Roughly 4% of all stands are not covered by a phase 1 plot (Böckmann, personal communication). Without input in addition to the phase 1 plots, it cannot be assumed that forest districts can be covered entirely by this predictor.

For beech, the synthetic estimator yields a lower RMSE than the Null-variant, but is less efficient than the Null-variant in spruce stands. In theory, the Null-variant should always be more efficient unless the employed additional information, that is, stocking density and mixture proportion, is imprecise. The better performance of the synthetic estimator in beech stands supports the assumption that stocking densities and mixture proportions are difficult to assess in mixed stands without measurements.

The quality of predictions obtained from **universal kriging** is affected by several factors like the quality of the employed auxiliary information, the plot alignment, and the plot size.

The concentric layout of the sample plots reduces their suitability for kriging purposes. The need to extrapolate characteristics of the smaller circle diminishes the spatial

autocorrelation between field plots. In general, it holds that the smaller the plot size the smaller the mean surveyed absolute plot volume is. Smaller mean absolute plot volumes imply smaller variances. If the plot size is reduced by half then in theory the variance per unit area remains constant as can be seen from $Var(abx) = Var(a \frac{b'}{a}x) = Var(b'x)$, where a is the plot inflation factor (here $\frac{1}{2}$), b and b' are constant multipliers to obtain a variance per unit area ($\frac{unit\ area}{plot\ size}$), and x the plot volume (a random variable).
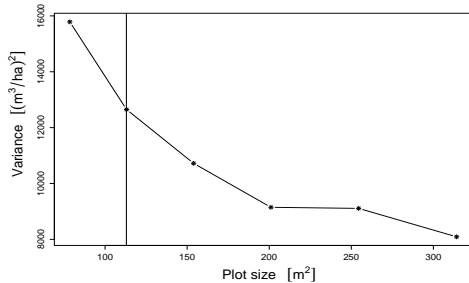

Figure 18.
Plot volume per unit area variability
vertical line runs at 113 m^2

However, empirical studies even show an increase of per unit area variance. This can be attributed to the spatial inhomogeneity of tree growth, an additional source of variation of plot volume per unit area. The resulting allometric relationship between plot size and variance of plot characteristics per unit area was first described by Smith (1938). An illustration based on a subset over all species of the evaluation stands is shown in Fig. 18.

The spatial alignment of the terrestrial sample plots will also affect their suitability for kriging purposes. The two-phase sampling design applied in Lower Saxony had been optimized for precise estimation of the mature stands at the entire forest district level. The use of the sampling grid for regionalization purposes was not taken into account (Saborowski and Dahm, 1996). Optimal sampling designs for regionalization purposes have been developed (McBratney et al., 1981a; McBratney et al., 1981b; Olea, 1984). These designs usually align plots on a regular systematic grid. Such an alignment would run contrary to the optimizations developed for the sampling design applied in Lower Saxony because any stand type would have a sampling density proportional to its total area.

Employing auxiliary information from the field assessment requires a join of these data to the surveyed field plots. The data sets do not necessarily refer to the same spatial locations, because a forest district will be re-delineated in the course of the forest management planning (cf. section 5.3.2). This problem in the data join adds variation to the prediction because for some stands characteristics will be based on averages comprising the main stand and the structural elements. Furthermore, variation within

stands expressed by auxiliary management units cannot be captured because these units are not spatially explicit.

Nevertheless, with the auxiliary information provided by the field assessment universal kriging performed surprisingly well. For beech, this approach yields the second lowest RMSE of 44 m^3, outperforming the other universal kriging approaches using auxiliary information obtained from the images. It is also more efficient than the Null-variant, which is based on the same auxiliary data source.

Results for spruce stands are worse. The RMSE of 132 m^3 is twice as large as the RMSE of the Null-variant although comparable to the UK_7 and UK_{15} approaches, but inferior to the UK_{20} approach (cf. Table 10 and 14, p. 57 and 65).

The more logical trend functions based on products (cf. section 5.3.2, p. 58) are inferior to the simple approach based on a linear combination of SI, stocking density, and age, each multiplied by mixture proportion. This can be attributed to the variability within the employed auxiliary variables. For example, in a multiplicative model if a stocking density enters the prediction with a value of 0.6 instead of 0.3 then the predicted volume will be doubled. With the obtained coefficients in a linear approach, the increase of the volume estimate will be less pronounced.

For the two investigated species there is an inverse relationship between the efficiency of the Null-variant and universal kriging based on field assessment information. If the Null-variant performs well then universal kriging produces bad results and vice versa. Based on this observed relationship it is recommended to regionalize beech stand volumes via universal kriging and spruce volumes via the Null-variant if the auxiliary data are derived from the field assessment.

The prediction quality of universal kriging based on auxiliary information derived from the digital images is affected by additional factors.

The quality of the auxiliary information is influenced by positional accuracies of the field plots and the spatial rectification of the images. Holopainen and Wang (1998) report mean positional errors of rectification for the ground-control points between 2.9 m and 4.8 m. Their images were rectified using 10 to 12 ground points measured in the field by GPS (global positioning system). In the present study the images are rectified simultaneously via bundle adjustment. The achieved root mean square error

of all ground-control points is below 2 m in easting and northing directions for both the Grünenplan and the Solling images, and is considered good.

Assuming that the phase 2 BI plots had been correctly located, these positional displacements imply that some of the field plots are only partly covered by the plots on the images. Unfortunately, the field plots are not always situated where they are believed to be. Their locations were measured by GPS but positioning errors of GPS measurements under closed canopies are a well known problem. So, plots on the images will match the terrestrial plots on average only.

Changes in illumination conditions caused, for example, by slopes or radial displacement on the images, are accounted for by mosaicing applied to the images at Grünenplan and band rationing applied at the Solling. A more sophisticated procedure of radiometric calibration based on regression analysis is proposed by Holopainen and Wang (1998). However, their approach is based on supervised procedures and the authors admit that they did not solve the calibration problem in practical forest inventory.

The techniques used to derive auxiliary information from the classified images are rather simple and easy to grasp by practitioners. The importance of the last point should not be underestimated because the transparency of a technique furthers its acceptance and hence implementation by practitioners.

Results of universal kriging applied to beech stand volumes are quite good. The RMSE of this approach is always less than 32% of the one of the Null-variant. The UK_{15} predictor yields the third lowest RMSE, only slightly larger than the one of universal kriging based on existing information (cf. Table 13 and 10). Results for spruce are less satisfactory, although the RMSE of UK_{20} with 115 m^3 is the fourth lowest out of eleven methods (cf. Table 14). No general recommendation on the number of classes ISODATA should classify can be made because UK_{20} produced the largest RMSE for beech within this group of universal kriging predictors (cf. Table 13 and 14).

Hoffmann (2001) found a good linear relationship between growing space and total surveyed plot volume. She used CIR images with a nominal scale of 1:5500, scanned at 21μm. The DSM was derived with a spatial resolution of 1 m$\times$ 1 m. The studied area was rather small, covering only 109 hectares, and is dominated by beech.

Her good results could not be reproduced with the given data, although growing space had the highest correlation with volume among all variables derived directly from the imagery. Explanations are the smaller scale of the images employed in this study and the approach to model species rather than total volume.

The lower the scale of the images the larger the errors of the derived DSM will be, ceteris paribus. Hoffmann (2001) reports that a height error of 2 m already translated into an under- or over-estimation of total plot volume by 10%. The 0.85 height quantile as a proxy of the h_G seems less sensitive to height errors and to the spatial resolution of the nDSM (Nuske, 2002). This serves as one explanation why the trend function employing this quantile proved to be more efficient than the one employing growing space.

In order to predict volume per species the latter must be identified. Tree species recognition from digital imagery is an active research area. Gong et al. (1999) used a feed-forward back-propagation neural network algorithm and consistently obtained better than 80% accuracies of discriminations of six conifer species in the Sierria Nevada mountains in California. However, the results are based on hyperspectral data, containing much more spectral information than the CIR images regularly employed in forest inventories.

Nevertheless, their results point out the potential images obtained by digital cameras have. These images will contain more spectral information than comparable digitized analogue CIR images because spectral information is lost during the digitizing process.

With digitized analogue CIR images classification of plots into habitat types seems a more promising approach. Holopainen and Wang (1998) employed digitized CIR images at nominal scales of 1:5000, 1:30000, and 1:50000 for the classification of 11 habitat classes in Finnland. Using a supervised regression radiometric calibration, they obtained stratification results better than 80%. Increasing the image scale did not remarkably increase the classfication accuracy as one would have assumed.

Their regression based calibration needs ground-truth measurements as a priori information. Such information is usually not provided by forest inventories. Furthermore, the area covered was 300 hectares only, leaving open the question of performance when multiple images are employed.

Effects of the different parameters derived from digital images on modeling spruce plot volume are illustrated in Fig. 19 (a) through (d). In Fig. 19 (a) plot volume is modeled by a product of stocking density, h_G, and mixture proportion, all computed from the field plot data. Mixture proportions are derived following the approach used in the Null-variant (cf. section 4.6, p.23).
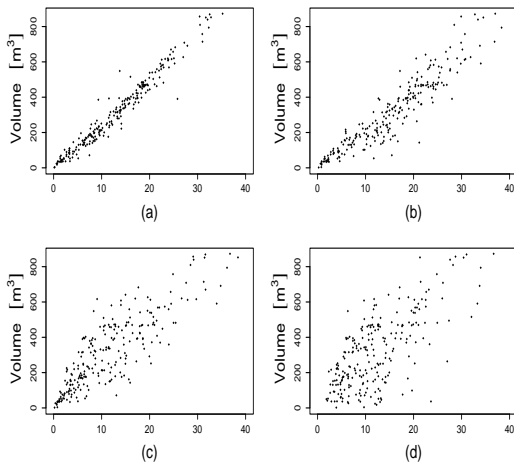

Figure 19.   Modeling of plot-volume

(a) Regressor as product of B°, h_G, and mixture; all from field plots
(b) Regressor as product of B° (field plots), 0.85 height quantile, and mixture (field plots)
(c) Regressor as product of B° (imagery), 0.85 height quantile, and mixture (field plots)
(d) Regressor as product of B° (imagery), 0.85 height quantile, and mixture (imagery)

There is a good linear relationship between volume and regressor. The relationship is still quite good when the h_G is substituted by the 0.85 height quantile, shown in Fig. 19 (b). The relationship gets worse when measured stocking density is substituted by an estimate modeled by mixture analysis (cf. Fig. 19 (c)). Finally, if mixture proportions are also substituted by an estimate from the imagery the relationship breaks down (cf. Fig. 19 (d)). The problem of capturing mixtures are another explanation of the severe over-prediction of the low spruce volume in "pure" beech stands (cf. Table 14).

The trend function universal kriging is based upon is a function in three variables, stocking density, height, and species proportion. These variables are modeled from the initial information contained in the spectral image and in the nDSM. Improving the prediction quality of the variables requires the improvement of the initial information and the models that transform the initial information into the said three variables.

Stocking density and height can already be modeled satisfactorily from the available data. A better universal kriging based prediction of volume per species hinges on a better model identifying species (cf. Fig. 19). The weak spatial autocorrelations encountered in the final universal kriging models (cf. Table 12, p.63), may not justify the application of this computational demanding technique. However, it is recom-

mended to employ universal kriging instead of a simple regression approach because the former is able to capture possible stronger spatial-autocorrelations (cf. Fig. 11, p. 38).

The negative effects of the concentric field plot layout and positional accuracies on the quality of the auxiliary information discussed in the context of universal kriging prediction are also valid for **most similar neighbor** prediction. The two methods differ with respect to how they are influenced by the alignment of field plots. The quality of the MSN predictor is influenced by the alignment of the field plots in the feature space only, that is, how the design attributes are distributed.

The most similar neighbor predictor makes direct use of the initial variables obtained from the classified spectral image and of the variables derived from the nDSM. Improvements in the quality of the initial auxiliary information discussed in the context of universal kriging will also improve the prediction quality of the MSN predictor. However, no transformation of the inital variables and no parametric models need to be specified as the canonical correlation analysis weights the contribution of the initial variables to the canonical variables by their respective covariances. The only additional modeling input necessary is a check of multicollinearity.

Let X be an $n \times m$ matrix of indicator or design attributes. Multicollinearity can be detected by examining the eigenvalues of $X'X$, where prime denotes transposition, by their condition number $\kappa = \sqrt{\frac{\lambda_1}{\lambda_m}}$, and condition indices $\kappa_j = \sqrt{\frac{\lambda_1}{\lambda_j}}$, $j = 1, \ldots, m$, where the λ_j are the eigenvalues in descending order. Condition numbers less than 100 indicate no serious problems with multicollinearity. κ in $[100, 1000]$ implies moderate to strong multicollinearity, and if κ exceeds 1000, severe multicollinearity is indicated (Montgomery et al., 2001). Severe multicollinearity can be caused by only one variable. Including growing space instead of the 0.85 height quantile in the approach based on 15 ISODATA clusters resulted in 33 out of 57 condition indices being larger than 1000 and a condition number of 26741. When the 0.85 height quantile was used the condition indices were all below 650.

The condition indices should also be analyzed because the canonical correlation can be misleading due to over-fit of the employed model. The problem of over-fit is also well known for parametric regression approaches. Consequences observed during model

evaluation are models with larger canonical correlations and larger root mean square errors than those of more stable competitors.

In addition to the check of multicollinearity the correlation matrix of X should be analyzed. Problems are indicated by large pairwise correlations and inconsistent signs of corresponding canonical coefficients; for example, if two variables are positively correlated but their canonical coefficients are of opposite sign. Correlation matrices of both the indicator and the design attributes need to be checked as the canonical coefficients of the latter influence the canonical coefficients of the former (cf. Eq. 36, p. 117).

The required modeling input to MSN prediction is still less than for a parametric regression approach. Apart from the sometimes time consuming variable selection, regression approaches also demand checks on the residuals in addition to checking for multicollinearity. Furthermore, checks of large condition indices and inconsistencies of pairwise correlations and corresponding canonical correlation coefficients can be handled by an application software and thus shielded from the less statistically inclined practitioner.

The most similar neighbor predictor performed better for both species when nearest neighbors are assigned out of all BI plots, not only the BI plots comprising the respective species. In the latter approach the lowest RMSE for beech is 48 m^3 by MSN$_{15}$, 133% of the final MSN$_7$ prediction looking for neighbors within all BI plots (cf. Table 15 and 17). The improvement for spruce is larger as the lowest RMSE by MSN$_7$ based on the BI plots comprising spruce is 170% of the final MSN$_7$ prediciton (cf. Table 16, and 17). The improvements can be attributed to the reduced over-prediction of stands with low volume ($\leq$ 50 m^3) of the respective species.

The following discussion refers to the MSN predictor assigning neighbors out of all BI plots. The final MSN$_7$ prediction yields the lowest RMSE of 36 m^3 over all investigated approaches applied to beech. The prediction systematically under-predicts beech stand volume but in 7 out of 10 stands where beech occurs predictions are inside the confidence interval of the surveyed volume.

The MSN_7 predictor yields the second lowest RMSE for spruce with 73 m^3 which is only 112% of the best prediction produced by the Null-variant (cf. Table 17 and 10). Predictions of MSN_7 are inside the confidence interval of the surveyed volume in 5 out of the 12 stands whereas the Null-variant predicts 9 stands inside the confidence interval. There is a systematic over-prediction by MSN_7 of stands with low spruce volume and an under-prediction of stands with high spruce volume (cf. Table 17).

If a regionalization of a forest enterprise should be carried out and digitized CIR images are available then, based on the above results, it is recommended to use a MSN_7 predictor employing information from both the spectral image and the nDSM.

The MSN predictor distinguishes surprisingly well between spruce and beech. Predictions by MSN_7 of stands with low beech volume are always within the confidence interval of the surveyed volume (cf. Table 17). For spruce, the "pure" beech stands with low spruce volume are captured rather well by the MSN_7 predictor, too. One prediction is inside the confidence interval and the Null-variant is the only other predictor yielding closer predictions for these stands (cf. Table 17 and 10, p. 71 and 57).

The two pure spruce stands without beech get volume for the latter species predicted by MSN_7, however, the predictions are moderately low with 13 m^3 and 14 m^3. Zero volume predictions are possible and the MSN_{15} predictor correctly assigns such a volume to one stand but predicts 53 m^3 for the other stand (cf. Table 17).

Good separation of deciduous and coniferous species is also reported by Haara and Nevalainen (2002). Their approach used a robust segmentation algorithm with subsequent classification into 7 classes, one for bare ground, pine, and deciduous trees, and four defoliation classes of spruce. Material employed were CIR images of nominal scale of 1:5000 and two stands, 0.6 ha and 0.4 ha in size. Haara and Nevalainen stress the importance of the quality of training data in their supervised approach, rendering this approach unsuitable for applications investigated in this study.

Recent applications of most similar neighbor techniques using canonical correlations and digitized CIR images as data source for auxiliary information are found in Muinonen et al. (2001) and Anttila (2002). In both studies predictions of total stand volume was based on k-nearest neighbors with respect to stands, not field plots

as is the case in the present study. Choosing entire stands can be explained by the small stand size, 1.87 and 1.9 ha respectively, and their homogeneity.

Muinonen et al. investigated 59 stands with a total area of 125 hectare covered by just one CIR image with scale 1:30,000. Total stand volume was the only design variable used. The indicator variables represented mean stand values, for example, mean brightness value in the red channel. Empirical semivariograms were calculated from the pixel values in the red band. The semivariogram values of the first 25 lags, lag distance being 1 m, were used as auxiliary information. The variogram information improved the volume predictions. However, the authors admit that if forest stands are large the heterogeneity within the stands increases, resulting in a possible decrease of prediction quality of their approach.

Their most efficient approach yielded a RMSE of 17.9% and a bias of -0.6%. The relative values are obtained by dividing the RMSE and BIAS by the (unweighted) mean volume of all stands. The RMSE is based on cross-validation. In the present study the RMSE is 27.7% for beech and 26.3% for spruce. The biases are -17.6% and -8%, respectively. The RMSEs obtained in this study are similar to the 29.9% reported by Anttila, which are also based on cross-validation. It should be noted that Muionen et al. and Antilla predicted total stand volume whereas in the present study the harder to predict species volume is investigated.

Antilla used three CIR images with scale 1:20,000 covering 577 field stands. The investigated stands must be rather homogeneous as less than 10 field plots with radius $\leq$ 10 m were sufficient to estimate stand volume reliably. Values of the empirical semivariogram were used as indicator variables, as in the approach by Muinonen et al. Old inventory information was also used as indicator variables. Excluding this information and relying only on information derived from the digital images increased the RMSE to 36.7%. The use of old forest inventory information is considered not operational in the context of the present study because dynamic events like storm breaks cannot be captured.

Both Muinonen et al. and Antilla report better predictions for $k > 1$. The best predictions were obtained for k=7 (Muinonen et al.) and k=15 (Antilla). These results could not be reproduced in the present study. Some predictors produce lower

root mean square errors when more than one neighbor is considered but no general increase in efficiency with increasing k can be observed (cf. Fig. 17).

Assuming that the variables to be predicted have some sort of bell-shaped or even equal distribution most similar neighbors techniques are known for problems of predicting values at the tails of the distribution. Over-prediciton of values at the lower tail and under-prediction of values at the upper-tail of the distribution are likely to occur. This phenomenon can be explained by the lower number of suitable neighbors at the tails. The magnitude of the bias will be inversely proportional to the canonical correlations of the indicator and design attributes, ceteris paribus.

If the bias can be quantified according to the encountered conditions then such quantification could be used to correct it. The bias of the final MSN_7 predictor is plotted over the surveyed stand volume in Fig. 20 (a) for beech and Fig. 20 (b) for spruce. Except for the stands without beech the stand volume of this species is systematically underestimated, but there is no discernible relationship between the bias and the surveyed volume.

(a) beech (b) spruce
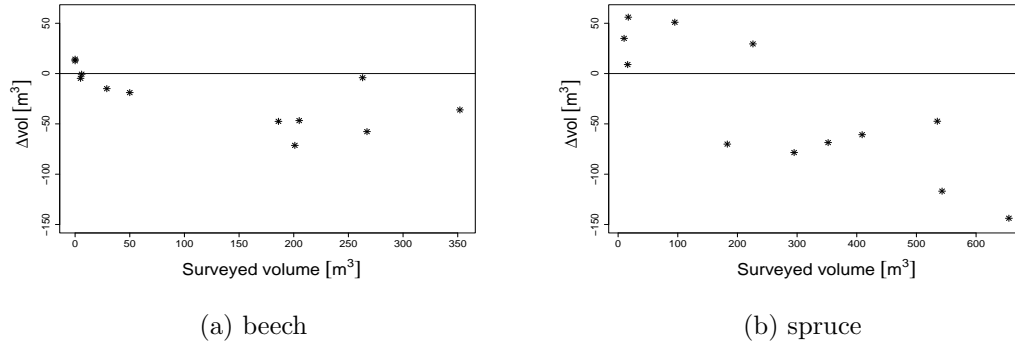
Figure 20.   Bias over surveyed volume of final MSN_7 predictor
Δvol: volume$_{predicted}$-volume$_{surveyed}$

There is a weak negative linear relationship between bias and surveyed spruce stand volume (cf. Fig. 20 (b)). It is conjectured that increasing predictions for both species above certain thresholds will improve overall efficiency. For example, a moderate increase of predictions in [200,300) by 25 m^3 and by 50 m^3 if predicted volume is $\geq$ 300 m^3 is considered conservative.

Validation of the conjecture requires further research comprising more stand types and more observations per stand type. The stand types should also include mixtures of coniferous and deciduous species, for example beech in oak stands or larch in beech stands. Furthermore, the species separation capabilities of MSN should be tested on forest districts comprising douglas fir and spruce because these species are hardest to distinguish on CIR images.

One problem of research on most similar neighbors is the demand this technique imposes on the data. Based on the experience gathered in this study it is recommended to investigate only entire forest districts. Performance of the MSN depends heavily on the number of suitable neighbors available. Regionalizations are applied at the forest district level and restricting model evaluation to smaller areas would distort prediction quality.

Furthermore, the suitability of cross-validation must be questioned because it cannot capture over-fit of models. Cross-validation is well established in geostatistics but taking spruce as an example the cross-validated error variance based on the BI-plots of the most efficient UK_{20} approach is 16530 whereas the cross-validated error variance of MSN_7 is more than double as large with 34041. However, the RMSE of the MSN_7 predictor is only 73% of the RMSE of UK_{20}.

An evaluation should be based on data resembling forest stands, although this would be difficult to accomplish with the BI data. One approach would be to pool BI plots of same stand type and age and use these plots only during model evaluation. One problem is the reduction of plots available for prediction. An evaluation stand usually comprises more than 25 *prediction plots* (cf. Table 15 and 16, p. 68 and 70). There are 373 BI plots comprising beech situated in the 8,000 hectares of the Solling investigated here. Pooling data to represent 5 stands would already significantly reduce the available number of suitable neighbors.

Despite all these obstacles and challenges further research on most similar neighbor techniques seems justified by its good performance. Furthermore, of the methods investigated, MSN is the only multivariate prediction technique. Variables, that are also influenced by the canopy and height structure, can be regionalized by MSN as a byproduct of the species volume regionalization without additional modeling input.

This features qualify MSN as a promising tool to fulfill the information requirements of modern muli-functional forest managment planning, a view supported by Anttila (2002) and Muinonen et al. (2001).

7 Summary

The introduction of continuous cover forestry systems and the difficult economic situation of the forestry sector lead to the replacement of classical surveys by sample based inventories. Affordable sampling densities enable statistically sound inference at the entire district level only. However, there is a demand for reliable information at the spatial scale of management activities: the forest stands. This demand can only be met by a regionalization of the sample plots. Any regionalization needs auxiliary information, provided by the already implemented inventory or by additional sources like digital airborne images.

In the most extensive comparison of techniques, known to the author, four different regionalization approaches have been investigated in this study in three different forest districts. The first two forest districts in Lower Saxony and Baden-Württemberg served a model selection guided by cross-validation. Auxiliary information has been derived from the field assessment as part of the regular inventory and digitized CIR images in Lower Saxony.

The digital images are an additional source of auxiliary information although analogue CIR images are part of the regular inventory in Lower Saxony. Spectral properties of the digital images and height information derived from digital stereo-pairs have been employed as auxiliary information.

A selection of regionalization techniques has been evaluated in the third forest district situated in Lower Saxony. The evaluation is based on the comparison of predicted volume to surveyed volume of 12 intensively sampled forest stands. The evaluation stands represent 4 different types at three different ages.

The intensive sample covers a total of 10 hectares in addition to the phase 2 BI plots located in the evaluation stands. Stands are much more inhomogeneous than expected. The youngest pure spruce stand already needs a minimum of 21 field plots of size 707 m^3 to obtain a half confidence interval length of 10% of the estimated mean at the α=0.05 level. Even setting the half confidence interval length to 20% some of the mixed stands cannot reliably be characterized by a sample but would require a total enumeration.

The evaluation stands were surveyed and predicted without regard of the canopy layer concept applied in Lower Saxony. The layer concept would further increase the variability of species volume between sample plots, rendering a regionalization much more difficult.

The methods investigated focus on one hand on the maximization of information extraction from existing sources. On the other hand focus is laid on approaches that rely on additional resources but produce similar or better results at lower cost. Techniques employing auxiliary information of both existing and additional sources are not supported owing to scarce economical resources.

The forest management planning of the forest service of Baden-Württemberg relies on a 1 phase sampling design. The only source of auxiliary information is field assessment. Regionalization of the sample plots using this auxiliary information proved difficult owing to the different spatial scales these two data sets refer to. A successful regionalization of the high resolution field plots demand auxiliary information of similar high resolution. Such information can be provided by 2 phase sampling designs. The acquisition of CIR images costs less than € 0.75 per hectare. 2 phase sampling designs have the potential to obtain results similar to 1 phase designs with less field plots, which are expensive and cost around € 30.-.

The most similar neighbor predictor, MSN, produces the most efficient predictions in terms of RMSE of the 12 evaluation stands over the two principal tree species beech and spruce. MSN performes better when neighbors are sought out of all BI plots, not only the plots comprising the respective species. The auxiliary information employed is derived from the digitized CIR images. The derivation is achieved using the commercial, but widely used, GIS Arc/Info and self-developed programs; is semi-automatic, and can handle multiple images.

The heuristic regionalization approach of the forest service of Lower Saxony, based on a modified synthetic estimator is the second most efficient predictor. Auxiliary information employed is provided by the field assessment and consists of occular assessments of stocking density and mixture proportion. The modified synthetic estimator produced the best predictions for spruce stands but also the worst predictions for beech stands. The investigated beech stands are more mixed than the spruce stands

and an occular assessment of stocking density and mixture proportion seems more difficult in such mixed stands.

The field assessment serves also the delineation of newly established structural elements and has controlling functions. The controlling function can also be fulfilled by visiting not all stands of a forest district but only a sample of stands. A sampling design can assure that all forest stands will be visited over time.

Assuming an area to be inventoried of 20,000 hectares annually, if the field assessment is scaled back by 50%, that is, only every other stand is visited, the cost saving potential of MSN is more than € 120,000. Furthermore, MSN is a multi-variate prediction technique and variables also influenced by the canopy structure can be regionalized simultaneously with species volume. This makes MSN more appropriate for the regionalization of multi-functional forest inventories than the Null-variant.

The structural elements are delineated in the course of the regular forest management planning on the CIR images and checked during the field assessment. Only a subset of the delineated elements are kept. With the current set-up the structural elements are of little use for regionalization purposes because characteristics like mixture proportions of the field assessment cannot unequivocally be linked to these elements. It is recommended either to assess them separately in the field, like the auxiliary management units, or to abandon the concept.

If the concept of structural elements is maintained, a field assessment based on a sample rather than covering the entire forest district, may still be a viable option. The elements not covered by a field check can be predicted by MSN and depending on the predicted value of the element compared to the surrounding stand it can be decided to keep the structural element or to dissolve it.

8   Zusammenfassung

Waldbauliche Zielsetzungen einer modernen Forstwirtschaft sehen eine Überführung von strukturarmen Reinbeständen in strukturreiche, nach Alter und Baumarten gemischte Bestände vor. Die resultierenden Wuchsdynamiken lassen sich nur noch unbefriedigend durch herkömmliche Ertragstafelmodelle abbilden. Daher werden klassische Forsteinrichtungsverfahren zunehmend durch stichprobenbasierte Verfahren ersetzt. Die schwierige wirtschaftliche Lage des forstlichen Sektors diktiert dabei Stichprobendichten, die verlässliche Aussagen nur auf der gesamten Betriebsebene ermöglichen.

Es besteht dabei weiterhin ein Bedarf an Information auf Bestandesebene, der räumlichen Ebene der waldbaulichen Maßnahmenumsetzungen. Dieser Informationsbedarf kann mittels einer Regionalisierung der Stichprobendaten befriedigt werden. Unter Regionalisierung wird die Übertragung der qualitativ hochwertigen Stichprobeninformation, die punktförmig vorliegt, in die Fläche verstanden.

Vier unterschiedliche Regionalisierungsansätze sind im Rahmen dieser Arbeit an drei Forstbetrieben untersucht worden. Der Schwerpunkt der Arbeit liegt auf Methoden, die entweder die Regionalisierung aufgrund von bestehenden Quellen von Hilfsinformation optimieren, oder aber auf Methoden, die auf alternativen Quellen von Hilfsinformation beruhen und zu ähnlichen oder besseren Ergebnissen bei geringeren Kosten führen.

Allen Ansätzen ist gemein, dass sie nahe bzw. ähnliche Nachbarn oder ein gewichtetes Mittel von nahen bzw. ähnlichen Nachbarn zur Vorhersage heranziehen. Auch benötigen alle Verfahren sogenannte Hilfsinformationen zur Vorhersage. Hilfsinformationen sind dabei Variablen, die sich kostengünstig an den Feldplots der Forstinventur sowie an allen Vorhersagepunkten erheben lassen, und die mit der Zielvariable, Volumen je Baumart (Buche und Fichte), in einem funktionalen Zusammenhang stehen müssen.

Quellen für Hilfsinformation sind der Bestandesbegang und Luftbilder. Der Bestandesbegang ähnelt den Aufnahmeverfahren der klassischen Forsteinrichtung, jedoch werden keine Messungen vorgenommen. Alle Angaben beruhen auf Experten-

einschätzung. Falschfarben–Infrarot-Luftbilder werden bereits im Rahmen der zweiphasigen Stichprobeninventur in Niedersachsen eingesetzt. Um eine großflächige und möglichst automatisierte Bearbeitung zu ermöglichen, wurden die Luftbilder digitalisiert. Hilfsvariablen wurden anschliessend mit selbstentwickelten Programmen aus den bearbeiteten digitalen Luftbildern abgeleitet.

Die vier Regionalisierungsansätze lassen sich in zwei Kategorien einteilen, die durch die jeweilige Definition von Nähe charakterisiert sind. Eine Kategorie definiert Nähe bezogen auf Merkmale. Zwei Verfahren dieser Kategorie verwenden einen synthetischen Schätzer. Im einfachsten Fall erfolgt die Vorhersage über die Zuordnung des Zielbestandes zu einem geeigneten Auswertestratum. Vorhergesagte Werte sind dann die für das Auswertestratum geschätzten Mittelwerte.

Das von der Niedersächsischen Forstverwaltung momentan angewandte Regionalisierungsverfahren beruht auf einem modifizierten synthetischen Schätzer und ist in die Untersuchung mit einbezogen. Der zur Vorhersage herangezogene Mittelwert je Stratum wird entsprechend dem Verhältnis von mittlerem Baumartenmischungsanteil und mittlerem Bestockungsgrad im Stratum zu den entsprechenden Werten im Zielbestand korrigiert.

In die gleiche Kategorie "Nähe bezogen auf Merkmale" gehören *nächste-Nachbarn*-Ansätze. Nachbarn werden über multivariate statistische Verfahren wie der kanonischen Korrelationsanalyse oder mittels heuristischer Modellierung bestimmt.

Die zweite Kategorie definiert Nähe durch die gewöhnliche räumliche Nähe. Ein Mass für die Ähnlichkeit ist dann durch die räumliche Selbstähnlichkeit oder räumliche Auto–Korrelation gegeben. Die Vorhersageverfahren sind Methoden der Geostatistik und werden als Kriging bezeichnet. Kriging basiert auf einigen Modellannahmen. Die Kriging-Verfahren können so zum Beispiel nicht auf die gemessenen Baumartenvolumen angewendet werden, da jene durch einen Alterstrend beinflusst werden.

Ein Ansatz der Trendeliminierung unter Anwendung des gewöhnlichen Krigings besteht in der Verwendung einer Trendfunktion, die extern bezüglich der Inventurdaten ist, also nicht aus letzteren geschätzt werden muss. Solch eine externe Trendfunktion ist durch eine Ertragstafel gegeben. Die beobachteten Inventurwerte müssen

auf vollbestockte Werte hochgerechnet und dann über die Ertragsklasse in die Tafel eingehängt werden. Die Differenz des hochgerechneten Volumens zur Ertragstafel wird dann in die Fläche übertragen und in den Zielbständen wieder entsprechend auf Baumartenvolumen zurückgerechnet.

Ein zweiter geostatistischer Ansatz, universelles Kriging, modelliert den Trend über Hilfsvariablen simultan zur Bestimmung der Residuen vom Trend zu den beobachteten Werten. Universelles Kriging ist sehr flexibel, stellt aber auch die höchsten Ansprüche an die Qualität der Hilfsinformation.

Die ersten beiden Forstbetriebe dienten der Modellselektion durch Kreuz-Validierung. In der Kreuz-Validierung wird sukzessive ein Stichprobenpunkt durch alle anderen vorhergesagt. Evaluierungskriterium ist dabei die Varianz des Fehlers, also der Differenz des beobachteten zum vorhergesagten Wert. Die Stichprobenpunkte sind durch starke Heterogenität selbst innerhalb von Abteilungen gekennzeichnet, da die Feldplots in dichte oder offene Teile eines Bestandes fallen können. Solche Punktwerte sind schwierig vorherzusagen, was sich auch in den grossen Fehlervarianzen der Kreuzvalidierung widerspiegelt. Für die Baumarten Fichte und Buche ergaben sich dabei geringere Fehlervarianzen für das niedersächsische Forstamt Grünenplan als für das zweite Forstamt im Schwarzwald.

Das Stichprobenverfahren in Baden-Württemberg ist einphasig und liefert somit keine hochauflösende Hilfsinformation auf Ebene der Feldplots. Zufallsbedingt konnten zeitnahe panchromatische Luftbilder im Maßstab 1:18.500 vom Vermessungsamt bezogen werden. Eine Ableitung von Hilfsinformationen auf Feldplotebene aus diesen Luftbildern scheiterte jedoch an der schlechten Qualität letzterer.

Die Bereitstellung von hochauflösender Hilfsinformation könnte flächendeckend aus digitalisierten CIR Luftbildern erfolgen, die im Rahmen von 2-phasigen Stichprobeninventuren ohnehin eingesetzt werden. Die Kosten für die Verwendung von analogen CIR Luftbildern in 2-phasigen Inventuren belaufen sich auf weniger als € 0,75 je Hektar. Bei gleichen Genauigkeitsansprüchen werden diese Kosten in der Regel durch den geringeren Bedarf an Feldplots, die mit ca. € 30,- teuer sind, mehr als kompensiert.

Universelles Kriging, beide synthetische Schätzer und "Most Similar Neigbor" (MSN),

ein *nächste Nachbarn* Ansatz basierend auf kanonischen Korrelationen, wurden an 12 intensiv beprobten Beständen in einem 8000 ha grossen Teilgebiet des Sollings evaluiert. Die Evaluierungsbestände umfassten vier Typen: Buche rein, Fichte rein, und Mischungen von Buche-Fichte und Fichte-Buche. Innerhalb jeden Types wurde drei verschiedene Altersstufen untersucht.

Die untersuchten Bestände erwiesen sich wesentlich inhomogener als erwartet. Für einen ca. 55 jährigen Fichtenreinbestand konnte eine halbe Vertrauensintervallänge von 10% des geschätzten Mittelwertes beim α=0.05 Level erst mit 21 Feldplots der Grösse 707 m³ (Radius=15 m) erzielt werden. Selbst eine halbe Vertrauensintervallänge von 20% des geschätzten Mittelwertes erfordert in einigen Mischbeständen Stichprobendichten, die eine Vollaufnahme des Bestandes bedeuten. Die Evaluierungsbestände wurden dabei ohne das in Niedersachsen angewandte Prinzip der Bestandesschichten inventarisiert, welches die Variabilität zwischen den Stichprobenplots nochmals erhöhen würde.

Die Trendfunktion im universellen Kriging wurde mit Hilfsvariablen, abgeleitet aus den digitalisierten Luftbildern und aus dem Bestandesbegang, modelliert. In beiden Fällen ergaben sich ähnliche Bestandesvorhersagen mit z.T. hohen Abweichungen für Fichtenbestände aber insgesamt zufriedenstellenden Ergebnissen für Buchenbestände. Das unbefriedigende Ergebnis für Fichten liegt in der extremen Überschätzung von geringen Fichtenvolumen in "reinen" Buchenbeständen begründet.

Der modifizierte synthetische Schätzer der Niedersächsischen Landesforstverwaltung produzierte die beste Vorhersage für Fichtenbestände über alle untersuchten Verfahren. Für Buchenbestände produzierte dieser Schätzer allerdings die schlechteste Vorhersage. Die verwendete Hilfsinformation wird bei diesem Ansatz durch den Bestandesbegang bereitgestellt.

Die Unterlegenheit des modifizierten synthetischen Schätzer gegenüber dem normalen synthetischen Schätzer, selbst bei Verwendung von Hilfsinformation auf Bestandesebene, ist auf die geringe Qualität letzterer zurückzuführen. Buchen sind vermehrt in Mischbeständen anzutreffen, was deren genaue Ansprache erschwert.

Der MSN-Prädiktor produzierte die effizienteste Vorhersage der 12 Evaluierungsbe-

stände über beide Baumarten Buche und Fichte. Die besten Ergebnisse wurden erzielt, wenn nächste Nachbarn aus dem Kollektiv aller Betriebsinventurplots ausgesucht wurden, nicht nur aus dem Kollektiv der Plots, die die jeweilige Baumart enthalten. Eine distanzgewichtete Verwendung von k-nächsten Nachbarn konnte dabei die Vorhersage nicht eindeutig verbessern.

Der "root mean square error" (RMSE), also die Wurzel des mittleren quadrierten Abweichungsquadrates, ist im Falle von Buchenbeständen am geringsten und beträgt weniger als 50% des RMSE des momentan angewandten modifizierten synthethischen Schätzers der Niedersächsischen Forstverwaltung. Im Falle der Fichtenbestände ist der MSN-Prädiktor nur diesem Schätzer unterlegen, allerdings ist der RMSE des MSN-Prädiktor nur um 12% grösser.

Der MSN-Prädiktor verwendet dabei ausschliesslich Hilfsinformation, die aus den digitalisierten CIR-Luftbildern der regulären Forsteinrichtung abgeleitet wurde. Unter der Annahme, dass die jährlich zu inventarisierende Fläche mindestens 20.000 Hektar beträgt, belaufen sich die Kosten für die Aufbereitung der Hilfsinformation aus den schon bestehenden Luftbildern auf ca. € 1,05 je Hektar. Die Kosten je Hektar für den Bestandesbegang liegen dabei bei ca. € 15,-.

Der Bestandesbegang dient nicht nur der qualitativen Beschreibung von Beständen, sondern ist auch ein administratives Kontrollinstrument. Zusätzlich werden die auf den Luftbildern ausgewiesenen waldbaulichen Strukturelemente auf die Richtigkeit ihrer Ansprache überprüft. Die administrative Kontrollfunktion könnte aber auch erfüllt werden, wenn anstatt des gesamten Forstamtes nur eine Stichprobe überprüft würde. Ein geeignetes Stichprobendesign stellt dabei sicher, dass alle Bestände im Laufe der Zeit im Feld aufgesucht werden.

Wird der Bestandesbegang nur noch auf 50% der Fläche ausgeführt, so beträgt das Einsparungspotenzial des MSN-Prädiktors unter der obigen Flächenannahme mehr als € 120.000,- jährlich. Zusätzlich ist der MSN-Prädiktor ein multi-variates Verfahren und eignet sich besser zur Regionalisierung von multi-funktionalen Forstinventuren als der angewandte modifizierte synthetische Schätzer. Variablen, die wie Speziesvolumen ebenfalls durch die Kronendachstruktur beeinflusst werden, können mit Speziesvolumen in einem Arbeitsgang vorhergesagt werden und bedürfen keiner weiteren

Modellierung.

Wird das Konzept der Strukturelemente beibehalten, so bleibt ein stichprobenba-
sierter Bestandesbegang in Kombination mit MSN-basierter Regionalisierung immer
noch eine mögliche Alternative. Die Strukturelemente, die nicht durch den Bestandes-
begang abgedeckt sind, könnten mittels des MSN-Prädiktors überprüft werden.

Bedingt durch die beobachteten systematischen Abweichungen in der Vorhersage von
Beständen mit grossem Volumen wird diesem Verfahren noch Potenzial zur Effizienz-
steigerung eingeräumt. Gelingt es, die systematischen Abweichungen in Abhängigkeit
der vorgefundenen Rahmenbedingungen zu quantifizieren, dann kann die Güte der
Volumenvorhersage nach Baumarten noch weiter gesteigert werden.

References

Akça, A., Dong, P. H., Beisch, T. (1993): "Zweiphasige Stichprobeninventur zur Holzvorrats- u. Zuwachsschätzung". In: Application of Remote Sensing in Forestry pp. 16–25. TU Zvolen

Anttila, P. (2002): Nonparametric estimation of stand volume using spectral and spatial features of aerial photographs. Can. J. For. Res. 32, 1849–1857

Beck, O. A. (1999): Inventur, Planung und Kontrolle im Forstbetrieb. Forst und Holz 54(22), 691–695

Bilodeau, M., Brenner, B. (1999): Theory of Multivarite Statistics. Springer Texts in Statistics

Biondi, F., Myers, D. A., Avery, C. C. (1994): Geostatistically Modelling Stem Size and Increment in Old-Growth Forest. Can. J. For. Res. 24, 1354–1368

Bitter, A., Merrem, M. (1997): Typenorientierte Kontrollstichprobe als Basis für die mittelfristige betriebliche Planung. In: Dt. Verband Forstlicher Forschungsanstalten, Sektion Forstliche Biometrie und Informatik-9 Tagung, TU Dresden

Bitter, A., Merrem, M. (1998): Typenorientierte Kontrollstichprobe. AFZ/Der Wald 23, 1414–1415

Blodgett, C., Jakubauskas, M., Price, M., Martinko, E. (2000): Remote Sensing-based Geostatistical Modeling of Forest Canopy Structure. In: ASPRS 2000 Annual Conference. Washington, D.C.

Böckmann, T., Saborowski, J., Dahm, S., Nagel, J., Spellmann, H. (1998): Inventur, Planung und Kontrolle im Forstbetrieb. Forst und Holz 53(8), 219–226

Brahms, M., Graulich, R. (2000): Vorreiterrolle für eine zukunftsfähige, nachhaltige und umweltgerechte Waldwirtschaft. Forstliche Mitteilungen 3

Brandtberg, T., Walter, F. (1998): Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. Machine Vision and Applications 11, 64–73

Brostuen, D., Sutton, S., Siddiqui, Y. (2001): Software Review: OrthoEngine. Photogrammetric Engineering and Remote Sensing 67(3), 245–247

Carlowitz, H. C. v. (1713): Sylvicultura Oeconomica oder Hausswirthliche Nachsicht und Naturmässige Anweisung zur wilden Baumzucht. Braun, Leipzig

Christensen, R. (1990): The Equivalence of Predictions from Universal Kriging and Intrinsic Random-Function Kriging. Math. Geology 22, 655–664

Cressie, N. (1991): Statistics for Spatial Data. John Wiley & Sons

Cressie, N., Hawkins, M. (1980): Robust Estimation of the Variogram: I. Math. Geology 12(2), 115–125

Cressie, N., Zimmerman, D. L. (1992): On the Stability of the Geostatistical Method. Math. Geology 24(1), 45–49

Dahm, S. (1995): Bundeswaldinventur- Auswertungsmodelle und Vorschläge zur Effektivitätssteigerung. Mitteilungen der Bundesforschungsanstalt für Forst- & Holzwirtschaft, Hamburg

Davis, B. M. (1987): Uses and absues of Cross-Validation in Geostatistics. Math. Geology 19(3), 241–248

Davis, K. P. (1966): Forest Management. Mc Graw-Hill Book Company, New York

Dees, M. (1996): Regressions- und Kleingebietsschätzung bei forstlichen Grossraumin- venturen unter Nutzung von Forsteinrichtungs- und Satellitendaten. Dissertation, Universität Freiburg

Diamond, P., Armstrong, M. (1984): Robustness of Variograms and Conditioning of Kriging Matrices. Math. Geology 16(8), 809–822

Dietrich, C. R., Osborne, M. R. (1991): Estimation of Covariance Parameters in Kriging via Restricted Maximum Likelihood. Math. Geology 23, 119–135

Dougherty, C. (1992): Introduction to Econometrics. Oxford University Press

Dralle, K., Rudemo, M. (1996): Stem number estimation by kernel smoothing of aerial photos. Can. J. For. Res. 26, 1228–1236

Flury, B. (1997): A First Course in Multivariate Statistics. Springer Text in Statistics

Fouquet, C., Mandallaz, D. (1993): Using Geostatistics for Forest Inventory with Air Cover: An Example. In: Geostatistics Troia '92 Vol. 2 pp. 875–886. Kluwer Academic Publisher

FVA Baden-Württemberg (1999):. Betriebsinventur: Benutzeranweisung für das Auswertungsprogramm. Forstliche Versuchs- und Forschungsanstalt Baden-Württ- temberg. Abteilung Biometrie und Informatik

Gill, S. J., Biging, G. S., Murphy, E. C. (2000): Modeling conifer tree crown radius and estimating canopy cover. Forest Ecology and Management 126, 405–416

Gong, P., Biging, G., Lee, S. M., Mei, X., Sheng, Y., Pu, R., Xu, B., Schwarz, K.-P., Mostafa, M. (1999): Photo Ecometrics for Forest Inventory. Geographic Information Science 5(1), 9–14

Goovaerts, P. (1997): Geostatistics for Natural Resource Evaluation. Oxford Univer-

sity Press, Inc

Gotway, C. A., Hartford, A. H. (1996): Geostatistical Methods for Incorporating Auxiliary Informantion in the Prediction of Spatial Variables. J. Agr., Bio., and Env. Stat. 1(1), 17–39

Haara, A., Nevalainen, S. (2002): Detection of dead or defoliated spruces using digital aerial data. Forest Ecology and Management 160, 97–1007

Hartung, J., Elpelt, B. (1984): Multivariate Statistik. Oldenbourg Verlag

Harville, D. A. (1997): Matrix Algebra from a Statistician's Perspective. Springer

Hoffmann, B. (2001): Untersuchungen zum Monitoring von Naturwäldern mit Hilfe von Geo-Informationssystemen, modernen Luftbildauswertungsverfahren und Geostatistik: dargestellt am Beispiel des nordrhein-westfälischen Naturwaldreservates Hellerberg. Dissertation, Universität Göttingen

Holmström, H. (2002): Estimation of single-tree characteristics using the kNN method and plotwise aerial photograph interpretations. Forest Ecology and Management 167, 303–314

Holopainen, M., Wang, G. (1998): Accuracy of Digitized Aerial Photographs for Assessing Forest Habitats at Plot Level. Scand. J. For. Res. 19, 499–508

Hundeshagen, J. C. (1826): Die Forstabschaetzung auf neuen wissenschaftlichen Grundlagen. Laupp, Tuebingen

Ihaka, R., Gentleman, R. (1996): R: A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics 5(3), 299–314

Jost, A. (1993): Geostatistische Analyse des Stichprobenfehlers systematischer Stichproben. Dissertation, Universität Freiburg

Journel, A. G., Huijbregts, C. J. (1978): Mining Geostatistics. Academic Press, London

Kätsch, C., Stöcker, M. (2000): Untersuchungen zur automatischen Ermittlung von Bestandeshöhen auf Luftbildern mit Hilfe der Digitalen Photogrammetry. Allg. Forst- u. J.-Ztg. 171(4), 74–80

Kitanidis, P. (1983): Statistical Estimation of Polynomial Generalized Covariance Functions and Hydrologic Applications. Water Resource Management 19(4), 909–921

Köhl, M., Gertner, G. (1992): Geostatistische Auswertungsmöglichkeiten für Waldschadensinventuren: Methodische Überlegungen zur Beschreibung räumlicher Verteilungen. Forstw. Cbl. 111, 320–331

Kraus, K. (1994): Photogrammetrie. Dümmler/Bonn 5'th edition. Band 1

Lillesand, T. M., Kiefer, R. W. (2000): Remote Sensing and Image Interpretation. John Wiley & Sons 4'th edition

Mahalanobis, P. C. (1936): On the generalized distance in statistics. In: Proceedings of the national Institute of Science, India Vol. 12 pp. 49–55

Maltamo, M., Kangas, A. (1998): Methods based on k-nearest neighbour regression in the estimation of basal area diameter distribution. Can. J. For. Res. 28(8), 1107–1115

Mandallaz, D. (1991): A unified approach to sampling theory for forest inventory based on infinite population and superpopulation models. Phd thesis no 9378, ETH Zürich, Chair of Forest Management and Planning

Mandallaz, D. (1996): Geostatistical Methods for double sampling schemes: application to combined forest inventories. Habilitationsschrift, ETH Zürich, Chair of Forest Management and Planning

McBratney, A. B., Webster, A. B., Burgess, T. M. (1981a): The Design of Optimal Sampling Schemes for Local Estimation and Mapping of Regionalized Variables I. Computers and Geosciences 7(4), 331–334

McBratney, A. B., Webster, A. B., Burgess, T. M. (1981b): The Design of Optimal Sampling Schemes for Local Estimation and Mapping of Regionalized Variables II. Computers and Geosciences 7(4), 335–365

Miguel-Ayanz, J. S., Biging, G. S. (1996): An interative classifcation approach for mapping natural resources from satellite imagery. Int. J. Rem. Sens. 17(5), 957–981

Moeur, M., Stage, A. R. (1995): Most similar neighbor: an improved sampling inference procedure for natural resource planning. For. Sc. 41(2), 337–359

Montgomery, D. C., Peck, E. A., Vining, G. G. (2001): Introduction to Linear Regression Analysis. John Wiley & Sons, 3rd edition

Muinonen, E., Maltamo, M., Hyppänen, H., Vainikainen, V. (2001): Forest stand characteristics estimation using a most similar neighbor approach and iamge spatial structure information. Rem. Sens. Env. 78, 223–228

Myers, D. E. (1982): Matrix Formulation of CoKriging. Math. Geology 7(3), 249–257

Nagel, J. (1999): Konzeptionelle Überlegungen zum schrittweisen Aufbau eines waldwachstumskundlichen Simulationssystems für Nordwestdeutschland. Schriften aus der Forstlichen Fakultät der Universität Göttingen und der Niedersächsischen

Forstlichen Versuchsanstalt. J.D. Sauerländers's Verlag, Frankfurt a. M.

Niedersächsische Landesregierung (1992): Niedersächsisches Program zur langfristigen ökologischen Waldentwicklung in den Landesforsten, 2. Auflage. Niedersächsische Landesregierung

Nieschulze, J., Saborowski, J. (2001): Regionalisation of Point Information: A Comparison of Parameter Estimation Techniques for Universal Kriging. In: K. Rennolls (ed.), Forest Biometry, Modelling and Information Science. IUFRO 4.11 Conference University of Greenwech, UK

Nuske, R. (2002): Automatisierte Höhenmessung aus digitalen Luftbildern mittels OrthoEngine. Institut für Forstliche Biometrie und Informatik, Georg-August Universität Göttingen, unveröffentlicht

Olea, R. A. (1984): Sampling Design Optimization for Spatial Functions. Math. Geology 16(4), 369–393

Olea, R. A. (1999): Geostatistics for Engineers and Earth Scientists. Kluwer Academic Publishers

Otto, H.-J. (1989): Langfristige ökologische Waldbauplanung für die Niedersächsischen Landesforsten. Aus dem Walde, Vol. 42, Hannover. Band 1

Otto, H.-J. (1991): Langfristige ökologische Waldbauplanung für die Niedersächsischen Landesforsten. Aus dem Walde, Vol.43, Hannover. Band 2

Ramirez-Maldonado, H. (1988): On the relevance of geostatistical theory and methods to forest inventory problems. Phd thesis, University of Georgia

Ripley, B. (1981): Spatial Statistics. John Wiley & Sons

Saborowski, J., Dahm, S. (1996): Optimiertes Stichprobenverfahren für die Kontrollstichprobe in Niedersachsen. Universität Göttingen. Unveröffentlichtes Fachgutachten im Auftrag des Niedersächsischen Forstplanungsamtes

Saborowski, J., Stock, R. (1994): Regionalisierung von Niederschlagsdaten im Harz. Allg. Forst- u. J.-Ztg. 165(7), 117–122

Samra, J. S., Gill, H. S., Bhatia, V. K. (1989): Spatial Stochastic Modeling of Growth and Forest Resource Evaluation. For. Sc. 35(3), 663–676

Särndal, C.-E., Swenson, B., Wretman, J. (1992): Model assisted survey sampling. Springer, New York

Schneider, T. W. (1995): Kriterien und Indikatoren für eine nachhaltige Bewirtschaftung der Wälder. Allg. Forst Zeitung 50(4), 184–186

Schober, R. (1995): Ertragstafeln wichtiger Baumarten. J.D. Sauerländers's Verlag

Frankfurt a.M., 4 Auflage

Smith, D. A. (1986): The practice of silviculture. John Wiley & Sons 8'th edition

Smith, F. (1938): An empirical law describing heterogenity in the yields of agricultural crops. J. Agri. Sci. 28, 1–23

Solow, A. R. (1990): Geostatistical Cross-Validation: A Cautionary Note. Math. Geology 22(6), 637–639

Stein, M. L., Handcock, M. S. (1989): Some Asymptotic Properties of Kriging when the Covariance Function is misspecified. Math. Geology 21(2), 171–180

Thompson, S. (1992): Sampling. John Wiley & Sons, Inc. New York

Tou, J. T., Gonzalez, R. C. (1974): Pattern Recognition Principles. Addison-Wesley Publishing Co.

Tzschupke, W. (1991): Betriebs- oder bestandesweise Waldzustandsinventurverfahren für die Forsteinrichtung. Allg. Forst- u. J.-Ztg. 162, 195–200

United Nations (1992): United Nations Conference on Environment and Development. United Nations, New York

Uuttera, J., Haara, A., Tokola, T., Maltamo, M. (1998): Determination of the spatial distribution of trees from digital aerial photographs. Forest Ecology and Management 110, 275–282

Venables, W. N., Ripley, B. D. (2002): Modern Applied Statistics with S. Springer, New York, 4'th edition

Warnes, J. J., Ripley, B. D. (1987): Problems with Likelihood Estiamation of Covariance Functions of Spatial Gauss Processes. Biometrika 74, 640–642

Watkins, A., Al-Bouthiahi, F. (1990): On Maximum Likelihood Estimation of Parameters in Incorrectly Specified Models of Covariance of Spatial Data. Math. Geology 24(2), 151–173

Webster, R., Oliver, R. A. (2001): Geostatistics for environmental scientists. John Wiley & Sons, Inc

Wensel, L. C. (1996): Wildland Resource Sampling. ESPM, College of Natural Resources, University of California, Berkeley; unpublished reader

Wildenhein, T. (1999): Geostatistische Analyse ertragskundlicher Parameter aus Stichprobendaten. Dipl. Arbeit, Universität Göttingen, Fakultät für Forstwissenschaften und Waldökologie

Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M. P. (1999): An Experimental Comparison of Ordinary and Universal Kriging and Distance Weighting. Math.

Geology 31(4), 375–390

Zimmerman, D. L., Zimmerman, M. B. (1991): A Comparison of Spatial Semivariogram Estimators and Corresponding Ordinary Kriging Predictors. Technometrics 33(1), 77–91

A Kriging

A.1 Ordinary Kriging

For further explanation and proofs of the following definitions see Olea (1999).

Definition I

Let $D \subset \mathbb{R}^2$ and Ω be a non empty set with elements ω representing realizations of a sample. A random function is then a set of random variables

$$\{Z(s, \omega) : s \in D, \omega \in \Omega\}. \tag{11}$$

For a given location $s_i \in D$, $Z(s_i, \omega)$ is a function of ω, that is, $Z(s_i, \omega)$ is a random variable. In contrast, for a specific $\omega_j \in \Omega$, $Z(s, \omega_j)$ is a deterministic function of s that represents a possible observation of the random field and is called a realization. Finally, $Z(s_i, \omega_j)$ is merely a number. Abbreviations are common, equation (11) is referred to by $Z(s)$ and $z(s_i)$ is used in place of $Z(s_i, \omega_j)$.

Kriging is a collection of generalized linear regression techniques. The generalization of classical regression theory comprises

- no assumption that the observations are independent and identical distributed

- the assumption of a random sample of classical statistics is not required

Kriging is based on multiple assumptions.

Assumption I:

The sample is a partial realization of a random function $Z(s)$.

Assumption II:

The random function honors the intrinsic hypothesis over the sampling domain, that is,

$$E\Big[Z(s)\Big] = m \tag{12}$$

$$Var\Big(Z(s) - Z(s+h)\Big) = E\Big[Z(s) - Z(s+h)\Big]^2 = 2\gamma(h) \tag{13}$$

where h is a distance vector, $E[\cdot]$ denotes the expectation operator, and $\gamma(\cdot)$ is the semivariogram.

It follows that the semivariogram is a statistic that measures the average decrease of similarity of two random variables over spatial distance. Semivariograms used in kriging-systems are continuous functions of distance. In order to carry out kriging a spatial structure analysis is required, semivariogram models have to be fit to the data.

A semivariogram is modeled here as $\gamma(h, \theta)$, with $\theta \in \Theta = \mathbb{R}^3_+$ being a vector of spatial-dependence parameters. The semivariogram models considered are the nugget, spherical, and exponential model where the nugget model can be nested with the two other ones. The range is the distance at which a semivariogram reaches a constant value, called sill. Let S be the sill and a be the range. The nugget model is of the form

$$\gamma(h) = \begin{cases} 0 & , \text{h=0} \\ S & , \text{h} > 0 \end{cases} \tag{14}$$

The spherical model is expressed by

$$\gamma(h) = \begin{cases} S(\frac{1.5h}{a}) - \frac{1}{2}(\frac{h}{a})^3 & , 0 \le h < a \\ S & , a \le h \end{cases} \tag{15}$$

The exponential model is expressed by

$$\gamma(h) = S(1 - e^{-(\frac{3h}{a})}) \tag{16}$$

where a practical definition of the range a is the distance where the semivariogram reaches $0.95S$. The models are fit to the data by empirical variograms. The classical estimator is expressed by

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \Big[z(s_i) - z(s_i + h) \Big]^2 \tag{17}$$

The estimator (17) is not robust with respect to outliers. In case of such data the following estimator

$$\hat{\gamma}_R(h) = \frac{\Big(\frac{1}{|N(h)|} \sum_{(i,j) \in N(h)} \sqrt{|z(s_i) - z(s_j)|}\Big)^4}{2(0.457 + \frac{0.494}{|N(h)|})} \tag{18}$$

is more appropriate (Cressie and Hawkins, 1980; Cressie, 1991). $N(h)$ denotes the number of pairs of points with a distance within an interval, a so called bin, with mean h. Modeling the semivariogram requires a large enough number of distance vectors with sufficient frequency of points pairs. Practitioners usually follow a recommendation by Journel and Huijbregts (1978), who recommend $N(h) \geq 30$. Because such extensive data sets are seldom available pairs of points with similar distances are grouped together into said bins. The partition of the bins influences the model fit and is done by expert knowledge.

Definition II

    Let Z be an intrinsic function. The krige predictor $\hat{Z}(s_0)$ at location s_0 is given by the following linear combination of random variables at locations $s_i,\ i = 1, ..., n$

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i) \tag{19}$$

subject to the constraint

$$\sum_{i=1}^{n} \lambda_i\ = 1 \tag{20}$$

Theorem I

    The prediction variance $\sigma^2(s_0) = Var\left(\hat{Z}(s_0) - Z(s_0)\right)$ is the objective function to optimize the weights λ_i. Let $\hat{Z}(s_0)$ and $\gamma(\cdot)$ be as introduced above. The prediction variance is expressed by

$$\sigma^2(s_0) = 2 \sum_{i=1}^{n} \lambda_i \gamma(s_i, s_0) - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \gamma(s_i, s_j) \tag{21}$$

Definition III

    Let $N \in I\!N$ and let $\{\alpha_1, ..., \alpha_N\}$ be a set of real or complex numbers and let $\{s_1, ..., s_N\}$ be a set of points in $\mathbb{R}^2$. The continuous function $\phi(s_i, s_j)$ is called negative semi-definite if the following holds

$$-\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \phi(s_i, s_j) \geq 0 \tag{22}$$

Definition IV

Let $\sigma^2(s_0)$ and λ_i be as in Theorem 1 and Definition 2 and let ψ be a Lagrange multiplicator. The Lagrange function for ordinary kriging is then given by

$$L(\lambda_1, ..., \lambda_n, \psi) = \sigma^2(s_0) + 2\psi(\sum_{i=1}^{n} \lambda_i - 1) \tag{23}$$

Theorem II

Let $\lambda_i, \gamma(\cdot)$ and ψ be as introduced above. If the semivariogram is negative semi-definite then the following statements hold

- the weights yielding the unique minimal prediction variance are the solutions to

$$\sum_{i=1}^{n} \lambda_i \gamma(s_i, s_j) - \psi = \gamma(s_j, s_0) \quad j = 1, ..., n \tag{24}$$

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{25}$$

- the kriging variance $\sigma_k^2(s_0) = E[Z(s_0) - \hat{Z}(s_0)]^2$ is never negative.

Ordinary kriging has the following properties:

$\star$ minimum mean square error among all linear unbiased predictors,

$\star$ estimation interval not restricted to the data interval,

$\star$ screen effect: with observations approximately in line and all on the same side of the prediction site, the closest collinear observation takes the normal amount of weight and all other observations behind it tend to be ignored,

$\star$ declustering ability: clusters of observations get weights assigned that collectively are close to the weight a single observation at the centroid of the cluster would get assigned,

$\star$ exact interpolation with zero-kriging variance,

$\star$ inability to handle duplicate sampling sites,

$\star$ dependence upon the sampling pattern,

$\star$ and independence of the kriging variance from individual observations.

A.2   Universal Kriging

In the universal kriging approach the random function $Z(s)$ is divided into a deterministic trend and a fluctuation, $Z(s) = m(s) + \delta(s)$, where $m(s) = \sum_{l=0}^{p} \beta_l f_l(s) = \vec{f}(s)'\vec{\beta}$ with $\vec{\beta}$ being a vector of $(p+1)$ scalars, $\vec{f}(s)' = \left(f_0(s), ..., f_p(s) \right)$, $f_0(s) = 1$, and prime denotes transposition.

Theorem III

The Krige predictor is given as in (19). The weights yielding the unique minimal prediction variance are the solutions to

$$\sum_{i=1}^{n} \lambda_i \gamma_{res}(s_i, s_j) - \psi_0 - \sum_{l=1}^{p} \psi_l f_l = \gamma_{res}(s_j, s_0) \quad j = 1, ..., n \qquad (26)$$

$$\sum_{i=1}^{n} \lambda_i = 1 \qquad (27)$$

$$\sum_{i=1}^{n} \lambda_i f_l(s_i) = f_l(s_0) \quad l = 1, ..., p \qquad (28)$$

where $\psi_l \quad l = 0, ..., p$ are $(p + 1)$ Lagrange multipliers and $\gamma_{res}(\cdot)$ is the semivariogram of the residuals, the difference of the estimated trend to the observed values.

The most common form of $f_l(s)$ is that of a monomial in the geographical coordinates and although each $f_l(s)$ has been written as a function of location s they could also be function of an explanatory variable associated with the datum at $s, s \in D$. The distinction between the trend m and the fluctuation δ is not clear cut. In general, m represents a large scale variation, regarded as fixed and δ a zero-mean intrinsically random process, that is $E[\delta(s)] = 0$ and $Var(\delta(s) - \delta(s + h)) = 2\gamma(h) = E[\delta(s) - \delta(s+h)]^2$. A stronger assumption is second order stationarity. The semivariogram can then be expressed as $\gamma(h) = C(0) - C(h)$, where $C(\cdot)$ denotes the covariance function of the random process. The krige prediction according to Eq. (19) of any location $s_0 \in D$ is re-written in vector form as the linear combination of random variables at locations $s_i, i = 1, ..., n$ as

$$\hat{z}(s_0) = \sum_{i=1}^{n} \lambda_i z(s_i) = \vec{\lambda}'\vec{z}_n \qquad (29)$$

subject to the constraints (27) and (28). The prediction variance $Var(\hat{Z}(s_0) - Z(s_0))$ is the objective function for the derivation of the optimal weights. Assume $\theta \in \Theta$ is known.

Let $\vec{c}_u = (\vec{c}, \vec{x}_0)' = (C(s_0 - s_1), ..., C(s_0 - s_n), f_0(s_0), ..., f_p(s_0))'$ be a vector of length $(n + p + 1)$, $\sum(\theta) = \sum = (C(s_i - s_j))$, $i, j = 1, ..., n$ be a $n \times n$ matrix, $X = (f_l(s_i)), i = 1, \ldots, n; l = 0, \ldots, p$ be a $n \times (p+1)$ matrix, and 0 be a $(p+1) \times (p+1)$ Null matrix.

Furthermore, let $\vec{\lambda}_u = (\vec{\lambda}, -\vec{m})'$, $\vec{\lambda} = (\lambda_1, \ldots, \lambda_n)'$ be the vector of kriging weights, $\vec{m} = (\vec{m}_0, ..., \vec{m}_p)'$ be the vector of Lagrange multipliers, and $\sum_u^{-1} = \begin{bmatrix} \sum & X \\ \sum' & 0 \end{bmatrix}^{-1}$.
Using matrix notation and the covariance of the random process the optimal weights are obtained from (Cressie, 1991)

$$\vec{\lambda}_u = \Sigma_u^{-1} \vec{c}_u \tag{30}$$

Following Harville (1997) (p.99) (30) can be rewritten as

$$
\begin{aligned}
\vec{\lambda}_u = \begin{pmatrix} \vec{\lambda} \\ -\vec{m} \end{pmatrix} &= \Sigma_u^{-1} \vec{c}_u = \begin{pmatrix} \sum & X \\ \sum' & 0 \end{pmatrix}^{-1} \begin{pmatrix} \vec{c} \\ \vec{x}_0 \end{pmatrix} \\
&= \begin{pmatrix} \Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1} & \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1} \\ (X'\Sigma^{-1}X)X'\Sigma^{-1} & -(X'\Sigma^{-1}X)^{-1} \end{pmatrix} \begin{pmatrix} \vec{c} \\ \vec{x}_0 \end{pmatrix}
\end{aligned} \tag{31}
$$

from which follows that

$$-\vec{m} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\vec{c} - (X'\Sigma^{-1}X)^{-1}\vec{x}_0 \quad \text{and}$$
$$\vec{\lambda} = (\Sigma^{-1} - \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1})\vec{c} + \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}\vec{x}_0.$$

Plugging the latter result into (29) yields

$$
\begin{aligned}
\hat{z}(s_0) &= \vec{c}\,'\Sigma^{-1}\vec{z}_n - \vec{c}\,'\Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}\vec{z}_n + \vec{x}_0\,'(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\vec{z}_n \\
&= \vec{c}\,'\Sigma^{-1}(\vec{z}_n - X(X'\Sigma^{-1}X)^{-1}\vec{z}_n + \vec{x}_0\,'(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\vec{z}_n
\end{aligned} \tag{32}
$$

Recalling that the parameter vector $\vec{\beta}_{GLS}$ of the deterministic trend is estimated by generalized least squares (GLS) as $\widehat{\vec{\beta}}_{GLS} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\vec{z}_n$ (32) is rewritten as

$$\hat{z}(s_0) = \vec{c}\,'\Sigma^{-1}(\vec{z}_n - X\widehat{\vec{\beta}}_{GLS}) + \vec{x}_0\,'\widehat{\vec{\beta}}_{GLS} = \vec{c}\,'\Sigma^{-1}\vec{\epsilon} + \vec{x}_0\,'\widehat{\vec{\beta}}_{GLS} \tag{33}$$

As seen from (33) prediction by universal kriging comprises the estimation of the deterministic trend and simple kriging of the residuals.

In practice θ is seldom known and has to be estimated. A circular problem arises since $\Sigma = \Sigma(\theta)$ is the covariance matrix of the residuals. The residuals can only be estimated once $\widehat{\vec{\beta}}_{GLS}$ is obtained but the latter is based on the covariance matrix.

One approach is based on ordinary least squares (OLS) and proceeds iteratively. First, get an approximation of $\widehat{\vec{\beta}}_{GLS}$ by OLS, then model the residuals and take their covariance matrix to refine $\widehat{\vec{\beta}}_{GLS}$ in the next step. Repeat until the procedure converges (Ripley, 1981). One disadvantage of such an approach is that the variogram of the residuals will be biased and underestimate the true variogram even if the true covariance matrix were known (see Cressie (1991), p.166 for an example). This matters because although most software rely for reasons of numerical stability on the covariance for solving kriging systems, spatial structure analysis is done by means of the semivariogram.

The kriging variance will be much more severely affected than the prediction itself. However, it is generally acknowledged that the bias is small at short distances (Cressie, 1991; Venables and Ripley, 2002), and kriging is usually carried out in local neighborhoods, thus the variogram will only be evaluated at the smaller lags.

Another iterative approach is based on maximum likelihood (ML) estimation. If both β and θ are unknown, then the maximum likelihood predictor is

$$\hat{E}[Z(s_0)|\vec{z}_n] = \vec{c}(\hat{\theta})'\Sigma(\hat{\theta})^{-1}(\vec{z}_n - X\widehat{\vec{\beta}}(\hat{\theta})) + \vec{x}_0{}'\widehat{\vec{\beta}}(\hat{\theta}) \tag{34}$$

(Cressie and Zimmerman, 1992) where $\widehat{\vec{\beta}}(\hat{\theta}) = (X'\Sigma(\hat{\theta})^{-1}X)^{-1}X'\Sigma(\hat{\theta})^{-1}\vec{z}_n$ and $\hat{\theta}$ maximizes the profile likelihood

$$(2\pi)^{\frac{-n}{2}}|\Sigma(\theta)|^{\frac{-1}{2}} \, exp\Big(-\frac{1}{2}(\vec{z}_n - X\widehat{\vec{\beta}}(\theta))'\Sigma(\theta)^{-1}(\vec{z}_n - X\widehat{\vec{\beta}}(\theta))\Big) \; over \; \theta \in \Theta.$$

The profile likelihood can be multi-modal (Warnes and Ripley, 1987). $\hat{\theta}$ need not be an unbiased estimator of θ for (34) to be an unbiased predictor (Cressie and Zimmerman, 1992). However, the bias of $\hat{\theta}$ can be a matter of concern in geostatistics where data sets often comprise less than 100 observations.

As a remedy, a linear transformation $T\vec{z}_n$ of the data is sought that filters out the mean, that is, $E[T\vec{z}_n] = TX\vec{\beta} = 0 \quad \forall\vec{\beta}$. This leads to the restricted maximum likelihood estimation (REML) (Kitanidis, 1983). The transformed observation vector

$T\vec{z}_n$ is assumed to be normally distributed with zero mean and covariance $TC(\theta)T'$ (Dietrich and Osborne, 1991). The REML estimate of the parameter vector θ is the solution to

$$\max_{\theta}\Big(-log|TC(\theta)T'| - \vec{z}_n\ 'T'(TC(\theta)T')^{-1}T\vec{z}_n\ '\Big).$$

The likelihood based algorithms require an initial estimate of the parameter vector θ. Most often such estimates are derived from the semivariogram of the residuals of an OLS fit. It is not always intuitively clear what kind of semivariogram model is appropriate but the maximum likelihood estimation seems rather robust to incorrect specifications (Watkins and Al-Bouthiahi, 1990).

B   Canonical Correlation

The objective of canonical correlation analysis is to get a simple description of the structure between subsets of variables. Assume that two subsets of variables $\vec{X} = (X_1, \ldots, X_p)'$ and $\vec{Y} = (Y_1, \ldots, Y_q)'$ have a joint normal distribution,

$$\begin{pmatrix} \vec{X} \\ \vec{Y} \end{pmatrix} \sim N\left(\begin{pmatrix} \vec{\mu}_X \\ \vec{\mu}_Y \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}\right).$$

Without loss of generality suppose $p \leq q < n$, with n being the number of observations. The analysis searches for weights $\vec{\alpha} = (\alpha_1, \ldots, \alpha_p)'$ and $\vec{\beta} = (\beta_1, \ldots, \beta_q)'$ that maximize the absolute (canonical) correlation of a pair of linear combinations $\vec{\alpha}'\vec{X}$ and $\vec{\beta}'\vec{Y}$ (Hartung and Elpelt, 1984; Bilodeau and Brenner, 1999). These linear combinations are called canonical variables.

The canonical correlation can be derived as the square root of the largest Eigenvalue of the matrix product $\Sigma_X^{-1}\Sigma_{XY}\Sigma_Y^{-1}\Sigma'_{XY}$, with

$$\Sigma_X = Cov(X_1, \ldots, X_p) = \begin{pmatrix} \sigma_{X_1}^2 & \sigma_{X_1 X_2} & \cdots & \sigma_{X_1 X_p} \\ \sigma_{X_2 X_1} & \sigma_{X_2}^2 & \cdots & \sigma_{X_2 X_p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_p X_1} & \sigma_{X_p X_2} & \cdots & \sigma_{X_p}^2 \end{pmatrix}, \quad \Sigma_Y = Cov(Y_1, \ldots, Y_q)$$

$$\text{and } \Sigma_{XY} = \begin{pmatrix} \sigma_{X_1Y_1} & \sigma_{X_1Y_2} & \cdots & \sigma_{X_1Y_q} \\ \sigma_{X_2Y_1} & \sigma_{X_2Y_2} & \cdots & \sigma_{X_2Y_q} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_pY_1} & \sigma_{X_pY_2} & \cdots & \sigma_{X_pY_q} \end{pmatrix}.$$

The square roots of the remaining $p-1$ Eigenvalues are also called canonical correlations and the corresponding linear combinations canonical variables. The second to largest canonical correlation is the maximum absolute correlation of all pairs of linear combinations that are perpendicular to the first linear combination. The remaining canonical correlations are derived accordingly.

The canonical correlations $\rho_{(X_1,\ldots,X_p),(Y_1,\ldots,Y_q)}$ can be estimated based on a sample of n observations out of a population. Measure the observable values $x_{1k},\ldots,x_{pk}$, $y_{1k},\ldots,y_{qk}$, $k = 1,\ldots,n$, of the features $X_1,\ldots,X_p,Y_1,\ldots,Y_q$ at all n objects. Estimate the matrices Σ_X, Σ_Y and Σ_{XY} by S_X, S_Y and S_{XY} by estimating for $i = 1,\ldots,p$ the variances $\sigma_{X_i}^2$ by $s_{X_i}^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_{ik} - \bar{x}_i)^2$, with $\bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ik}$,

the variances $\sigma_{Y_j}^2$ $(j = 1,\ldots,q)$ and the covariances $\sigma_{X_iX_j}$, $\sigma_{X_iY_j}$, and $\sigma_{Y_iY_j}$ accordingly; for example

$s_{X_iY_j} = \frac{1}{n-1}\sum_{k=1}^{n}(x_{ik} - \bar{x}_i)(y_{jk} - \bar{y}_j)$. Calculate

$$Q = S_X^{-1}S_{XY}S_Y^{-1}S_{XY}' \tag{35}$$

and use its Eigenvalues for the derivation of the empirical canonical correlations.

The vector $\vec{\alpha}$ of canonical coefficients can be estimated by any Eigenvector $\widehat{\vec{\alpha}}$ of the respective Eigenvalue of the matrix Q. The vector $\hat{\beta}$ is estimated by $\widehat{\vec{\beta}}$ according to

$$S_Y^{-1}S_{XY}'\widehat{\vec{\alpha}} = \widehat{\vec{\beta}} \tag{36}$$

C   Lebenslauf

Angaben zur Person

Name:	Jens Nieschulze
Geburtstag und Ort:	31 Mai 1971 in Uelzen
Vater:	Horst Nieschulze, Altbauer
Mutter:	Hiltrud Nieschulze, geb. Niebel, Hausfrau
Geschwister:	Vier ältere Brüder und eine jüngere Schwester
Familienstand:	Ledig
Nationalität:	Deutsch

Schulische Ausbildung

1977- 1981	Grundschule, 29562 Suhlendorf
1981- 1983	Orientierungsstufe, 29562 Suhlendorf
1983- 1990	Lessing Gymnasium Uelzen: Abitur am 26. Mai

Berufliche Ausbildung und Tätigkeiten

Jan 1991- Dez 1992	Bundeswehr (Z2)
Apr 1993- Sep 1993	Praktikum in der Gräflich Bernstorff'sche Forstverwaltung, Gartow/Elbe
Sep 1993- März 1994	Praktikum im Freiherrlich von Adelsheim'schen Rentamt und Forstei, Adelsheim, Baden-Württemberg
Apr 1994- März 1996	Universität Göttingen. Grundstudium der Forstwissenschaften und Waldökologie. Vordiplom im März 1996
Mai 1996- Juli 1996	Hospitation bei der Gesellschaft für Technische Zusammenarbeit (GTZ) in The Gambia, Westafrika

Aug 1996- Mai 1997	University of California, Berkeley, USA. Non-degree non-fee Austausch-Programm mit der Universität Göttingen. Spezialisierung: Stichprobentheorie, Forstliche Biometrie und Fernerkundung
Mai 1997- März 1999	Universität Göttingen. Hauptstudium der Forstwissenschaften und Waldökologie. Spezialisierung: Ökologie. Diplom-Forstwirt im März 1999. Dipl. Arbeit: 'Entwicklung und Erprobung von angepaßten Naturwaldinventurverfahren in The Gambia, Westafrika'
Okt 1998- März 1999	Universität Göttingen. Mathematische Fakultät. Einführung in Lineare Algebra und Differentialrechnung
Mai 1999- Sep 1999	Inventurtätigkeit bei der Mendocino Redwood Company, Ukia, Kalifornien, USA
Mai 1999- Okt 1999	Department of Environmental Science, Policy and Management of the University of California, Berkeley: Graduate Student Researcher (GSR) (wissenschaftlicher Assistent)
Nov 1999- Jan 2000	Tätigkeit in dem Freiherrlich von Adelsheim'schen Rentamt und Forstei, Adelsheim, Baden-Württemberg
Feb 2000- Apr 2000	Universität Göttingen. Institut für Forstliche Arbeitswissenschaft und Verfahrenstechnologie: Wissenschaftlicher Mitarbeiter im Qualitätsmanagement
Seit Mai 2000	Universität Göttingen. Promotion in der Forstlichen Biometrie und Informatik gefördert durch ein Stipendium der Deutschen Bundesstiftung Umwelt (DBU)
Okt 2000- März 2001	Universität Göttingen. Mathematische Fakultät. Einführung in Stochastik
Seit Nov 2002	Wissenschaftlicher Mitarbeiter im SFB 552 (Stabilität von Randzonen tropischer Regenwälder in Indonesien): Aufbau und Betreuung eines Informationssystemes