

Biometrie Workshop
6.-8. November 2013 in Freising
an der Bayerischen Landesanstalt für Wald und Forstwirtschaft

mit einem Tutorial von
Andrew Finley (Michigan State University)
zu **Univariater und Multivariater Raum-Zeit-Modellierung**

organisiert von der
Internationalen Biometrischen Gesellschaft – Deutsche Region
Arbeitsgruppen
Ökologie und Umwelt, Bayes Methoden und Räumliche Statistik
zusammen mit der
Sektion „Forstliche Biometrie und Informatik“ im Deutschen Verband
Forstlicher Forschungsanstalten (DVFFA)

Programm

Biometry workshop
November 6-8, 2013 in Freising
at the Bavarian State Institute of Forestry

with a Tutorial by
Andrew Finley (Michigan State University)
on **Univariate and Multivariate Spatio-temporal Modeling**

organized by the
International Biometric Society (IBS) - German Region (DR)
Working groups
Ecology and Environment, Bayes Methods and Spatial Statistics
together with the
Unit Forest Biometrics of the
German Association of Forest Research Stations (DVFFA).

Programme

LWF Bayerische Landesanstalt
für Wald und Forstwirtschaft


ZENTRUM WALD FORST HOLZ
WEIHENSTEPHAN


DVFFA



Seminar Room: Hanskarl-Goettling-Saal (LWF)

How to get there: <http://www.lwf.bayern.de/ueber-uns/31061/index.php>

Wednesday, November 6

| | | |
|---------|-------------------------------------|--|
| 1.00 pm | Nothdurft, Falk, Dreesman, König | Introduction |
| 1.15 pm | Andrew Finley | Tutorial (Introduction I) |
| 2.30 pm | | Coffee break |
| 3.00 pm | Christoph Staubach | Bayesian modeling of factors potentially influencing the distribution of <i>Echinococcus multilocularis</i> in foxes |
| 3.30 pm | Thomas Waldhoer | On the combination of difference and equivalence tests in spatial maps |
| 4.00 pm | Thomas Waldhoer | Relevance of the type III error in epidemiological maps |
| 4.30 pm | Juliane Vogt | The importance of conspecific facilitation in promoting recruitment and regeneration: a case study from degraded mangrove forests |
| 5.00 pm | | Geschäftssitzung DVFFA-Sektion Biometrie |
| 7.00 pm | | Dinner at Bräustüberl Weihenstephan http://www.braeustueberl-weihenstephan.de/#nogo22 |

Thursday, November 7

| | | |
|----------|----------------------|---|
| 8.30 am | Andrew Finley | Tutorial (Introduction II) |
| 10.00 am | Manuela Zucknick | Risk-prediction modelling in cancer with multiple genomic data sets: a Bayesian variable selection approach |
| 10.30 am | | Coffee break |
| 11.00 am | Gerhard Nehmiz | Introduction to Bayesian non-parametric survival analysis |
| 11.30 am | Thomas Welchowski | Identification of tolerance limits of forest tree species based on species occurrences and binary quantile regression |
| 12.00 am | Peter Biber | Beziehungen zwischen Substrat, Oberflächeneigenschaften und Vegetation in einem künstlichen Wassereinzugsgebiet |
| 12.30 am | | Lunch |

| | | |
|---------|-------------------------|--|
| 1.30 pm | Philip Beckschäfer | Using RapidEye imagery and the RandomForest algorithm for spatial modelling of Leaf Area Index values – Methodological Considerations and a Case Study from tropical China |
| 2.00 pm | Sebastian Schoneberg | Extension of k-most similar neighbours methods by local polynomial regression |
| 2.30 pm | Philip Mundhenk | Model uncertainty in ALS-assisted forest inventories |
| 3.00 pm | | Coffee break |
| 3.30 pm | Andrew Finley | Tutorial (Computer work on own laptops) |
| 5.00 pm | | Geschäftssitzung AG Räumliche Statistik |
| 5.20 pm | | Geschäftssitzung AG Bayes Methoden |
| 5.40 pm | | Geschäftssitzung AG Ökologie und Umwelt |
| 7.00 pm | | Dinner at Weihenstephaner am Dom http://www.weihenstephaner-am-dom.net/ |

Friday, November 8

| | | |
|----------|-------------------------|--|
| 8.30 am | Andrew Finley | Tutorial (Computer work on own laptops and wrap-up) |
| 10.00 am | | Coffee break |
| 10.30 am | Susanne Brandl | A site index model for Norway spruce in Bavaria |
| 11.00 am | Hooman Latifi | A naïve Bayes model to describe natural forest ground vegetation by waveform LiDAR |
| 11.30 am | Johannes Breidenbach | Estimating mean timber volume on stand level using unit-level and area-level small area estimators |
| 12.00 am | Ambros Berger | European Forestry Dynamics Model (EFDM) |

Contents

| | |
|--|-----------|
| Bayesian modeling of factors potentially influencing the distribution of <i>Echinococcus multilocularis</i> in foxes <i>Staubach, C., Schwarz, S., Sutor, A., Hoffmann, L., Tackmann, K., Schmid, V. and Conraths, F.J.</i> | 1 |
| On the combination of difference and equivalence tests in spatial maps <i>Harald Heinzl and Thomas Waldhoer</i> | 2 |
| Relevance of the type III error in epidemiological maps <i>Thomas Waldhoer and Harald Heinzl</i> | 3 |
| The importance of conspecific facilitation in promoting recruitment and regeneration: a case study from degraded mangrove forests <i>Juliane Vogt.....</i> | 4 |
| Risk-prediction modelling in cancer with multiple genomic data sets: a Bayesian variable selection approach <i>Manuela Zucknick.....</i> | 5 |
| Introduction to Bayesian non-parametric survival analysis <i>Gerhard Nehmiz</i> | 6 |
| Testing Bayesian techniques and quantile regression to identify limiting responses of tree species <i>Thomas Welchowski, Felix Klug, Karl Mellert and Helmut Küchenhoff.....</i> | 7 |
| Beziehungen zwischen Substrat, Oberflächeneigenschaften und Vegetation in einem künstlichen Wassereinzugsgebiet <i>Peter Biber.....</i> | 8 |
| Using RapidEye imagery and the RandomForest algorithm for spatial modelling of Leaf Area Index values. Methodological considerations and a case study from tropical China <i>Philip Beckschäfer, Lutz Fehrmann, Rhett D. Harrison, Jianchu Xu and Christoph Kleinn</i> | 9 |
| Extension of k-most similar neighbours methods by local polynomial regression <i>Sebastian Schonberg</i> | 10 |
| Model uncertainty in ALS-assited forest inventories <i>Philip Mundhenk.....</i> | 11 |
| A site index model for Norway spruce in Bavaria <i>Susanne Brandl.....</i> | 12 |
| A naïve Bayes model to describe natural forest ground vegetation by waveform LiDAR <i>Hooman Latifi, Viktor Gaub, Mrcó Heurich and Stefan Dech.....</i> | 13 |
| Estimating mean timber volume on stand level using unit-level and area-level small area estimators <i>Johannes Breidenbach and Ramus Astrup.....</i> | 14 |
| European Forestry Dynamics Model <i>Ambros Berger.....</i> | 15 |

Biometry Workshop Network Information



Dear workshop participant,

in the following rooms there is WLAN (802.11b/g) Internet Access.
(No validation necessary):

- Workshop-Room (Hanskarl-Goettling-Saal, LWF); if signal is low:
- Foyer TU München (building opposite to LWF)

Access is open during congress days from 8 o'clock to 17 o'clock.

Connection properties:

SSID con

Network authentication Open

Data encryption No encryption

TCP/IP Obtain an IP address automatically

Please configure your WLAN adapter with SSID „**con**“. This network name is not always broadcasted, enter it manually if necessary.

Do not specify Network Authentication or any encryption. For TCP/IP Properties choose „**Obtain an IP address automatically**“. You get a worldwide valid IP Address, no Proxy Server is needed.

All Internet applications except Microsoft File and Printer sharing should work.

Dinner on Wednesday, November 6
7:00 pm
Bräustürberl Weihenstephan
Weihenstephaner Berg 10
<http://www.braeustueberl-weihenstephan.de/>

Dinner on Thursday, November 7
7:00 pm
Weihenstephaner am Dom
Domberg 5a
<http://www.weihenstephaner-am-dom.net/>

Bayesian modeling of factors potentially influencing the distribution of *Echinococcus multilocularis* in foxes

Staubach, C.¹, Schwarz, S.¹, Sutor, A.¹, Hoffmann, L.², Tackmann, K.¹, Schmid, V.³ and Conraths, F.J.¹

¹Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Germany
christoph.staubach@fli.bund.de

²Thuringian State Authority for Food Safety and Consumer Protection, Germany

³Department of Statistics, Ludwig-Maximilians-University, Germany

Abstract. Alveolar echinococcosis is considered the most dangerous autochthonous parasitic zoonosis in central Europe. The red fox (*Vulpes vulpes*) represents the main definitive host in Europe, and various species of rodents are involved in the parasitic cycle as intermediate hosts. Since soil condition, temperature, and desiccation can effectively reduce the infectivity of the tapeworm eggs, microclimate and habitat may influence the spatial distribution of the parasite. To investigate the influence of environmental factors on the spatial epidemiology of *E. multilocularis*, 42,861 foxes were sampled in two Federal States of Germany, tested for the parasite and the results linked to a GIS. The landscape composition per spatial unit was derived from a high-resolution land-survey vector database and supplemented by a digital elevation model. Data were analyzed using a hierarchical Bayesian model. Although we used Markov Chain Monte Carlo (MCMC) techniques to reduce simulation error and to save time, the model runs to explore the parameter space were time consuming. It has been shown that Bayesian inference using integrated nested Laplace approximation (INLA) might reduce computing time dramatically without increasing the error of the estimates. We therefore applied and compared both methods. The study confirmed results of a previous publication, which utilized exact locations and micro-habitat data of foxes on a much larger scale and lower resolution. Furthermore, the preference of infected foxes for open landscapes with pasture was demonstrated in both regions despite the different landscape characteristic. Bayesian modeling of surveillance data might be also useful for ecological regression analysis, carefully considering the assumptions and limitations of the approach.

On the combination of difference and equivalence tests in spatial maps

Harald Heinzl and Thomas Waldhoer.

Medical University of Vienna, Austria
thomas.waldhoer@meduniwien.ac.at

Abstract. A choropleth map provides a proven method for presenting spatially distributed health data. Quite often local health authorities and the general public are concerned whether the health situation in certain areas can be considered different or equivalent to a reference value. Highlighting statistically significant areas enables the statement that the health situation in these areas differs from the reference value. However, this approach does not allow concluding that the situation in some areas is sufficiently close to the reference value. Since such conclusions are of interest for health policy making as well, a combined integration of statistical difference and equivalence tests into choropleth maps is suggested. The approach will be graphically exemplified with gestational age data of 60,303 newborns with Austrian mothers. The data are from 2008 and cover all 121 administrative Austrian districts. Choropleth maps, which only show the variable of interest and their corresponding difference test results, may easily misguide local health authorities and the general public. The adding of equivalence test results can lead to a better understanding of regional health care results as the issue of relevance is explicitly addressed through a pre-defined equivalence range.

References

Waldhoer, T. and Heinzl, H. 2011. Combining difference and equivalence test results in spatial maps. *International Journal of Health Geographics*. **10**,3.

Relevance of the type III error in epidemiological maps

Thomas Waldhoer and Harald Heinzl

Medical University of Vienna, Austria
thomas.waldhoer@meduniwien.ac.at

Abstract. Background: In classical hypothesis testing a type III error arises, when one side is erroneously favored although the true effect actually resides on the other side. The relevance of this grave error in decision making is studied for epidemiological maps. Methods: A simulation study in the R-package INLA using infant mortality data from Austria for the time period 1984 to 2008 was done in order to investigate the effect of spatial smoothing using unstructured as well as structured models. In addition to the type III error the q-value (Heinzl, H. and Waldhoer, T., 2012) i.e. the conditional probability for a type III error, provided a statistically significant result has been observed, was calculated. Results: Theoretical considerations for the standardized crude mortality ratio confirm that a type III error may be unacceptably large for regions with small numbers of expected cases even when no spatial smoothing has been performed. The study results reveal that especially spatial smoothing may increase the risk of type III errors to an unreasonable extent and lead to q-values larger than 50%. Conclusions: The occurrence of a type III error should be taken into account when interpreting results presented in epidemiological maps, particularly with regard to sparsely populated regions and spatial smoothing.

References

Heinzl, H. and Waldhoer, T. 2012. Relevance of the type III error in epidemiological maps. *International Journal of Health Geographics*. 11,34. DOI: 10.1186/1476-072X-11-34.

The importance of conspecific facilitation in promoting recruitment and regeneration: a case study from degraded mangrove forests

Juliane Vogt

Forest Biometrics / Systems Analysis, University of Dresden, Germany

vogt@forst.tu-dresden.de

Abstract. Often the theoretical concepts of recruitment processes focus only on competitive interaction disregarding facilitation both interspecific and conspecific. Little information is available considering the balance between facilitation and competition during recruitment and establishment along stress gradients. Facilitation promotes spatial aggregation even in the presence of competition. In this study different plant parameters including height, crown elongation and explicit spatial position were detected at mangrove seedlings and saplings on a degraded regenerating site. The calculated interaction intensity index reflects the aggregated distribution of the spatial stem pattern of seedlings. This aggregation remained even for saplings even though competition is increasing. This indicates that facilitation is still important and self-thinning is not the main driving factor being not powerful enough to change that pattern.

Risk-prediction modelling in cancer with multiple genomic data sets: a Bayesian variable selection approach

Manuela Zucknick

German Cancer Research Center, Germany

m.zucknick@dkfz-heidelberg.de

Abstract. Risk prediction based on clinical and molecular information is fundamental in translational cancer research. The statistical analysis of high-throughput omics data allows the identification of prognostic and predictive biomarkers, which can aid the development of targeted therapies. We present a Bayesian hierarchical regression model with variable selection (BVS) for risk prediction modelling based on high-dimensional genomics input data. A common application in cancer research is the prognosis of patient survival or the prediction of therapy response with simultaneous feature selection. In addition to gene expression data, high-throughput technologies are also available for many other types of genomic data, and today clinical researchers are routinely collecting genome-wide data from various sources on the DNA- and RNA-level. If data from several sources are available for the same set of biological samples, they can be analysed together in an integrative manner, with the aim of providing a more comprehensive picture of the disease biology and improving the performance of feature selection for risk prediction models. BVS models are very flexible in their setup and are naturally well-suited to extensions allowing the integration of additional data sources. We will present such an extension for the integration of copy number variation information into gene expression-based risk prediction models. We investigate model behaviour and the influence of prior specifications through simulation studies. The model is illustrated in applications in translational oncology.

Introduction to Bayesian non-parametric survival analysis

Gerhard Nehmiz

Boehringer Ingelheim Pharma GmbH & Co. KG, Germany
gerhard.nehmiz@boehringer-ingenelheim.com

Abstract. Non-parametric models, or models with infinitely many parameters, define prior and posterior probabilities on function spaces. Specific questions are: What does it mean that elements of the function space are neighbored to each other? When can a probability on a function space be re-interpreted as a family of probabilities, indexed by the range of definition of the given functions? Such an indexed family of probabilities is called a stochastic process. Specifically for survival times ([1]), whose range of definition is $[0, \infty)$ and which can be censored, relevant questions are: What is a reasonable prior process for the cumulative hazard function? How can the information generated by the observations be summarized into a likelihood function? Which relative weights should the prior process and the likelihood have in the posterior process ([2])? How robust are the results relative to implicit assumptions? How certain can we be that e.g. a given prior process is really uninformative ([3])? An example of piecewise Cox regression (nonparametric for event times, parametric for treatment effect) from the literature will be brought ([4], [5]).

References

1. Hellmich M. 2001. Bayes'sche Untersuchung von zensierten Daten [Bayesian investigation of censored data]. Presentation at the Workshop of the Bayes WG: "Hierarchical models and MCMC methods", Homburg/Saar 2001-03-19. <http://www.imbei.uni-mainz.de/bayes/TutoriumHOM.htm>
2. Susarla V. and Van Ryzin J. 1976. Nonparametric Bayesian Estimation of Survival Curves from Incomplete Observations. *J.A.S.A.* **71**, 897-902.
3. Berger J.O. 2000. Bayesian Analysis: A Look at Today and Thoughts of Tomorrow. *J.A.S.A.* **95**, 1269-1276.
4. Laud P.W., Damien P. and Smith A.F.M. 1998. Bayesian Nonparametric and Covariate Analysis of Failure Time Data. In: Dey D., Müller P. and Sinha D. (eds.): Practical Nonparametric and Semiparametric Bayesian Statistics. New York / Heidelberg: Springer Verlag, 213-225.
5. Lunn D., Jackson C., Thomas A., Best N. and Spiegelhalter D. 2012. The BUGS Book. Boca Raton/FL: Chapman&Hall / CRC.

Testing Bayesian techniques and quantile regression to identify limiting responses of tree species

Thomas Welchowski¹, Felix Klug¹, Karl Mellert², Helmut Küchenhoff¹

¹Institute of Statistics, University of Munich, Germany

Thomas.Welchowski@gmx.net

²Bavarian State Institute of Forestry, Germany

Abstract. Species distribution models (SDM) are used as empirical models to evaluate the possible impact of climate change on key species in nature conservation, agriculture, and forestry. Basically, SDMs use presences and/or absences to estimate the habitat suitability for a given species in a certain landscape. For assessing the risk to grow forest tree species in Bavaria under climate change, the identification of species climatic tolerance limits is crucial. For the identification of species limits a number of methods had been proposed such as stepwise linear regression, regression trees and boundary line modelling. Specifically, quantile regression is well suited to model the limiting response (upper quantile) of species along environmental gradients. But this technique requires a metric target variable, which was not available in the European data set we used for calibration of our SDMs. Therefore, we tested methods to identify the limiting response involving presence/absence data. Some approaches to solving this problem are presented and discussed.

Beziehungen zwischen Substrat, Oberflächeneigenschaften und Vegetation in einem künstlichen Wassereinzugsgebiet

Peter Biber

Leibniz-Rechnerzentrum, Germany
Peter.Biber@lrz.tu-muenchen.de

Abstract. Anhand von Daten zu Geländeoberfläche, Vegetation, Substrat und Grundwasser aus dem künstlichen Wassereinzugsgebiet "Hühnerwasser" (bei Cottbus, Niederlausitz) wurden verschiedene Hypothesen zur Entwicklung initialer Ökosysteme geprüft. Mithilfe von GAMM's und GLMM's konnten sowohl aus der Theorie zu erwartende Zusammenhänge als auch z.T. überraschende Spezifika des Untersuchungsgebietes nachgewiesen werden.

Using RapidEye imagery and the RandomForest algorithm for spatial modelling of Leaf Area Index values. Methodological considerations and a case study from tropical China

Philip Beckschäfer¹, Lutz Fehrmann¹, Rhett D. Harrison², Jianchu Xu², Christoph Kleinn¹

¹Chair of Forest Inventory and Remote Sensing, Georg-August-Universität Göttingen, Germany

philip.beckschaefer@gmail.com

²World Agroforestry Center, ICRAF-Kunming Office, Kunming, Yunnan 650201, China

Abstract. Leaf Area Index (LAI) is frequently used as a both physiologically meaningful and ecologically important characteristic of vegetation classes, including forests; and landscape-scale maps of LAI may be relevant to the understanding of larger area ecosystem processes. Optimized towards vegetation analyses, RapidEye satellite imagery bears the potential to facilitate LAI mapping throughout large and remote areas. We derived 59 image features from RapidEye imagery (6 vegetation indices, 16 texture indices calculated at 3 spatial scales and the 5 genuine spectral image bands) and investigated the data's potential to predict LAI values for a landscape in the uplands of Xishuangbanna, Yunnan Province, China. As ground reference we used field observations of LAI derived from hemispherical photographs taken on 252 locations along a gradient, ranging from heavily disturbed shrub land to mature montane rainforests. For predictive mapping a RandomForest model was used. Model calibration was based on the reference data. Boruta analysis revealed that only 22 out of the 59 image features were relevant for mapping LAI. Among these, vegetation indices were particularly important. Nevertheless, maps had a significantly higher accuracy - quantified by the mean absolute error obtained from a 10-fold cross validation - if not only vegetation indices but also texture features were used for mapping. Allowing the user to gain a better understanding of the uncertainties associated to the map, we provided a second map depicting the spatial distribution of the goodness of fit of the model – quantified by the mean absolute error (MAE). From this “uncertainty map” an area weighted MAE = 0.35 was calculated and compared to the unweighted MAE of 0.29.

Extension of k-most similar neighbours methods by local polynomial regression

Sebastian Schoneberg

Northwest German Forest Research, Germany

Sebastian.Schoneberg@NW-FVA.de

Abstract. Nearest neighbors techniques are used to predict timber volume per ha on forest compartment level (Abteilungen) based on sample plot data from a regional inventory and different sets of predictor or auxiliary variables. The auxiliary variables were derived from colour infrared (CIR) data. As spatial prediction methods k-most similar neighbours (k-MSN) were considered, as well as a new so-called local polynomial regression (LPR) approach. Actually, the latter is a k-nearest neighbour method with varying k and distance dependent weights for the nearest neighbours. It turned out that the LPR with Gaussian or uniform kernel function led to moderately higher precision than the k-MSN.

Model uncertainty in ALS-assisted forest inventories

Philip Mundhenk

Chair of Forest Inventory and Remote Sensing, Georg-August University of Göttingen,
Germany

pmundhe@gwdg.de

Abstract. The past twenty years have demonstrated a great potential for Airborne Laser Scanning (ALS) technologies to improve the accuracy and efficiency of forest inventories. ALS applications in forestry have targeted the estimation of aboveground biomass in Mg ha^{-1} , timber volume ha^{-1} , and basal area ha^{-1} , to name just a few. ALS data can be integrated into estimation in either a model-dependent (MD), or model-assisted (MA) framework. Both approaches require the formulation of a model that relates field observations of a target variable to so-called ALS metrics. The number of available metrics, i.e., explanatory variables, is often large in ALS applications, and an analyst has to decide which variables to include in the model. Statistical variable selection procedures as, for example, the AIC(c), maximum R^2 , or best-subset selection, are commonly used in practice to identify the final working model. However, the form of the model is likely to change from sample to sample. As, for MA approaches, the precision of an estimate depends on the model's residuals, the pursuit of “best possible fit” of a regression model to a single sample may lead to an overestimation of precision of a target parameter. We used data from forest inventories in western Alberta, Canada, and south-eastern Norway, to evaluate the impact of different variable selection procedures on MA inference. Synthetic populations were simulated using C-vine copulas. These populations were subsequently used in a Monte-Carlo experiment. Results show that standard variable selection procedures lead to an overestimation of precision of above-ground biomass estimates by up to 25%. Multi-model inference and Bayesian model-averaging approaches coupled with the integration of information from previous studies provide more reliable estimates of precision, i.e., differences of about 2% between empirical and mean estimated standard errors. We, therefore, recommend their use in ALS-assisted forest inventories when the model form is unknown prior to sampling/estimation.

A site index model for Norway spruce in Bavaria

Susanne Brandl

Bavarian State Institute of Forestry
Susanne.Brandl@lwf.bayern.de

Abstract. The growth potential of Norway spruce is modeled with environmental conditions as explanatory variables. The project focuses on modeling site index (SI). In addition the growth of biomass is modeled as well. a) Environmental dependent modeling of site index: aim of the project is the development of site index maps, which allow estimations of growth and comparisons of different sites. The data of the national forest inventories BWI 1, BWI 2 and BWI 3 are available. The modeling is based on an approach of Albert and Schmidt (2009). The relationship between height and diameter is described by the function of Korf. The parameters of this function can vary with environmental conditions. In order to determine site index at age 100 diameter at age 100 must be known. Therefore diameter is modeled dependent on age, allowing this relationship to vary with environmental conditions (Albert et al. 2012). By combining these two models SI can be estimated. Alternatively SI is directly calculated from BWI-data and subsequently used as response variable in a regression model with environmental conditions as explanatory variables (Nothdurft et al. 2012). It is intended to compare the results of both variants. b) Environmental dependent modeling of biomass growth: As sites with the same SI can still differ in biomass growth, it is of interest not only to model site index depending on environmental variables but also to model the relationship between biomass growth and environmental conditions. Therefore annual biomass increment is calculated from the BWI-data for each plot and explained by environmental variables. Special attention is paid to the selection and integration of environmental variables. Different ways to obtain indices of water balance from daily climatic data are investigated. Another aim is to determine threshold values of soil nutrients, below which nutrient supply can be limiting.

References

- Albert und Schmidt. 2009. Ein longitudinales Höhen-Durchmesser-Modell für Fichte in Deutschland. Tagungsband der Jahrestagung der Sektion Ertragskunde im DVFFA. 69-82.
- Albert et al. 2012. Konzept zur standort- und behandlungssensitiven Modellierung des Durchmesserzuwachses. Tagungsband der Jahrestagung der Sektion Ertragskunde im DVFFA. 130-139.
- Nothdurft et al. 2012. Konzept zur standort- und behandlungssensitiven Modellierung des Durchmesserzuwachses. *Forest Ecology and Management*, 279, 97-111.

A naïve Bayes model to describe natural forest ground vegetation by waveform LiDAR

Hooman Latifi, Viktor Gaub, Mrcó Heurich and Stefan Dech

University of Würzburg/DLR, Germany

hooman.latifi@uni-wuerzburg.de

Abstract. Though understory layer in forest (including the herbal species) contributes minimally to the total biomass available within the stands, it is one of the most important representatives of species richness as well as a widely-accepted criterion for characterizing plant communities. Furthermore, the ground vegetation is a main source of nutrients used as forage by ungulate animals, as well as a potential provider of contributing fuel to forest fire hazard. As a result of being masked by the overstory species, the most types of remote sensing data are basically unable to penetrate down to the forest ground layer, which in turn hinders their use for studying the forest surface material. We applied 3D understory metrics from a recent waveform Light Detection and Ranging (LiDAR) survey to describe the dominant cover type classes of ground vegetation within a portion of Bavarian Forest National Park (BFNP) in Germany. We assumed the independence of input predictors (e.g. vertical distribution ratio, understory height, shrub ($n(\text{height}<0.5\text{m}) / n(\text{height}<5\text{m})$), and cover ($n(\text{height}<2\text{m}) / n(\text{height}<60\text{m})$ for all laser returns) by an arbitrary selection as well as by applying a Stepwise forward predictor selection using the Wilk's Lambda. The conditional posterior probabilities were computed by a Naïve Bayes classifier incorporating a Laplace smoothing. The computed posterior was used for an area-based prediction of ground vegetation cover. The preliminary results show that the class-based conditional distributions along the training dataset can be fairly used for making new predictions. Further experiments will presumably unveil the use of naïve Bayes rule for describing ground forest vegetation under its more or less strict underlying assumptions. The focus is currently being kept on this. Keywords: Conditional posterior, Ground vegetation, LiDAR, naïve Bayes, predictive modelling, understory.

Estimating mean timber volume on stand level using unit-level and area-level small area estimators

Johannes Breidenbach and Ramus Astrup

Norwegian Forest and Landscape Institute, Germany

job@skogoglandskap.no

Abstract. Estimating the properties of small domains such as forest stands is a traditional field of use of airborne laser scanning (ALS) in the Nordic countries. The statistical methods used in this field of application are subsumed under the term small area estimation (SAE). Two important classes of model-based SAE methods are unit-level and area-level approaches. In the unit-level approach, observed (field) data need to be exactly co-registered with the auxiliary data (e.g., ALS covariates) in order to derive a regression model. Exact coordinates of the field data (e.g., sample plots) are needed in this case. The regression model is used to map the response variable, such as biomass or timber volume. The mean of the estimated map pixels within a stand is known as the synthetic estimate which is currently the most commonly used SAE technique in Nordic inventory practise. In the area-level approach, direct estimates of the response variable need to be available for some domains in order to derive a regression model on the domain level. In forest inventories, these data can be obtained by clustering sample plots within some stands. Exact coordinates of the sample plots are not required. Estimates are made directly on the stand level using the regression model. Model-based variance estimators are defined for both, the unit-level and area-level approach. Stand-level mean timber volume on stand level was estimated using ALS covariates for a study site in southern Norway. The field data consisted of 184 sample plots with known coordinates clustered in 30 stands. The estimated stand-level variances of the area-level EBLUP estimator were more variable but in the same range as the variances of the unit-level EBLUP estimator. Both estimators were in almost all cases considerably more efficient than direct estimates. While the unit-level approach may be more intuitive for a forester with experience in ALS-based inventories, area-level SAE methods can be expedient in areas where exact coordinates of the field sample plots are difficult to obtain or where stand-level estimates are available anyways.

European Forestry Dynamics Model (EFDM)

Ambros Berger

Federal Research and Training Centre for Forests, Natural Hazards and Landscape, Austria
ambros.berger@bfw.gv.at

Abstract. The EFDM is being developed by Ola Sallnäs from the Southern Swedish Forest Research Centre, SLU, in Alnarp. The aim of this project is not only to model forest growth but also to predict the effects of management decisions. The EFDM is basically a Markov chain with transition probabilities from one class to another. These classes are for example standing stock volume, stems per ha, or age and the model can be stratified according to sea level or dominant tree species. A high number of classes and strata results in many different possible transitions and thus very little data for the estimation of each transition. This is countered by a Bayesian method for the calculation of the transition probabilities which at first ignores the stratification completely and then step by step introduces the different strata. The results from the previous step are used as prior for the next. A first international test (participating countries: Austria, Finland, France, Portugal, Sweden) that did not yet include management activities showed promising results in predicting forest development for 100 years. Acknowledgements: Dr. Ola Sallnäs, SLU, 230 53 Alnarp, Sweden, Dr. Tuula Packalen, Metla, Yliopistokatu 6, FI-80100 Joensuu, Finland and Dr. Thomas Gschwantner, BFW, Seckendorff-Gudent-Weg 8, 1131 Wien, Austria.