Applied Statistics in Ecological and Environmental Sciences

Workshop

including a Tutorial by
Simon Wood, University of Bath
on Generalized Additive Models (GAMs)

organized by
Working group Ecological and Environment
International Biometric Society - German Region

November 7–11, 2012
Forest Research Institute Baden-Württemberg
Freiburg, Germany

Wednesday, November 7

1.00 - 2.30 pm	Tutorial, Simon Wood
2.30 - 3.00 pm	Coffee break
3.00 - 3.30 pm	Yvonne Morgenstern Analysing spatial and temporal patterns of soil water in small forest areas by using GAMs
3.30 - 4.00 pm	Carina Sucker Impact of forest management practices on stream water quality: Results of generalized additive mixed models (GAMMs)
4.00 - 4.30 pm	Karl Mellert Spatial Modeling of forest (site) properties - some practical experiences with GAM
4.30 - 5.00 pm	Stefan Lang Multilevel Structured Additive Regression
7.00 pm	Dinner

Thursday, November 8

9.00 -9.30 am	Nicole H. Augustin Space-time modelling for trend estimation of natural resources
9.30 -10.00 am	Alex Griffiths Soil structure and defoliation in Norway spruce
10.00 -10.30 am	Matthias Schmidt Modelling the grub density of the May beetle (Melolontha hippocastani) in the area of the Hessisch Ried
10.30 -11.00 am	Coffee break
11.00 - 11.30 am	Arne Nothdurft Spatio-temporal prediction of site index based on forest inventories and climate change scenarios
11.30 -12.00 am	Matthias Albert Using shape constrained additive models (SCAM) to quantify climate and site effects on forest productivity
12.00 am -1.00 pm	Lunch
1.00 -2.00 pm	Tutorial, Simon Wood
2.00 -5.00 pm	GAM exercises, group work with own laptops
7.00 pm	Dinner

Friday, November 9

9.00 -10.00 am	Tutorial, Simon Wood
10.00 -10.30 am	Coffee break
10.30 -11.00 am	Sebastian Schoneberg Nonlinear mixed modelling of tree height growth
11.00 -11.30 am	Tim Ritter Correcting the nondetection bias of angle count sampling
11.30 -12.00 am	Ronny Peters Revealing the invisible half: model based identification of below-ground competition in plant communities

Contents

Analysing spatial and temporal patterns of soil water in small forest areas by using GAMs

Yvonne Morgenstern

Forest Research Institute Baden-Württemberg
yvonne.morgenstern(at)forst.bwl.de

Abstract. Models for the simulation of the soil water dynamics almost always need data for a model calibration. However, soil moisture data are usually not available in the time and space resolution desired for a thorough model calibration. We used the measuring concept of "Randomised Moving Plots" (RMP) to observe the spatial and temporal variability of soil moisture in four forested areas. In each area (8-10 ha) 12 plots (circle with radius of 40 meters) with different properties of forest stand, slope and aspect positions were chosen. Soil moisture was measured in each of these plots simultaneously at 31 positions over a period of 14 days in half an hour time steps. Measuring positions were changed randomly every 14 days. We observed a total of 2356 14 day-time series of soil moisture from 2008 to 2010. 1781 datasets can be used for statistical analysis. For each measuring position, various site parameters (e.g. morphology, soils, vegetation) were described. Site parameters which significantly influence the spatial patterns of the soil moisture were included in a multiple linear regression model which can describes 35% of the spatial variability of soil moisture. One continuously measuring soil moisture probe in each investigated area was used to relate the spatial variability observed in the 14 day-time series to the overall temporal dynamics of soil moisture. Also time series of climate properties were measured continuously in each region. With this information, a continuous statistical time-space model of the soil moisture in our investigated areas is obtained. If the seasonal development (temporal dynamics) and the spatial patterns (observed in the RMPs) are combined, a space-time-model of the soil moisture will be developed. So, a general additive model (GAM) seems to be a good analysing tool for such a space-time model. I will present the dataset and first analysis with the use of a GAM at this workshop.

Impact of forest management practices on stream water quality: Results of generalized additive mixed models (GAMMs)

Carina Sucker

Forest Research Institute Baden-Württemberg
carina.sucker(at)forst.bwl.de

Abstract. Forested catchments are considered to guarantee a high quality of surface and drinking water. However, during the past decades, anthropogenic deposition has led to a continuous acidification of forest soils which has negative effects on stream water quality. The acidification processes can be influenced by forest management practices (liming and stand structure) as well as natural disturbances of the stand structure (storms and bark beetle infestation). The aim of my research is to determine trends of stream acidification depending on management practices, deposition trends and site conditions on the basis of statistical analyses. These should allow for a comparatively quick and straightforward evaluation of the effectiveness of the forest measures in catchments where no intensive field observations of soil and water chemistry are available. The analyses are carried out over a broad group of 86 water catchments in Germany and France where various forest management measures have been applied. Within the observation period, which starts 1980 and ends 2009 I modeled stream water quality data from a broad database (descriptive catchment factors, soil characteristics, geology, meteorology, deposition, run-off information and potential forest management drivers). I used generalized additive mixed models incorporating scale invariant tensor product smooths of the space-time dimensions. At the biometry workshop, I will present the results of GAMMs which help to identify the contribution of influencing factors on water quality.

Multilevel Structured Additive Regression

Stefan Lang

University of Innsbruck
stefan.lang(at)uibk.ac.at

Abstract. Models with structured additive predictor provide a very broad and rich framework for complex regression modeling. They can deal simultaneously with nonlinear covariate effects and time trends, unit- or cluster-specific heterogeneity, spatial heterogeneity and complex interactions between covariates of different type. In this paper, we propose a hierarchical or multilevel version of regression models with structured additive predictor where the regression coefficients of a particular nonlinear term may obey another regression model with structured additive predictor. In that sense, the model is composed of a hierarchy of complex structured additive regression models. The proposed model may be regarded as an extended version of a multilevel model with nonlinear covariate terms in every level of the hierarchy. The model framework is also the basis for generalized random slope modeling based on multiplicative random effects. Inference is fully Bayesian and based on Markov chain Monte Carlo simulation techniques. We provide an in depth description of several highly efficient sampling schemes that allow to estimate complex models with several hierarchy levels and a large number of observations within a couple of minutes (often even seconds). We demonstrate the practicability of the approach in a complex application on childhood undernutrition with large sample size and three hierarchy levels.

Space-time modelling for trend estimation of natural resources

Nicole H. Augustin

Department of Mathematical Sciences, University of Bath, Bath, UK
n.h.augustin(at)bath.ac.uk

Abstract. We present two applications of the use of Generalized Additive Mixed Models (GAMMs) with a space time interaction represented via a tensor product of different basis functions most suitable for the dimensions of space and time. The space-time smoother allows separate smoothing parameters and penalties for the space and time dimensions and hence avoids the need to make arbitrary or ad hoc choices about the relative scaling of space and time. The first application is for fisheries stock management. We model catch data from the blue ling fishery off the northwest coast of Scotland, using GAMMs with a space time interaction represented via a tensor product of a soap film smooth of space with a penalized regression spline of time. The use of soap film smoothers avoids imposing correspondences between spatially adjacent areas that are in fact separated by the stock boundary. The second application is for forest health monitoring. We use a GAMM to model defoliation of beech trees in Baden-Württemberg. The temporal trend of defoliation differs between areas because of site characteristics and pollution levels, making it necessary to allow for space-time interaction in the model.
This is joint work with Stefan Meining, Klaus von Wilpert, Verena Trenkel, Pascal Lorance and Simon Wood.

References

Augustin, N.H., Musio, M., von Wilpert, K., Kublin, E., Wood, S. & Schumacher, M. 2009. Modelling spatio-temporal trends of forest health monitoring data. *Journal of the American Statistical Association*. **104**(487), 899–911.

Soil structure and defoliation in Norway spruce

Alex Griffiths

University of Bath, UK
aeg32(at)bath.ac.uk

Abstract. The aim of this work was to determine the effect of soil structure and composition on defoliation in Norway spruce (Picea abies), using data collected since 1983 as part of the Terrestrial Crown Condition Inventory in Baden-Württemberg. We use a generalized additive mixed model (GAMM) to account for spatial and temporal correlations (Augustin, et al. 2009). Some soil characteristics were measured at several depths, and we extend the model to allow for functional covariates, permitting the effect size to vary with depth. We find that higher carbon content had a protective effect when present in lower layers of the soil (60-90cm), though not in surface layers. In addition, in areas which had been limed in response to high defoliation, a strong protective effect begins to appear after about 30 years. Other soil characteristics Ű such as clay, silt and sand content and depth of soil development Ű have little or no effect on defoliation.

References

Augustin, N.H., Musio, M., von Wilpert, K., Kublin, E., Wood, S. & Schumacher, M. 2009. Modelling spatio-temporal trends of forest health monitoring data. *Journal of the American Statistical Association*. **104**(487), 899–911.

Modelling the grub density of the May beetle (Melolontha hippocastani) in the area of the Hessisch Ried

Matthias Schmidt and Rainer Hurling

Nordwestdeutsche Forstliche Versuchsanstalt (NW-FVA)
matthias.schmidt@nw-fva.de

Abstract. Die Wälder im Ballungsraum Rhein-Main gehören zu den forstlichen Brennpunkten in Mitteleuropa. Die auSSergewöhnlichen Ansprüche an den Wald und die Nähe zum Rhein-Main-Gebiet sind durch einen überdurchschnittlich hohen Flächenverbrauch, hohe Schadstoff-immissionen und durch einen stetig ansteigenden Wasserbedarf gekennzeichnet. Die abiotischen Belastungen und hier insbesondere die gravierenden Grundwasserabsenkungen haben die Waldökosysteme mittlerweile auf groSSer Fläche soweit geschwächt, dass in der Folge massive biotische Schäden durch den Waldmaikäfer, Borkenkäfer und Schwammspinner auftreten. Zur Abschätzung eines räumlich differenzierten Gefährdungspotentials durch den Waldmaikäfer führte die Abteilung Waldschutz der Nordwestdeutschen Forstlichen Versuchsanstalt 2009 eine flächendeckende und systematische Stichprobe von Engerlingsgrabungen im 'Hessischen Ried' durch. Auf der Basis dieser Punktinformationen wurden verschiedene räumlich explizite Regressionsmodelle für Zähldaten parametrisiert, die sich bezüglich der unterstellten Verteilungsannahme unterscheiden. In einem Modellvergleich auf der Basis von randomisierten Quantilsresiduen wurde abschlieSSend ein negativ Binomial Regressionsmodell als am besten geeignet ausgewählt. Das Modell ermöglicht eine flächendeckende Regionalisierung der Engerlingsdichte unter gleichzeitiger Berücksichtigung von Standortseffekten. So weisen der Flurabstand und der Anteil von Tonschichten eines Standortes signifikante Effekte auf die Engerlingsdichte auf. Die Regionalisierung der Engerlingsdichte liefert in Kombination mit Bestockungsinformationen eine Grundlage für die räumliche Optimierung von Waldbau- und insbesondere ForstschutzmaSSnahmen. Die verwendete Methodik kann bei Vorliegen von Wiederholungsinventuren relativ einfach um einen Zeittrend erweitert werden.

Spatio-temporal prediction of site index based on forest inventories and climate change scenarios

Arne Nothdurft[1], Thilo Wolf[1], Andre Ringeler[2], Jürgen Böhner[2], and Joachim Saborowski[3]

[1] Forest Research Institute Baden-Württemberg
arne.nothdurft(at)forst.bwl.de
[2] University of Hamburg
[3] University of Göttingen

Abstract. A methodological framework is provided for quantification of climate change effects on site index. Spatio-temporal predictions of site index are derived for six major tree species in German state of Baden-Württemberg using simplified universal kriging based on large data sets from forest inventories and a climate sensitive site-index model. It is shown by simulations that, with large sample sizes, residual kriging using ordinary least squares estimates of the mean function leads to an approximately unbiased spatial predictor. Moreover, the coverage probabilities of resulting prediction intervals are close to the required level. B-spline regression techniques are applied to model nonlinear cause-and-effect curves for estimating site indexes at existing inventory plots dependent on retrospective climate covariates. The spatially structured error is modeled by exponential covariance functions. The mean model is applied to climate projection data to spatially predict the relative changes of site index under perturbed climate conditions. Applying climate projections based on IPCC emission scenarios A1B and A2, it is found that site index of all tree species would be decreased in lowland areas, and may increase in mountainous regions. Silver fir and common oak stands would also show increased site indexes in mountainous regions, but further extended to lower elevation levels. Site conditions in the Alpine foothills may remain highly productive for growth of Norway spruce. Whereas site index of common beech and Douglas-fir may decrease to almost the same relative amount and on nearly the same sites as Norway spruce, site index of Scots pine may be less affected by future climate change.

References

Nothdurft, A., Wolf, T., Ringeler, A., Böhner, J. & Saborowski, J. 2012. Spatio-temporal prediction of site index based on forest inventories and climate change scenarios. *Forest Ecology and Management*, 279, 97–111. http://dx.doi.org/10.1016/j.foreco.2012.05.018.

Using shape constrained additive models (SCAM) to quantify climate and site effects on forest productivity

Matthias Albert and Matthias Schmidt

Northwest German Forest Research Institute, Göttingen
matthias.albert(at)nw-fva.de

Abstract. We use shape constrained additive models (SCAM) to describe site-productivity relationships of tree species using climate and site variables. In contrast to approaches using generalized additive models (GAM) the partial effects of the climate variables temperature sum during growing season and aridity index during growing season on the dependent variable site index show biologically plausible characteristics. Nonetheless, also SCAM technology cannot prevent implausible effects when parameterising a model using all desirable explanatory variables in one step. Thus, we use a two step approach: firstly we fit a model using only the variables temperature sum and aridity index and secondly we use their offsets together with the remaining variables soil nutrients, available soil moisture, nitrogen deposition and a spatial trend to fit the final model. The applied method is discussed and results are presented.

Correcting the nondetection bias of angle count sampling

Tim Ritter

University of Göttingen
tim.ritter(at)forst.uni-goettingen.de

Abstract. The well known angle count sampling (ACS) proved to be an efficient sampling technique and has been applied in forest inventories for many decades. However, ACS assumes total visibility of objects, any violation of this assumption leads to a nondetection bias. We present a novel approach, in which the theory of distance sampling is adopted to traditional ACS in order to correct for the nondetection bias. Two new estimators were developed based on expanding design-based inclusion probabilities by model-based estimates of the detection probabilities. The new estimators were evaluated in a simulation study as well as in a real forest inventory. It is shown, that the nondetection bias of the traditional estimator is up to -52.5%, whereas the new estimators are approximately unbiased.

Revealing the invisible half: model based identification of below-ground competition in plant communities

Ronny Peters, Lin Yue and Uta Berger

Institute of Forest Growth and Computer Sciences, TU Dresden
ronny.peters(at)forst.tu-dresden.de

Abstract. Competition is a common basis for plant ecology and therefore a crucial concept of individual based modelling of plant populations. The mode of competition is the measure of how neighbouring individuals share the available resources. For example, the aboveground competition for light is assumed to be size-asymmetric since the taller plant is not affected by the smaller one. For below-ground competition, however, it is hardly possible to measure the competition mode up to now. The future goal of our research is to find a way to estimate the below-ground competition mode (BGCM) by indirect measurements. As a first step, we identified relevant variables, which fulfil the two principal constraints: firstly, they must be in relationship with the BGCM. It must be possible to estimate the BGCM with these variables. And secondly, they should be somehow easy to measure at an experimental site. These analyses were carried out with the individual based model PI (PI stands for Plant Interactions, Lin et al. 2012) and self-organizing feature maps (SOM). The PI model is based on metabolic energy use of each individual plant. Simulating the dynamics of a mono-specific plant population, the model was used to generate datasets of several parameter constellations and four different BGCM as model inputs and a multitude of output variables. Based on these datasets, SOM were used to set up empirical models for the relationship between those variables. A strong relation between two model-inputs nutrient limitation and BGCM and two model outputs Clark-Evans-Index and death rate was found. This provided a successful forecast of unknown BGCM based on information about nutrient availability, spatial plant distribution, and plant mortality for 'blind datasets'. The flexibility of this approach as well as its lack of predefined functional limits turned out to be very conducive for our goal. Its transfer to empirical systems is on the agenda.