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], Joachim Saborowski[3]

[1]Forest Research Institute Baden-Württemberg, Department of Biometrics and Informatics
[2]University of Hamburg, Institute of Geography
[3]University of Göttingen, Department of Ecoinformatics, Biometrics and Forest Growth and Department of Ecosystem Modelling

7-9 November 2012, Freiburg

Motivation: Spatio-temporal prediction of site index

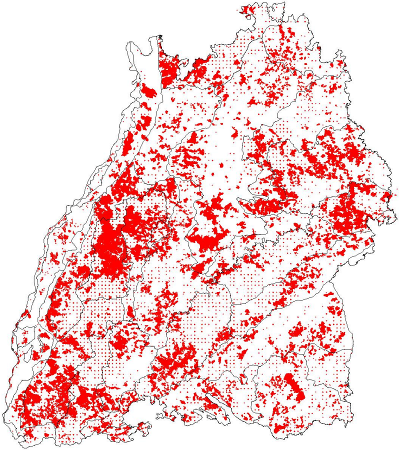
Definition (Site index)

A relative measure of forest site quality based on the height of the dominant trees at a reference age.

Goals:

- Reformulation of the existing site-productivity model of Schöpfer & Moosmayer (1972).
- Application of comprehensive data from regional and national forest inventories.
- To provide site-index predictions based on climate scenario data.

Dominant height measurements on forest inventory sample plots



Species	Sample plots
Spruce	91,642
Fir	33,987
Douglas-fir	9,356
Pine	31,559
Oak	28,230
Beech	81,241

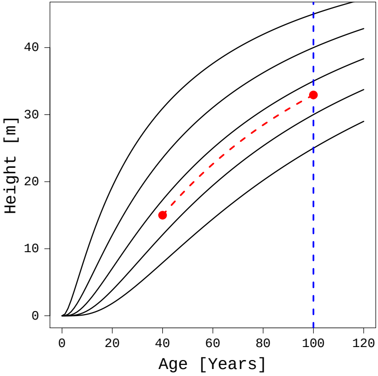
Climate data

We use

1. retrospective climate data for regression modeling,
2. climate scenario projections for spatio-temporal predictions.

Dominant height growth: Sloboda's (1971) differential equation

$$h\left(t_o|h(t),t\right) = \psi_1 \left(\frac{h(t)}{\psi_1}\right)^{\exp\left[\frac{\psi_2}{(\psi_3-1)t_o^{(\psi_3-1)}} - \frac{\psi_2}{(\psi_3-1)t^{(\psi_3-1)}}\right]}$$



The spatial process of site index

The spatial process of site index is defined for any arbitrary location s as

$$y(s) = x(s)'\beta + b(s) + \epsilon(s) \ ,$$

and is composed of

(i) a fixed linear trend dependent on site and climate variables in $x(s)$,

(ii) a spatially correlated error in terms of a Gaussian process $b(s)$, with 0 mean, variance θ^2, and spatial autocovariance function $c(h)$, h being the distance between two locations s_i and s_j,

(iii) and an unstructured error $\epsilon(s) \sim N\left(0, \sigma^2\right)$.

$\Rightarrow$ universal kriging (UK) model

Reformulation of the UK model as mixed model

The UK model can be reformulated as mixed model

$$y = X\beta + Ub + \epsilon$$

having covariance matrix

$$\mathbb{C}\mathrm{ov}(y) = \Sigma = \theta^2 URU' + \sigma^2 I_n$$

and correlation function

$$\mathbb{C}\mathrm{orr}\left[b(s_i), b(s_j)\right] = R = \rho(s_i - s_j) \ .$$

Spatial autocorrelation

The exponential correlation function

$$\rho\left(h\right) = \exp\left(-h/\alpha\right)$$

leads to the covariance function

$$c\left(h\right) = \begin{cases} \theta^2 \exp\left(-\frac{h}{\alpha}\right), & \text{for } h > 0 \\ \sigma^2 + \theta^2, & \text{for } h = 0 \end{cases}$$

and to the parametric semi-variogram model

$$\gamma\left(h\right) = \begin{cases} \sigma^2 + \theta^2 \left[1 - \exp\left(-\frac{h}{\alpha}\right)\right], & \text{for } h > 0 \\ 0, & \text{for } h = 0 \; . \end{cases}$$

Estimation: iteratively re-weighted generalized least squares (IRWGLS)

(I) Perform prior OLS $\hat{\beta}_{OLS} = (X'X)^{-1} X'y$ to estimate the mean model

(II) Iterate between (i) and (ii) until convergence is achieved

(i) Fit the parametric semi-variogram model to the empirical counterpart

$$\min \arg \sum_{u=1}^{K} \left\{2\gamma^{\sharp}(h(u)) - 2\gamma(h(u);\theta,\sigma,\alpha)\right\}^2$$

$$\hat{\gamma}^{\sharp}(h) = \frac{1}{2|N(h)|} \sum_{N(h)} [e(s_i) - e(s_j)]^2$$

$$N(h) \equiv \{(s_i;s_j) : s_i - s_j = h;\ i,j = 1,\ldots,n\}$$

$$\text{with } e = \begin{cases} y - X\hat{\beta}_{OLS}, & \text{in the first iteration} \\ y - X\hat{\beta}_{GLS}, & \text{otherwise} \end{cases}$$

(ii) Perform GLS $\hat{\beta}_{GLS} = \left(X'\hat{\Sigma}^{-1}X\right)^{-1} X'\hat{\Sigma}^{-1}y$

Estimation

Dilemma of large n

- Autocovariance matrix is high-dimensional and becomes intractable.
- Iteratively re-weighted generalized least squares (IRWGLS) is not feasible.

Hypotheses (given large data exists)

- Kriging of OLS residuals in moving 100 point neighborhoods leads to an unbiased spatial predictor.
- The error of the mean model can be neglected for constructions of prediction intervals.
- B-spline regression techniques provide robust estimates of biologically plausible cause-and-effect curves.

Prediction

Site-index predictions are spatially kriged at any location s_0 in moving 100 point neighborhoods by

$$\hat{y}_0 = x(s_0)'\hat{\beta}_{OLS} + \hat{c}_-'\hat{\Sigma}_-^{-1}\left(y_- - X_-\hat{\beta}_{OLS}\right) ,$$

"$-$" indicates the simplification of the universal kriging predictor in terms of a local predictor and $\hat{\Sigma}_-$ has dimension 100×100.
For interval prediction, the universal kriging variance

$$\begin{aligned}\mathbb{E}\left(y_0 - \hat{y}_0\right)^2 =& \sigma^2 + \theta^2 - c'\Sigma^{-1}c \\ &+ \left[x(s_0)' - c'\Sigma^{-1}X\right]\left[X'\Sigma^{-1}X\right]^{-1}\left[x(s_0)' - c'\Sigma^{-1}X\right]'\end{aligned}$$

is approximated by

$$\widehat{\mathbb{V}\mathrm{ar}}\left(\hat{y}_0\right) = \hat{\sigma}^2 + \hat{\theta}^2 - \hat{c}_-'\hat{\Sigma}_-^{-1} ,$$

in which the error of the estimator of β is neglected.
Constructions of 95%-intervals are obatained by

$$\hat{y}_0 \pm 1.96\sqrt{\widehat{\mathbb{V}\mathrm{ar}}\left(\hat{y}_0\right)} .$$

Mean model based on B-Spline regression techniques

A B-spline function is defined for the range of a vector of knots

$$\kappa = (\kappa_1, \ldots, \kappa_{l+1}, \ldots, \ldots, \kappa_{d-l}, \ldots, \kappa_d)^T \quad ,$$

and consists of d basis functions $\left(B_1^l(x), \ldots, B_d^l(x)\right)$ of degree l.
A single basis functions is defined for the range between $l+2$ consecutive knots and overlaps the $2l$ neighboring basis functions.
With given κ a basis function of degree $l = 0$ is defined by

$$B_v^0 = \mathbb{1}_{[\kappa_v, \kappa_{v+1})}(x) = \begin{cases} 1 & \text{for } \kappa_v \leq x < \kappa_{v+1} \\ 0 & \text{otherwise,} \end{cases} \qquad v = 1, \ldots, d-1 \ ,$$

and basis functions of higher order are recursively defined by

$$B_v^l(x) = \frac{x - \kappa_v}{\kappa_{v+l} - \kappa_v} B_v^{l-1}(x) + \frac{\kappa_{v+l+1} - x}{\kappa_{v+l+1} - \kappa_{v+1}} B_{v+1}^{l-1}(x) \ .$$

To guarantee that $\sum_{v=1}^{d} B_v^l(x) = 1$ holds for every $x \in [\kappa_1, \kappa_d]$, boundary knots are replicated l times, so that $\kappa_1 = \ldots = \kappa_{l+1}$ and $\kappa_{d-l} = \ldots = \kappa_d$.

Mean model based on B-Spline regression techniques

The entire regression matrix

$$X = (X_1, \cdots, X_p)$$

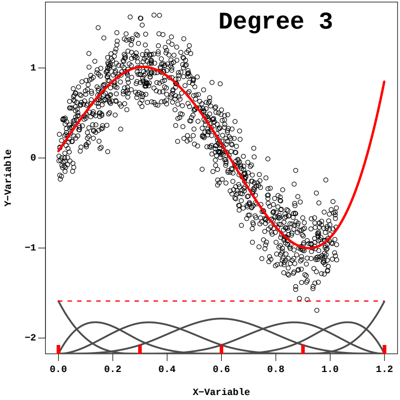
is composed of the p sub-matrices

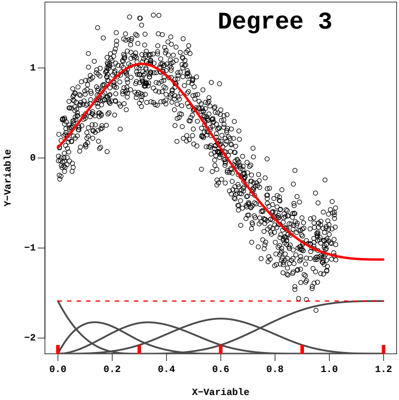
$$X_z = \begin{pmatrix} B_1^l(x_{z1}) & \dots & B_d^l(x_{z1}) \\ \vdots & & \vdots \\ B_1^l(x_{zn}) & \dots & B_d^l(x_{zn}) \end{pmatrix} .$$

Note: B-spline regression technique is nothing else than a multiple linear regression model!

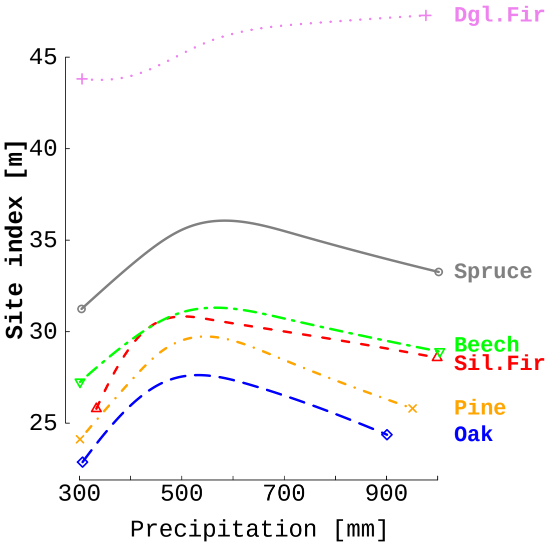
Mean model based on B-Spline regression techniques

Mean model based on B-Spline regression techniques and with natural boundary constraints.

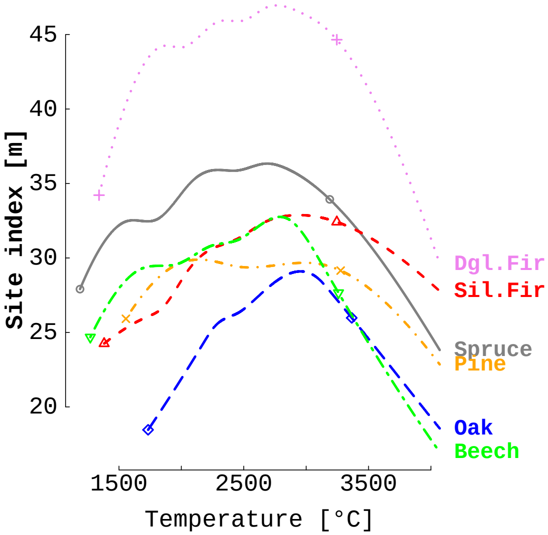




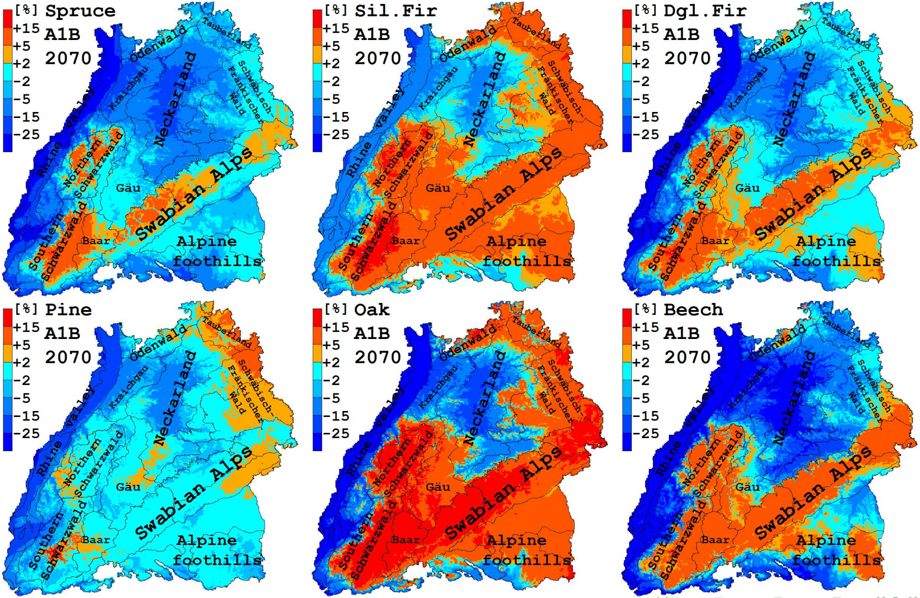
Mean model based on B-Spline regression techniques



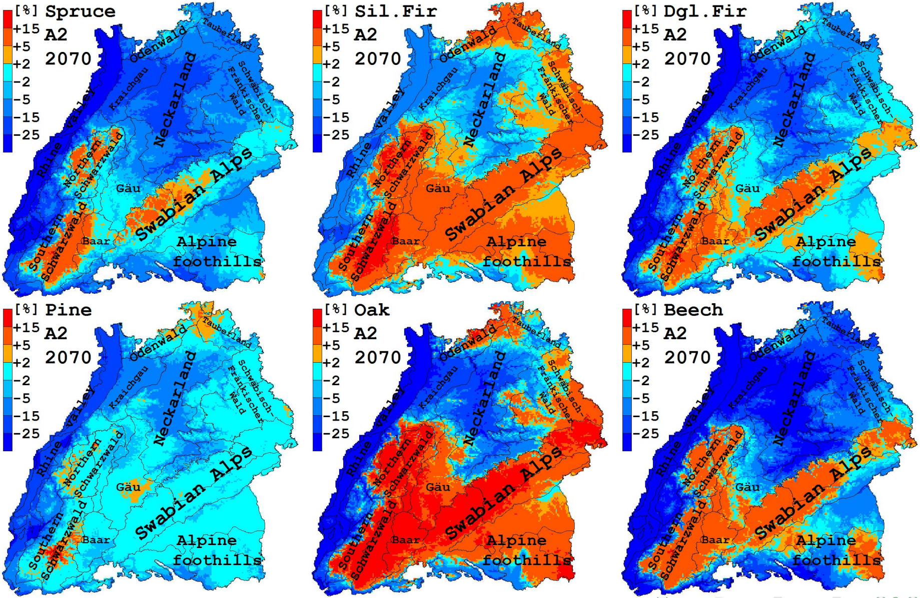
Mean model based on B-Spline regression techniques



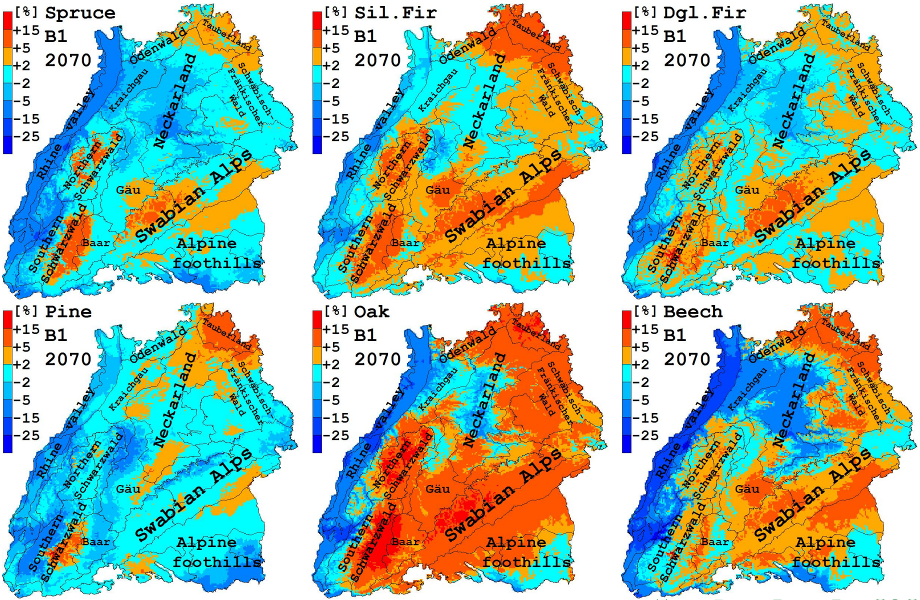
Predictions of relative changes of site index



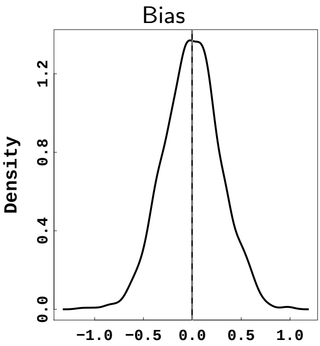
Predictions of relative changes of site index

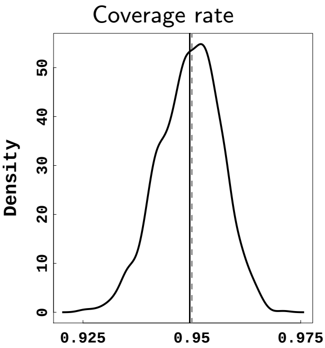


Predictions of relative changes of site index

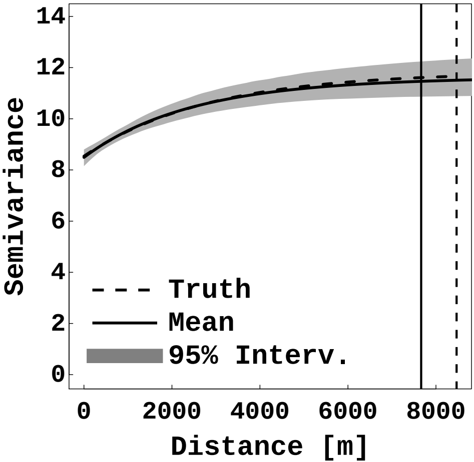


Validation based on 1000 simulations of Gaussian random fields





Validation based on 1000 simulations of Gaussian random fields



Résumé

- Site index decreases in low elevation areas and increases in mountainous regions.
- Silver fir and oak stands show increased site index also on lower elevation sites.
- Site index of Scots pine is less affected by a changing climate.
- Site conditions in the Alpine foothills region remain highly productive for growth of Norway spruce.

Résumé

If a universal kriging model is applied to a large forest inventory data set, then

- OLS is sufficient for the estimation of the mean and of the spatial covariance,
- spatial mean plus kriged OLS-residuals in moving 100 point neighborhoods is a practically unbiased predictor, and
- approximating the UK variance by neglecting the error of the mean yields a quasi-exact interval predictor.

Ende

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.

Thank you for your attentation !

Retrospective climate data

1. Lateral boundary conditions: NCAR/NCEP reanalysis series in resolution $2.5^\circ \times 2.5^\circ$ (lat./long.).
2. Dynamical downscaling in two steps: non-hydrostatic regional climate model (RCM) Weather Research and Forecasting (WRF) $\rightarrow$ 5 km $\times$ 5 km resolution.
3. Statistical downscaling: observations from the met-station network of the German National Meteorological Service $\rightarrow$ 50 m $\times$ 50 m resolution.
4. Bias correction and smoothing.

30-year long-term means (between 1978 and 2007) as regressor covariates:

- total precipitation during the growing season (PGS)
- total of average daily temperatures during the growing season (TGS)

Climate projection data

1. Existing runs of the hydrostatic RCM REMO on initiative of the German Federal Environment Agency
    - Lateral boundary conditions: general circulation model (GCM) ECHAM5-MPIOM.
    - $10\,\mathrm{km} \times 10\,\mathrm{km}$ horizontal grid resolution.
    - Scenarios A1B, A2 and B1 from Special Report on Emissions Scenarios (SRES).
2. Statistical downscaling $\rightarrow 1\,\mathrm{km} \times 1\,\mathrm{km}$: see retrospective climate data.
3. Spatio-temporal bias correction.