RESEARCH   **Open Access**

CrossMark

Incorporating shape constraints in generalized additive modelling of the height-diameter relationship for Norway spruce

Natalya Pya[1]* and Matthias Schmidt[2]

Abstract

Background: Measurements of tree heights and diameters are essential in forest assessment and modelling. Tree heights are used for estimating timber volume, site index and other important variables related to forest growth and yield, succession and carbon budget models. However, the diameter at breast height (dbh) can be more accurately obtained and at lower cost, than total tree height. Hence, generalized height-diameter (h-d) models that predict tree height from dbh, age and other covariates are needed. For a more flexible but biologically plausible estimation of covariate effects we use shape constrained generalized additive models as an extension of existing h-d model approaches. We use causal site parameters such as index of aridity to enhance the generality and causality of the models and to enable predictions under projected changeable climatic conditions.

Methods: We develop unconstrained generalized additive models (GAM) and shape constrained generalized additive models (SCAM) for investigating the possible effects of tree-specific parameters such as tree age, relative diameter at breast height, and site-specific parameters such as index of aridity and sum of daily mean temperature during vegetation period, on the h-d relationship of forests in Lower Saxony, Germany.

Results: Some of the derived effects, e.g. effects of age, index of aridity and sum of daily mean temperature have significantly non-linear pattern. The need for using SCAM results from the fact that some of the model effects show partially implausible patterns especially at the boundaries of data ranges. The derived model predicts monotonically increasing levels of tree height with increasing age and temperature sum and decreasing aridity and social rank of a tree within a stand. The definition of constraints leads only to marginal or minor decline in the model statistics like AIC. An observed structured spatial trend in tree height is modelled via 2-dimensional surface fitting.

Conclusions: We demonstrate that the SCAM approach allows optimal regression modelling flexibility similar to the standard GAM but with the additional possibility of defining specific constraints for the model effects. The longitudinal character of the model allows for tree height imputation for the current status of forests but also for future tree height prediction.

Keywords: Height-diameter curve, Norway spruce, Shape constrained additive models, Impact of climate change, Varying coefficient models

---

*Correspondence: natalya.pya@nu.edu.kz
[1]Department of Mathematics, School of Science and Technology, Nazarbayev University, 53 Kabanbay Batyr Avenue, Astana, Kazakhstan
Full list of author information is available at the end of the article

Background

Two of the main questions of forest management planning concern the current status of forests and how forests will develop in future. To estimate forest stock and assortment from sample forest inventories, for example, in forest districts or federal states, single tree volumes have to be predicted and then summed up to get timber volume estimates for a considered forest area. A tree volume estimate is usually based on three parameters: tree species, tree diameter and tree height. Since measuring tree diameter at breast height (1.3 m) (dbh), is relatively cheap, but measuring tree height is cost intensive, it is desirable to model tree height as a function of tree species, tree diameter, tree age and other possible stand- and site-specific parameters. An important feature of the height-diameter (h-d) relationship is that it develops over time and varies from stand to stand (Curtis 1967; Lappi 1997; Mehtätalo 2004). In Mehtätalo (2005) it is noted that trees reach maturity at different ages depending on site conditions. Hence, asymptotic height and the height that is reached at any particular age differ significantly among sites. The poorer the site conditions are, the lower the tree height will be for a certain age and dbh, with the dbh itself depending on age, stand and site conditions, but also on silvicultural treatments. Height of particular trees of a stand at predefined ages of usually 50 or 100 years is used as a measure for site quality and is denoted as 'site index'.

In this paper we develop site-sensitive longitudinal h-d models for forests in Lower Saxony, Germany, with the main focus on modelling fixed effects via unconstrained (GAM) and shape constrained generalized additive models (SCAM). Since climate change has already affected forests in Central Europe and much heavier impact is anticipated in the future, the models should be applicable for prediction of future tree height development and able to quantify the impact of climate change. Therefore, to achieve the necessary higher causality we use a combination of causal and proxy site parameters as predictors.

Many studies of forest research have been devoted to model the height-diameter relationship (see, e.g., Jayaraman and Lappi 2001; Eerikäinen 2003; Mehtätalo 2004; Sharma and Parton 2007; Schmidt et al. 2011). Several approaches are now available for height predictions. Those studies differ in the type of underlying principal h-d model used: linear (Lappi 1997; Eerikäinen 2003) or non-linear (Huang et al. 1992; Calama and Montero 2004; Castedo-Dorado et al. 2006; Sharma and Parton 2007). The principal h-d models also vary on how the model coefficients are being interpreted, which is especially important if they are then modelled as smooth functions of predictors. The approaches differ also in terms of the specification of the model effects. The effects are either assumed to be strictly linear or allowed for non-linear patterns for which spline techniques are commonly

applied (e.g., Schmidt et al. 2011). Finally, there are different procedures to account for spatial autocorrelation. This can be modelled via dummy fixed effects or uncorrelated random effects on the level of territorial units and stands (Jayaraman and Lappi 2001), Kriging methods (Nanos et al. 2004), a Markov random field smoother for estimating correlated random effects on the level of territorial units, or 2-dimensional smooth terms of the geographic location of the stands or sample plots (Schmidt et al. 2011).

In this study a general underlying modelling approach of a reparameterized version of the Korf-function, that was developed by Lappi (1997) is used as the principal model. The reason for using this model is that the model parameters considered there are less correlated and have biological meaning. Moreover, a heuristic fixation of the 'non-linear' parameters applied in this case linearizes the model, which makes the generalized additive model approach reasonable to use for the estimation of the covariate effects on the original parameters. The model is then extended to include some tree-specific and site-specific variables. As some of the covariate effects are supposed to be monotone, a shape constrained additive modelling (SCAM) approach (Pya and Wood 2015) is applied to account for influence of such variables as tree age, relative diameter at breast height and altitude among others, and also of site variables that will partially alter with expected climate change.

Data

The data analyzed here are observations from 23 145 sample plots of 29 324 Norway spruce trees [*Picea abies (L.) Karst.*] and some site-specific variables from the first cycle of the state forest enterprise inventories (district sample plot inventories) conducted by the Lower Saxony forest planning agency. Norway spruce is the most common and by far the most economically important species in Europe. Lower Saxony is the second largest federal state of Germany and is located in the north-western part. Every year two or three state owned forest districts are inventoried. The data come from inventories in the time interval 1996 – 2008. There are almost no consecutive inventories during this period (no longitudinal data), but all forest districts are inventoried, with the exception of a small area of the "Nationalpark Harz".

Two types of covariates are considered: tree-specific and stand- and site-specific. The tree-specific variables are tree diameter at breast height (dbh), tree age (age) and relative diameter at breast height (rel.dbh). The relative diameter at breast height is calculated as

$$\text{rel.dbh} = \text{dbh/mqd},$$

where mqd is the mean quadratic diameter of a sample plot. The rel.dbh is a measure of the rank of a tree

within all trees in a stand. A similar covariate is used by Eerikäinen (2003) who used the tree's `dbh` in relation to the `dbh` of a stand's dominant tree as predictor.

The second type of covariates, site-specific, can be differentiated into causal and proxy site variables. The proxy variables include altitude (`alt`), topex index (`topex.sw`), and geographic location, easting (`east`) and northing (`north`) in Gauß-Krüger coordinates referring to the 3rd meridian. The topex index describes topographic exposure and terrain morphology in the South-West direction. It is calculated as a sum of topographic exposure indices in the directions to the West, South-West and South using a distance limit of 250 meters (see, e.g., Scott and Mitchell 2005). A digital terrain model (DTM) with a resolution of 90 meters by 90 meters was used for topex calculation. A tree located on a summit is highly exposed resulting in a negative topex index. Positive topex indices belong to sites such as depressed areas or valleys rectangular-orientated in the direction of the topographic exposure. Topex indices of trees growing along the flat areas would be near zero. Since exposure to the South-West might result in drought stress, the topex index is used as a proxy for drought stress. Moreover, extra exposed sites will usually show a lower capacity of available soil water due to higher percentage of rocks and lower depth to parent rock.

The additional causal site (climate) explanatory variables are temperature sum of daily mean temperature during vegetation period (growing season) (`temp.veg`), and De Martonne's aridity index (`ari`). The aridity index is a fraction of annual precipitation in millimetres over mean annual temperature in degrees Centigrade plus ten (P/(T+10)) (De Martonne 1926; Thornthwaite 1931). The aridity index is calculated for the entire year, since the precipitation during winter (non-growing season) could be partially stored by the soil. `Temp.veg` and `ari` are retrospective simulation means (Spekat et al. 2007) of the normal climate period 1961–1990 that were regionalized from weather stations of the German weather service (DWD) using GAM with model effects for the geographic location and altitude. Table 1 summarizes the data under study.

Methods

A difficulty with the h-d relationship is that it is not constant but rather varies from stand to stand and develops over time (Lappi 1997; Mehtätalo 2004). In this paper we use an approach to modelling the longitudinal h-d relationship proposed by Schmidt (2010) that combines the principal h-d-model of Lappi (1997) with (unconstrained) generalized additive model technology as a starting point. The development of the h-d model consists of three steps: 1) initial specification of the h-d relationship as a log-linear mixed model with random stand effects, 2) 'a priori'

Table 1 Characteristics of Norway spruce trees and site parameters from the first cycle of all state forest enterprise inventories in Lower Saxony. 29 324 Norway spruce trees from 23 145 sample plots were observed

	Min	25 % qu.	Median	75 % qu.	Max
Tree height [m]	3.7	14.6	21.8	27	47.3
dbh [cm]	7	16.8	30.5	37.9	104
Tree age [years]	20	41	54	77	199
Altitude [m]	0	90	307	475.2	947
Sum of topographic exposure indices [°× 1000] (DTM 90 m × 90 m resolution)	–84 560	–3108	1489	8135	89 208
Temperature sum during the vegetation period [°C]	833.6	1716.4	1996.6	2196.5	2456.8
Aridity index	24.8	37	44.8	54.6	87.5

determination of non-linear model parameters, and 3) developing unconstrained and shape constrained generalized additive models for investigating potential tree and site specific effects on the original parameters of the modified Korf function (Lappi 1997).

The initial steps, 1) and 2), of the model development are briefly described in the following subsection.

Initial model development

A data base for the whole of Germany was applied for this 'a priori' estimation of specific model parameters. As a starting point, the following height-diameter model known as the Korf function is used for the description of the relationship between tree height and diameter (Lappi 1997):

$$\log(\mu_{ki}) = A_k - B_k \, (\text{dbh}_{ki} + \lambda)^{-C}, \qquad (1)$$

where $\mu_{ki} = E(H_{ki})$ and H_{ki} is a height of tree i on sample plot k, dbh_{ki} is the diameter at breast height of tree i on sample plot k; H_{ki} follows a Gaussian distribution; A_k, B_k, λ, and C are parameters of the model. The preliminary modelling showed that Gaussian models with the log link function performed better in terms of the Akaike information criterion (AIC) than Gamma models. Height-diameter curves differ for different plots and for different points of time, however, the measurement occasion effect was not included in the considered model. The reason behind it was the lack of computer memory as the whole data base contains several thousands of sample plots with on average only very few height measurements per measurement occasion. Therefore, the model parameters vary only over plots. Since parameters A_k and B_k are

highly correlated, it is suggested to reparameterize `dbh` as follows (Lappi 1997):

$$x_{ki} = \frac{(\text{dbh}_{ki} + \lambda)^{-C} - (30 + \lambda)^{-C}}{(10 + \lambda)^{-C} - (30 + \lambda)^{-C}}.$$

The model (1) can now be written as

$$\log(\mu_{ki}) = A_k - B_k x_{ki}, \tag{2}$$

where A_k and B_k are not highly correlated and have biological meanings. A_k is the expected value of the log height of trees with `dbh` = 30 cm for sample plot k; and B_k is the expected value of the difference in the $\log(H_{ki})$ between trees of `dbh` = 30 cm and 10 cm for sample plot k. These interpretations are important since the parameters will be described as functions of additional tree, stand and site-level covariates in the second step of the model development.

The model (2) is linear with respect to A_k and B_k. Taking into consideration the random stand effect, these parameters can be represented at the first stage as $A_k = A + \alpha_k, B_k = B + \beta_k$, where A and B represent fixed effects which have to be estimated; α_k and β_k are random stand level effects with zero means and constant variance. It may be noted that (2) is overparameterized. Moreover, a model of that specification cannot be linearized with respect to the parameters λ and C. Therefore, it is suggested firstly to estimate λ and C. These parameters were selected by testing a variety of combinations of λ and C when fitting a linear mixed model

$$\log(\mu_{ki}) = A - B x_{ki} + \alpha_k + \beta_k x_{ki},$$

The combination of the parameters with the lowest error variance was $\lambda = 7$ and $C = 1.225$. There were no clear trends found in λ and C over different mean stand age and the models were not very sensitive to the value C.

Additive model for tree height

One of the model requirements is to predict actual and future tree heights of a forest stand. Since every stand has different characteristics, effects of site and stand variables should be incorporated into the h-d model in combination with an age effect that describes the developmental stage of the trees within a stand. Since the proportion of structured and multi-aged stands in Lower-Saxony is constantly increasing we use single tree age as a covariate. The additional tree- and site-specific effects on the original parameters A and B of the Korf function that are partially sensitive to climate change, are assumed to be non-linear. Then, based on the principal h-d model

$$\log(\mu_{ki}) = A - B x_{ki}, \tag{3}$$

where the mean tree height can be modelled as a function of tree age and additional tree and site parameters using GAM (Hastie and Tibshirani 1990; Wood 2006a)

Model h1: unconstrained additive model

$$\begin{aligned}
\log(\mu_{ki}) = {}& \alpha_0 + f_{1a}(\text{age}_{ki}) + f_{2a}(\text{rel.dbh}_{ki}) \\
& + f_{3a}(\text{topex.sw}_k) + f_{4a}(\text{temp.veg}_k) \\
& + f_{5a}(\text{ari}_k) + f_{6a}(\text{east}_k, \text{north}_k) \\
& + p_{0b} \times x_{ki} + p_{1b} \times \text{age}_{ki} \times x_{ki} + p_{2b} \\
& \times \text{alt}_k \times x_{ki},
\end{aligned} \tag{4}$$

where x_{ki} is the re-parameterized `dbh` of tree i on sample plot k introduced at the initial step of the h-d model development, α_0 is the model intercept, p_{0b}, p_{1b} and p_{2b} are model coefficients. H_{ki} is assumed to follow a Gaussian distribution. The model terms f_{1a}–f_{5a} are unknown smooth functions of the corresponding predictor variables. We also added a spatial smooth function $f_{6a}(\text{east}, \text{north})$ of easting and northing, since there is a spatial correlation in the residuals. This unconstrained model assumes a linear combination of the covariate effects and due to the log-link, the effects act multiplicative exponentially on tree height.

In the above mentioned case the effects of age and altitude on the slope B of the h-d curve were assumed to be linear. Now, suppose that both predictors have non-linear effects on B. Then the following model may be considered:

Model h2: GAM with varying coefficients

$$\begin{aligned}
\log(\mu_{ki}) = {}& \alpha_0 + f_{1a}(\text{age}_{ki}) + f_{2a}(\text{rel.dbh}_{ki}) \\
& + f_{3a}(\text{topex.sw}_k) + f_{4a}(\text{temp.veg}_k) \\
& + f_{5a}(\text{ari}_k) + f_{6a}(\text{east}_k, \text{north}_k) \\
& + p_{0b} \times x_{ki} + f_{1b}(\text{age}_{ki}) \times x_{ki} + f_{2b}(\text{alt}_k) \\
& \times x_{ki},
\end{aligned} \tag{5}$$

where the non-linear effects of age and altitude are represented by the smooth functions $f_{1b}(\text{age})$ and $f_{2b}(\text{alt})$. Model h2 is referred to as a 'variable coefficient model' (Hastie and Tibshirani 1993; Wood 2006a).

The drawback of modelling with GAM is that it may result in insufficiently smooth effects of the covariates. Moreover, it is biologically plausible to expect that the effects of such covariates as `age`, `rel.dbh`, `topex.sw`, `temp.veg` and `ari` on the original parameter A will be monotone under the current growth conditions of Lower Saxony, which is not guaranteed for the GAM fit. Therefore, we propose to impose additional constraints on the univariate smooth terms by applying a SCAM approach (Pya and Wood 2015) described in the next subsection.

Modelling non-linear effects using SCAM

The first shape constrained model (model h3) considered is simply h1 as given in (4) with monotonicity restrictions described below on univariate smooth components,

Model h3: shape constrained additive model

$$\begin{aligned}
\log(\mu_{ki}) = {} & \alpha_0 + m_{1a}(\texttt{age}_{ki}) + m_{2a}(\texttt{rel.dbh}_{ki}) \\
& + m_{3a}(\texttt{topex.sw}_k) + m_{4a}(\texttt{temp.veg}_k) \\
& + m_{5a}(\texttt{ari}_k) + f_{6a}(\texttt{east}_k, \texttt{north}_k) \\
& + p_{0b} \times x_{ki} + p_{1b} \times \texttt{age}_{ki} \times x_{ki} + p_{2b} \\
& \times \texttt{alt}_k \times x_{ki}.
\end{aligned} \tag{6}$$

To distinguish from unconstrained smooths, smooth terms under monotonicity constraints are denoted by m_{ja}. The effect of age on the original parameter A in (3) is supposed to be increasing, since for any constant vector of model predictors, the level of the h-d curve, that is the expected $\log(H_{ki})$ of a tree with dbh = 30 cm, is assumed to be increasing with increasing age. The effect of rel.dbh on the original parameter A is expected to be monotone decreasing, since lower values of the rel.dbh correspond to a lower rank of a tree within a stand. Within the same stand a tree with a lower rank has on average a greater competition pressure compared to a tree with a higher rank. While struggling for the light, suppressed trees have to invest more into height than diameter growth. Hence, trees will be taller with the value of rel.dbh decreasing given fixed values of dbh, age and the additional covariates. Trees with high values of rel.dbh are dominant trees that are usually more exposed to the wind and consequently, they have to invest more into diameter than height growth for stability reason. Therefore, given any fixed covariate vector tree height is assumed to decrease with increasing values of rel.dbh. The effect of topex.sw on the original parameter A should be monotone increasing, since an exposure to the South West might result in drought stress as it was explained previously. We assume a monotone increasing netto assimilation with increasing temp.veg under the climatic conditions of Lower Saxony (if not limited by the deficit of other resources). The lower site indices of Norway spruce, that are partially observed on warmer sites of Lower Saxony, are, for instance, assumed to result from limited water and lower nutrient supply. The effect of temp.veg must not be confused with optimum curves that are observed under varying temperature values in experiments. Hence, no temperature optimum is assumed to be present under the current climatic conditions of Lower Saxony. The effect of ari on the original parameter A is expected to increase with increasing humidity. The lower site indices of Norway spruce that are partially observed on very humid sites in higher altitudes of the uplands, are assumed to be a result of limited temperature sums. Hence, ari and temp.veg are both assumed to have monotone increasing effects on the original parameter A, hence on the level of the h-d curve.

Next, we consider the shape constrained version of the variable coefficient model h2 as model h4.

Model h4: SCAM with varying coefficients

$$\begin{aligned}
\log(\mu_{ki}) = {} & \alpha_0 + m_{1a}(\texttt{age}_{ki}) + m_{2a}(\texttt{rel.dbh}_{ki}) \\
& + m_{3a}(\texttt{topex.sw}_k) + m_{4a}(\texttt{temp.veg}_k) \\
& + m_{5a}(\texttt{ari}_k) + f_{6a}(\texttt{east}_k, \texttt{north}_k) \\
& + p_{0b} \times x_{ki} + m_{1b}(\texttt{age}_{ki}) \times x_{ki} + m_{2b}(\texttt{alt}_k) \\
& \times x_{ki},
\end{aligned} \tag{7}$$

where the non-linear effects of age and alt on the slope B are represented by the smooth functions $m_{1b}(\texttt{age})$ and $m_{2b}(\texttt{alt})$. Increasing effects of both $m_{1b}(\texttt{age})$ and $m_{2b}(\texttt{alt})$ on the h-d relationship are assumed in this model. It is well known that the slope of the h-d relationship increases with the developmental stage of a stand (e.g., Mehtätalo 2004). In our investigation age serves as a covariate that describes the developmental stage of a stand. Therefore, when fitting a varying coefficient model for the age effect on B, it should be monotone increasing. However, the gradient of the actual tree heights that are predicted in applications is also affected by the dbh values that are used to initialize the model. The direction of the monotonicity of effect $m_{2b}(\texttt{alt})$ remains unspecified at this point and will be defined later based on the results of the unconstrained model variant. Moreover, for all the monotonicity constraints a validation of the assumptions will be conducted based on the corresponding unconstrained model effects.

When fitting model with monotonicity constraints on the effects of temp.veg and of ari, we noticed some possibly artificial sharp changes in the corresponding estimated smooths (see sec. 4.2). To avoid these limitations the shape constrained model is enhanced by concavity constraints on the smooth terms of temp.veg and of ari. We propose model h5 as a variable coefficient model since the performance of model h4 was shown to be better than of model h3 in terms of AIC and GCV scores.

Model h5: SCAM with concavity constraints

$$\begin{aligned}
\log(\mu_{ki}) = {} & \alpha_0 + m_{1a}(\texttt{age}_{ki}) + m_{2a}(\texttt{rel.dbh}_{ki}) \\
& + m_{3a}(\texttt{topex.sw}_k) + mc_{4a}(\texttt{temp.veg}_k) \\
& + mc_{5a}(\texttt{ari}_k) + f_{6a}(\texttt{east}_k, \texttt{north}_k) \\
& + p_{0b} \times x_{ki} + m_{1b}(\texttt{age}_{ki}) \times x_{ki} + m_{2b}(\texttt{alt}_k) \\
& \times x_{ki},
\end{aligned} \tag{8}$$

where now mc_{4a}, mc_{5a} are subject to both monotone increasing and concavity constraint.

The following basic initial model with only age effect on the original parameters A and B was used as a reference model which all the considered models were compared with.

Model h.ref:

$$\log(\mu_{ki}) = \alpha_0 + f_{1a}(\text{age}_{ki}) + p_{0b} \times x_{ki} + p_{1b} \\ \times \text{age}_{ki} \times x_{ki}. \tag{9}$$

Model estimation

To estimate the SCAM models (6), (7) and (8) we employ the penalized regression spline approach which can be split into two stages: representation of smooth model terms via penalized unconstrained and constrained regression splines along with specification of the smoothness/wiggliness penalty followed by model coefficients estimation by penalized log likelihood maximization along with smoothness parameter selection by minimization of a prediction error criterion such as AIC or GCV. Shape COnstrained P-splines (SCOP-splines) (Pya and Wood 2015) were used for representation of the shape constrained smooth model terms. Since the bivariate function $f_{6a}(\text{east}, \text{north})$ is a function of geographic coordinates, it was represented by a thin plate regression spline (Wood 2006a).

Combining the model matrices of each smooth column-wise into one model matrix and absorbing identifiability constraints result in the following expression of the SCAM model

$$\log(\mu_{ki}) = \mathbf{X}_{ki}\boldsymbol{\beta}, \tag{10}$$

where $\mathbf{X}$ is the combined model matrix of strictly parametric model components and smooth basis functions and $\boldsymbol{\beta}$ is a vector of unknown coefficients. After setting the penalties on each smooth model term which are expressed as quadratic forms of the full coefficient vector, $\boldsymbol{\beta}$, the penalized log likelihood maximization can be written as

$$l_p(\boldsymbol{\beta}) = l(\boldsymbol{\beta}) - \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta}/2,$$

where $l(\boldsymbol{\beta})$ is the log likelihood of the model, $\mathbf{S} = \sum_k \lambda_k \mathbf{S}_k$, and $\mathbf{S}_k$ are the smooth penalty matrices enlarged by zeros to be expressed in terms of the full vector of the model coefficients, λ_k are smoothing parameters. The model coefficients, $\boldsymbol{\beta}$, are estimated by $l_p(\boldsymbol{\beta})$ maximization given the values of the vector of smoothing parameters, λ. Optimization of the $l_p(\boldsymbol{\beta})$ is achieved by a Newton method which shares several features with a penalized iteratively re-weighted least squares scheme standard for GLM estimation. The smoothing parameter vector λ is estimated by minimizing the generalized cross validation score (GCV),

$$\mathcal{V}_g = nD(\hat{\boldsymbol{\beta}})/(n - \tau)^2,$$

where $D(\hat{\boldsymbol{\beta}}) = 2\left\{l_{\max} - l(\hat{\boldsymbol{\beta}})\right\}\sigma^2$ is the model deviance, $l_{\max}$ is the saturated log likelihood, $\hat{\boldsymbol{\beta}}$ is the vector of the model parameters estimates, and τ is the effective degrees of freedom.

Confidence intervals for the model smooth terms are obtained through the distributional results for $\hat{\boldsymbol{\beta}}$. The Bayesian approach to interval estimates for the smoothing spline models proposed by Wahba (1983) and Silverman (1985) was extended to generalized additive models by Lin and Zhang (1999) and Wood (2000). SCAM adopts this approach with an addition for establishing the approximate distribution of the exponentiated $\boldsymbol{\beta}$, denoted as $\tilde{\boldsymbol{\beta}}$, resulting in the normal distribution $\tilde{\boldsymbol{\beta}}|\mathbf{y} \sim N(\hat{\tilde{\boldsymbol{\beta}}}, \mathbf{V}_{\tilde{\boldsymbol{\beta}}})$, where the expression for the covariance matrix $\mathbf{V}_{\tilde{\boldsymbol{\beta}}}$ as well as all tedious details of the model parameters estimation can be found in Pya and Wood (2015). The SCAM approach is implemented in an R package `scam` available at http://CRAN.R-project.org/.

To fit the unconstrained models h1 and h2 we use the penalized regression spline approach (Wood 2006a). The univariate functions f_{2a}–f_{5a} of (4) and (5) and also the unconstrained effects f_{1b} and f_{2b} of model h2 (5) are represented by P-splines (Eilers and Marx 1996) whereas an isotropic two dimensional thin plate regression spline (Wood 2006a) was used for representation of f_{6a}. The standard penalized iteratively re-weighted least squares (PIRLS) scheme is applied for the model parameter estimation. The multiple smoothing parameter is selected by minimizing the GCV score in outer iterations. The Newton method is used for optimizing the GCV to update the smoothing parameter. The interval estimates for the component smooth functions of models h1 and h2 are obtained using the Bayesian approach to uncertainty estimation (Wahba 1983; Silverman 1985; Wood 2006b).

Results and discussion

Model selection

All covariates considered in the h-d models revealed their relevance to the tree height modelling. In addition we estimated possible submodels, where one at a time smooth effects were dropped. Table 2 presents the model fitting results (to keep the paper short, the results on the submodels are shown only for the models with one dropped smooth effect). The adjusted r^2 and GCV scores are included into the table. The last column of the table shows the percentage of improvement in the Akaike information criterion (AIC.diff.perc) in comparison with the reference model, h.ref, calculated as follows

$$\text{AIC.diff.perc} = \frac{\text{AIC}_{h.ref} - \text{AIC}_{h_j}}{\text{AIC}_{h.ref}} \times 100,$$

where $\text{AIC}_{h.ref}$ is the AIC of the reference model and AIC_{h_j} of the model under consideration. The best selected model in terms of the AIC is the shape constrained varying coefficients model h4 with all initial smooth effects included. The measures of the model performance of the model h2 are only slightly worse than those of h4. Adding the variable coefficients proposed in the GAM model h2

Table 2 Comparison of statistics for different height-diameter-models including a base model with only age effects (h.ref), the unconstrained additive model (h1), unconstrained additive model with varying coefficients (h2), shape constrained additive model (h3), shape constrained additive model with varying coefficients (h4), additive model with concavity constraints (h5). For all models the result of dropping single model effects on different model statistics are presented

Model	adj r^2	GCV	AIC.diff.perc
h.ref	.885	7.309	0
h1	.909	5.798	4.79
$h1-f_{2a}$	.907	5.883	4.49
$h1-f_{3a}$	.908	5.846	4.63
$h1-f_{4a}$	.908	5.842	4.64
$h1-f_{5a}$	.908	5.848	4.62
$h1-f_{6a}$	.900	6.324	2.99
h2	.909	5.784	4.85
$h2-f_{2a}$	.908	5.87	4.54
$h2-f_{3a}$	.908	5.832	4.68
$h2-f_{4a}$	.908	5.83	4.68
$h2-f_{5a}$	.908	5.837	4.66
$h2-f_{6a}$	.901	6.311	3.04
$h2-f_{1b}$	.907	5.916	4.38
$h2-f_{2b}$	.909	5.811	4.75
h3	.909	5.805	4.77
$h3-m_{2a}$	.907	5.887	4.48
$h3-m_{3a}$	.908	5.851	4.61
$h3-m_{4a}$	.901	6.290	3.11
$h3-m_{5a}$	.908	5.866	4.55
$h3-f_{6a}$	.901	6.316	3.02
h4	.909	5.778	4.87
$h4-m_{2a}$	.908	5.867	4.55
$h4-m_{3a}$	.909	5.812	4.75
$h4-m_{4a}$	.902	6.2582	3.21
$h4-m_{5a}$	.908	5.838	4.66
$h4-f_{6a}$	.899	6.382	2.81
$h4-m_{1b}$	.907	5.895	4.45
$h4-m_{2b}$	.907	5.914	4.39
h5	.907	5.877	4.52
$h5-m_{2a}$	.906	5.93	4.33
$h5-m_{3a}$	.907	5.865	4.56
$h5-m_{c4a}$	.901	6.302	3.07
$h5-m_{c5a}$	.907	5.860	4.58
$h5-f_{6a}$	.900	6.406	2.73
$h5-m_{1b}$	.906	5.96	4.22
$h5-m_{2b}$	.908	5.86	4.58

improves the unconstrained model h1, although to a lesser extent that it does in case of the SCAMs. Dropping either of the effects from any of the five considered models increases the AIC, with the exception of the three cases of the model h5 where the AIC slightly decreases. The other measures of the model performance such as the GCV and adjusted r^2 also give worse results than those of the full models h1-h5, when dropping any single effects. The spatial effect improves the model significantly: e.g., the models without spatial effect result in much higher GCV than the corresponding full model (about 24 % difference in the GCV in case of h2). Introducing stricter concavity constraints in model h5 leads to a slight increase in AIC and GCV, and correspondingly to a poorer model fit. It should be noted that there are only marginal differences in the performance criteria between the unconstrained GAM models h1 and h2, and their constrained counterparts, SCAM models h3-h5. The estimates and the corresponding standard errors of the coefficients of the linear part of the unconstrained model h1 and the shape constrained version h3 are shown in Table 3.
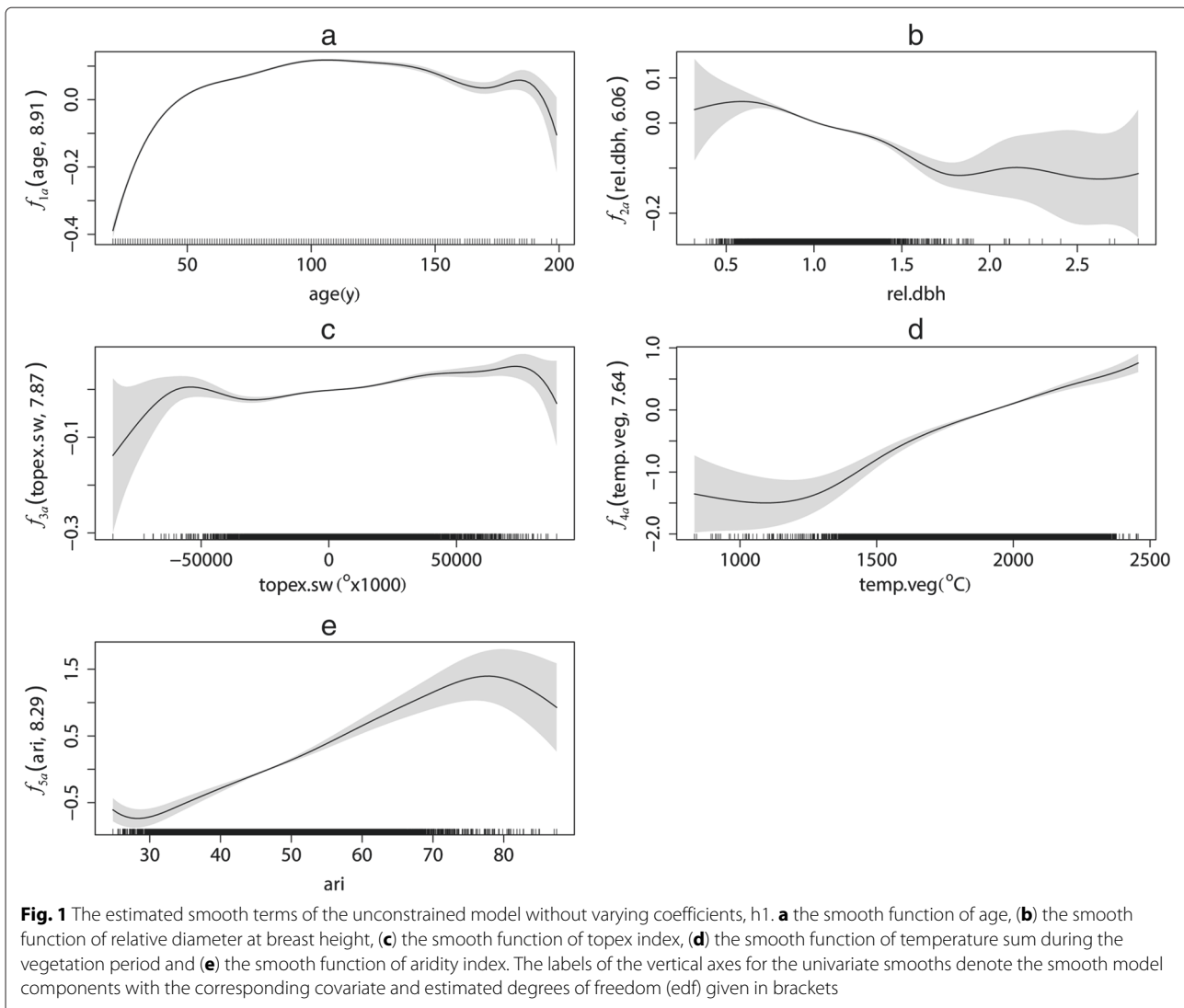
Interpretation of unconstrained effects and validation of their monotone counterparts

Overall, the monotonicity constraints on the univariate smooth terms result in less wiggly pattern compared to the unconstrained effects (see Fig. 2 versus Fig. 1). It should be noticed that the estimated effects of the shape constrained smooths are not centered as they are in the case of the unconstrained GAM, as different identifiability constraints were applied.

The estimated unconstrained effect of `age` on the original parameter A of model h1 is increasing with a decreasing gradient for almost the whole data range (Fig. 1a). However, for high ages, above 150 years, the effect is implausibly decreasing. This pattern probably occurred due to an unbalanced data structure for the combination of site index and age. It is typical for forests and especially managed forests that 'old stands grow on poor sites', since trees need longer production periods to reach merchantable timber dimensions. The proposed h-d models cover some site factors, e.g. `temp.veg`. However, a certain proportion of the variability in site quality probably

Table 3 Estimates of the coefficients of the linear parts of models h1 and h3. The corresponding standard errors are given in brackets

	Model h1	Model h3
Intercept	3.095(.0011)	-1.907(.399)
p_{0b}	.5654(.0084)	.606(.0072)
p_{1b}	.00354(1.42×10^{-4})	.00276(1.1×10^{-4})
p_{2b}	1.23×10^{-4}(1.31×10^{-5})	1.22×10^{-4}(1.32×10^{-5})

Fig. 1 The estimated smooth terms of the unconstrained model without varying coefficients, h1. **a** the smooth function of age, (**b**) the smooth function of relative diameter at breast height, (**c**) the smooth function of topex index, (**d**) the smooth function of temperature sum during the vegetation period and (**e**) the smooth function of aridity index. The labels of the vertical axes for the univariate smooths denote the smooth model components with the corresponding covariate and estimated degrees of freedom (edf) given in brackets

remains unquantified, which presumably leads to the implausible decreasing effect for high ages. The effect of `age` of model h3 is assumed to be monotone increasing, so that at high ages the estimated smooth tends to a constant guaranteeing a plausible pattern over the whole data range (Fig. 2a).

The estimated unconstrained effect of `rel.dbh` of model h1 (Fig. 1b) supports the imposition of a monotone decreasing constraint on the function $f_{2a}(\texttt{rel.dbh})$ when constructing model h3. The confidence intervals of f_{2a} near both boundaries of the data range are very wide which suggest that the minor deviates of the estimated smooth from monotonicity are not significant. The monotone effect of `rel.dbh` of model h3 is linear with a negative slope which fulfills the imposed monotone decreasing constraint (Fig. 2b). The effect of `topex.sw` on the original parameter A is not very strong, which might be because the digital terrain model

used for the topex calculation has a low resolution of 90 m × 90 m (Fig. 1c). At the upper boundary of the range of `topex.sw` the estimated smooth is considerably decreasing, but has a wide confidence interval. Hence, the assumption of a monotone increasing effect made in model h3 need not to be rejected. Although there is an increasing effect of `topex.sw` near the lower boundary of the covariate range, this effect is much stronger (the gradient of the function is very steep) in comparison with the overall pattern. The corresponding confidence intervals are wide which might be due to the small amount of data available in that range. Therefore, the resulting linearity of the constraint effect could be validated as feasible also for this data range of `topex.sw`. (Fig. 2c).

The unconstrained effects of `temp.veg` and `ari` of model h1 (Fig. 1d, e) are both increasing over almost the whole data ranges except for the boundaries with not many data available. The results of the `temp.veg`

Fig. 2 The estimated shape constrained univariate smooths of model h3. **a** the smooth of age, (**b**) the smooth of relative diameter at breast height, (**c**) the smooth of topex index, (**d**) the smooth of temperature sum during the vegetation period and (**e**) the smooth of aridity index. The labels of the vertical axes denote the smooth terms with the corresponding covariate and edf given in brackets

effect are mainly in accordance with findings of Albert and Schmidt (2009) who describe a monotone increasing effect with declining rate of mean temperature in growing season on site index for Norway spruce in Lower Saxony. In contradiction, Nothdurft et al. (2012) found an optimum curve with a slight tendency of a decreasing effect for high values of temperature sum in growing season for Norway spruce in Baden-Württemberg. This might be a result of the warmer climate of Baden-Württemberg which is located in Southwest Germany. However, an investigation for the whole of Germany (Schmidt 2010) showed monotone increasing effects of temperature sum in growing season and aridity index. These partially differing results might be due to the collinearity of climatic covariates which hinders the estimation of robust causal effects especially for the upper boundaries of the data ranges. From our point of view the scam approach offers a possible solution to the problem by integrating expert knowledge. Even if the modelling procedure includes a more subjective component, we argue that predictions from our scam models are more reliable than their unconstrained counterparts, because of limited extreme

data values. However, future model building should use extended data bases with a specific focus on warm-dry site conditions. The corresponding constrained effect of temp.veg of model h3 (Fig. 2d) is monotone increasing with a weak effect below temp.veg = 1400, a stronger effect above 1500 and with a slight tendency of a decreasing gradient. The constrained effect of ari (Fig. 2e) is approximately linear with a steep slope below the value of ari around 70 and nearly constant above that value, indicating almost no further impact of increasing humidity. Compared to the other shape constrained effects the constraint effects for temp.veg and ari might be thought as still implausible to a certain extent. The weak effect of temp.veg at its small values can be considered as implausible, since the marginal utility of a unit increase of the temperature sum should be high especially under the condition of low temperature. Furthermore, the sharp change in the gradient of m_{4a}(temp.veg) at around 1400 seems to be artificial. The plateau part of the estimated effect of ari (Fig. 2e) is observed at very humid site conditions only which also could be validated as implausible. Additionally, the sharp change
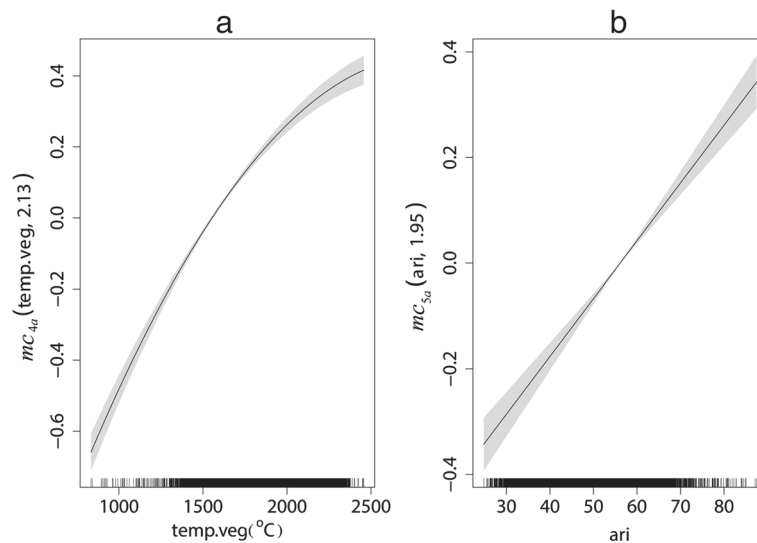
Fig. 3 Model h5: the estimated smooth terms with both monotonicity and concavity constraints: (**a**) the smooth of temperature sum during the vegetation period and (**b**) the smooth of aridity index

in the gradient seems to be spurious. Figure 3 shows the estimated effects of the two terms with both monotone increasing and concavity constraints, mc_{4a} and mc_{5a}. This figure reveals now more convincing and reasonable smooth curves of the sum of daily mean temperature during vegetation period and aridity index. The other smooth terms of model h5 have similar effect to those of model h3.

The estimated varying coefficients smooths of the unconstrained h2 and shape constrained h4 models, are illustrated in Figs. 4 and 5 correspondingly. From an expert view the unconstrained non-linear structure of the effects of age and altitude on the original parameter B is too flexible (Fig. 4). The unconstrained effect of age supports the assumption of an increasing slope of the h-d curve with increasing developmental stage, since generally the effect of age on B is increasing. Only for high ages the effect is decreasing. The unconstrained effect of altitude, $f_{2b}(\mathtt{alt})$, shows a weak increasing tendency, and the overall amplitude of the effect is small in comparison with the age effect. The corresponding confidence intervals are very large.

However, the two plots of the constrained version (Fig. 5) show the plausible monotone effects of age and altitude, although the non-linear structure of $m_{2b}(\mathtt{alt})$ is not very strong. Additional information about monotonicity of the effects narrowed the confidence intervals. The variability of the smooth estimates decreased as our beliefs in the shape of the effects were appended to the h-d relationship.

Figure 6 shows the spatial effect of the model h5. The effect was similar for the other considered models. The spatial smooth can be interpreted as a proxy of additional

predictors such as available water capacity of the soil, nutrient supply of the soil, etc., which were not at our disposal. The southern medium mountain area has better soil condition, therefore the trees are taller and slender in this part (light grey), compared to the worser conditions in the flat lands (silver) which have mainly glacial (sandy) type of soil. The conditions are even worse in terms of height growth near the North Sea coast (dark grey) due to the higher wind speed.

Conclusions

The presented framework and software allow the inclusion of a combination of shape constrained and unconstrained smooth terms of one or more covariates as well as inclusion of strictly parametric model components and varying coefficient terms. The smoothing parameter selection is integrated with the SCAM parameter estimation procedure which is a great advantage. The model estimation scheme also provides interval estimates of the smooth terms which does not incur any additional simulations.

The previous approach that was used as a starting model (Schmidt 2010) used unconstrained GAM for modelling fixed effects on tree height development which resulted in some non-monotonic effects that are scientifically implausible. Based on the foregoing justification for the monotonicity of such model components, it is claimed that the observed non-monotonicity is a result of unmeasured and unknown covariates and insufficient observations and collinearity of covariates. Not only does this limit the interpretability and usage of the scientific model, but it also leads to underestimating the variation associated with prediction of tree height. The specification of
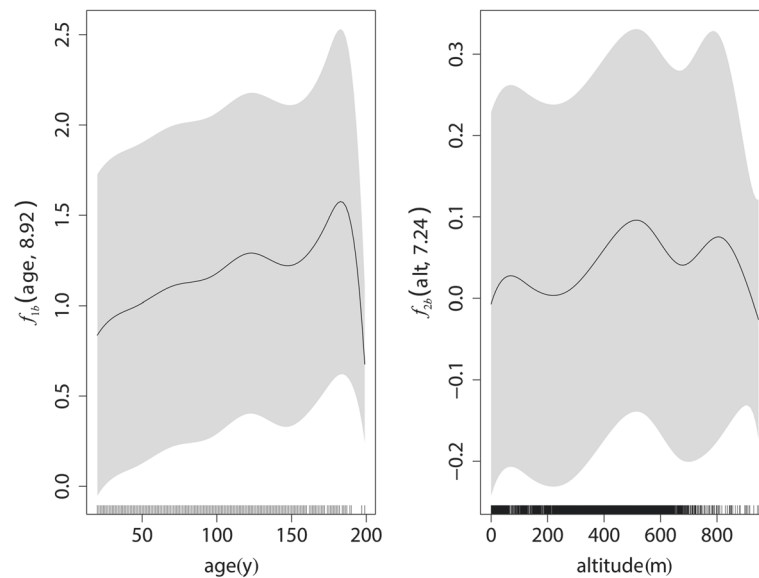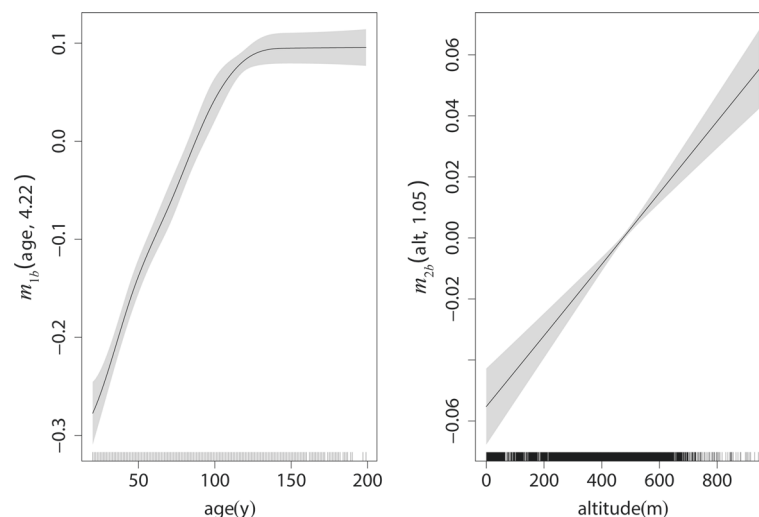
Fig. 4 The estimated varying coefficients effects of the unconstrained model h2. Left panel: estimate of f_{1b}(`age`); right panel: of f_{2b}(`alt`). The labels of the vertical axes denote the smooth terms with the corresponding covariate and edf given in brackets
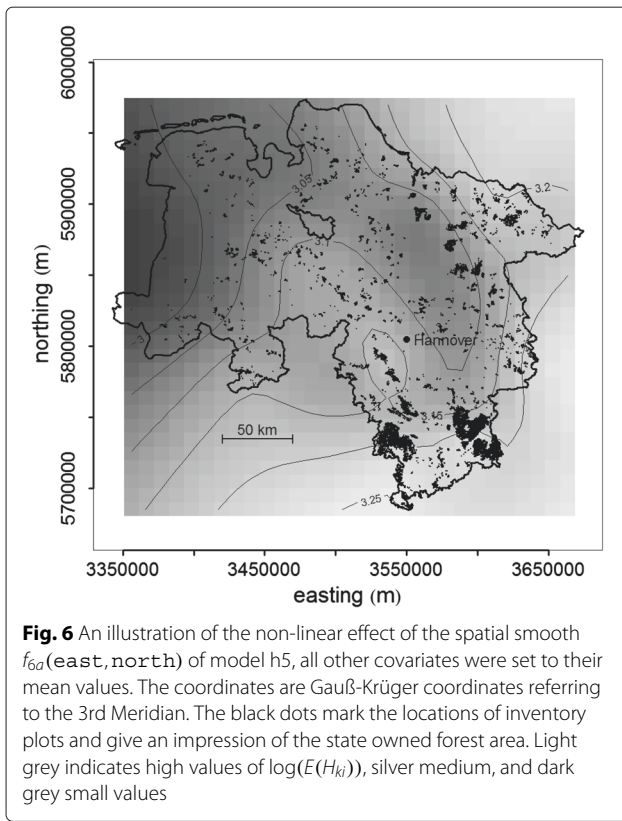
appropriate monotonicity constraints allows for an optimal combination of flexibility and expert knowledge to guarantee for a more robust modelling. This is especially useful in models using causal covariates applied to the prediction of future forest status.

The properties of the finally selected model (h5) can be summarized as follows:

1) The model comprises significant non-linear effects of covariates.

2) The plausibility of non-linear effects of covariates is enforced by the integration of monotonicity constraints.

3) The plausibility of some non-linear effects of covariates is enforced by the additional integration of concavity constraints.

4) The implementation of expert knowledge via constraints is enabled because the original parameters of the principal h-d model have a biological meaning.

Fig. 5 The estimated shape constrained varying coefficients effects of model h4. Left panel: estimate of m_{1b}(`age`); right panel: of m_{2b}(`alt`). The labels of the vertical axes denote the smooth terms with the corresponding covariate and edf given in brackets

Fig. 6 An illustration of the non-linear effect of the spatial smooth f_{6a}(`east,north`) of model h5, all other covariates were set to their mean values. The coordinates are Gauß-Krüger coordinates referring to the 3rd Meridian. The black dots mark the locations of inventory plots and give an impression of the state owned forest area. Light grey indicates high values of $\log(E(H_{ki}))$, silver medium, and dark grey small values

5) The present autocorrelation in the large scale data base is covered by a 2-dimensional surface fitting as a function of coordinates.

6) The causality and generality of the model for prediction purposes is improved by use of causal site variables like sum of daily mean temperature during vegetation period and index of aridity.

None of the height-diameter-models referenced in the introduction chapter cover all these aspects simultaneously. Most models assume linear effects of covariates (e.g., Lappi 1997; Eerikäinen 2003; Calama and Montero 2004; Mehtätalo 2004). However, sometimes transformations of covariates are employed to achieve approximately linear effects (Eerikäinen 2003). At least in our case some of the estimated effects are significantly non-linear which would lead to biased predictions if disregarded. Moreover, there is a qualified need for constraining the non-linear effects because particularly at the boundaries of data ranges effect pattern resulted that conflict with expert knowledge. Hofner et al. (2011) presented a structured additive regression model for ordered categorical data of the breeding distribution of Red Kite that employs monotonic penalized splines. As in our application they emphasize the optimal combination of flexibility and expert knowledge that is enabled by use of the

monotone P-Splines. Schmidt et al. (2011) modelled nonlinear effects of covariates via penalized regression splines but monotonicity resulted directly from the model fit without specifying constraints. Moreover, since the original parameters of their principal height-diameter model ("Näslund function", see e.g. Kangas and Maltamo 2002) have no clear biological meaning, there would not be biological expert knowledge that could be included in the model selection as in our case. Data from large scale forest inventories typically show spatial autocorrelation of residuals that could not be related to fixed effects when conducting regression analyses. In h-d-modelling often a mixed model approach is used to assess between-plot covariance structures (Jayaraman and Lappi 2001; Mehtätalo 2004). However, in this approach it is disregarded that random effects of sample plots are usually not spatially independent themselves, but show some similarity due to effects of unobserved covariates like soil properties. As a solution to the problem (Brezger and Lang 2006) separate the overall spatial trend into a spatially correlated (structured) and an uncorrelated (unstructured) effect. The latter one accounts for local correlation, in the case of h-d modelling of trees of the same sample plot or stand. Only the unstructured spatial effect should be modelled by uncorrelated random effects. Structured spatial effects can be modelled via a Gaussian Markov random field, i.e. spatially correlated random effects are estimated for discrete spatial units (Kammann and Wand 2003) or via 2-dimensional surface fitting by applying specific generalized additive models based on e.g. penalized regression splines with thin plate basis (Wahba 1990; Wood 2006a). We use the latter approach since our observations are exactly localized via coordinates. More simple approaches for describing structured spatial effects in h-d-models are dummy variables for territorial units (Huang et al. 2000; Jayaraman and Lappi 2001; Calama and Montero 2004) or univariate linear effects of coordinates (Hökkä 1997; Mehtätalo 2004). However, these approaches disregard either the large scale autocorrelation between units or would assume at least in our case unrealistically simple pattern of the structured spatial effect (Fig. 6). A more detailed analysis is presented by Nanos et al. (2004), who fitted ordinary mixed models but applied Kriging methods to the estimated random effects to account for spatial correlation. Hence, a structured spatial effect is modeled but in a 2 step procedure. We did not model random effects on plot level to account for local, hence unstructured spatial effects because for most sample plots only one height was measured (Table 1). Causal site variables have not been widely used as predictors in h-d modelling. Many approaches use no site variables at all or only proxy site variables like altitude or coordinates Hökkä (1997). Huang et al. (2000) use ecoregions as a proxy for large scale site conditions. Mehtätalo (2004)

combined causal variables like a longtime mean cumulative temperature sum and a soil type classification with proxy site variables as we did. The advantage of proxy site variables is that they are usually known (like coordinates of stand centroids) or can be easily calculated with high accuracy (like altitude from high resolution digital terrain models). Causal site variables like continuous climatic and soil variables are usually unknown for forest stands or inventory plots and have to be predicted from auxiliary models. Thus they include a prediction error that will affect the height-diameter modelling also. However, our decision to use causal site variables is based on the following reasons. 1) Our model should be able to predict future tree heights under projected changeable climatic conditions. 2) The integration of expert knowledge via monotonicity constraints is much more evident for causal covariates since proxy variables usually subsume several causal variables with differing effects. 3) The combination of causal covariates and monotonicity constraint improves the generality of the model in predictions.

The approach of SCOP-splines is an additional extension of the variety of smoothing techniques incorporated in the R-library `mgcv` (Wood 2006a). For this specific application of modelling the height-diameter relationship of Norway spruce, we have shown that the implementation of shape constrained smooths ensures a robust biologically meaningful interpretation with only marginal loss of prediction accuracy and no increase in prediction bias.

Authors' contributions
NP developed the R-package `scam` that was used for the model development. Both authors contributed to the model building, validation and writing of the manuscript. Both authors read and approved the final manuscript.

Author details
[1]Department of Mathematics, School of Science and Technology, Nazarbayev University, 53 Kabanbay Batyr Avenue, Astana, Kazakhstan. [2]The Northwest German Forest Research Station, Department of Forest Growth, Grätzelstr. 2, 37079 Göttingen, Germany.

References
Albert M, Schmidt M (2009) Climate-sensitive modelling of site productivity relationships for Norway spruce (Picea abies (L.) Karst.) and common beech (Fagus sylvatica L.) For Ecol Manag 259:739–749

Brezger A, Lang S (2006) Generalized structured additive regression based on Bayesian P-splines. Comput Stat Data Anal 50(4):967–991

Calama R, Montero G (2004) Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain. Can J For Res 34:150–163

Castedo-Dorado F, Diéguez-Aranda U, Barrio Anta M, Sánchez Rodríguez M, von Gadow K (2006) A generalized height-diameter model including random components for radiata pine plantations in northwestern Spain. For Ecol Manag 229(1-3):202–213

Curtis RO (1967) Height-diameter and height-diameter-age equations for second-growth douglas-fir. For Res 13(4):365–375

De Martonne E (1926) Une nouvelle fonction climatologique: l'indice d'aridité. La Météorologie (1942) 21:449–458

Eerikäinen K (2003) Predicting the height-diameter pattern of planted Pinus kesiya stands in Zambia and Zimbabwe. For Ecol Manag 175:355–366

Eilers PH, Marx BD (1996) Flexible smoothing with B-splines and penalties. Stat Sci 11:89–121

Hastie T, Tibshirani R (1990) Generalized Additive Models. Chapman & Hall, Florida

Hastie T, Tibshirani R (1993) Varying-coefficient models. J R Stat Soc Ser B 55(4):757–796

Hofner B, Müller J, Hothorn T (2011) Monotonicity-constrained species distribution models. Ecology 92(10):1895–1901

Hökkä H (1997) Height-diameter curves with random intercepts and slopes for trees growing on drained peatlands. For Ecol Manag 97:63–72

Huang S, Price D, Titus SJ (2000) Development of ecoregion-based height-diameter models for white spruce in boreal forests. For Ecol Manag 129:125–141

Huang S, Titus SJ, Wiens DP (1992) Comparison of nonlinear height-diameter functions for major Alberta tree species. Can J For Res 22:1297–1304

Jayaraman K, Lappi J (2001) Estimation of height-diameter curves through multilevel models with special reference to even-aged teak stands. For Ecol Manag 142:155–162

Kammann EE, Wand MP (2003) Geoadditive models. J R Stat Soc Ser C 52:1–18

Kangas A, Maltamo M (2002) Anticipating the variance of predicted stand volume and timber assortments with respect to stand characteristics and field measurements. Silva Fennica 36(4):799–811

Lappi J (1997) A longitudinal analysis of height/diameter curves. For Sci 43(4):555–570

Lin X, Zhang D (1999) Inference in generalized additive mixed models by using smoothing splines. J R Stat Soc Ser B 61:381–400

Mehtätalo L (2005) Height-diameter models for Scots pine and birch in Finland. Silva Fennica 39(1):55–66

Mehtätalo L (2004) A longitudinal height diameter model for norway spruce in finland. Can J For Res 34(1):131–140

Nanos N, Calama R, Montero G, Gil L (2004) Geostatistical prediction of height/diameter models. For Ecol Manag 195(1-2):221–235

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. For Ecol Manag 279:97–111

Pya N, Wood SN (2015) Shape constrained additive models. Stat Comput 25(3):543–559

Schmidt M (2010) Ein standortsensitives, longitudinales Höhen-Durchmesser-Modell als eine Lösung für das Standort-Leistungs-Problem in Deutschland. Deutscher Verband Forstlicher Forschungsanstalten Sektion Ertragskunde: Beiträge zur Jahrestagung 2010:131–152. http://sektionertragskunde.fvabw.de/band2010/Tag2010_14.pdf

Schmidt M, Kiviste A, Gadow K (2011) A spatially explicit height-diameter model for Scots pine in Estonia. Eur J For Res 130:303–315

Scott R, Mitchell S (2005) Empirical modelling of windthrow risk in partially harvested stands using tree neighbourhood and stand attributes. For Ecol Manag 218:193–209

Sharma M, Parton J (2007) Height-diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. For Ecol Manag 249:187–198

Silverman BW (1985) Some aspects of the spline smoothing approach to nonparametric regression curve fitting. J R Stat Soc Ser B 47:1–52

Spekat A, Enke W, Kreienkamp F (2007) Neuentwicklung von regional hoc haufgelösten wetterlagen für Deutschland und Bereitstellung regionaler Klimaszenarien mit dem Regionalisierungsmodell WETTREG 2005 auf der Basis von globalen Klimasimulationen mit ECHAM5/MPI-OM T63L31 2010 bis 2100 für die SRES Szenarios B1, A1b und A2. Endbericht CEC-Potsdam GmbH, Im Auftrag des Umweltbundesamts, Dessau; 148. https://www.umweltbundesamt.de/sites/default/files/medien/publikation/long/3133.pdf

Thornthwaite CW (1931) The climates of North America: according to a new classification. Geogr Rev 21(4):633–655

Wahba G (1983) Bayesian confidence intervals for the cross validated smoothing spline. J R Stat Soc Ser B 45:133–150

Wahba G (1990) Spline models for observational data. SIAM, Philadelphia

Wood SN (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties. J R Stat Soc Ser B 62:413–428

Wood SN (2006a) Generalized Additive Models. An Introduction with R. Chapman and Hall/CRC, Boca Raton, Florida

Wood, SN (2006b) On confidence intervals for generalized additive models based on penalized regression splines. Aust N Z J Stat 48(4):445–464