Sensitivity of phyto-geocentric site productivity models to spatial extent and climate data aggregation

Dissertation

to attain the doctoral degree Dr. forest.
of the Faculty of Forest Sciences and Forest Ecology
Georg-August-Universität Göttingen

submitted by
Levent S. Burggraef
born on March 20, 1987
in Cologne, Germany

Göttingen, July 2023

1. Referee: **Prof. Dr. Jürgen Nagel**
2. Referee: **Prof. Dr. Christian Ammer**
Date of oral examination: 2023-09-04

Acknowledgement

I would like to thank my parents, Tilman and Andrea Burggraef, for their patience and support and for moving from the urban sprawl of Cologne to the Eifel region right after my birth. The forested environment and some amount of family history have certainly led me to pursue an education in forest science. Thanks to my siblings, Maria, Lena, Antonia and Elias, especially those who finish or have finished their advanced degrees in this very year, for several long phone calls. I want to thank Timeea Bucur for her love, and for staying around in spite of the multiple iterations of the old "started making it-had a breakdown-bon appetite"-cycle, that had to occur until I finally took on, and completed, this thesis. With regard to mental health, I want to thank the wall-gazers and meditators at Zendo Göttingen and the Dharma Sangha for existential grounding, as well as my clique of local musicians and BWA Harz, because Zen and performance art are more similar than it seems.

Moving on to my academic support system, I first want to thank Horst Reinecke, who threw me off the deep end of the `R`-ocean some time in 2011. And Dr. Robert Nuske who mentored me through my bachelor and master thesis and to some degree beyond.

Special thanks go to my supervisor Dr. Matthias Schmidt for his guidance, for granting me lots of time to finish this thesis and also for telling me about some hidden musical gems. And of course to Prof. Dr. Jürgen Nagel for taking on the role of the main supervisor and applying just the right amounts of pressure and patience. Thanks to Prof. Dr. Christian Ammer for agreeing to be the second referee and being available with advice in spite of a tight schedule, and to Prof. Dr. Kerstin Wiegand for agreeing to be part of the examination committee.

I want to thank all my colleagues from the Northwest German Forest Research Institute and the Department of Forest Growth for a great working atmosphere that allows for in-depth academic debate as well as as the occasional lively exchange on trash cinema. I also want to put some emphasis on thanking the more traditionally and practically educated colleagues for answering my questions about the actual forest that exists outside of `R`. Special thanks to Robert Larkin and Dr. Holger Sennhenn-Reulen for proof reading and help with the English language as well as the intricacies of model equations.

Finally, I would like to list and thank all the more or less famous role models that somehow manage to walk between the worlds of academia, art and entertainment. Since the complete list is too long, I'll name Greg Graffin and Austin Matelson, whose Wikipedia articles and interviews have inspired me ever so often. In this spirit, I also want to dedicate these pages to the recently deceased Sarah Sottile, neuroscientist, exotic dancer and meme queen. And to the attention seeking felines who have stepped over my keyboard during the creation of this thesis and are responsible for all remaining typography errors.

Contents

Abstract

The impact of climate warming on forestry in Germany and Europe has become more and more visible in recent years and decades. This has lead to an increased demand for predictions on the development of forest stands with regard to timber production, CO2 sequestration and conservation related properties, such as biodiversity. Phyto-geocentric climate sensitive site productivity models are helpful tools for estimating trends in the development of timber and CO2 stocks. Traditional site productivity models are phytocentric, relying solely on the properties of stands or trees. Adding climate sensitivity requires building geocentric or phyto-geocentric models, which are based on or include environmental covariates. Ideally, these models would be based on real time series. However, due to the long time periods of forest growth processes, these rarely cover climate gradients required for model development. Hence, they are often replaced by false time series. This approach is also known as Space-for-Time substitution. Phyto-geocentric climate sensitive site productivity models following this approach are based on two data components: forest inventory data, which gives information on site productivity, and climate data from which the necessary independent covariates are selected. The thesis presented here consists of two studies, each dealing with one component affecting Space-for-Time based site productivity models: (1) the spatial extent of the underlying forest inventory data and (2) the aggregation period used for the climate data. In the first study, height-diameter models for common and sessile oak (*Quercus robur/petraea*), European beech (*Fagus sylvatica*), Norway spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*) were fitted based on a pan-European forest inventory dataset, supplemented with climate and soil data. In the second study, height-age models for European beech and Scots pine were developed based on three different aggregation scenarios for climate data. Soil data was included in both studies to analyze the potential of soil covariates for increasing the geocentric component of site productivity models. Both studies were compared with regard to (1) the potential of continental scale forest inventories as a basis for climate sensitive site productivity models compared to national scale inventories, (2) the differences in covariate effects and predictions resulting from different aggregation periods of the underlying climate data and (3) differences in the selection of soil covariates. Results showed that forest inventories on a continental scale may be better suited for site productivity model fitting. Compared to the German National Forest Inventory used in the second study, the identification of plausible effects proved easier, especially at the extreme edges of the data. However, similar studies show that supplementing national scale inventories with fine scale regional inventory data may yield equally plausible effects. With regard to the analyzed aggregation periods, dynamic aggregation of climate data over the tree or stand age is clearly recommended. Static aggregation periods might lead to over- or underestimation

or even contrasting directions of productivity changes when projecting into the future. The identification of plausible soil covariate effects proved difficult in both studies, with most parameters returning implausible or insignificant effects. However, reasonable effects were identified for C:P ratio and pH value.

Zusammenfassung (German)

Der Einfluss der Klimaerwärmung auf alle Bereiche der Forstwirtschaft in Deutschland und Europa ist in den letzten Jahren und Jahrzehnten immer deutlicher geworden. Daraus entsteht ein gesteigerter Bedarf an Vorhersagen der Entwicklung von Waldbeständen in Bezug auf Holzproduktion, CO2-Bindung und naturschutzbezogenenen Eigenschaften wie Biodiversität. Phyto-geozentrische klimasensitive Standort-Leistungs-Modelle sind hilfreiche Werkzeuge um Trends in der Entwicklung von Holzaufkommen und CO2-Vorräten einzuschätzen. Klassische Standort-Leistungs-Modelle sind phytozentrisch, d.h. sie basieren lediglich auf Eigenschaften von Bäumen oder Beständen. Um Klimasensitivität zu erreichen, müssen geozentrische oder phyto-geozentrische Modelle gebildet werden, welche auf Umwelt-Kovariablen basieren oder diese einbeziehen. Unter idealen Voraussetzungen würden diese Modelle auf Basis echter Zeitreihen entwickelt. Durch die langen Zeiträume forstlicher Wachstumsprozesse decken diese jedoch i.d.R. nicht die für die Modellentwicklung erforderlichen Klimagradienten ab. Daher werden sie meist durch falsche Zeitreihen ersetzt. Diese Methode wird auch als Space-for-Time Substitution bezeichnet. Auf diesem Ansatz basierende klimasensitive Standort-Leistungs-Modelle greifen auf zwei Teildatensätze zurück: Waldinventurdaten, welche Informationen über die Standort-Leistungs-Beziehung liefern und Klimadaten, aus denen die unabhängigen Kovariablen selektiert werden. Die vorliegende Arbeit besteht aus zwei Studien, die sich mit jeweils einer Komponente beschäftigen, welche Space-for-Time basierte Standort-Leistungs-Modelle beeinflusst: (1) der räumlichen Ausdehnung der zugrunde liegenden Waldinventurdaten und (2) dem zur Aggregierung der Klimadaten genutzten Zeitfenster. Im Rahmen der ersten Studie wurden Höhen-Durchmessermodelle für Stiel- und Traubeneiche (*Quercus robur/petraea*), Buche (*Fagus sylvatica*), Fichte (*Picea abies*) und Kiefer (*Pinus sylvestris*) auf Basis eines europaweiten, um Klima- und Bodendaten erweiterten Waldinventurdatensatzes entwickelt. In der zweiten Studie wurden Höhen-Altersmodelle für Buche und Kiefer auf Basis von drei verschiedenen Aggregierungszenarien für Klimadaten entwickelt. In beiden Studien wurden Bodendaten einbezogen, um ihr Potenzial zur Verstärkung der geozentrischen Komponente von Standort-Leistungs-Modellen zu analysieren. Beide Studien wurden bezüglich folgender Kriterien analysiert: (1) des Potenzials von Waldinventuren auf kontinentaler Ebene gegenüber nationalen Waldinventuren als Grundlage klimasensitiver Standort-Leistungs-Modelle, (2) der Unterschiede in den

Kovariableneffekten und Modellvorhersagen, basierend auf den Zeitfenstern der Klimadaten-Aggregierungen und (3) der Unterschiede in der Selektion von Boden-Kovariablen. Die Ergebnisse zeigten, dass Waldinventurdatensätze auf einem kontinentalen Stichprobennetz zur Modellentwicklung potenziell besser geeignet sind als nationale Waldinventuren. Im Vergleich zur in der zweiten Studie genutzten Bundeswaldinventur erwies sich die Identifikation plausibler Effekte als einfacher, besonders an den Datenrändern. Jedoch zeigen ähnliche Studien, dass die Ergänzung nationaler Waldinventuren mit Daten aus regionalen Inventuren mit engmaschigem Aufnahmeraster zur Identifizierung ähnlich plausibler Effekte führen kann. Bezüglich der Aggregierung von Klimadaten für die Entwicklung von Standort-Leistungs-Modellen ist die dynamische Aggregierung über dem Baum- bzw. Bestandesalter deutlich zu empfehlen. Statische Aggregierungszeiträume können zu Über- oder Unterschätzungen oder gar gegenläufigen Produktivitätsänderungen in Projektionen der Standortleistung führen. Die Identifizierung plausibler Kovariableneffekte für die Bodenparameter erwies sich in beiden Studien als schwierig, wobei die meisten Parameter unplausible oder nicht signifikante Effekte zeigten. Es konnten jedoch plausible Effekte für das C:P-Verhältnis sowie den pH-Wert identifiziert werden.

Declaration

Section 2 of this thesis is based on a manuscript that was submitted for publication in the Annals of Forest Science with the authors being Levent Burggraef, Paul Schmidt-Walter, Matthias Schmidt and Henning Meesenburg. With regard to § 12 paragraph 1 c) + Appendix 2 of the doctoral degree regulations for the Graduate School of Forest and Agricultural Sciences, it is hereby declared that all analyses presented here were conducted by Levent Burggraef. All sections have been written by Levent Burggraef with the exception of of the text in section 2.2.3, which has been drafted by Paul Schmidt-Walter.

1 General introduction

The impact of climate warming on forestry in Europe and Germany has become more and more visible in recent years. Climate projections for Europe based on the fourth IPCC assessment report predict higher temperatures, especially in winter, as well as a decrease in precipitation in summer and an increase in winter (IPCC 2007; Jacob et al. 2012). Moreover, extreme weather events, such as the 2003 heat wave and heavy storms in 1990, 1999, 2007 and 2018, have had a major impact on forests and forestry in Europe, leading to massive amounts of damaged wood (Schelhaas et al. 2003; Albrecht et al. 2012; Udali et al. 2021) and reduced production and carbon storage. This may have resulted in forest ecosystems becoming carbon sources rather than sinks (Ciais et al. 2005; Eggers et al. 2008). Heat waves and floods are more likely to occur in the future (Schiermeier 2018). While it is not possible to connect specific events to climate change, research from the field of *attribution studies* concludes that the 2018 heat wave (World Weather Attribution 2018; Vogel et al. 2019; Yiou et al. 2020), as well as the 2019 heat wave (Van Oldenborgh et al. 2019) have been caused by human-induced climate change with *virtual certainty*. The 2018 heat wave lead to severe wildfires in Scandinavia and Greece.

This has lead to an increased demand for predictions on the development of forest stands with regard to timber production, CO2 sequestration and nature conservation.

Predictions on the future development and restoration of forest stands can be made based on qualitative reasoning (Bolte et al. 2010; McDowell et al. 2020) or quantitative analyses (Löf et al. 2019; Vauhkonen et al. 2019).

Quantitative analyses usually come in the shape of more or less complex models.

Models dealing with future forest development with regard to climate change can be divided into three different branches: site productivity models, risk/survival models and presence-absence models. Site productivity models deal with the development of forest biomass in the form of net primary production or classical silvicultural measures of timber production (Battaglia and Sands 1997; Palahí et al. 2004; Albert and Schmidt 2010). Risk and survival models deal with the probability of calamities like forest fires, storms or pests (Jaiswal et al. 2002; Schmidt et al. 2010; Overbeck and Schmidt 2012). qnd presence-absence models focus on changes in the geographical regions which may be suitable for tree species (Mauri et al. 2017; Mellert et al. 2018; Pecchi et al. 2019).

These three modeling approaches can each take the shape of statistical (Albert and Schmidt 2010; Overbeck and Schmidt 2012; Mellert et al. 2018) and process based models (Battaglia and Sands 1998; Morin et al. 2007; Hoffman et al. 2018).

Site productivity models are traditionally phytocentric, while climate sensitive models necessarily have to be geocentric or phyto-geocentric. Phytocentric models are based exclusively

on tree or stand properties, such as traditional yield tables, or models which are only used, for example, for height imputation (Bontemps and Bouriaud 2013), while geocentric models include environmental variables.

This thesis deals with site productivity models, which can be subdivided into two approaches: longitudinal models, which are based on real time series obtained from long term experimental plots (Yue et al. 2016). This is the straight forward approach to correlate climate observations to the growth of trees and stands.

Due to the long growth times of trees in contrast to, for example, agricultural crops, it is often impossible to analyze the results derived from experimental plots within the scope of one scientific study. Also, the climatic gradients covered during the lifetime of a tree or stand or the time passed since the introduction of an experimental plot may be too small to identify meaningful climate effects, or the available climate data may not reach back to the germination year of the tree or stand.

Thus, researchers often have to resort to false or artificial time series, substituting the temporal climate gradient by a spatial gradient observed across trees and stands located at sites with varying climatic conditions and at different ages (Pretzsch 2009, Figure 1.1). Such studies are also labeled Space-for-Time (SFT) or gradient studies
(Pickett 1989; Fukami and Wardle 2005; Rustad 2008; Lester et al. 2014) and, in spite of widespread ambitions to transition to real time series, they are still the rule in forest productivity modeling.

Climate sensitive site productivity models following the SFT-approach are usually based on two data components: forest inventory data, which gives information on site productivity, and climate data, which provides the necessary independent covariates. Additional environmental covariates, such as nitrogen deposition or soil physiological, chemical or hydrological covariates are often added to investigate which covariates have the most explanatory power and to move the model further towards geocentricity.

Inventory data is mostly obtained from forest inventories on the national, state or enterprise level. National forest inventories are common in Europe (Kändler 2006; Vidal et al. 2007) as well as North America (Gillis et al. 2005; Tinkham et al. 2018) and Asia (Kitahara et al. 2009; Zeng et al. 2015), with varying capabilities in tropical countries (Romijn et al. 2015). In Germany, fine scale enterprise-level inventories are an established management tool (Böckmann et al. 1998). Inventories differ in sampling method, grid resolution and spatial extent. Since the SFT approach is based on a spatial gradient, the larger the spatial extent, the more likely it is to get reasonable covariate effects. Finer grid resolution can increase the probability of including plots representing rare site conditions.

Climate data can be obtained from different sources, mostly being taken from regionalized climate station data. Regionalized data can be taken from public sources like the global

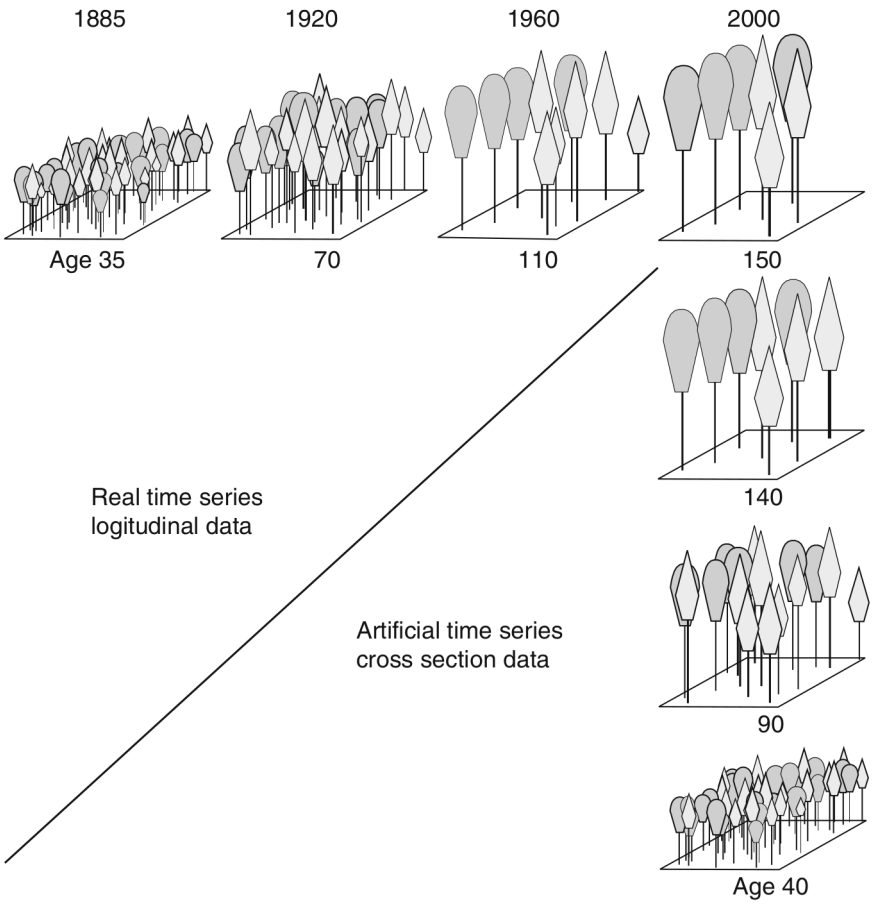


Figure 1.1: Visual representation of the concepts of real time series vs. false time series used in SFT studies. The stand in the row at the top is sampled in 1885, 1920, 1960 and 2000, thus providing a real time series. In the column on the right, four stands at ages 40, 90, 140 and 150 years are sampled in the same year, returning a false time series. Figure taken from Pretzsch (2009), reuse with kind permission from the author.

WorldClim dataset (Hijmans et al. 2005). It is also often supplied project-based by specific institutions (Zimmermann et al. 2007; Dietrich et al. 2019) or in-house.

Climate data differs not only in the spatial extent for which it is available and the quality based on regionalization methods.

The time period for which it is available and the aggregation of the climate data to sums or averages, which are to be correlated with forest inventory data, also have to be considered. The climate periods covered by site productivity models are often restricted to 30 year baseline periods (Daly et al. 2008; Albert and Schmidt 2010; Brandl et al. 2018) and rarely cover longer periods above 40 or even 100 years (Sharma et al. 2012; Bouwman et al. 2020). Since those real time series of climate data are used to supplement false time series data from forest inventories, they have to be aggregated to one value for each sample plot. They are usually averaged to annual means or sums or values for specific months. Values aggregated over modeled vegetation periods are also very popular (Nuske 2017).

The presented thesis consists of two studies which dealt with two components of site-productivity models based on a SFT approach: (1) the spatial extent of the underlying inventory and (2) the aggregation of climate data from real time series for false time series inventory data.

In the first study, height-diameter (h-d) models were fitted based on a pan-European dataset supplemented with climate and soil data. The selected covariate effects were analyzed and projected site indices (SI) were calculated for 100 year old stands all over Europe.

In the second study, height-age (h-a) models were developed based on three different aggregation scenarios for climate data. Climate data was aggregated over two 30 year baseline periods and dynamically over stand age. Soil data was included to further analyze the potential of soil covariates for phyto-geocentric models. SIs were calculated based on climate grids across Germany for the years 2012 and 2050.

In a final synthesis, the results of both studies were compared and the following questions were discussed:

1. Do continental-scale forest inventories as a basis for site productivity models yield more plausible and/or significant climatic effects as opposed to national-scale inventories?
2. Is the dynamic aggregation of climate parameters over stand age advisable in contrast to aggregation over static 30 year baseline scenarios?
3. Do continental scale forest inventories yield better soil related covariate effects in contrast to national scale inventories?

2 Climate-sensitive height-diameter models based on large-scale forest inventories

2.1 Introduction

The idea of phytocentric site productivity models goes back to traditional yield tables. They list various quantifiers of productivity based on Eichhorn's rule and its continuation in the work of Assman (Skovsgaard and Vanclay 2008). Relations of stand volume to height and height to age are assessed to identify classes of productivity (e.g. yield classes, yield levels). Thus, productivity is quantified through plant properties alone. Site quality is indirectly described by those classes, which are established for specific species, regions and thinning regimes. But environmental properties concerning soil, climate or radiation, which causally determine site quality, are not directly addressed and not assumed to change.

As has been discussed in the general introduction, rapidly changing site conditions require moving from the phytocentric to a geocentric or phyto-geocentric approach, including predictors which represent a changing environment (Dănescu et al. 2017).

This requires data sets which include site productivity information in connection with climate data and, if desired, additional environmental parameters.

Due to the difficulty of identifying large enough climate gradients, site productivity information from experimental plots has rarely been used to develop climate sensitive site productivity models (Kahn 1995; Fries et al. 2000).

On the other hand, forest inventories have already been used extensively for building site-productivity models, making use of the aforementioned SFT-approach (Albert and Schmidt 2010; Sharma et al. 2012; Brandl et al. 2014, 2019).

Large scale statistical models based on inventory data have covered average scales from 20.000 to $2.7 \cdot 10^6\,\mathrm{km}^2$ with resolutions, i.e. average distances between selected plots, varying mostly around 30 km (Ung et al. 2001; Seynave et al. 2005, 2008; Bontemps and Bouriaud 2013). Only one study found in current literature was based on a resolution below 2 km (Albert and Schmidt 2010). Larger scales have only been covered by process modeling studies, which rely on a given process model and do not always require actual tree data. Process modeling studies reach from geographically large states like the U.S.A. (Kautz et al. 2018) and Australia (Coops et al. 1998) and geological units like the Amazon basin (Galbraith et al. 2010; Zhang et al. 2015; Rammig et al. 2018) to the global level (Malhi et al. 2011; Fisher et al. 2014).

For this study, a European-scale data set was compiled from three different forest inventories, supplemented with climate and soil data. The dataset provides a combination of stand data

as well as parameters for climate and soil on an international scale and extent that has not yet been used for statistical site-productivity modeling.

The objectives of this study were to (1) develop site productivity models for common and sessile oak (*Quercus robur/petraea*), European beech (*Fagus sylvatica*), Norway spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*) based on the compiled data set, (2) check for the plausibility of the identified effects and (3) discuss the potential and limitations of the data set for site productivity modeling.

2.2 Materials and methods

2.2.1 Forest inventory data

Forest inventory data was composed of three different inventories: (1) the German national forest soil inventory (NFSI), (2) the BioSoil inventory at the ICP Forests Level I grid and (3) the plots of the French national forest inventory (French NFI) (Figure 2.1). The German NFSI was sampled on an 8 km × 8 km raster (Hilbrig et al. 2014), the BioSoil inventory on a 16 km × 16 km (Bastrup-Birk et al. 2007; De Vos and Cools 2011) and French NFI on a 1.41 km × 1.41 km (= 2 km^2) raster (Vidal et al. 2007).

All three inventories used concentric plot designs with different radii and caliper thresholds. Consequently, they provided different numbers of tree diameter and height observations (Table 2.1).

Only European beech and Scots pine from the main stand were used for model fitting, assuming that effects of competition and shading are more homogenous than between different stand layers.

The quadratic mean diameter (`qmD`) was calculated for each plot (Kramer and Akça 2008).

Table 2.1: Statistics on the stand inventory data, differentiated by surveys (NFSI , BioSoil, French NFI) and aggregated over the full data set.

Species	Survey	N trees	N plots	N trees/plot	N heights/ plot	dbh range (5%-95% quantile) [cm]	qmD range (5%-95% quantile) [cm]	Height range (5%-95% quantile) [m]
Oak	NFSI	3102	395	8	3	10.3 - 57.4	10.6 - 53.4	10.2 - 32.8
Beech	NFSI	8419	690	12	4	10 - 58.1	11.4 - 49.6	10.3 - 36.6

Spruce	NFSI	11991	715	17	4	11 - 52.8	12.9 - 47.5	11.2 - 36.6
Pine	NFSI	13833	584	24	4	10.4 - 43.7	11.1 - 40.6	12.8 - 32
sum	NFSI	37345	2384	16	4			
Oak	BioSoil	1071	137	8	1	8 - 59.5	9 - 50.5	7.6 - 30
Beech	BioSoil	3628	203	18	2	3 - 61	8 - 43.2	13 - 37.1
Spruce	BioSoil	10898	614	18	2	4.4 - 51.3	7.5 - 42.9	11 - 37
Pine	BioSoil	8684	506	17	2	6.7 - 35	7.3 - 32.6	7.5 - 29
sum	BioSoil	24281	1460	17	2			
Oak	NFI France	86338	19036	5	4	8.9 - 61.8	11.4 - 57.6	8 - 28.8
Beech	NFI France	60649	11881	5	4	8.3 - 60.2	10.8 - 52	8 - 31.4
Spruce	NFI France	33246	4881	7	6	9.5 - 57	12.7 - 49.7	7.2 - 33
Pine	NFI France	36050	6487	6	5	9.2 - 48.4	12.4 - 43.6	5.4 - 25.2
sum	NFI France	216283	42285	5	4			
Oak	all	90511	19568	5	4	8.9 - 61.8	11.3 - 57.3	8.1 - 28.9
Beech	all	72696	12774	6	4	8.3 - 59.8	10.8 - 51.5	8 - 32
Spruce	all	56135	6210	9	5	9.5 - 55	10.7 - 48	7.4 - 33.8
Pine	all	58567	7577	8	5	9.5 - 46.7	11.3 - 41.9	5.5 - 26.6
sum	all	277909	46129	6	4			

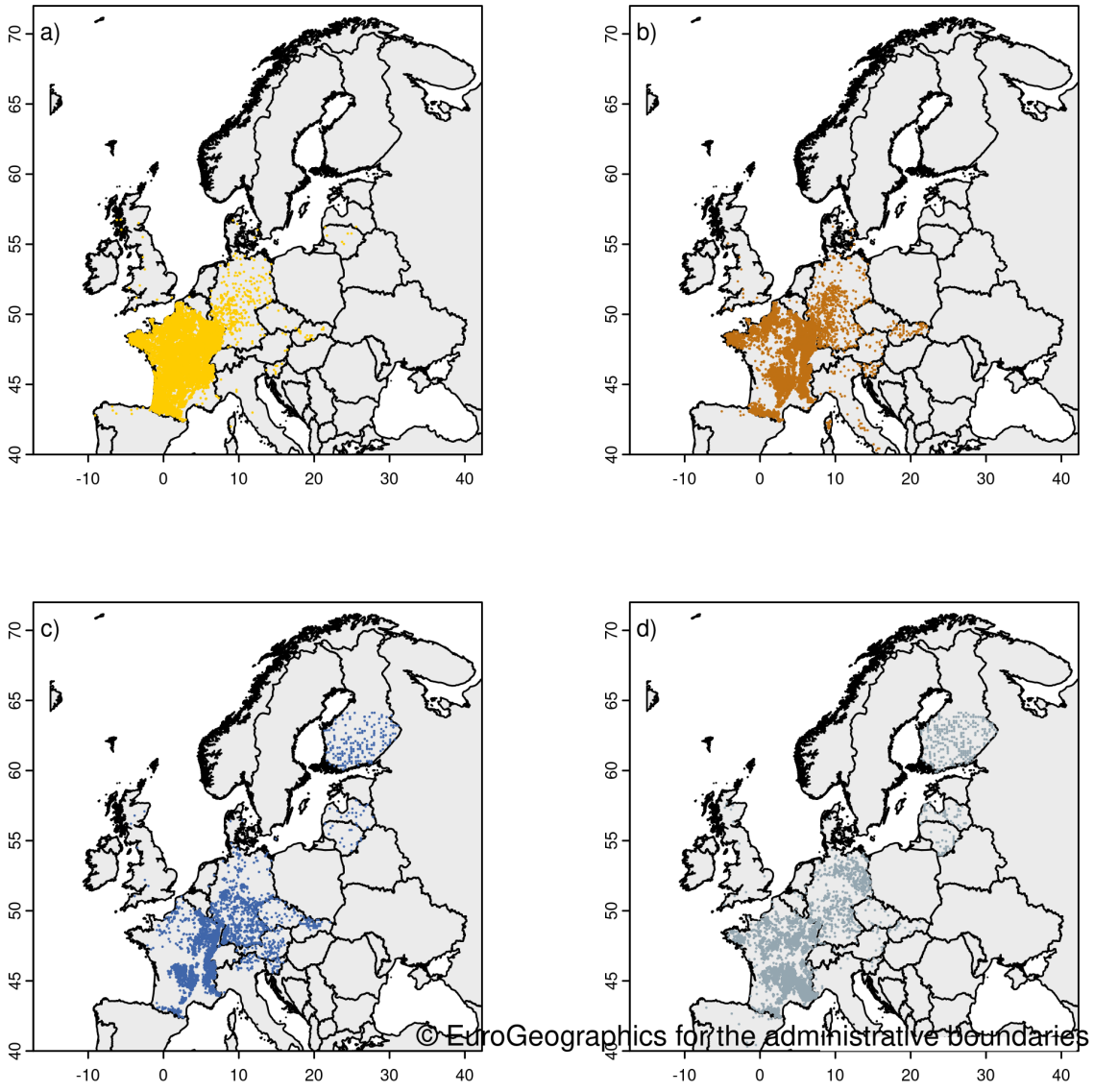


Figure 2.1: Map of oak (a), beech (b), spruce (c) and pine (d) inventory plots.

2.2.2 Climate data

Climate data was obtained from the WorldClim data set (Hijmans et al. 2005). It was provided as a global grid with a resolution of 30 arc seconds (ca. $1\,\text{km} \times 1\,\text{km}$) of regionalized temperature and precipitation measurements.

Data for temperature and precipitation has been provided in the form of different aggregates based on daily means. Annual aggregates, as well as aggregates for the warmest month, coldest month and warmest quarter, were taken from the BIOCLIM dataset provided by the WorldClim project (Bioclim 2021). The warmest quarter includes the warmest full three months, which in Europe is mostly June to August, in rare cases July to September. Aggregates for January and July were computed specifically for the WP-KS-KW (Waldproduktivität, Kohlenstoffspeicherung, Klimawandel; engl: Forest productivity, carbon sequestration, climate change; see Mette and Kölling (2015)). See Appendix A.1 for a listing of all available parameters. All climate parameters were averaged for a baseline period from 1971 to 2000.

Table 2.2: Parameters from the WorldClim-Dataset used as covariates. Covariate effects were tested for significance and plausibility during model selection in step I.2. Parameters were aggregated from daily values, i.e. means were aggregated from daily means, minima from daily minima etc. Parameters without BioClim parameter name were aggregated directly from the WorldClim dataset. For detailed species-level information, see Appendix A.1.

Description	Code	BioClim parameter	5% quantile	95% quantile
Annual mean temperature [° C]	tmn	BIO1	4.50	11.70
Mean temperature of January [° C]	t01		-5.70	4.60
Mean temperature of July [° C]	t07		14.30	19.80
Mean temperature May to September [° C]	t.5to9		11.80	17.60
Mean temperature June to August [° C]	t.678		13.40	19.00
Maximum temperature of warmest month [° C]	tmax.wm	BIO5	19.40	25.70
Minimum temperature of coldest month [° C]	tmin.cm	BIO6	-9.10	1.10

Mean temperature of warmest quarter [° C]	tmn.wq	BIO10	13.40	19.00
Annual precipitation sum [mm]	psum	BIO12	590.00	1223.00
Precipitation sum May to September [mm]	p.5to9		271.00	519.00
Precipitation sum June to August [mm]	p.678		153.00	317.00
Precipitation sum of warmest quarter [mm]	psum.wq	BIO18	154.00	317.00

2.2.3 Soil data

Soil data was available for all three inventories with varying levels of differentiation.

The NFSI and the BioSoil inventory contain detailed physical and chemical measurements for different soil layers that were aggregated to represent soil depths (0-10, 0-30, 0-100) and the entire soilprofile (Höhle et al. 2018). See Appendix A.1 for detailed by-species information on data and aggregation depths.

Different from the NFSI and BioSoil, the French NFI contains only a rough description of site ecology and soil properties. The description included estimations of soil texture and volume fraction of coarse fragments (stones) of the topsoil and the subsoil, along with information on total soil depth that was provided for estimating available soil water capacity (`awc`) (Vidal et al. 2007).

For all three datasets, `awc` was estimated using tabulated pedotransfer functions (Wessolek et al. 2009). The tables provide volumetric `awc` values based on soil texture, bulk density and soil organic matter (SOM) content classes. However, `awc` estimation was solely based on soil texture class, as information on bulk density and SOM was not available for the French NFI plots. The obtained volumetric values of the individual soil layers or the topsoil and subsoil compartments were reduced for the volume fraction of coarse fragments and summed up to represent `awc` of 0-100 cm soil depth. Water logged sites (bogs/fens, stagnosols and gleysols) were excluded from the analysis.

Table 2.3: Soil parameters for the three inventories used as covariates. Effects of each covariate were tested for significance and plausibility during model selection in step II.2, except for *awc* which was available for the full dataset and included in step I.2. For detailed species-level information, see Appendix A.1.

Parameter	Aggregation depth	Unit	Comment	Code	Dataset	5% quantile	95% quantile
Effective soil depth	Max. 100 cm	cm			BioSoil, NFSI, NFI France		
Available water capacity (AWC)	0 – effective soil depth	mm		awc	BioSoil, NFSI, NFI France	7.50	194.85
Minimum air capacity of soil layers	0 – effective soil depth	%		aircap	BioSoil, NFSI	4.00	29.00
organic matter stock of the forest floor	Forest floor	t/ha		fbv	BioSoil, NFSI	6.00	147.31
Carbon stock	Forest floor, mineral soil (0-10, 0-40, 0-ET)	t/ha		oc	BioSoil, NFSI	2.84	151.08
Nitrogen stock	Forest floor, mineral soil (0-10, 0-40, 0-ET)	t/ha		ton	BioSoil, NFSI	0.16	13.41
Phosphorous stock	Forest floor, mineral soil (0-10, 0-40, 0-ET)	t/ha		p	BioSoil, NFSI	41.49	2611.17
C:N ratio	Forest floor, mineral soil (0-10, 0-20)	-	mass ratio, dimensionless	cn	BioSoil, NFSI	11.36	33.48

C:P ratio	Forest floor, mineral soil (0-10, 0-20)	-	mass ratio, dimension-less	cp	BioSoil, NFSI	36.77	663.24
pH value	Forest floor, mineral soil (0-10, 0-20, 0-40, 40-80)	-		ph	BioSoil, NFSI	3.20	7.18
Base saturation	Forest floor, mineral soil (0-10, 0-40, 40-80, 0-ET)	%		basesat	BioSoil, NFSI	6.32	98.89
Cation exchange capacity	Forest floor, mineral soil (0-10, 0-40, 0-ET)	molc/ha	moles of electronic charge per ha	cec	BioSoil, NFSI	0.02	13.28
Exchangeable calcium stock	Forest floor, mineral soil (0-10, 0-40, 0-ET)	kg/ha		ca	BioSoil, NFSI	19.55	22847.14
Exchangeable magnesium stock	Forest floor, mineral soil (0-10, 0-40, 0-ET)	kg/ha		mg	BioSoil, NFSI	3.84	2442.62
Exchangeable potassium stock	Forest floor, mineral soil (0-10, 0-40, 0-ET)	kg/ha		k	BioSoil, NFSI	6.24	1303.46

2.2.4 Models

The height-diameter (h-d) relationship was used as a response variable to describe site productivity. Site index (SI) would have been preferable since it is less dependent on management, easy to interpret and commonly used by practitioners. Unfortunately, SI could not be obtained since no precise age estimate was available from the BioSoil inventory.

In order to describe the h-d relation of individual trees, a reparameterized version of the Korf-function $H = 1.3 + A'e^{-bD^{-C}}$ (Lappi 1997) was used. The reparameterized version replaces the `dbh` by a transformed diameter (see Equation 2.1), while tree height as the response variable is replaced by the natural logarithm ln(.) of tree height. The resulting model includes the parameters A, B, λ and C which are henceforth called first-order parameters. Second-order parameters encompass all phytocentric or geocentric covariates which describe A or B.

The reparameterization has three advantages over the original function: (1) Three of the four first-order parameters can be interpreted in terms of tree allometry. Parameter A signifies the expected logarithmic height E[ln(h)] of a tree with 30 cm `dbh` and determines the level (position on the y-axis) of the h-d relationship. λ represents the expected difference between the diameter at ground level and at breast height (Mehtätalo 2004). Parameter B signifies the difference in expected logarithmic height between a tree with 30 cm and 10 cm `dbh` and determines the slope of the h-d relationship. (2) For constant first-order parameters λ and C, the reparameterized function is linear. This allows for a straightforward description of the phytocentric first-order parameters A and B using geocentric covariates and effects. (3) The first-order parameters show only little correlation (see also Schmidt et al. (2018)).

The method of describing variations in first-order parameters A and B, which was used here, differed from the original study by Lappi (1997). The original study was based on experimental plots with repeated measurements. A and B were thus estimated on the level of plot and sampling date, with age trends allowing for the level and slope of the h-d relationship to change over time. In the present study, only one sampling date was available per plot. Also, no precise age estimate was given in the BioSoil data set (as mentioned above). Hence, different stages of stand development and environmental conditions leading to variations in A and B were represented by resorting (1) to an extensive pool of sampling plots over a large spatial scale and (2) substituting age by `qmD`.

The final geocentric models for oak, beech, spruce and pine were selected within a six-step framework creating a two-step model.

- Model step **I**
  - **I.1** First, a *phytocentric* model was selected for each species, containing only the transformed diameter $x_{k,i}$ and `qmD` as covariates (Equation 2.3). It was fitted as a

Generalized Additive Model (GAM). In some cases, it was necessary to smooth implausible effect patterns, especially at the edges of the data range. In that case, constraints for e.g. monotony or concavity were added, refitting the model as a Shape Constrained Additive Model (SCAM). Software R (R Core Team 2020, version 3.4.4.) packages `scam` (Pya 2019) and `mgcv` (Wood 2017) were used. The effect of `qmD` was included, described as a linear coefficient on $x_{k,i}$ and as penalized basis-function approach using a thin-plate spline basis, in the following short *spline* f(`qmD`), (see Wood 2017 for documentation on splines). Note that covariate effects *multiplied* on $x_{k,i}$ affect the *slope* of the h-d relation, while *addition* of covariate effects affects the *level*. In this step, the optimal combination of first-order parameters λ and C was selected using an iterative grid search. A range of values was given for both parameters, which was increased if the grid search reached the limits of the range. Typical values for the grid ranged between 1 to 50 for λ in integer steps and 0.1 to 2 for C in steps of 0.1. The grid contained all possible combinations of both parameter ranges. The model was fitted in a loop with values for λ and C taken from the grid until changes in the Akaike Information Criterion became marginal (i.e. were only observed in the decimal range). With constant λ and C, a linear model which allowed for easy addition of geocentric covariates was obtained. A log-link function was used to ensure positive expected values. It also led to covariate effects acting in a multiplicative-exponential way on tree height, thus accounting for positive interaction effects by the non-linearity of the link-function. More complex interaction terms described by varying coefficients (Hastie and Tibshirani 1993; Wood 2017) had been part of preliminary selections, but lacked significance and/or plausibility. For this and all following modeling steps, it was assumed that the response follows a Gamma distribution. Note that this method differs from Lappi (1997) where the residual variance was modeled explicitly.

– **I.2** Then, one model was fitted for each climate parameter from the WorldClim data set in order to analyze the effect of each parameter individually (Equation 2.4). The models were based on the stand data from all inventories. However, as the French NFI contained no soil parameters except for `awc`, no other soil parameters were included in this model step. Each model was fitted as a GAM or SCAM (depending on step I.1) with the geocentric covariate effect being estimated as an unconstrained spline. The effects of geocentric covariates were evaluated based on model fit (R^2), significance (p-value) and the plausibility of the shape of the spline. R^2 for a GAM is computed as the proportion of the null deviance explained (Wood 2019). Plausible effect shapes were considered to be: (1) linear increasing

or decreasing, (2) parabolic, i.e. characterized by an optimum for the respective covariate, or (3) asymptotic, i.e. increasing or decreasing monotonously above or below value of saturation (Loehle 2000). A more intuitive way of analyzing covariate effects is to look at the relative amplitude of each effect for a specific range of the corresponding covariate. Amplitudes were calculated as the difference between maximum and minimum value of the covariate *term* (i.e. the spline or the linear coefficient multiplied with the covariate value) for the 5%-95% quantile range of the *covariate* (Equation 2.7). Each amplitude was divided by the sum of the amplitudes of all terms to get a relative amplitude, thus quantifying the effect on logarithmic height variation in comparison to all other terms. Geocentric covariates were only included by addition in the present study, thus affecting the *level* of the h-d relation.

  - **I.3** A final model was fitted by combining the selected geocentric covariates in one model (Equation 2.8). Again, geocentric covariate effects were first estimated as unconstrained splines and evaluated based on R^2, significance and plausibility. At least one covariate was selected from parameter groups for temperature and water (precipitation and `awc`), which, together with radiation and nutrients, are assumed to be the main drivers of tree growth (Kahn 1995; Pretzsch and Kahn 1995; Gadow 2003). After finishing the covariate selection, the original parameters λ and C were optimized a second time and modifications were added depending on wiggliness and plausibility. This included adding constraints as well as adjusting the number of basis dimensions k of splines. Limiting basis dimensions reduces the degrees of freedom of a spline and produces a less wiggly shape.

- Model step **II**

  - **II.1.** For the second model step, the covariate effects of the final model of step I were summarized to give estimates of A and B (Equation 2.5). Subsequently, a Generalized Linear Mixed Model (GLMM) was fitted including $x_{k,i}$, the estimates of A and B ($\hat{A}$ and $\hat{B}$), and random effects for the plot level. Random effects have the additional advantage of serving as a filter for unexplained plot-level variation. Thus, they smooth artifacts which otherwise may result in implausible wiggliness of the splines, which were added in the following step (II.2). Random effects were multiplied on $\hat{A}$ and $\hat{B}$, thus affecting both level and slope of the h-d relation. Again, the first-order parameters λ and C were optimized via iterative grid search and the optimal combination was selected based on the Akaike Information Criterioin (AIC). GLMMs were fitted using R package `lme4` (Bates et al. 2015).
  - **II.2** As in step I.2, one model was fitted for each parameter individually. Again,

the covariate effect was estimated as an unconstrained spline (Equation 2.6). GAMs with mixed effects (GAMMs) at the plot level were created with a reduced inventory data set. Data from the French NFI was excluded. Instead, covariate effects were estimated from the parameters for soil chemistry and physiology which was available for the NFSI and BioSoil inventories. Parameters were evaluated based on p-value, R^2, and plausibility respectively. The R package `gamm4` was used to fit GAMMs (Wood and Scheipl 2020).

– **II.3** Finally, candidate parameters selected in step II.2 were combined. Different covariate effect combinations were estimated as unconstrained splines and tested for effect significance (p-value) and plausibility.

$$\ln(h_{k,i}) = A_{k,i} - B_{k,i}x_{k,i} + \epsilon_{k,i} \tag{2.1}$$

with

$$x_{k,i} = \frac{(\texttt{dbh}_{k,i} + \lambda)^{-c} - (30 + \lambda)^{-c}}{(10 + \lambda)^{-c} - (30 + \lambda)^{-c}} \tag{2.2}$$

and:

$h_{k,i}$: height of tree i at sample plot k;

$x_{k,i}$: transformed `dbh` of tree i at sample plot k;

$\texttt{dbh}_{k,i}$: `dbh` for tree i at sample plot k;

$A_{k,i}, B_{k,i}, C, \lambda$: first-order parameters of the h-d model at sample plot k;

ln(.): natural logarithm

$\epsilon_{k,i}$: error term for tree i at sample plot k with $\epsilon_{k,i} \sim \text{Normal}(0, \sigma^2)$

$$\ln(\mu_{k,i}) = \beta_{I,1} + \beta_{I,2}x_{k,i} + \beta_{I,3}x_{k,i}\texttt{qmD}_k + f(\texttt{qmD}_k) \tag{2.3}$$

with

$\ln(\mu_{k,i})$ natural logarithm of the conditional expectation $\text{E}(h_{k,i}|x_{k,i}, \texttt{qmD}_k)$ of tree height.

$h_{k,i} \sim \text{Gamma}(\mu_{k,i}, \theta)$ with expected value μ and scale parameter θ;

$f(\texttt{qmD}_k)$ one-dimensional spline, describing a non-linear effect of the quadratic mean diameter on the level of the h-d relationship (first-order parameter A);

$\beta_{I.}$: coefficients describing fixed linear effects on level (first-order parameter A) or slope (first-order parameter B) of the h-d relationship;

$x_{k,i}$: transformed `dbh` of tree i at sample plot k (see Equation 2.1);

$\texttt{qmD}_k$: factor of quadratic mean diameter for sample plot k on the slope of the h-d relationship (first-order parameter B);

$$\ln(\mu_{k,i}) = \beta_{I,1} + \beta_{I,2}x_{ki} + \beta_{I,3}x_{ki}\texttt{qmD}_k + f(\texttt{qmD}_k) + f_{I,1}(x_{I,1,k}) + f_{I,2}(x_{I,2,k}) + ... + f_{I,n}(x_{I,n,k}) \tag{2.4}$$

with:

$x_{I.k}...x_{I,n,k}$: covariates with potentially non-linear one-dimensional effects on the h-d relationship;

$f_{I,1}(x_{I,1,k})...f_{I,n}(x_{I,n,k})$: one-dimensional splines affecting the level of the h-d relationship (first-order parameter A).

$$\begin{aligned} \hat{A}_k &= \beta_{I,1} + f_{I,1}(x_{I,1,k}) + f_{I,2}(x_{I,2,k}) + ... + f_{I,n}(x_{I,n,k}) \\ \hat{B}_k &= \beta_{I,2}x_{k,i} + \beta_{I,3}x_{k,i}\texttt{qmD}_k \\ \ln(\mu_{k,i}) &= \beta_{II,1}\hat{A}_k + \gamma_{k,1} + x_{k,i}(\beta_{II,2}\hat{B}_k + \gamma_{k,2}) \end{aligned} \tag{2.5}$$

with

$h_{k,i} \sim \text{Gamma}(\mu_{k,i}, \theta)$ with expected value μ and scale parameter θ;

$\hat{A}_k, \hat{B}_k$: estimates of first-order parameters A_k and B_k (Equation 2.1);

$\beta_{II,1}, \beta_{II,2}$: coefficients describing linear fixed effects on the estimates of first-order parameters A and B;

$x_{II,1,k}...x_{II,m,k}$: covariates with one-dimensional effects on the h-d relationship;

$$\ln(\mu_{k,i}) = \beta_{II,1}\hat{A}_k + \gamma_{k,1} + x_{k,i}(\beta_{II,2}\hat{B}_k + \gamma_{k,2}) + f_{II,1}(x_{II,1,k}) + f_{II,2}(x_{II,2,k}) + ... + f_{II,m}(x_{II,m,k}) \tag{2.6}$$

with:

$f_{II,1}(x_{II,1,k}) + ...f_{II,m}(x_{II,m,k})$: one- dimensional penalized splines describing the level of the h-d relationship;

$\gamma_{k,1}, \gamma_{k,2}$: random effects for sample plot k

$$\text{amp}(f_j(x_j)) = \frac{\max(f_j(x_j)) - \min(f_j(x_j))}{\sum_{\tilde{j}=1}^{n} \max(f_{\tilde{j}}(x_{\tilde{j}})) - \min(f_{\tilde{j}}(x_{\tilde{j}}))} \tag{2.7}$$

with:

$\text{amp}(x_j)$: relative amplitude for covariate j

$x_j...x_n$: terms for covariates $j...n$, meaning $\beta_j x_j$ for linear effects or $f_j(x_j)$ for splines.

As an exemplary application, projections were calculated based on the models of step `I.3`. Temperature and precipitation projections were obtained from the WorldClim dataset based on the model MPI-ESM RCP 4.5. Parameters were selected as 30 year averages for a baseline period from 1961 to 1990 and a future period from 2051 to 2080. `qmD` was obtained

from German yield tables for all four species for a 100 year old stand (Schober 1975). Since traditional yield tables tend to underestimate current tree and stand dimensions, `qmD` was corrected using species-specific calibration functions (Staupendahl and Schmidt 2016). For simplicity, `dbh` was assumed to be equal to `qmD`.

2.3 Results

2.3.1 Model Step I

Phytocentric models (Step I.1) based only on the transformed diameter $x_{k,i}$ and `qmD` (Equation 2.3) showed R^2 above 50% with estimated residual standard deviation close to 4 m for all species. Monotony constraints were added for the splines of all species to avoid declining effects at higher `qmD`s. Grid optimization of the first-order parameters yielded λ of 10 for oak and spruce and 13 and 22 for beech and pine respectively. C was set to 2.4 (oak), 1.6 (beech), 1.9 (spruce) and 3.5 (pine).

In step I.2, all individual models yielded high significances ($p < 2 \cdot 10^{-16}$) for the respective geocentric covariate.

For the final model (step I.3), two climate parameters were selected as covariates for oak, beech, spruce and pine: the temperature mean for the warmest quarter (from here on denoted as `T`) and the precipitation sum for the warmest quarter (`P`) (Appendix A.1). These specific covariates were selected for the following reasons:

1. Aggregations for the warmest quarter yielded plausible effect shapes for the four species.
2. Including the same two covariates for each species for inter-species comparability.
3. The warmest quarter is a dynamic date range which may change with changing climate conditions and adds flexibility for projections.

In addition, `awc` was selected as a geocentric covariate from the soil parameter set for all species (Equation 2.8, Figure 2.2). Compared to step I.1, the final grid optimization of first-order parameters lead to little to no change in λ and C except for spruce. λ stayed the same for oak and pine, while increasing to 14 for beech. For spruce, λ was reduced to 2. C stayed at 2.4 for oak and increased by 0.2 for beech and pine (to 1.8 for and 3.7 respectively). For spruce, it decreased to 1.0. The estimate of the scale (θ) of the Gamma distribution which was assumed for the response was 0.043 for oak, 0.050 for beech, 0.036 for spruce and 0.076 for pine.

$$
\begin{aligned}
\ln(\mu_{k,i,oak})\text{oak}]) &= \beta_{I,1}\beta_{I,2}x_{k,i} + \beta_{I,3}x_{k,i}\texttt{qmD}_k\beta_{I,4}\tilde{\texttt{P}} + m(\texttt{qmD}_k) + f(\texttt{T}_k) + mc(\texttt{awc}_k) \\
\ln(\mu_{k,i,\text{beech}}) &= \beta_{I,1} + \beta_{I,2}x_{k,i} + \beta_{I,3}x_{k,i}\texttt{qmD}_k + m(\texttt{qmD}_k) + f(\texttt{T}_k) + m(\texttt{P}_k) + m(\texttt{awc}_k) \\
\ln(\mu_{k,i,\text{spruce}}) &= \beta_{I,1} + \beta_{I,2}x_{k,i} + \beta_{I,3}x_{k,i}\texttt{qmD}_k + m(\texttt{qmD}_k) + f(\texttt{T}_k) + m(\texttt{P}_k) + m(\texttt{awc}_k) \\
\ln(\mu_{k,i,\text{pine}}) &= \beta_{I,1} + \beta_{I,2}x_{k,i} + \beta_{I,3}x_{k,i}\texttt{qmD}_k + m(\texttt{qmD}_k) + c(\texttt{T}_k) + m(\texttt{P}_k) + m(\texttt{awc}_k)
\end{aligned}
\tag{2.8}
$$

with:

$f(.)$ spline; to reduce wiggliness, the spline base for $f(T_k)$ was reduced to 6 for beech.

$m(.)$ spline with monotony constraint

$c(.)$ spline with concavity constraint

$\tilde{\texttt{P}}_k = \begin{cases} \texttt{P}_k - 150, & \text{if } \texttt{P}_k \leq 150, \\ 0, & \text{if } \texttt{P}_k > 150 \end{cases}$: segmented effect of precipitation sum for the warmest quarter.

T mean temperature for the warmest quarter plot k.

$\texttt{P}_k$ precipitation sum for the warmest quarter for plot k.

$\texttt{awc}_k$ available water capacity for plot k.

Statistics on splines are given in the form of estimated degrees of freedom (*edf*), which indicate the wiggliness of its shape (Table 2.5). An *edf* value of 1 corresponds to a perfectly linear shape.

Splines showed parabolic as well as asymptotic shapes (Figure 2.2). qmD showed a clear asymptotic shape for oak, beech and spruce and an asymptotic trend for pine, but with a continuously positive slope. Parabolic shapes were observed for $f(T)$. The parabolic shape is most pronounced for pine, with a steep decrease after the optimum. $f(P)$ shows a monotonous increase for beech, spruce and pine, trending towards asymptotes above the 95% quantile of precipitation data. A different effect description for P has to be applied for oak, since splines proved to be unstable in connection with T and awc, in spite of an obvious negative effect of P below 150 mm (see Appendix A.2). Hence, a segmented linear effect is fitted, with a coefficient for values below 150 mm and values above 150 mm set to zero. $f(\texttt{awc})$ also increases monotonously for all species, trending towards an asymptote. Note that qmD has both a linear effect on slope and a nonlinear effect on the level of the h-d relation (Figure 2.3).

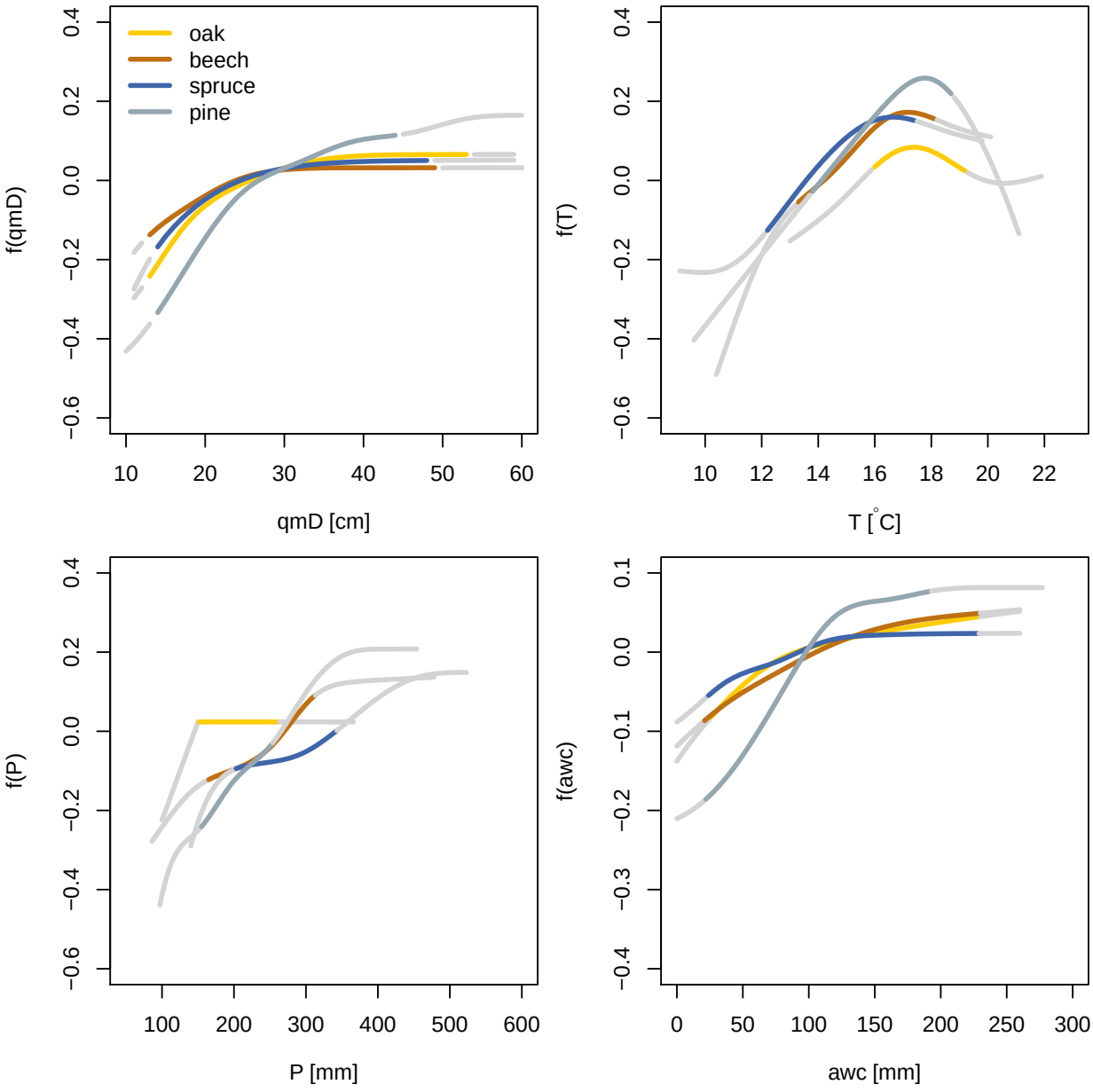


Figure 2.2: Covariate effects for `qmD`, `T`, `P` and `awc` for oak, beech, spruce and pine (see Equation 2.8). Colored line segments identify the inter-quantile range, light grey segments the extremes of the data range. Note the varying y axis ranges. See Appendix A.2 for single splines with confidence limits.

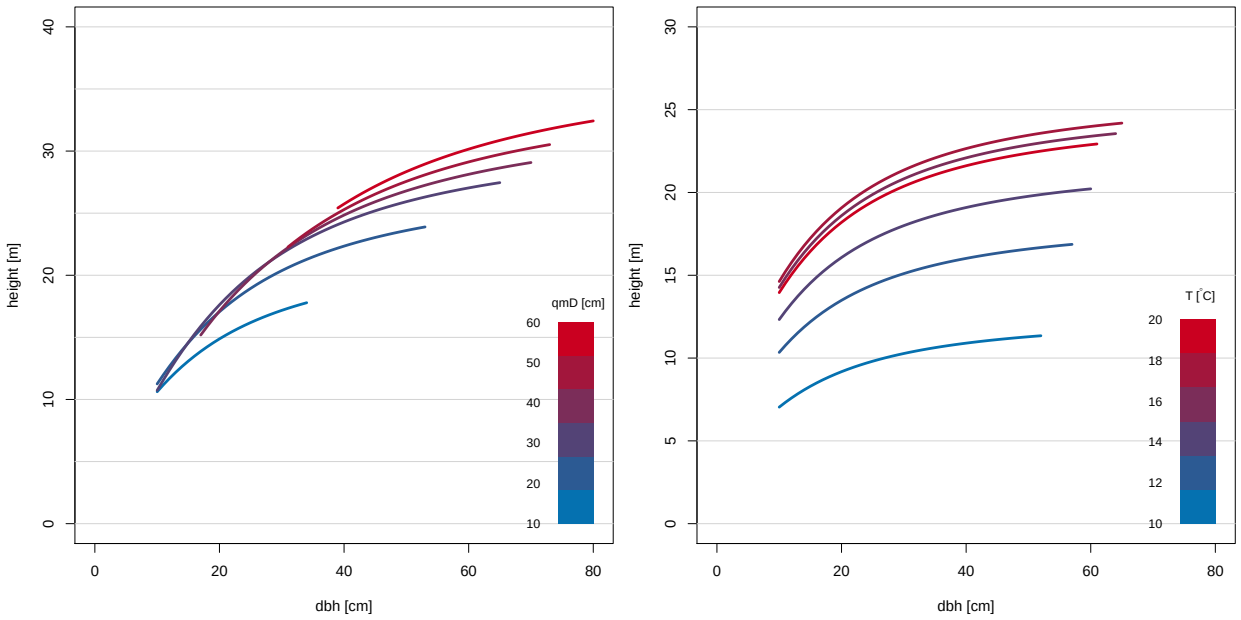


Figure 2.3: Shift in the level of the h-d curve for beech based on the effect of `qmD` and `T` (see Figure 2.2). Curves were plotted by varying the respective covariate and setting all other covariates to mean.

2.3.2 Model Step II

In step II.1, the exclusion of the French NFI from the data set reduced the number of observations by 26,000 to 70,000, leaving 1,007 observations for oak, 2,119 for beech, 2,965 for spruce and 2,591 for pine (Table 2.4).

In comparison to step I.1, first-order parameters λ and C increased for beech, oak and spruce while decreasing for pine. The estimate for the scale θ of the Gamma distribution was 0.009 for oak and 0.005 for pine respectively.

Table 2.4: Summary statistics of the final models of step II (step II.3, see equation 2.9).

Species	n	λ	C	R^2 [%]	Est. residual standard deviation [m]	R^2 (random effects) [%]	Standard error [m] (random effects)
Oak	1007	13	2.50	65.40	3.86	93.37	1.68
Beech	2259	19	2.50	63.73	4.71	93.77	1.93
Spruce	2965	17	2.00	77.07	3.78	96.69	1.43
Pine	2591	9	2.00	58.53	4.01	95.46	1.33

R^2 (for fixed effects) decreased slightly for beech and pine (Table 2.4). It increased marginally for oak and stayed the same for pine. Estimated residual standard deviation increased a few cm for all species. Random effects increased R^2 above 90% while reducing estimated residual standard deviation below 2 m for all species.

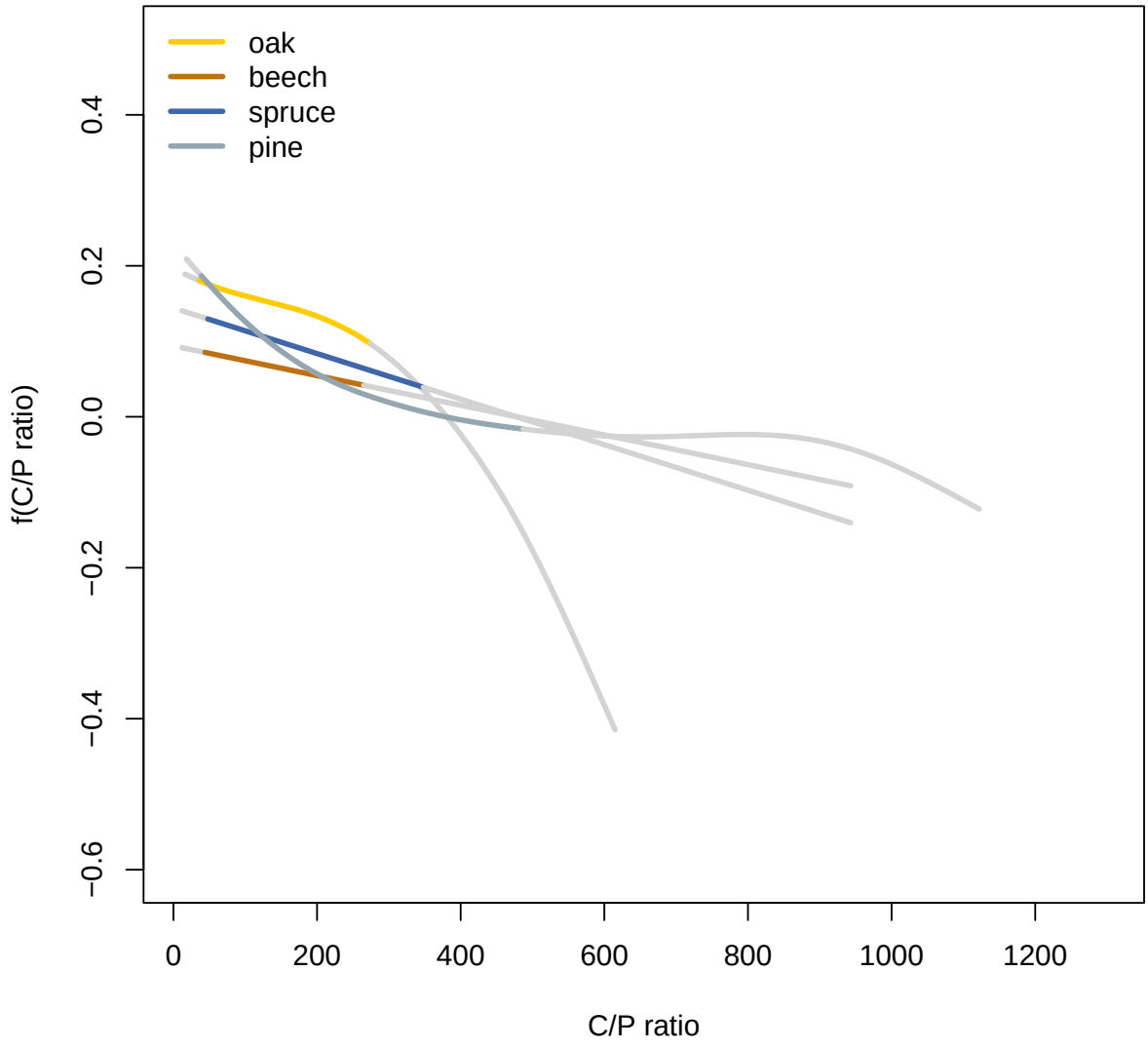


Figure 2.4: Penalized regression splines for `CP`, aggregated for 0-10 cm soil depth for oak, beech, spruce and pine for model step II.3. Lines are colored for the 5%-95% quantile range and lightgray for the edges of the data range.

Within step II.2, pH value, base saturation, potassium stock and air capacity were among the candidate covariates for all species. However, C/P ratio (CP) for 0-10 cm soil depth turned out to be the only covariate that was significant and plausible for all species. Hence, it was selected as a covariate for the final models of step II.

$$\ln(\mu_{k,i}) = \beta_{II,1}\hat{A}_k + \gamma_{k,1} + x_{k,i}(\beta_{II,2}\hat{B}_k + \gamma_{k,2}) + f(\texttt{CP}_k) \tag{2.9}$$

with:

$f(\texttt{CP})$ spline for the CP of plot k, aggregated for 0-10 cm soil depth.

Table 2.5: Statistics on model coefficients, terms and covariates for the final models of step I and II (steps I.3 and II.3, equations 2.8 and 2.9). Estimated degrees of freedom (EDF) indicate the wiggliness of the effect, with 1 specifying a linear effect. * Note that relative amplitudes (eq. 2.7) of the linear covariates of step II contain the amplitudes of step I.

Species	Coefficient/ Spline	Estimate	Standard error	P-value	EDF	Term	Relative amplitude [%]
Oak	$\beta_{I,0}$	2.34	0.06	0.00		$\beta_{I,0}$	
Oak	$\beta_{I,1}$	0.20	0.01	0.00		$\beta_{I,1}x_{k,i}$	8.4
Oak	$\beta_{I,2}$	0.02	0.00	0.00		$\beta_{I,2}x_{k,i}\texttt{qmD}_k$	36.3
Oak	$\beta_{I,4}$	0.01	0.00	0.00		$\beta_{I,4}\tilde{\texttt{P}}_k$	0
Oak	$m(\texttt{qmD}_k)$			0.00	4.89	$m(\texttt{qmD}_k)$	12.3
Oak	$f(\texttt{T}_k)$			0.00	3.81	$f(\texttt{T}_k)$	15.7
Oak	$m(\texttt{awc}_k)$			0.00	3.77	$m(\texttt{awc}_k)$	8.9
Oak	$\beta_{II,1}$	1.02	0.00	0.00		$\beta_{II,1}\hat{A}_k$	26.5 *
Oak	$\beta_{II,2}$	1.01	0.05	0.00		$\beta_{II,2}\hat{B}_k x_{k,i}$	55.1 *
Oak	$f(\texttt{CP}_k)$			0.00	1.78	$f(\texttt{CP}_k)$	18.4
Beech	$\beta_{I,0}$	1.99	0.21	0.00		$\beta_{I,0}$	
Beech	$\beta_{I,1}$	0.37	0.01	0.00		$\beta_{I,1}x_{k,i}$	21.4
Beech	$\beta_{I,2}$	0.01	0.00	0.00		$\beta_{I,2}x_{k,i}\texttt{qmD}_k$	25.7
Beech	$m(\texttt{qmD}_k)$			0.00	3.99	$m(\texttt{qmD}_k)$	5
Beech	$f(\texttt{T}_k)$			0.00	4.85	$f(\texttt{T}_k)$	3.5
Beech	$m(\texttt{P}_k)$			0.00	4.69	$m(\texttt{P}_k)$	25.4
Beech	$m(\texttt{awc}_k)$			0.00	4.11	$m(\texttt{awc}_k)$	10.5
Beech	$\beta_{II,1}$	1.01	0.00	0.00		$\beta_{II,1}\hat{A}_k$	50.9 *

Beech	$\beta_{II,2}$	1.25	0.04	0.00		$\beta_{II,2}\hat{B}_k x_{k,i}$	40.6 *
Beech	$f(\texttt{CP}_k)$			0.01	1.00	$f(\texttt{CP}_k)$	8.5
Spruce	$\beta_{I,0}$	1.13	0.24	0.00		$\beta_{I,0}$	
Spruce	$\beta_{I,1}$	0.62	0.01	0.00		$\beta_{I,1}x_{k,i}$	30.7
Spruce	$\beta_{I,2}$	0.01	0.00	0.00		$\beta_{I,2}x_{k,i}\texttt{qmD}_k$	20.9
Spruce	$m(\texttt{qmD}_k)$			0.00	4.78	$m(\texttt{qmD}_k)$	5.5
Spruce	$f(\texttt{T}_k)$			0.00	5.71	$f(\texttt{T}_k)$	5.8
Spruce	$m(\texttt{P}_k)$			0.00	4.08	$m(\texttt{P}_k)$	12.9
Spruce	$m(\texttt{awc}_k)$			0.00	3.87	$m(\texttt{awc}_k)$	3
Spruce	$\beta_{II,1}$	1.01	0.00	0.00		$\beta_{II,1}\hat{A}_k$	40.6 *
Spruce	$\beta_{II,2}$	1.04	0.02	0.00		$\beta_{II,2}\hat{B}_k x_{k,i}$	38.3 *
Spruce	$f(\texttt{CP}_k)$			0.00	1.00	$f(\texttt{CP}_k)$	21.1
Pine	$\beta_{I,0}$	0.35	0.55	0.52		$\beta_{I,0}$	
Pine	$\beta_{I,1}$	0.26	0.02	0.00		$\beta_{I,1}x_{k,i}$	6.5
Pine	$\beta_{I,2}$	0.02	0.00	0.00		$\beta_{I,2}x_{k,i}\texttt{qmD}_k$	15
Pine	$m(\texttt{qmD}_k)$			0.00	5.95	$m(\texttt{qmD}_k)$	18.9
Pine	$c(\texttt{T}_k)$			0.00	4.13	$c(\texttt{T}_k)$	7.7
Pine	$m(\texttt{P}_k)$			0.00	5.00	$m(\texttt{P}_k)$	10.8
Pine	$m(\texttt{awc}_k)$			0.00	4.15	$m(\texttt{awc}_k)$	18.2
Pine	$\beta_{II,1}$	1.07	0.00	0.00		$\beta_{II,1}\hat{A}_k$	51.6 *
Pine	$\beta_{II,2}$	0.84	0.03	0.00		$\beta_{II,2}\hat{B}_k x_{k,i}$	25.4 *
Pine	$f(\texttt{CP}_k)$			0.00	2.45	$f(\texttt{CP}_k)$	23

Species	Covariate	5% quantile	95% quantile	Minimum	Maximum
Oak					
Oak	dbh	9.87	57.93	7.40	127.32
Oak					
Oak	$\texttt{P}_k$	151.00	263.00	100.00	366.00
Oak	qmD	12.71	53.10	10.00	60.00
Oak	$\texttt{T}_k$	15.90	19.30	13.00	21.90

Oak	awc	25.80	230.00	0.00	260.00
Oak	$\hat{A}$	2.68	3.06	2.48	3.08
Oak	$\hat{B}$	0.40	1.27	0.28	2.07
Oak	CP	35.00	273.12	15.62	615.70
Beech					
Beech	dbh	8.59	58.25	7.00	174.12
Beech					
Beech	qmD	12.15	49.07	10.01	60.00
Beech	T_k	13.20	18.20	10.40	20.20
Beech	P_k	164.00	313.00	86.00	478.00
Beech	awc	20.80	230.00	0.00	260.00
Beech	$\hat{A}$	2.88	3.26	2.50	3.36
Beech	$\hat{B}$	0.50	0.99	0.42	3.04
Beech	CP	43.57	264.52	11.70	943.64
Spruce					
Spruce	dbh	10.19	57.30	7.00	106.95
Spruce					
Spruce	qmD	13.95	48.75	10.00	59.90
Spruce	T_k	12.10	17.50	9.10	19.80
Spruce	P_k	202.00	343.00	140.00	523.00
Spruce	awc	23.25	228.56	0.00	260.00
Spruce	$\hat{A}$	2.82	3.28	2.21	3.43
Spruce	$\hat{B}$	0.72	1.10	0.63	1.33
Spruce	CP	47.14	347.94	11.28	943.64
Pine					
Pine	dbh	9.55	48.70	7.00	91.04
Pine					
Pine	qmD	13.14	44.23	10.00	60.00
Pine	T_k	13.70	18.80	9.60	21.20
Pine	P_k	154.00	254.00	97.00	454.00
Pine	awc	21.50	192.50	0.00	277.20
Pine	$\hat{A}$	2.32	3.15	1.98	3.44

Pine	$\hat{B}$	0.45	1.03	0.27	1.47
Pine	CP	38.05	487.00	17.40	1122.73

2.3.3 Projections

Projections for all species showed potential productivity losses in southern and increases in northern Europe (Figure 2.5). In central Europe, productivity decreased at lower elevations, while increases were visible at higher elevations in the Alps and central European mountain ranges. Large areas will be shifted into the extrapolation range of models in southern regions in the future, while areas in Norway and Finland will move into the range of the model data.

2.4 Discussion

In the study presented here, climate sensitive, phyto-geocentric h-d models for oak, beech, spruce and pine were built. Models were based on a unique set of large scale, one-time forest inventory data using an SFT approach (Bontemps and Bouriaud 2013).

Model development started with fitting a phytocentric model, using qmD as a longitudinal covariate. qmD has been successfully used as an indicator of stand development for non-even aged stands that lack precise age estimates (Mehtätalo 2004; Schmidt et al. 2018). Since it is based on dbh, it is easier to obtain and less subject to error compared to age. Another advantage over age is that it can be easily calculated, even in mixed forests, where age is hard to define. One disadvantage is that tree height at a specific point in time can't be directly predicted with the qmD based models. A possible workaround would be to obtain qmD values for a specific age from a yield table. Also, it can't be known to which degree the effect of environmental covariates is hidden in the effect of qmD. Comparing covariate effects from the given models to an age based model will be the subject of a follow up study.

The original model by Lappi (1997) contains the parameters λ and C (Equation 2.1), which was optimized using an iterative grid search. Values for λ and C are consistently higher than in the original study on jack pine, where λ was set to 1 cm and C to 0.6-1.6 depending on age. The values were also mostly higher than in a preceding study on Norway spruce, where λ was set to 7 cm (Mehtätalo 2004) and C to 1.64. However, λ and C were very similar to a study on spruce, pine and birch, which also employed an iterative grid search and identified values between 16 and 20 cm for λ and 2.4 and 2.5 for C (Schmidt et al. 2018). It was pointed out in that study that values were often located at the upper limits of the search grid. This supports the higher values of λ=22 cm and C=3.5 for pine in the present study, where the search grid was adjusted if λ and C reached the limits of the grid.

The selection of covariates in step I.2 and the subsequent covariate combinations in step

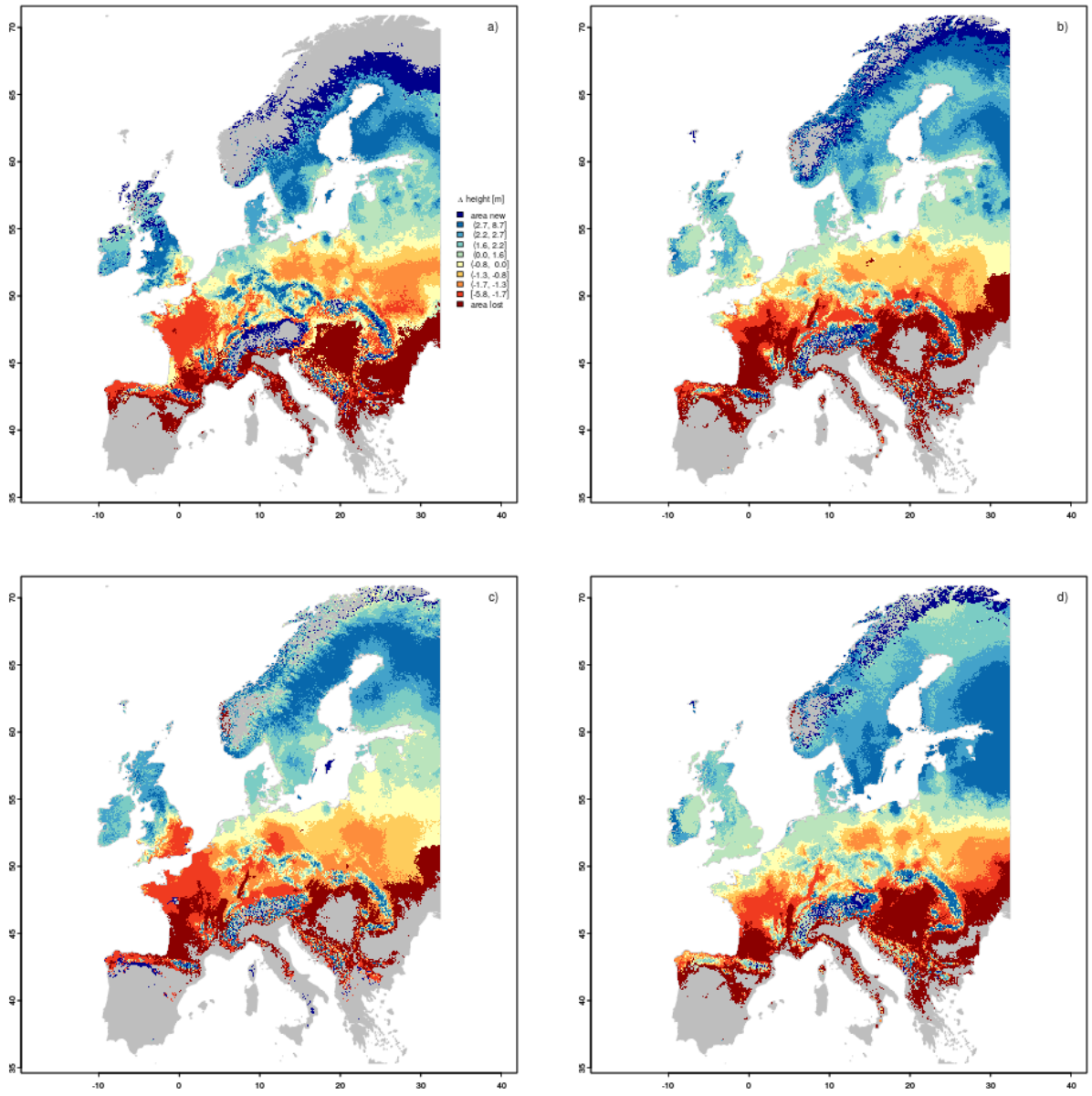


Figure 2.5: Change in height for oak (a), beech (b), spruce (c) and pine (d) at age 100 from a baseline period 1960-1990 to 2050-2080, based on RCP 4.5. Dark red areas (areas lost) were part of the model's data range in the baseline period, but enter the extrapolation range in the projection period. Vice versa, dark blue areas were not part of the model's data range during the baseline period, but enter the range in the projection period.

I.3 showed that the effects of temperature and precipitation are mostly plausible (Appendix A.2), with temperature often showing parabolic shapes.

Responses of plant-growth related covariates are often assumed to be parabolic rather than asymptotic. However, it has been stated that this assumption is rarely backed by data. Rather, it is predefined by fitting parabolic functions, e.g. polynomials. It has been suggested that an asymptotic response is closer to ecological reality and should be implemented in forest simulators as opposed to a parabolic response (Loehle 2000). The data-based identification of parabolic temperature effects is a clear benefit of the large spatial extent of the data set.

The effects of `P` displayed higher relative amplitudes than the effects of `T` (Table 2.5) except for oak. There was a strong tendency towards asymptotic shapes for all species, however, `P`-distributions were heavily right-skewed (Figure 2.2) and asymptotes were located above the 95% quantile of `P` observations, while steep declines were observed below the 5% quantile. With regard to the `P` effects, it was apparent that more data in the extreme spectrum (dry and wet) is needed.

For `awc`, asymptotic effects were based on evenly distributed data for all species. Parameters describing the ability of the soil to supply plant available water are regularly used to fit site productivity models (Bergès et al. 2005; Friedrichs et al. 2009; Brandl et al. 2014, 2019; Yener et al. 2018; Mellert et al. 2018).

For pine, the `awc`-effect displayed an exceptionally high amplitude, while being based on a smaller data range than for beech. High sensitivity and a small environmental gradient indicated a strong response to a limiting resource. Indeed, a study on tree species distribution based on soil water storage has pointed out a pronounced sensitivity of pine to `awc` (Mellert et al. 2018).

Results of the covariate selection in step II.2 were more ambivalent compared to step I.2 (Appendix A.3). Soil parameters were available for different subsets of data, with high variation in the number of observations. Distributions were often clustered and R^2 displayed higher variations; increase in variation was even more pronounced for the relative amplitude of effects. The selection of soil covariates showed that splines have to be evaluated based on expert knowledge, since favorable model statistics do not indicate plausibility. An example was the effect of cation exchange capacity for the 0-40 cm layer of the mineral soil (`ca.minst.040`), which displayed an amplitude of 41% but also an implausible steep decline for increasing cation exchange capacity at values below 5000 molc/ha (Appendix A.3).

`CP` was the only soil parameter that was selected as covariate for the final models in step II.3. A decreasing effect on logarithmic height with increasing `CP` was observed for all species, indicating lower tree heights at potentially more P-limited sites.

Frequently, nitrogen has been identified as a nutrient limiting productivity, but phosphorous may be limiting, especially at sites which are already nitrogen saturated (Talkner et al. 2015).

Phosphorous has been identified as limiting nutrient for spruce in the Bavarian Alps (Mellert and Ewald 2014). Although a marginal effect of `CP` on beech was observed, phosphorous limitation of sites and the effect of `CP` opposed to nitrogen stock, C/N ratio and absolute phosphorous stock have been observed for European beech forests (Talkner et al. 2015; Van Sundert et al. 2019).

Projections for beech and pine based on models of step `I.3` and `T` and `P` projections under the RCP 4.5 scenario displayed a shift in productivity from south to north and from lower to higher altitudes (Figure 2.5). Mountaineous sites have been described as similar to northern latitude sites with respect to the trees' ecological niche (Uhl et al. 2021). Overall patterns for beech were similar to a previous study based on the RCP 4.5 scenario and data for Germany and France (Brandl et al. 2018). The results extended further to the North due to the European data set, but the area covered in southern Europe was very similar. A possible improvement of productivity for beech in mountaineous areas has also been predicted in a site index modeling study for Lower Saxony (Albert and Schmidt 2010).

Finally, a few points concerning assumptions made within the scope of data aggregation and model building have to be discussed.

First, the method of identifying covariate effects needs to be addressed. Covariate identification was conducted by the evaluation of an unconstrained spline, and subsequent setting of constraints if deemed necessary. Of course, constraints should only be set if a generally plausible trend is already visible (Schmidt et al. 2018). It should also be noted that, if an ecological response is already well known and established functions exist, using those functions may be preferable to a GAM/SCAM approach. This was taken into account by using the linearized Korf function as a basis describing the h-d relation, with only the slope being modified by splines for `qmD`, and level by splines for geocentric covariates. Thus, in contrast to an exclusively spline based, free estimate of tree height, series of h-d curves can always be obtained (Figure 2.3).

It needs to be stressed that climate parameters were aggregated over a 30 year period without accounting for tree age, while `awc` and `CP` were assumed to be constant. Thus, the same effect of a climate covariate was assigned to, for example, a 30 year old stand and a 100 year old stand. The former might have been subjected to more recent and likely warmer climate and the latter to a colder climate period, compared to the given 30-year average. Climate covariate effects might be different in shape, amplitude and significance given the actual climate period experienced by each tree, which was not available for the analyses presented here.

The study presents a typical example of space-for-time substitution. This approach includes some important assumptions on how trees or stands are affected by their environment over time. Caveats may be rooted in the lack of knowledge of events that have affected a stand in

its history, their chronology and spatial distribution and other potentially hidden assumptions (Yue et al. 2016)). For example, similar climate conditions might be expected to act differently if they arise early or late in the life of a tree.

Another caveat which may lead to a bias in the model fit could lie in different management strategies being applied based on site quality, ownership structures and size of the managed area. Poor sites are likely to be less intensively managed. Ownership structures encompass private, state-owned and communal forest. Small private forests are likely to be managed differently compared to large private forests, while management strategies by the state may vary between national sub-units such as federal states. Within the model selection approach, those differences in thinning methods may lead to unbalanced shifts in `qmD` and height, which would confound the correct identification of geocentric effects. For example, stands at poor sites may be mostly thinned from below, while stands at very productive sites may be thinned from above. The first method would shift `qmD` to higher, the latter to lower values. However, due to the large number of observations and the international scale of the data set, coverage of a large, balanced set of management strategies and ownership types, could be assumed.

2.5 Conclusion

The goal of the presented study was the creation of phyto-geocentric site productivity models for oak, beech, spruce and pine, based on a large-scale pan-European set of forest inventory data and data on climate, soil physiology and chemistry. A linearized version of the Korf function (Lappi 1997) was used for describing a basic h-d relation and added linear covariates using a GAM/GAMM/SCAM framework to check for shape and plausibility. Plausible effects for temperature and precipitation for the warmest quarter, available water capacity and C/P ratio, were identified. Temperature tended towards parabolic shapes, which are often assumed but rarely supported by data. Covariate effects of precipitation tended to be asymptotic. Based on effect shapes at the edges of the data range, more extensive sampling at more extreme, especially warmer sites is recommended to further validate a decreasing effect on site productivity at high temperatures. It was shown that a large scale SFT approach can provide a good basis for the identification of geocentric covariate effects.

3 Height-age models based on dynamic and static aggregation of climate data

3.1 Introduction

Climate sensitive forest productivity models are usually based on regionalized climate data. Climate data can encompass different aggregation methods of temperature and precipitation, like sums or means for a year, specific months or growing season (Albert and Schmidt 2010; Sharma et al. 2012; Brandl et al. 2018; Campbell et al. 2021). Other possible methods of aggregation include, e.g., the aggregation over the warmest quarter of each year or the seasonality defined as standard deviation of the temperature multiplied with 100 for temperature or the coefficient of variation for precipitation (Brandl et al. 2019). More extensive datasets combine climate and soil information to obtain hydrological parameters like soil water capacity (Bergès et al. 2005) or water balance during growing season (Brandl et al. 2014). If daily climate values are available, time frames during which values cross certain thresholds, like growing degree days (eg. days with air temperature $> 5°C$), can be aggregated (Brandl et al. 2014).

Regionalized climate data is based on specific measurements from climate stations. In rare cases, climate stations are located in the vicinity of the plots from which forest productivity information is obtained (Bouwman et al. 2020). Usually, they are located at greater distance to forest inventory or experimental plots and measurements are only valid for the point at which they are measured (Apaydin et al. 2004). In those cases, data is interpolated using statistical methods like regression approaches (inverse distance weighting, geographically weighted regression), spline based approaches or kriging (Hutchinson 1989; Tveito et al. 2001; Apaydin et al. 2004; Bergès et al. 2005; Hijmans et al. 2005; Schulla and Jasper 2007; Zimmermann et al. 2007; Daly et al. 2008; Dietrich et al. 2019). Selection of the best interpolation method depends on the characteristics of the data set as well as the aim of the interpolation. An important characteristic is the spatial distribution of the target climate variable, especially in mountainous regions where data is sparse and varies on small spatial scales (Apaydin et al. 2004). Scales of the respective regionalizations can range from very fine scales for regional models ($50\,\mathrm{m} \times 50\,\mathrm{m}$, Zimmermann et al. 2007) to the $1\,\mathrm{km} \times 1\,\mathrm{km}$ scale for global climate grids (Hijmans et al. 2005).

Climate measurements are restricted to specific time periods, depending on the beginning of measurements by the available stations. Time frames for the interpolation of past climate station data typically reach from the 1960s or 1970s to the 1990s or the beginning of the millennium (Bergès et al. 2005; Zimmermann et al. 2007; Sharma et al. 2012; Wang et al. 2016; Dietrich et al. 2019).

Aggregation periods for climate data rarely cover the lifetimes of actual trees or stands. Often, the time frame for climate interpolation differs from the sampling time frame of the underlying productivity data. E.g., forest inventory data from 2012 may be combined with climate parameter means for 1971 to 2000 (Brandl et al. 2014, 2018). In some studies, the available climate data does cover the sampling years of stand data (Dănescu et al. 2017; Brandl et al. 2019). Yet even in those cases, climate variables are averaged over a static time frame ranging from 5 to 30 years.

Consequently, static climate aggregates may not be related to the actual climate which trees or stands have been subjected to during their lifetime. Especially using a SFT approach, a static climate aggregation over a fixed time period assigns climate from the same period to trees or stands of different ages.

In this study, site-productivity models were fitted, based on data from the third German National Forest Inventory and climate values aggregated for two static periods and dynamically over stand age.

Based on the fitted models, the following questions were investigated:

1. Do climate values aggregated dynamically over stand age yield different covariate combinations and effects during model selection?

2. Do site-productivity models based on dynamic and static climate data aggregations yield different predictions?

3. Which recommendations can be made for future modeling studies based on observed differences in covariate selection and climate data aggregation?

3.2 Materials & Methods

3.2.1 Forest inventory data

Forest inventory data was taken from the third German National forest inventory (NFI). This was conducted in 2012 on a stratified grid. The basic nation wide grid had a density of $4\,\text{km} \times 4\,\text{km}$ (16 km^2). Local strata were intensified to a double ($2.83\,\text{km} \times 2.83\,\text{km}$, 8 km^2) or quadruple ($2\,\text{km} \times 2\,\text{km}$, 4 km^2) grid (Riedel et al. 2017). Sampling plots were set up as quadratic *tracts* with an edge length of 150 m. Trees with diameter at breast height (`dbh`) of 7 cm and above were sampled using angle count sampling with a counting factor of four at each of the four tract corners (Bitterlich 1947). This yielded a number of 4 to 6 measured trees per tract for beech and pine (Table 3.1). Beech data included more trees of higher age, height and `dbh` , while also including more low `dbh` trees compared to pine (Figure 3.1).

Table 3.1: Statistics on the stand inventory data of the third German NFI.

Species	N trees	N tracts	N trees/tract	age min
Beech	92115	22801	4.04	9
Pine	92726	17177	5.40	8
Species	age quantile 5% [cm]	age quantile 95% [cm]	age max [cm]	dbh min [cm]
Beech	33	170	464	0.1
Pine	27	141	281	0.1
Species	dbh quantile 5% [cm]	dbh quantile 95% [cm]	dbh max [cm]	N heights/tract
Beech	0.100	68.7	200.0	4.04
Pine	8.625	51.2	112.1	5.40
Species	height min [m]	height quantile 5% [m]	height quantile 95% [m]	height max [m]
Beech	2.3	12.1	37.5	48.7
Pine	1.6	12.6	31.9	45.9

3.2.2 Climate data

Climate data was provided in the form of raster grids for annual temperature sum (`T`) and precipitation sum (`P`). Both parameters were aggregated for a dynamic vegetation period (Nuske 2017). `T` grids were provided with $50\,\mathrm{m} \times 50\,\mathrm{m}$ and `P` grids with 100×100m resolution. Retrodictions were available for all years from 1900 to 2019. Predictions were available for 2020 to 2100 based on different projection models. Retrodictions and projections were concatenated to aggregate climate for different periods (see below), using the projections from the MIROC5 model (Watanabe et al. 2010).

Table 3.2: Statistics on the climate data.

Species	Temperature sum min [°C]	Temperature sum quantile 5% [°C]	Temperature sum quantile 95% [°C]	Temperature sum min [°C]
Beech	505	1803	2647	3224
Pine	835	2020	2751	3224
Species	Precipitation sum min [mm]	Precipitation sum quantile 5% [mm]	Precipitation sum quantile 95% [mm]	Precipitation sum min [mm]
Beech	99	227	662	1539
Pine	88	189	505	1442

3.2.3 Soil data

Soil data was regionalized from the German National Forest Soil Inventory (NFSI), which was sampled on an 8 km × 8 km grid (Höhle et al. 2018). The regionalization took place within the scope of the project *Forest Productivity-Carbon Sequestration-Climate Change* (Waldproduktivität-Kohlenstoffspeicherung-Klimawandel, WP-KS-KW, Mette and Kölling 2015) using the NFSI data and additional soil profiles. Methods included Ordinary Least Squares, Boosted Regression Trees and Random Forests. Regionalization models were applied to large strata which were based on landscape properties (Wilpert et al. 2017 for overwiev; Zirlewagen and Wilpert 2010 for detailed method description; Wilpert et al. 2016). Parameters were available for aggregation depths 0-30 cm and 30-60 cm (Table 3.3).

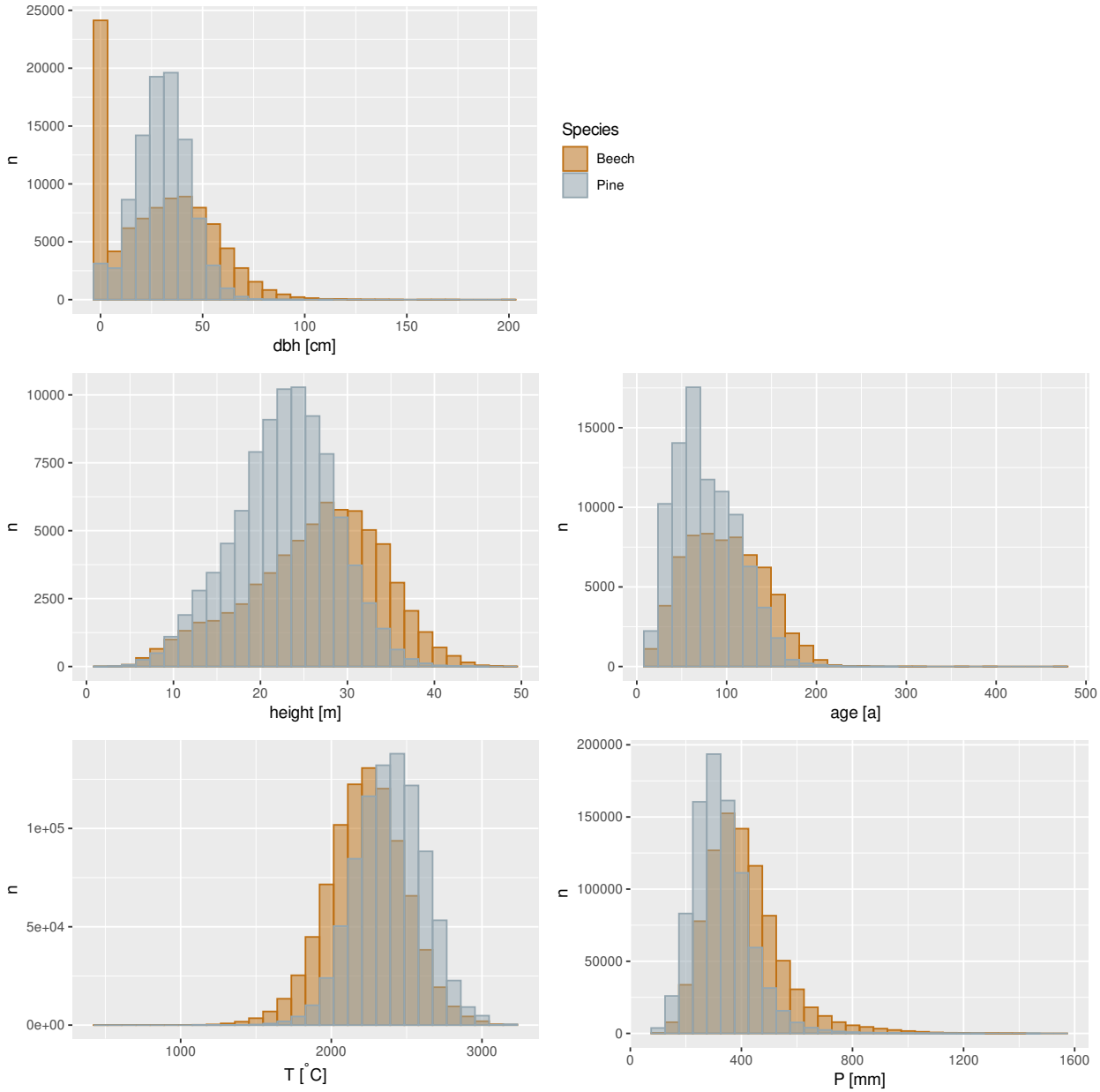


Figure 3.1: Histograms of diameter at breast height (`dbh`), height, age and temperature/precipitation (`T`/`P`) sums for the vegetation period for single tree data.

Table 3.3: Statistics on soil parameters available for covariate selection in step II.2 of the h-a model.

Species	Parameter	Unit	Code	Aggregation depth	Min	5% quantile	95% quantile	Max
Beech	Cation exchange capacity	μeq/g	`ake`	0-30	1.98	4.67	120.00	532.78
Pine	Cation exchange capacity	μeq/g	`ake`	0-30	1.90	5.12	70.26	226.70
Beech	Cation exchange capacity	μeq/g	`ake`	30-60	2.27	6.78	107.71	467.90
Pine	Cation exchange capacity	μeq/g	`ake`	30-60	2.27	5.49	45.96	211.11
Beech	Base saturation	%	`bs`	0-30	0.68	7.27	100.00	100.00
Pine	Base saturation	%	`bs`	0-30	0.68	2.89	41.28	100.00
Beech	Base saturation	%	`bs`	30-60	0.00	7.37	100.00	100.00
Pine	Base saturation	%	`bs`	30-60	0.00	3.94	44.90	100.00
Beech	Carbon-nitrogen ratio	1/1	`cn`	0-30	9.17	12.59	24.02	36.73
Pine	Carbon-nitrogen ratio	1/1	`cn`	0-30	9.70	17.79	33.08	39.27
Beech	Carbon-nitrogen ratio	1/1	`cn`	30-60	1.60	6.70	21.92	40.35
Pine	Carbon-nitrogen ratio	1/1	`cn`	30-60	1.60	9.69	32.43	43.43
Beech	Organic carbon content	g/kg	`corg`	0-30	0.38	8.77	72.83	456.71
Pine	Organic carbon content	g/kg	`corg`	0-30	0.38	0.38	30.94	456.71

Beech	Organic carbon content	g/kg	`corg`	30-60	0.04	0.30	19.57	241.85
Pine	Organic carbon content	g/kg	`corg`	30-60	0.04	0.06	7.36	158.00
Beech	Proportion of bulk soil	%	`gba`	0-30	0.00	1.07	39.05	73.33
Pine	Proportion of bulk soil	%	`gba`	0-30	0.00	0.18	13.76	73.33
Beech	Proportion of bulk soil	%	`gba`	30-60	0.00	2.07	69.68	95.00
Pine	Proportion of bulk soil	%	`gba`	30-60	0.00	0.14	26.70	95.00
Beech	Soil depth	cm	`gruend`		2.50	43.74	110.30	209.21
Pine	Soil depth	cm	`gruend`		2.50	48.45	157.87	209.21
Beech	Available water capacity	%	`nfk`		10.84	16.56	28.06	32.91
Pine	Available water capacity	%	`nfk`		10.17	12.30	23.59	31.28
Beech	pH value	log10	`ph`	30-60	2.88	3.77	7.07	8.24
Pine	pH value	log10	`ph`	30-60	2.46	3.77	4.88	8.24
Beech	Proportion of sand	%	`sand`	0-30	1.13	5.42	85.06	99.70
Pine	Proportion of sand	%	`sand`	0-30	3.07	37.38	94.34	99.70
Beech	Proportion of sand	%	`sand`	30-60	0.00	5.32	85.61	99.68
Pine	Proportion of sand	%	`sand`	30-60	2.86	38.78	95.93	99.68
Beech	Proportion of silt	%	`schluff`	0-30	0.11	11.05	65.23	81.21

Pine	Proportion of silt	%	`schluff`	0-30	0.11	3.04	44.36	73.52
Beech	Proportion of silt	%	`schluff`	30-60	0.30	9.75	59.42	77.37
Pine	Proportion of silt	%	`schluff`	30-60	0.00	2.14	39.88	70.98
Beech	Proportion of clay	%	`ton`	0-30	0.17	3.79	39.36	54.08
Pine	Proportion of clay	%	`ton`	0-30	0.15	1.47	18.36	53.18
Beech	Proportion of clay	%	`ton`	30-60	0.00	3.79	44.95	64.13
Pine	Proportion of clay	%	`ton`	30-60	0.00	0.28	22.85	54.39
Beech	Density of dried soil	g/cm^3	`trd`	0-30	0.10	0.66	1.36	1.53
Pine	Density of dried soil	g/cm^3	`trd`	0-30	0.10	1.05	1.48	1.61
Beech	Density of dried soil	g/cm^3	`trd`	30-60	0.10	0.86	1.54	1.88
Pine	Density of dried soil	g/cm^3	`trd`	30-60	0.10	1.23	1.64	2.00

3.2.4 Models

3.2.4.1 Quantifying productivity German NFI data provides limited options for the quantification of site productivity.

Site Index (SI) is a desirable measure of productivity, since it is only marginally dependent on density (and thus management). It can also be intuitively interpreted and is common among practitioners.

However, SI requires dominant height, which was not measured within the scope of the NFI.

Tree heights were measured according to the following criteria: two height measurements from the medium to upper `dbh` range in the main stand for the most frequent species group (spruce and other coniferous species, fir, douglas fir, pine, larch, oak, beech and other deciduous species). One additional height measurement for each additional species group in the main stand. One height measurement for deciduous and coniferous trees in the understory (Polley 2011).

As a substitute for dominant height, the height of Weise's mean diameter `hw` is modeled as the response variable. It is defined as the height corresponding to the quadratic mean diameter `dw` of the 20% thickest trees of the respective species in the stand (Kramer and Akça 2008). Consequently, the calculation of `hw` would have required a full sample of all tree `dbh`s in a stand.

Since this was not available, a kernel density estimation (KDE) was used to model a full `dbh` distribution for each tract and species. Only tracts with at least 80% of the sampled trees being identified as beech or pine were used. A `dbh` vector was created with each measured `dbh` being repeated according to to its expansion factor. Expansion factors were shipped with the NFI dataset for the calculation of the number of trees per hectare. The KDE was then fitted for the modeled `dbh` vector with a bandwidth of 3 cm being used for all tracts. Different bandwidths were not tested, since this should have no effect on the model fit and the modeled `dw` is merely a representation of a plausible dominant diameter of each stand. Those modeled `dbh` values above the 80% quantile were queried from the KDE of each tract (Figure 3.2).

A h-d model was used to predict `hw` for the `dw` of each tract. The core model was based on a transformed version of the Korf function (Lappi 1997) (Equation 3.1). For fixed values of λ and c, the model is linear, which makes it easy to fit. λ and c were optimized in previous analyses for both species. λ was set to 5 and 14, and c to 1.1 and 1.6 for beech and pine, respectively. h-d models were fitted as linear mixed-effects models (LME) in R version 3.6.3 (R Core Team 2020). The *gamm* function from package *mgcv* (Wood 2011; Wood 2017) was used for LME fitting, since the package was also needed for the following selection of h-a models. H-d models were fitted for both species based on the full data set of measured trees. A random effect was added in order to model the height variation between tracts (Equation 3.2).

In order to model the h-a relation, the age corresponding to `hw` (from here on referred to as `aw`) also needed to be calculated. This was achieved by creating an age vector based on the measured age of each tree and its expansion factor, following the method of the `dw` calculation. A KDE with bandwidth 3 cm was fitted over the resulting frequency distribution of ages. The mean of the modeled ages located above the 80% `dbh` quantile was calculated for each tract.

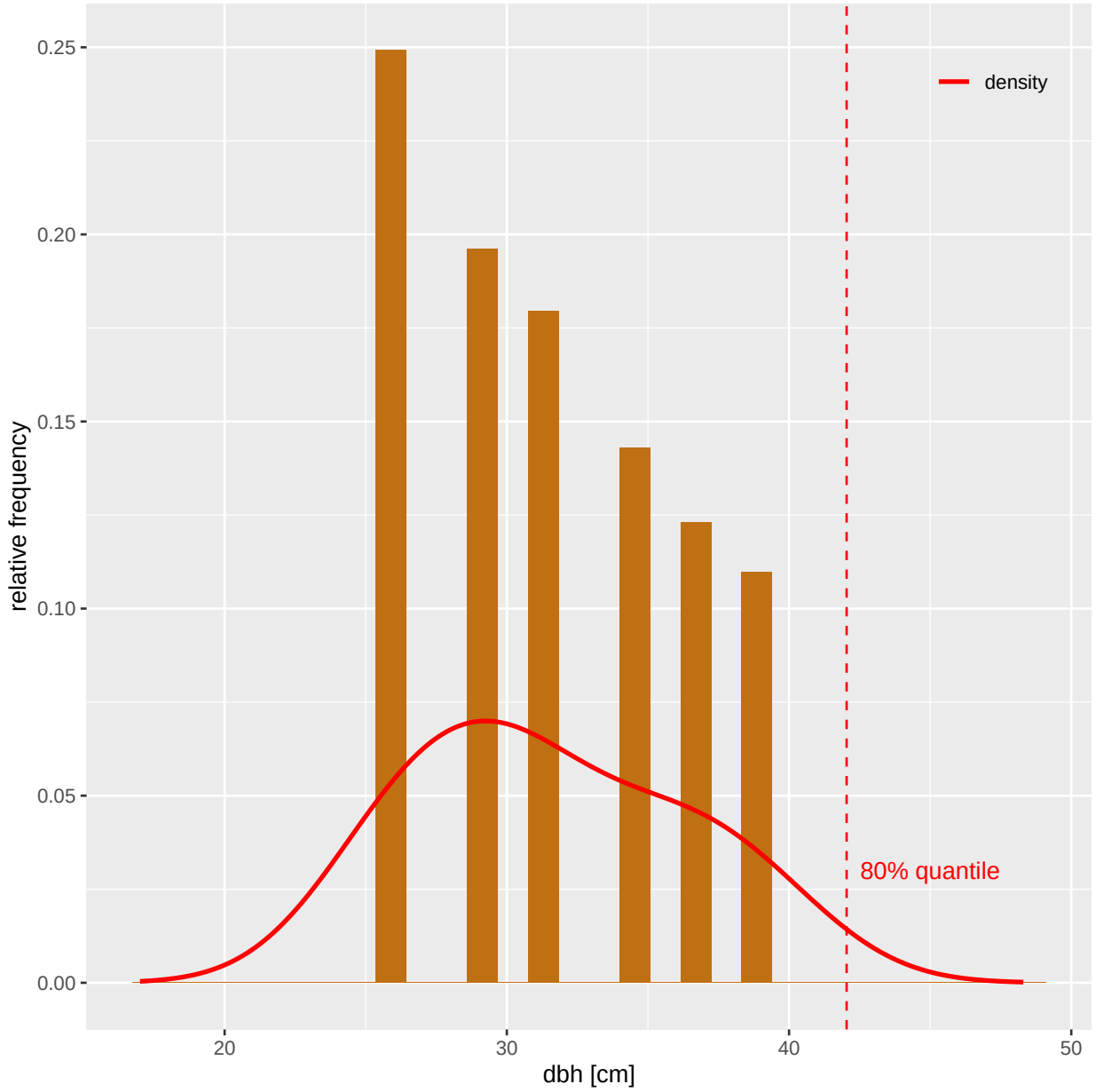


Figure 3.2: Example for a continuous `dbh` distribution, modeled using a kernel density estimation (KDE, red line). The KDE was based on the displayed histogram which was generated from multiplying `dbh`s measured at an NFI tract (beech on tract 1, plot 60598) with their expansion factors. The dashed red line denotes the 80% quantile. The mean quadratic diameter above the 80% quantile (`dw`) was used to model Weise's dominant height (`hw`).

$$\ln(h_{k,i}) = A_{k,i} - B_{k,i}x_{k,i} + \epsilon_{k,i} \tag{3.1a}$$

$$x_{k,i} = \frac{(\texttt{dbh}_{k,i} + \lambda)^{-c} - (30 + \lambda)^{-c}}{(10 + \lambda)^{-c} - (30 + \lambda)^{-c}} \tag{3.1b}$$

and:

$\ln(.)$:	natural logarithm;
$h_{k,i}$:	height of tree i at tract k;
$A_{k,i}, B_{k,i}, c, \lambda$:	first-order parameters of the h-d model at tract k;
$x_{k,i}$:	transformed `dbh` of tree i at tract k;
$\epsilon_{k,i}$:	error term for tree i at tract k with $\epsilon_{k,i} \sim \text{Normal}\,(0, \sigma^2)$.
$\texttt{dbh}_{k,i}$:	`dbh` for tree i at tract k;

$$\ln(\mu_k) = \gamma_k + \beta_0 + \beta_1 x_k \tag{3.2}$$

with:

$\ln(\mu_k)$:	natural logarithm of the conditional expectation $\text{E}(\texttt{hw}_k \vert x_k)$ of $\texttt{hw}_k \sim \text{Gamma}(\mu_k, \theta)$ with expected value μ and scale parameter θ at tract k;
$\beta_.$:	coefficients describing fixed linear effects on level (first-order parameter A) or slope (first-order parameter B) of the h-d relationship;
γ_k:	coefficient of random effect on intercept of the h-d relation tract k.

3.2.4.2 Climate data aggregation `T` grids were aggregated to 100 m × 100 m resolution in order to match the resolution of the `P` grids. Grids for retrodictions and projections were stacked for the full range of 1901 to 2100. `T` and `P` were then extracted for the coordinates of each tract. Averages were calculated for three different periods for each tract:

1. For a static climate baseline period from 1961 to 1990 (`nrm.1`).
2. For a static climate baseline period from 1991 to 2020 (`nrm.2`).
3. Dynamically over `aw` (`dyn`).

Since about 21% of the tracts were established before 1900, additional retrodictions had to be calculated for those tracts. For T, an LME was fitted based on T over year with a random effect at the tract level. The full range of retrodictions from 1900 to 2012 was used for the LME. For P, the mean of P from 1900-1930 of each tract was set for years before 1900. These simple extrapolations are not as accurate as actual retrodictions, but they can be assumed to correctly distinguish the tracts established before 1900 from younger tracts with respect to their colder climate. The retrodictions for T were also roughly in accordance with the famous *hockeystick curve* (Mann et al. 1999).

$$\mathrm{E}(\mathtt{T}_k|\mathtt{a}) = \gamma_k + \beta_1 + \beta_2\mathtt{a} \tag{3.3}$$

with:

$\mathrm{E}(\mathtt{T}_k\|\mathtt{a})$:	conditional expectation of temperature sum $\mathtt{T} \sim \mathrm{Normal}\,(0, \sigma^2)$at tract k;
$\beta_.$:	coefficients describing fixed linear effects on intercept and slope of the T-a relation;
a:	the year;
γ_k:	coefficient of random effect on intercept of the T-a line on tract k.

3.2.4.3 Model selection The height-age model was developed from the linearized version of the Korf-function, which has already been used for the hw prediction (Equation 3.1). Instead of dbh, a transformed version of aw is used as predictor (Equation 3.4, cmp. Equation 3.1b, Equation 3.4b).

The constants used in the equation for the diameter transformation were changed accordingly. The constant of 10 cm was changed to 50 years and 30 cm was changed to 100 years. This led to z_k taking a value of zero at an aw of 100 years. Also, the parameter B now represented the difference in hw between a stand at aw 50 and 100.

$$\ln(\mathtt{hw}_k) = A_k - Bz_k + \epsilon_k \tag{3.4a}$$

$$z_k = \frac{(\mathtt{aw}_k + \lambda)^{-c} - (100 + \lambda)^{-c}}{(50 + \lambda)^{-c} - (100 + \lambda)^{-c}} \tag{3.4b}$$

While climate data was available for the full 2 km × 2 km to 4 × 4 km grid of the NFI data, soil data was only available for a subset of the tracts.

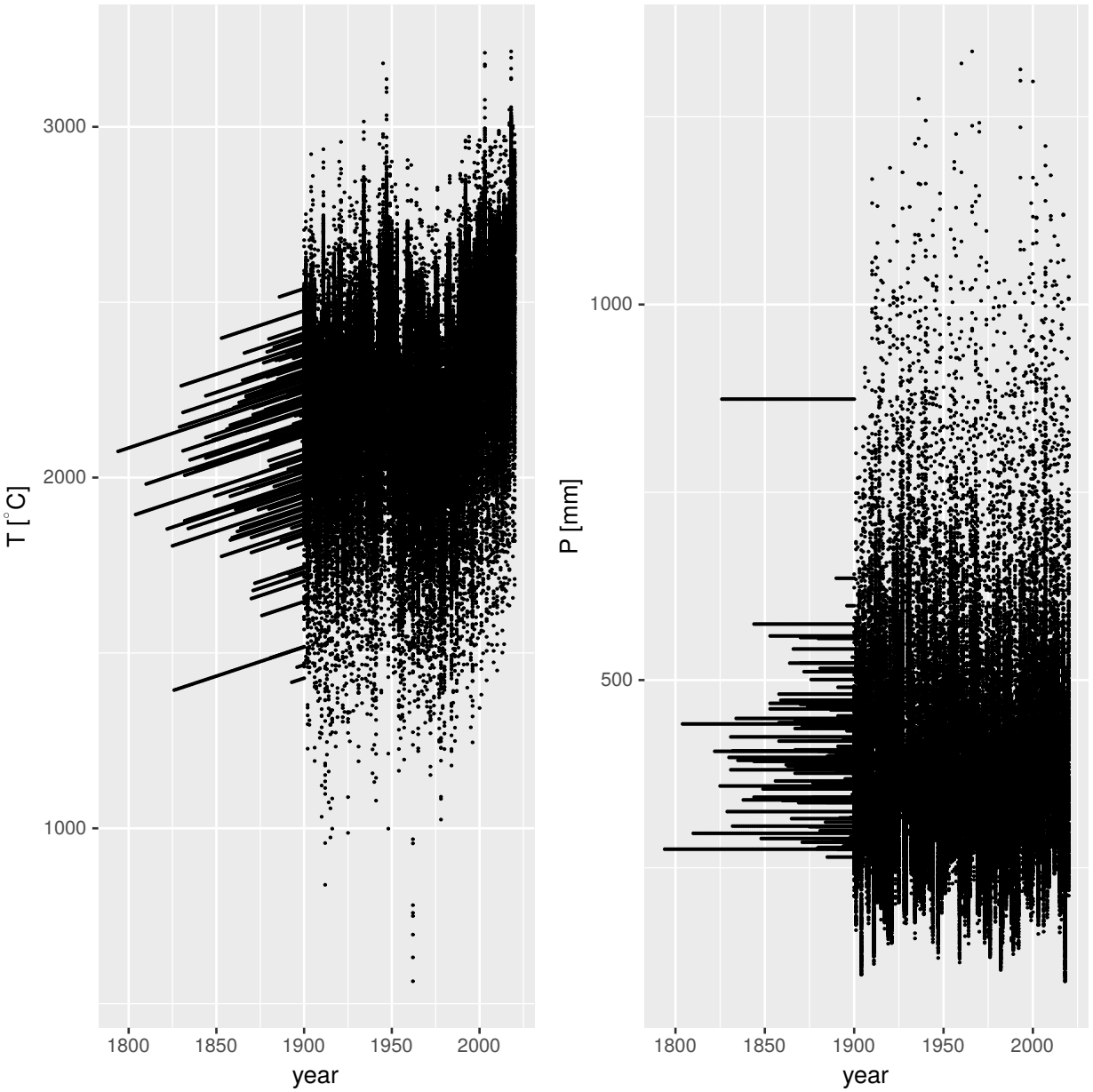


Figure 3.3: Retrodictions for sums of temperature `T` and precipitation `P` during vegetation period for NFI tracts. Retrodictions for tracts established before 1900 were extrapolated using different approaches for `T` and `P`. `T` was extrapolated using an LME for `T` over year with a random effect on tract level. For `P`, the mean of 1900-1930 for each tract was set for years before 1900.

To be able to use both climate and soil data, a two step model framework was used. Climate sensitive models were selected based on the full set of inventory data.

In a second step, the dataset was reduced to the tracts for which soil data was available. Soil covariates were added for the reduced dataset.

The detailed process was as follows:

- Step I
  - **I.1** Climate sensitive models were fitted as Generalized Additive Models (GAMs) (Equation 3.5) using package *mvcv* (Wood 2011; Wood 2017). The models contained transformed age z_k and nonlinear effects for `T` and `P` described by a penalized basis-function approach using a thin-plate spline basis (from here on referred to as *spline*). Models were fitted for different combinations of the first-order parameters λ and c. First, a grid of all possible combinations within a specific range for each parameter was initialized. Then a model was calculated for each grid line. The optimal parameter combination was identified by selecting the model with the lowest Akaike Information Criterion (AIC, Akaike 1974). If the parameter combination yielding the lowest AIC was located at the minimum or maximum of one parameter in the grid, the calculation was repeated with a larger grid until decreases in AIC were only observed in decimal places.
  - **I.2** One final model was fitted for each species and aggregation period (`nrm.1`, `nrm.2`, `dyn`). The same values for λ and c were used for all three periods since previous analyses showed no differences in AIC. If necessary, models were refitted as Shape Constrained Additive Models (SCAMs). SCAMs provide the option to add constraints to splines in order to avoid wiggliness of effects at the edges of data ranges. They were fitted using `R` package `scam` (version 1.2-5, Pya 2019).
- Step II
  - **II.1** In this step, the dataset was reduced to that subset of the NFI grid for which soil data was available. This reduced the number of measured trees from 92,115 trees (22,801 plots) to 42,718 trees (10,285 plots) for beech. The number was reduced from 92,726 trees (17,177 plots) to 47,162 trees (8,479 plots) for pine. Water logged plots were excluded in order to avoid interactions with precipitation or soil parameters, which would lead to implausible covariate effects. Estimates of the first-order parameters A and B were calculated from the terms of the final models from step **I.2**. A basic model was fitted containing the estimates of first-order parameters $\hat{A}$ and $\hat{B}$ as well as the transformed `aw` (z_{ki}) as predictors (Equation 3.6a). As in step **I.1**, first-order parameters λ and c were optimized

using an iterative grid search. Again, optimization was only conducted for period `nrm.1` for each species, since there was no difference in results between periods.

- **II.2** Then, one model was calculated for each available soil parameter. Each soil parameter was added to each of the final models from step **II.1** as an unconstrained spline. The soil covariates were then evaluated based on coefficient of determination (R^2), AIC, relative amplitude (Equation 3.7) and visual evaluation of plausibility. The relative amplitude quantifies the contribution of a covariate to the complete range of the prediction when increasing the covariate between the 5% and 95% quantile. Amplitudes of step **I** were multiplied with the sum of the amplitudes for $\hat{A}$ and $\hat{B}$ of step **II**, in order to be able to compare amplitudes of both steps.
- **II.3** The final model for each species and period was selected by finding the soil parameter which is best suited as a covariate based on the results of step **II.2**. The same covariate was selected for each aggregation period. Parameters which were usable as covariate for both species were given preference to enable inter-species comparisons. Constraints were added if necessary.

$$\ln(\mu_{z,k}) = \beta_{I,1} + \beta_{I,2} z_k + f_1(\mathtt{T}_k) + f_2(\mathtt{P}_k) \tag{3.5}$$

with:

$\ln(\mu_{z,k})$: natural logarithm of the conditional expectation $\mathrm{E}(\mathtt{hw}_k | z_k, \mathtt{T}_k, \mathtt{P}_k)$ of $z_k \sim \mathrm{Gamma}(\mu_{z,k}, \theta)$ with expected value μ_z and scale parameter θ at tract k;

$\beta_I, .$: coefficients describing linear effects on level (first-order parameter A) or slope (first-order parameter B) of the h-a relationship in model step I;

$f_1(\mathtt{T}_k)$: one-dimensional spline, describing a non-linear effect of `T` on the level of the h-a relationship (first-order parameter A);

$f_2(\mathtt{P}_k)$: : one-dimensional spline, describing a non-linear effect of `P` on the level of the h-a relationship (first-order parameter A);

$$\hat{A}_k = \beta_{I,1} + f_1(\mathtt{T}_k) + f_2(\mathtt{P}_k) \tag{3.6a}$$

$$\hat{B}_k = \beta_{I,2} z_k \tag{3.6b}$$

$$\ln(\mu_{z,k}) = \beta_{II,1} \hat{A}_k + \beta_{II,2} \hat{B}_k z_k \tag{3.6c}$$

$\hat{A}_k, \hat{B}_k$: estimates of first-order parameters A_k and B_k of the h-a relationship (equation 3.4);

$\beta_{II,1}, \beta_{II,2}$: coefficients describing linear fixed effects on the estimates of first-order parameters A and B in model step **II.1**.

$$\text{amp}(f_j(x_j)) = \frac{\max(f_j(x_j)) - \min(f_j(x_j))}{\sum_{\tilde{j}=1}^{n} \max(f_{\tilde{j}}(x_{\tilde{j}})) - \min(f_{\tilde{j}}(x_{\tilde{j}}))} \tag{3.7}$$

with:

$\text{amp}(x_j)$: relative amplitude for covariate j;

$x_j ... x_n$: terms for covariates $j...n$, meaning $\beta_j x_j$ for linear effects or $f_j(x_j)$ for splines.

3.2.5 Projection

In order to analyze the effect of the three aggregation periods, `hw` was projected for all three aggregation periods for the year 2050.

To get correct projections for each aggregation period, the period for which the climate data for the model fit was aggregated had to be moved 38 years into the future (as moving the year of the NFI, 2012, to 2050).

Climate data for the projections was then aggregated for the respective windows. For projections for the `dyn` model, two different aggregations were created. One using the periods 1912 to 2012 and 1950 to 2050 to predict 100 year old stands, and one using 1982 to 2012 and 2020 to 2050 to predict 30 year old stands.

This resulted in the final configuration of aggregation scenarios as follows (see also Figure 3.4):

- `nrm.1`: 1961 to 1990 (for 2012) and 1999 to 2028 (for 2050)
- `nrm.2`: 1991 to 2020 (for 2012) and 2029 to 2058 (for 2050)
- `dyn.1`: 1912 to 2012 (for 2012) and 1950 to 2050 (for 2050)
- `dyn.2`: 1982 to 2012 (for 2012) and 2020 to 2050 (for 2050)

Projections were created for all of Germany on a 100 m × 100 m grid. `hw` differences from 2012 to 2050 were then calculated for each scenario. Additionally, areas which entered or left the climate range covered by the models from 2012 to 2050 were identified as *new* or *lost* respectively.

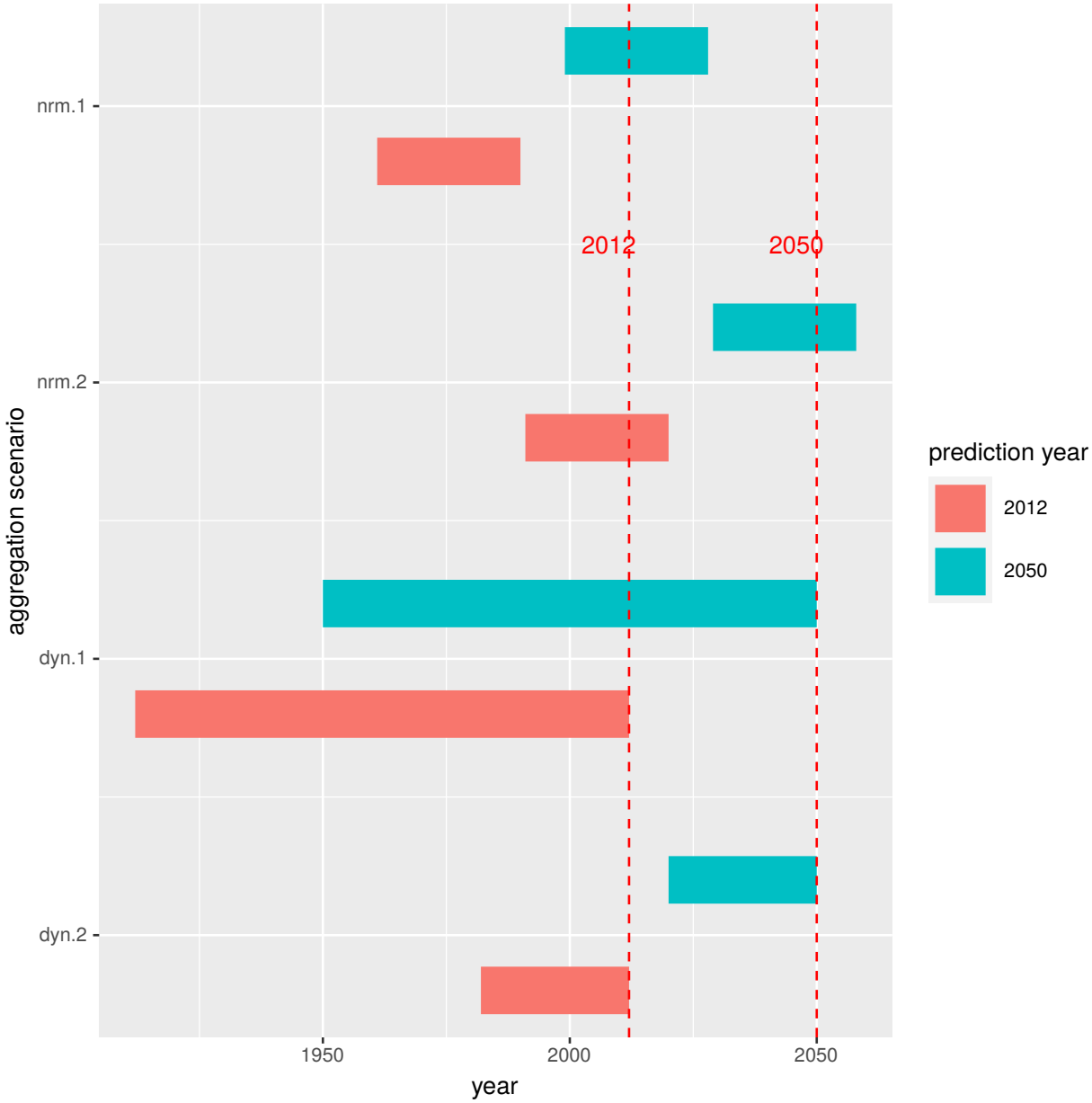


Figure 3.4: Time periods of the different projection scenarios `nrm.1`, `nrm.2`, `dyn.1` and `dyn.2` for 2012 and 2050 respectively (vertical dashed lines denote the target year)

3.3 Results

3.3.1 Quantifying productivity

The `dw` calculated from KDEs for each tract displayed a range of 14.24 to 138.42 cm for beech and 14.24 to 95.19 cm for pine.

`hw` displayed a distribution of 4.87 m to 47.96 m for beech and 3.02 m to 51.72 for pine, while `aw` ranged from 17 to 429 years for beech and 15 to 218 years for pine (Table 3.4). SIs were modeled based on the h-a pairs for each tract as additional information on site-productivity (Figure 3.5). R package `et.nwfva` was used to calculate SIs. Resulting SIs ranged from 8.5 to 48.5 m for beech and 9.20 to 45 m for pine.

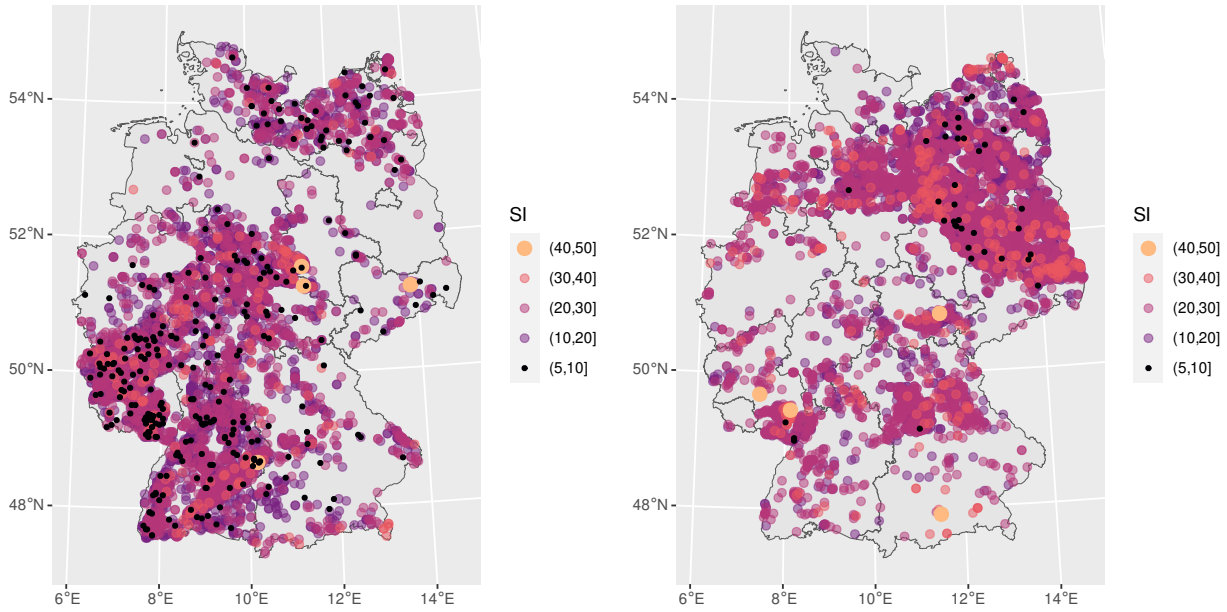


Figure 3.5: Maps of SIs modeled for beech and pine.

Table 3.5: Statistics on the mixed h-d model (eq. 3.2) used to predict `hw` from `dw`, including the scale parameter for the Gamma distribution of the response θ, estimated residual standard deviations (SD) for the fixed and random effects and the values for first-order parameters λ and c

Species	R^2	θ	SD fixed effects [m]	SD random effects [m]	λ	c
Beech	0.58	0.0026	4.80	1.15	5	14.0
Pine	0.58	0.0025	3.49	0.96	14	1.6

Table 3.4: Statistics on dominant diameter `dw`, dominant height `hw`, mean age of the dominant trees `aw` and SI calculated for beech and pine.

Species	Statistic	`dw`	`hw`	`aw`	SI
Beech	Minimum	14.24	4.87	17.20	8.50
Beech	5% quantile	20.39	14.00	32.20	21.40
Beech	95% quantile	86.69	39.33	181.20	40.20
Beech	Maximum	138.42	47.95	429.49	48.50
Pine	Minimum	14.24	3.02	15.20	9.20
Pine	5% quantile	23.78	15.04	32.20	20.80
Pine	95% quantile	60.75	32.04	137.20	36.20
Pine	Maximum	95.19	51.72	218.22	45.00

Table 3.6: Statistics on coefficients of the mixed h-d model used to predict `hw` from `dw` (eq. 3.2).

Species	Coefficient	Estimate	Standard error	P value
Beech	β_0	3.22	0.0031	$< 10^{-15}$
Beech	β_1	0.45	0.0016	$< 10^{-15}$
Pine	β_0	3.08	0.0021	$< 10^{-15}$
Pine	β_1	0.48	0.0012	$< 10^{-15}$

The h-d model used to model `hw` from `dw` displayed an R^2 of 58% for beech and pine. Residual standard deviations for the fixed effects were 4.8 m for beech and 3.49 m for pine (Table 3.5). Adding random effects reduced the standard deviation of residuals to 1.15 m for beech and 0.96 for pine. Parameter λ was set to 5 for beech and 14 for pine and c was set to 1.1 for beech and 1.6 for pine. All parameter coefficients yielded p values of $p < 10^{-15}$ (Table 3.6).

3.3.2 Climate aggregation

Table 3.7: Ranges of T and P values aggregated for the three aggregation periods used for fitting models in step **I.1**

Parameter	Aggregation period	Min	5% quantile	95% quantile	Max
T	nrm.1	1368.20	1917.90	2463.40	2694.63
T	nrm.2	1585.37	2161.86	2691.07	2928.40
T	dyn	1302.98	1996.70	2549.31	2913.52
P	nrm.1	239.50	255.63	465.17	971.03
P	nrm.2	279.60	293.83	485.04	1045.10
P	dyn	261.26	277.49	479.42	1007.68

Climate aggregations for the periods nrm.1, nrm.2 and dyn resulted in T ranging from 1368 to 2681 °C for nrm.1, 1585 to 2918 °C for nrm.2 and 1302 to 2774 °C for dyn. T values for the dynamic period dyn are located between the static aggregations nrm.1, and nrm.2 (Table 3.7, Figure 3.6). P ranged from 247 to 971 mm for nrm.1,from
280 to 1045 mm for nrm.2 and from 272 to 1008 mm for dyn. Scenario nrm.1 showed more plots below 300 mm compared to nrm.2 and dyn.

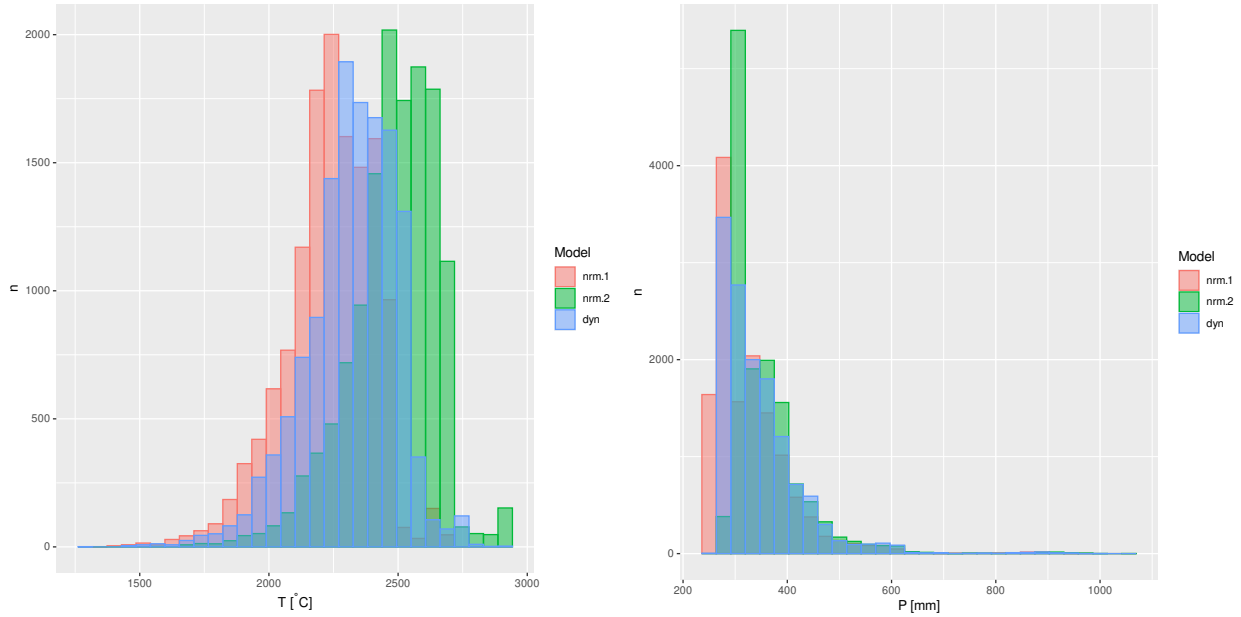


Figure 3.6: Histograms of temperature and precipitation sums for the vegetation period (`T` and `P`) for tracts of the third German NFI, averaged over the three different aggregation periods.

3.3.3 Model selection

Grid optimization in step **I.1** yielded λ of 293 for beech and 36 for pine and c of 10 for beech and 36 for pine.

Table 3.8: Statistics on the final models of step **I.2** based on the three aggregation periods `nrm.1`, `nrm.2` and `dyn` for beech and pine. Including estimated standard deviation of residuals (SD), Akaike Information Criterion (AIC) and the scale parameter of the Gamma distribution θ

Species	Model	R^2	SD [m]	AIC	θ
Beech	`nrm`	0.65	4.43	30467	0.027
Beech	`nrm`	0.65	4.43	30472	0.027
Beech	`dyn`	0.65	4.43	30464	0.027
Pine	`nrm`	0.50	3.63	45173	0.025
Pine	`nrm`	0.50	3.63	45180	0.025
Pine	`dyn`	0.50	3.64	45207	0.025

The final models in step **I.2** yielded R^2 between 64.6% and 64.7% for the three aggregation periods `nrm.1`, `nrm.2` and `dyn` for beech. R^2 ranged from 49.7% to 49.9 for pine. Estimated residual standard deviation stayed at about 4.43 m for beech and at 3.6 m for pine for the three aggregation periods. The scale θ of the Gamma distribution assumed for the response stayed at ca. 0.027 for beech and 0.025 for pine. `T` displayed distinct parabolic shapes for beech and pine. Concavity constraints were added for `T` for both species to avoid wiggliness of the effects. The decrease in the effect on `hw` above and below the maximum was within the inter-quantile range of the data for beech, while the data was located mostly above the maximum for pine (Figure 3.7, Figure 3.8).

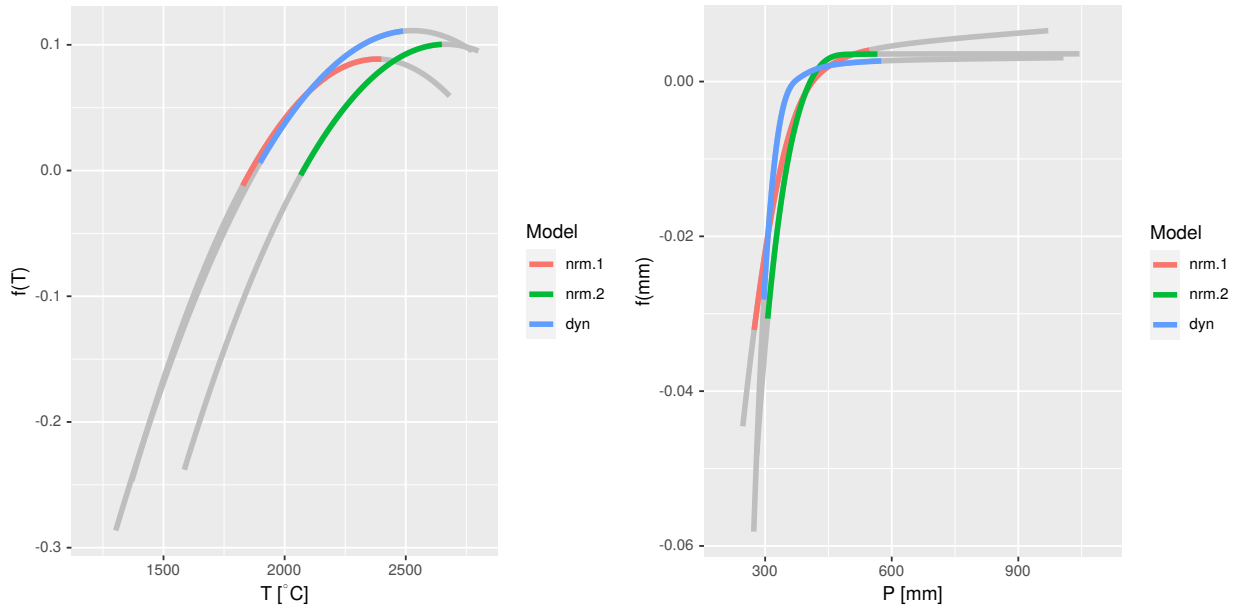


Figure 3.7: Splines for `T` (left) and `P` (right) for the final beech models of step **I.2**, for all three aggregation periods

For the `P` effect in the beech model, a monotony constraint was added to avoid a decrease in the effect at high values which is typical for plots located in mountainous areas. Problems arose with the `P` effect for pine, which was significant, but collapsed when the necessary monotony constraint was added. Hence, the effect was included as a segmented linear effect, with a positive linear coefficient below a specific `P` threshold and a coefficient of zero above that threshold. Thresholds were set at 245 mm for `nrm.1`, 300 mm for `nrm.2` and 280 mm for `dyn`.

Grid optimization, after filtering of plots for soil parameters (step **II.1**), led to a change in λ from 293 to 296 for beech and 36 to 80 for pine, while c changed from 10 to 9.5 for beech and 3.7 to 5.6 for pine.

Reducing the data and adding a soil parameter as covariate to model `nrm.1` in step **II.2**

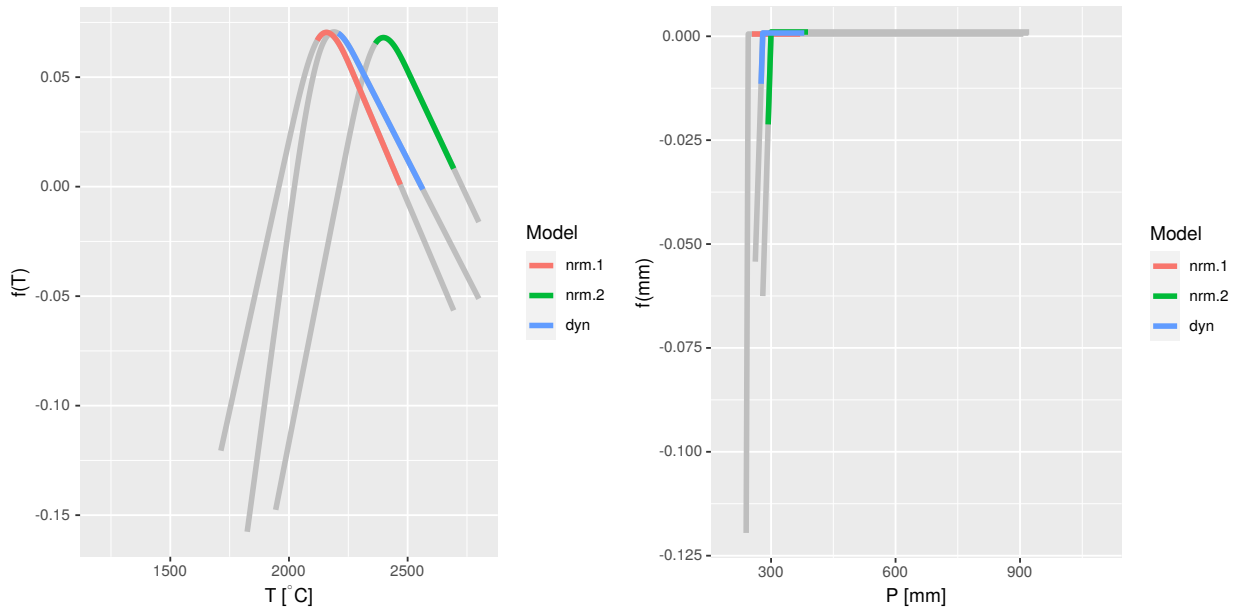


Figure 3.8: Splines for `T` (left) and `P` (right) for the final pine models of step **I.2**, for all three aggregation periods.

reduced R^2 by about 4% for beech and 9% for pine. It increased the estimated residual standard deviation slightly by 2-3 cm for beech and reduced it by 1 cm for pine.

In the final step **II.3**, the pH value for the 30-60 cm aggregation depth (`ph`) was identified as a soil covariate for both beech and pine (Equation 3.8, tbl. 3.9). See Appendix A.4 for splines for all soil parameters. The `ph` effect showed a pronounced parabolic shape for beech, with the inter-quantile range of data being located around the maximum of the effect (Figure 3.9). A parabolic shape could be identified for pine as well, with the inter-quantile range being clustered within a small data range. A concavity constraint was added on the `ph` effect for both species.

$$\ln(\mu_{k,i}) = \beta_{II,1}\hat{A}_k + x_{k,i}\beta_{II,2}\hat{B}_k + f(\texttt{ph}_k) \tag{3.8}$$

with:

$f(\texttt{ph})$ spline for the `ph` of plot k, aggregated for 30-60 cm soil depth.

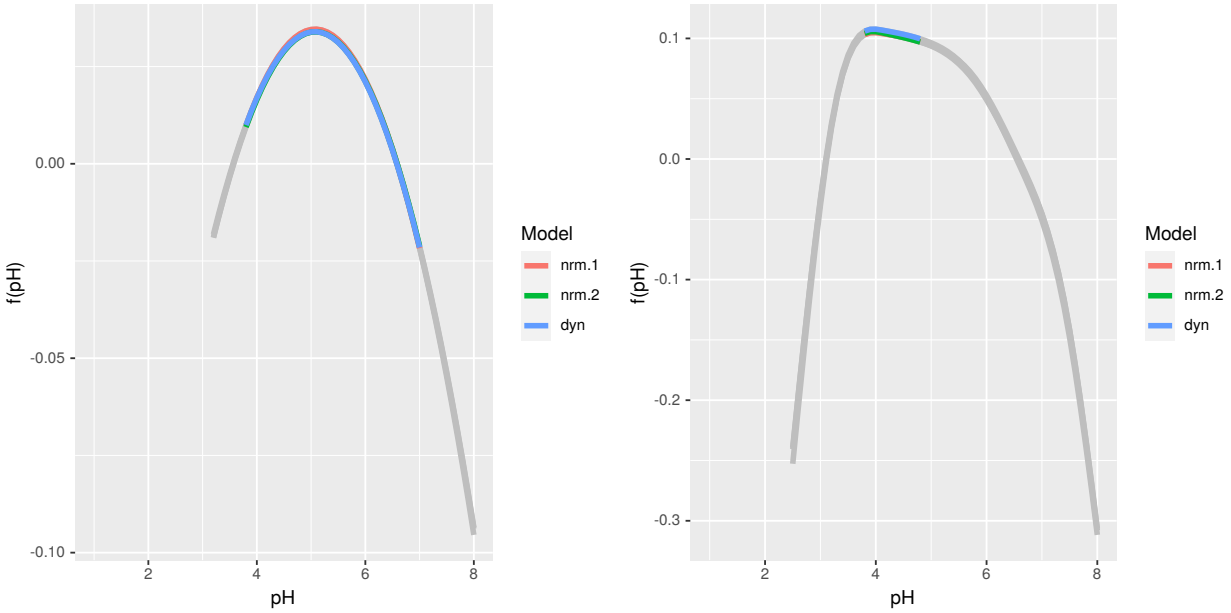


Figure 3.9: Spline for `ph` for beech (left) and pine (right) for all aggregation periods

Table 3.9: Statistics on final models of step I.2 and II.3. For the terms, which include each coefficient with its respective covariate or spline, the estimated degrees of freedom (EDF) and relative amplitudes (see equation 3.7) are given. Coefficients for both steps contain subscripts for model steps I and II respectively.

Species	Model	Model step	Coefficient	Estimate	Standard error	P-value	Term	EDF	Amplitude
Beech	nrm.1	I	$\beta_{I,0}$	2.89	0.09	0.00	$\beta_{I,0}$		
Beech	nrm.1	I				0.00	$cv(\mathtt{T})$	2.02	0.10
Beech	nrm.1	I				0.00	$m(\mathtt{P})$	1.77	0.04
Beech	nrm.1	II	$\beta_{II,1}$	0.98	0.01	0.00	$\beta_{II,1}\hat{A}$		0.13
Beech	nrm.1	II	$\beta_{II,2}$	1.01	0.01	0.00	$\beta_{II,2}\hat{B}$		0.81
Beech	nrm.1	II				0.00	$f(\mathtt{ph})$	2.01	0.07
Beech	nrm.1	I	$\beta_{I,1}$	0.37	0.00	0.00	$\beta_{I,1}z_{ki}$		0.80
Beech	nrm.2	I	$\beta_{I,0}$	2.83	0.16	0.00	$\beta_{I,0}$		
Beech	nrm.2	I				0.00	$cv(\mathtt{T})$	2.02	0.10

Beech	nrm.2	I				0.00	$m(\texttt{P})$	2.14	0.03
Beech	nrm.2	II	$\beta_{II,1}$	0.98	0.01	0.00	$\beta_{II,1}\hat{A}$		0.13
Beech	nrm.2	II	$\beta_{II,2}$	1.01	0.01	0.00	$\beta_{II,2}\hat{B}$		0.81
Beech	nrm.2	II				0.00	$f(\texttt{ph})$	2.00	0.07
Beech	nrm.2	I	$\beta_{I,1}$	0.37	0.00	0.00	$\beta_{I,1}z_{ki}$		0.80
Beech	dyn	I	$\beta_{I,0}$	2.59	0.13	0.00	$\beta_{I,0}$		
Beech	dyn	I				0.00	$cv(\texttt{T})$	2.01	0.10
Beech	dyn	I				0.00	$m(\texttt{P})$	1.95	0.03
Beech	dyn	II	$\beta_{II,1}$	0.98	0.01	0.00	$\beta_{II,1}\hat{A}$		0.13
Beech	dyn	II	$\beta_{II,2}$	1.01	0.01	0.00	$\beta_{II,2}\hat{B}$		0.81
Beech	dyn	II				0.00	$f(\texttt{ph})$	2.00	0.06
Beech	dyn	I	$\beta_{I,1}$	0.37	0.00	0.00	$\beta_{I,1}z_{ki}$		0.80
Pine	nrm.1	I	$\beta_{I,0}$	3.06	0.05	0.00	$\beta_{I,0}$		
Pine	nrm.1	I	$\beta_{I,2}$	0.02	0.01	0.03	$\beta_{I,3}\texttt{P}$		0.00
Pine	nrm.1	I				0.00	$cv(\texttt{T})$	2.05	0.12
Pine	nrm.1	II	$\beta_{II,1}$	0.78	0.03	0.00	$\beta_{II,1}\hat{A}$		0.12
Pine	nrm.1	II	$\beta_{II,2}$	0.95	0.02	0.00	$\beta_{II,2}\hat{B}$		0.86
Pine	nrm.1	II				0.00	$f(\texttt{ph})$	5.00	0.02
Pine	nrm.1	I	$\beta_{I,1}$	0.18	0.00	0.00	$\beta_{I,1}z_{ki}$		0.86
Pine	nrm.2	I	$\beta_{I,0}$	3.02	0.04	0.00	$\beta_{I,0}$		
Pine	nrm.2	I	$\beta_{I,2}$	0.00	0.00	0.00	$\beta_{I,3}\texttt{P}$		0.04
Pine	nrm.2	I				0.00	$cv(\texttt{T})$	2.01	0.10
Pine	nrm.2	II	$\beta_{II,1}$	0.77	0.03	0.00	$\beta_{II,1}\hat{A}$		0.13
Pine	nrm.2	II	$\beta_{II,2}$	0.95	0.02	0.00	$\beta_{II,2}\hat{B}$		0.85
Pine	nrm.2	II				0.00	$f(\texttt{ph})$	4.96	0.02
Pine	nrm.2	I	$\beta_{I,1}$	0.18	0.00	0.00	$\beta_{I,1}z_{ki}$		0.84

Pine	dyn	I	$\beta_{I,0}$	2.97	0.05	0.00	$\beta_{I,0}$		
Pine	dyn	I	$\beta_{I,2}$	0.00	0.00	0.00	$\beta_{I,3}$P		0.02
Pine	dyn	I				0.00	cv(T)	1.97	0.12
Pine	dyn	II	$\beta_{II,1}$	0.77	0.03	0.00	$\beta_{II,1}\hat{A}$		0.12
Pine	dyn	II	$\beta_{II,2}$	0.95	0.02	0.00	$\beta_{II,2}\hat{B}$		0.85
Pine	dyn	II				0.00	f(ph)	4.99	0.02
Pine	dyn	I	$\beta_{I,1}$	0.18	0.00	0.00	$\beta_{I,1}z_{ki}$		0.83

Species	Model	Model step	Covariate	Minimum	5% quantile	95% quantile	Maximum
Beech	nrm.1	I	T	1368.20	1826.07	2399.37	2681.47
Beech	nrm.1	I	P	246.50	273.90	547.92	971.03
Beech	nrm.2	I	T	1585.37	2065.93	2649.71	2918.83
Beech	nrm.2	I	P	279.63	305.34	566.75	1045.10
Beech	dyn	I	T	1302.98	1897.25	2489.53	2774.78
Beech	dyn	I	P	272.51	296.00	575.36	1007.68
Beech		II	ph	3.13	3.77	7.08	8.24
Beech		I	aw	17.20	32.20	181.20	429.49
Beech		I	hw	4.87	13.98	39.31	47.76
Pine	nrm.1	I	P	239.50	253.10	370.33	908.00
Pine	nrm.1	I	T	1712.97	2118.80	2469.47	2694.63
Pine	nrm.2	I	P	279.60	292.07	389.29	922.00
Pine	nrm.2	I	T	1943.40	2363.69	2694.55	2928.40
Pine	dyn	I	P	261.26	275.40	380.28	920.46
Pine	dyn	I	T	1823.73	2206.08	2564.51	2913.52
Pine		II	ph	2.46	3.78	4.89	8.24

Pine	I	aw	15.20	32.20	137.01	218.22
Pine	I	hw	3.02	15.00	32.03	51.72

z_{ki} showed the highest amplitude for both species, model steps and aggregation periods. z_{ki} contributed 80% to the relative amplitude for the beech models and 83%-85% for pine. Effects of T showed the second highest amplitudes with 10% for beech and 9%-11% for pine. The effects of P displayed low amplitudes from 3% to 4% for beech. The values for pine ranged from 0% to 4%. The 0% value for nrm.1 was observed due to the whole amplitude being located outside of the interquantile range of the P data. The ph effect added in model step **II** showed higher amplitudes compared to P for beech with 6-7% and partly for pine with 4% for all models.

3.3.4 Projection

The 100 m × 100 m pixels of the Germany-wide predictions showed an increase of T from 2012 to 2050 for all four projection scenarios (Figure 3.10, see also Figure 3.13, Figure 3.14). The increase was lowest for dyn.1, with increases up to 233 °C (sum for vegetation period), followed by nrm.2, dyn.2 and nrm.1. To better understand the differences, the following example might be considered: the difference in T between the aggregation for years 1961 to 1990 and 1999 to 2028 (nrm.1) is higher than the difference in T between aggregation windows for years 1991 to 2020 and 2029 to 2058 (nrm.2). This is due to a steeper T increase between the time periods of nrm.1 (see also Figure 3.4).

P mostly decreased from 2012 to 2050 for aggregation scenarios nrm.2, dyn.1 and dyn.2 and mostly increased for scenario nrm.1. The decrease was highest for nrm.2, lowest for dyn.1 and located between those scenarios for dyn.2.

Table 3.11: Measures *both, lost, new* and *null* denote the proportions of 100×100 m cells which were located inside or outside of the `T` and `P` range of the model fit for the four projection scenarios. *both* denotes cells which were located in the model data range in both prediction years, 2012 and 2050. *lost* cells were located within the model data range for 2012 but were located in the extrapolation range in the prediction for 2050. *new* cells were located in the extrapolation range for 2012, but within the range of the model data in 2050. *null* cells were located in the extrapolation range for both years. *freq.hw* rows contain the proportions of positive and negative `hw` differences from 2012 to 2050 for the respective projection scenario. Following rows contain the minima, maxima and 5%/95% quantiles of the changes in `hw`, calculated for positive and negative changes separately.

Species	Measure	`nrm.1`	`nrm.2`	`dyn.1`	`dyn.2`
Beech	both [%]	80.81	69.32	89.06	62.67
Beech	lost [%]	18.15	29.83	8.19	35.25
Beech	new [%]	0.54	0.20	0.08	0.09
Beech	null [%]	0.50	0.65	2.67	1.98
Beech	freq. `hw` < 0 [%]	48.67	62.03	37.20	68.79
Beech	freq. `hw` > 0 [%]	51.33	37.97	62.80	31.21
Beech	min `hw` < 0 [m]	-9.18	-11.42	-7.59	-15.09
Beech	qu.05 `hw` < 0 [m]	-1.46	-1.41	-0.60	-2.19
Beech	qu.95 `hw` < 0 [m]	-0.05	-0.07	-0.03	-0.09
Beech	max `hw` < 0 [m]	-0.00	-0.00	-0.00	-0.00
Beech	min `hw` > 0 [m]	0.00	0.00	0.00	0.00
Beech	qu.05 `hw` > 0 [m]	0.05	0.04	0.04	0.04
Beech	qu.95 `hw` > 0 [m]	2.40	2.21	0.91	1.96
Beech	max `hw` > 0 [m]	5.06	5.98	2.90	5.07
Pine	both [%]	82.14	69.00	96.66	79.33
Pine	lost [%]	16.50	29.28	1.37	19.55
Pine	new [%]	0.83	1.08	0.81	0.81
Pine	null [%]	0.54	0.64	1.16	0.31

Pine	freq. hw < 0 [%]	88.66	87.87	87.50	94.71
Pine	freq. hw > 0 [%]	11.34	12.13	12.50	5.29
Pine	min hw < 0 [m]	-5.66	-8.13	-9.88	-16.96
Pine	qu.05 hw < 0 [m]	-2.12	-4.58	-1.39	-3.75
Pine	qu.95 hw < 0 [m]	-0.83	-0.72	-0.40	-1.18
Pine	max hw < 0 [m]	-0.00	-0.00	-0.00	-0.00
Pine	min hw > 0 [m]	0.00	0.00	0.00	0.00
Pine	qu.05 hw > 0 [m]	0.14	0.13	0.09	0.15
Pine	qu.95 hw > 0 [m]	3.64	4.30	2.38	5.66
Pine	max hw > 0 [m]	5.46	4.72	2.46	6.05

In order to identify causes for the increases and decreases in hw, the splines of T and P were also plotted above histograms of the respective effects (Figure 3.11, Figure 3.12).

nrm.1 for beech displayed balanced increases as well as decreases in hw, with an almost 50-50% distribution between both (Table 3.11). About 18% of the cells were located in the extrapolation range of the model in 2050 (from here on called *lost* cells). The distribution of decreases and increases was right skewed, displaying higher variance of increases while decreases stayed closer to zero. Histograms showed that hw decreases were caused by the increase in T, with the major proportion shifting to the section of the spline above the optimum. Increases in hw were caused by (locally very high) increases in P, moving the distribution of P in the increasing direction of the spline. This made nrm.1 the most optimistic scenario for beech.

nrm.2 displayed a similar pattern in hw differences, with a higher proportion of decreasing hw (62.03%). Also, a much higher percentage (29.83%) were lost compared to nrm.1 . Decreases and increases in hw were otherwise distributed similar to nrm.1. The T development looked similar, but contributed less to the decrease since the T spline showed less decrease after the optimum. However, more of the cells experienced a decrease in P compared to nrm.1, moving P more into the decreasing direction of the spline and into the extrapolation range. Increases were still observed due to a certain amount of cells with increasing P (Figure 3.10).

dyn.1 for beech displayed a very balanced amount of decrease and increase in hw, with little variation around 0. It was also the only projection scenario with a distinctly higher proportion of increasing compared to decreasing hw. Moreover, it was the projection scenario

with the lowest proportion of lost cells. Compared to the other scenarios, hw differences were distributed closely around zero. Those patterns were in accordance with relatively small changes in T and P from 2012 to 2050. Histograms showed that T shifted to slightly higher values, which stayed mostly within in the data range of the beech model, while P changed very little. The histograms of climate differences for all scenarios (Figure 3.10) showed that P mostly decreased, but stayed close to zero. Due to a certain amount of cells with increasing P, and the increase in T being located in the increasing area of the T effect before the optimum, a high amount of hw increases was observed.

dyn.2 displayed the highest number of decreases in hw for beech and the highest proportion of lost cells. The decreases were also more extreme than the increases. This was caused by high increases in T and decreases in P. As in the other projection scenarios, small proportions of increases of T below the optimum together with local P increases still led to a small amount of cells with increasing hw.

For pine, the proportion of decreases in hw was far larger than for beech, with 88% to 95% of cells showing hw decreases from 2012 to 2050. At the same time, there were less lost cells compared to beech, except for nrm.2. This could have been due to the locally larger data ranges for T and P in the model fit for pine. Changes were also more extreme in both directions compared to beech. This was likely caused by steeper T effects, with stronger decreases before and after the optimum.

nrm.1 for pine displayed increases on few cells. They were still possible due to the same effect observed in beech, with a small amount of cells experiencing an increase in T below the optimum of the T effect. The steeper effect, though, led to higher values compared to beech. In contrast to beech, pine could not profit as much from the pronounced increase in P, due to the increase occurring mostly in the parallel to the x axis of the segmented P effect. Decreases were also mostly caused by the T increase. A higher amount of decreases compared to beech was also rooted in the lower optimum of the effect compared to beech. Lost cells were also caused by T values above the range of the model fit. The upper limit of T values observed for the pine model was marginally higher than that of the beech model, leading to the slightly lower percentage of lost cells.

nrm.2 was the only projection scenario with a percentage of lost cells as high as the corresponding beech model. It also displayed the most extreme decreases of all four scenarios. Increases were again possible due to T and P increases below the optima of the corresponding effects. Decreases were caused by a pronounced increase in T and an especially pronounced decrease in P. Those observations were also responsible for the lost cells, especially P, which also caused large amounts of lost cells in the corresponding beech model.

dyn.1 yielded a similar proportion of decreases as observed for nrm.2. The projection scenario also contained the least amount of lost cells. Increasing and decreasing hw were

closest to zero, as also observed in beech. Increases and decreases were both likely caused by increases in T. They moved below and above the T effect optimum with P decreases being less pronounced. The percentage of lost cells was especially low compared to beech due to the T data for the model ranging to higher values.

dyn.2 displayed the highest amount of hw decreases and a medium amount of lost cells. At the same time it showed the highest hw increases. Decreases were caused by the pronounced increase in T and decrease in P. T moved almost exclusively along the decreasing covariate effect after the optimum, while the reduction in P moved along the decreasing P effect at lower precipitation. Again, the percentage of lost cells was lower compared to beech due to the larger T range in the model fit data.

Cells which moved into the range of the data used for model fit (coded as new areas in Figure 3.13, Figure 3.14 and Table 3.11) were rarely observed, making up proportions of less than 1 % for all projection scenarios.

Projection maps for beech showed increases in hw for northern, central and south Germany, while decreases were observed between central and northern and especially east Germany as well as in the South-West (Figure 3.13). Areas which left the climate range of the model fit were mostly observed in east Germany, west Germany and the South-West. Areas with a decrease in hw and areas with lost pixels were largest for aggregation scenario dyn.2, closely followed by nrm.2. Increases in hw were most often observed in scenario dyn.1. Scenario nrm.1 showed a balanced coverage of areas with hw increase and decrease/loss of areas. The range of increase/decrease was also larger for nrm.1 while changes for dyn.1 were less extreme. New areas which entered the range of climate data of the model fit were only observed for a few pixels in north-east Germany for nrm.1.

hw almost exclusively decreased for pine, with the strongest decreases observed in the same areas as for beech, namely east, west and south-west Germany (Figure 3.14). Scenario dyn.2 showed only areas of decreasing hw except for very few pixels with slightly increasing hw. Decrease in hw was less extreme for dyn.1 and most extreme for nrm.1.

3.4 Discussion

This study presented the creation of climate sensitive h-a models based on three different periods of climate data aggregation and forest inventory data.

Productivity was quantified by modeling Weise's dominant height hw over dominant age aw. hw was modeled using a h-d model and using dominant diameter dw as a predictor. dw and aw were both estimated from a KDE for each tract.

The ranges of dw modeled for beech and pine can be compared to recently published

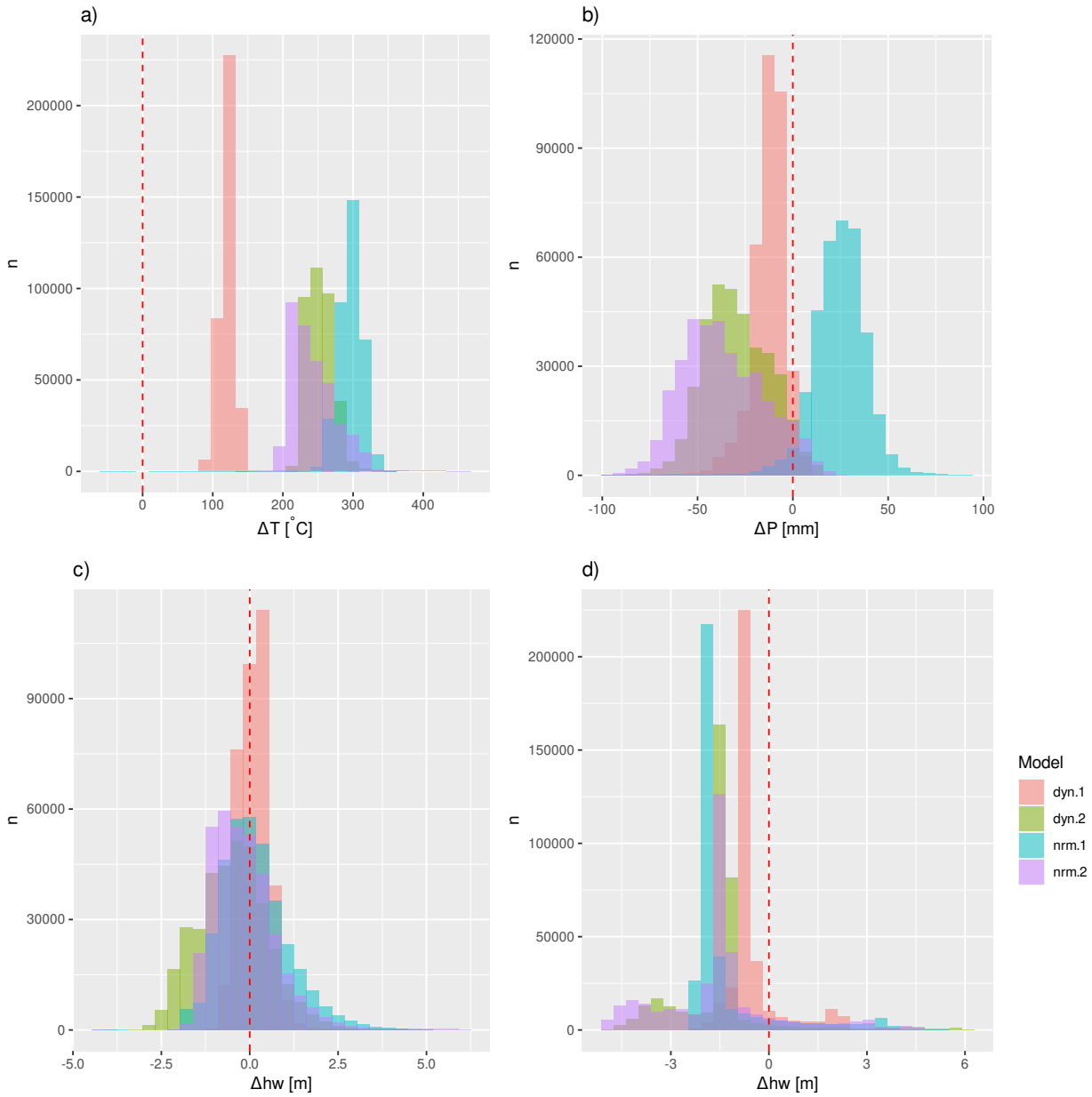


Figure 3.10: Histograms of changes in `T`, `P` and `hw` from 2012 to 2050 for the four projection scenarios for beech and pine.

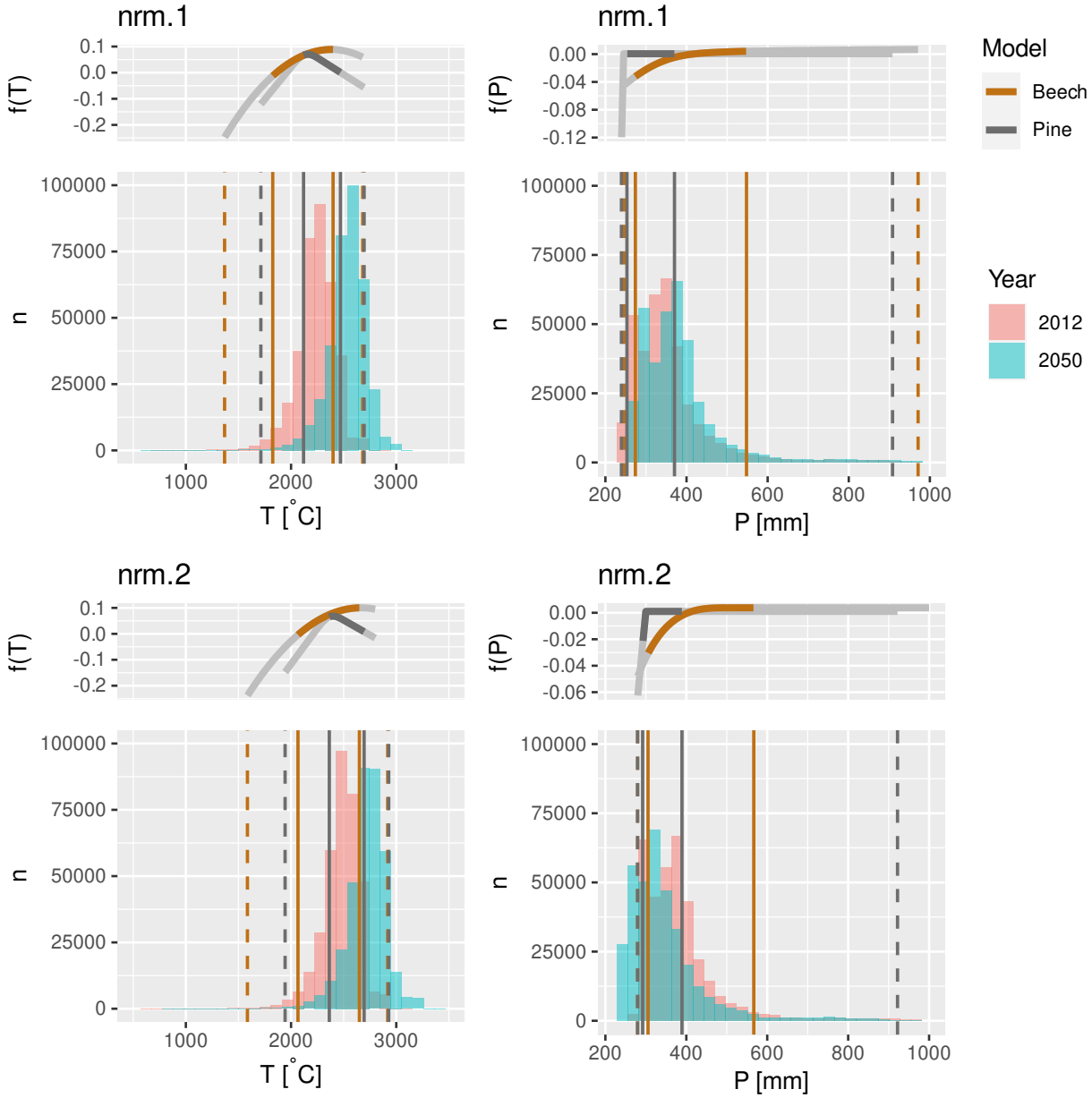


Figure 3.11: Histograms of projections for `T` and `P` based on scenarios `nrm.1` and `nrm.2` with splines for beech and pine. Colored segments in the splines denote the inter-quantile range (5%-95%) of the data used for the fit, grey segments the areas outside of the inter-quantile range. Solid vertical lines in the histograms show the inter-quantile ranges of the data from the fit, dashed lines the full data range.

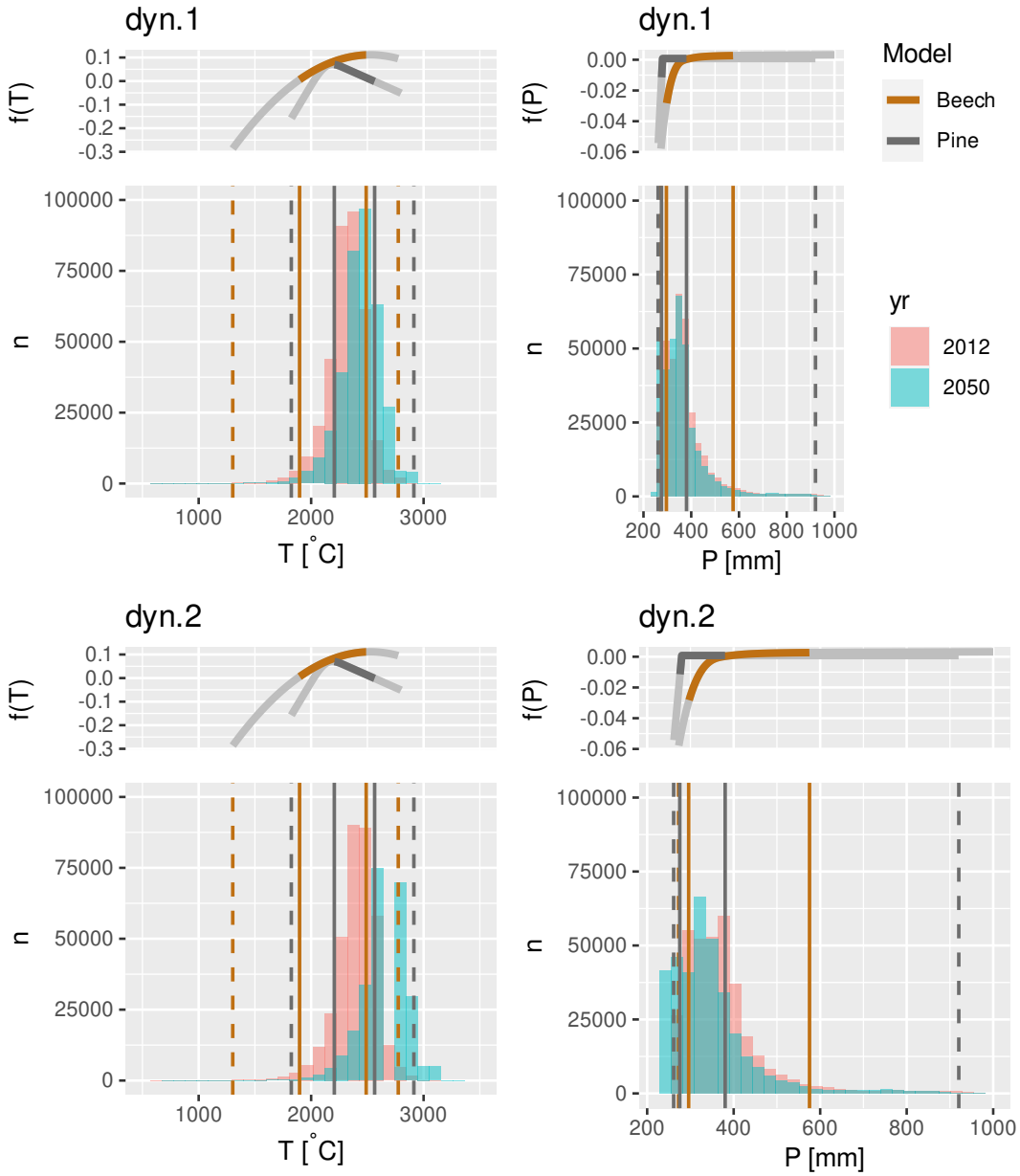


Figure 3.12: Histograms of projections for `T` and `P` based on scenarios `dyn.1` and `dyn.2` with splines for beech and pine. Colored segments in the splines denote the inter-quantile range (5%-95%) of the data used for the fit, grey segments the areas outside of the inter-quantile range. Solid vertical lines in the histograms show the inter-quantile ranges of the data from the fit, dashed lines the full data range.

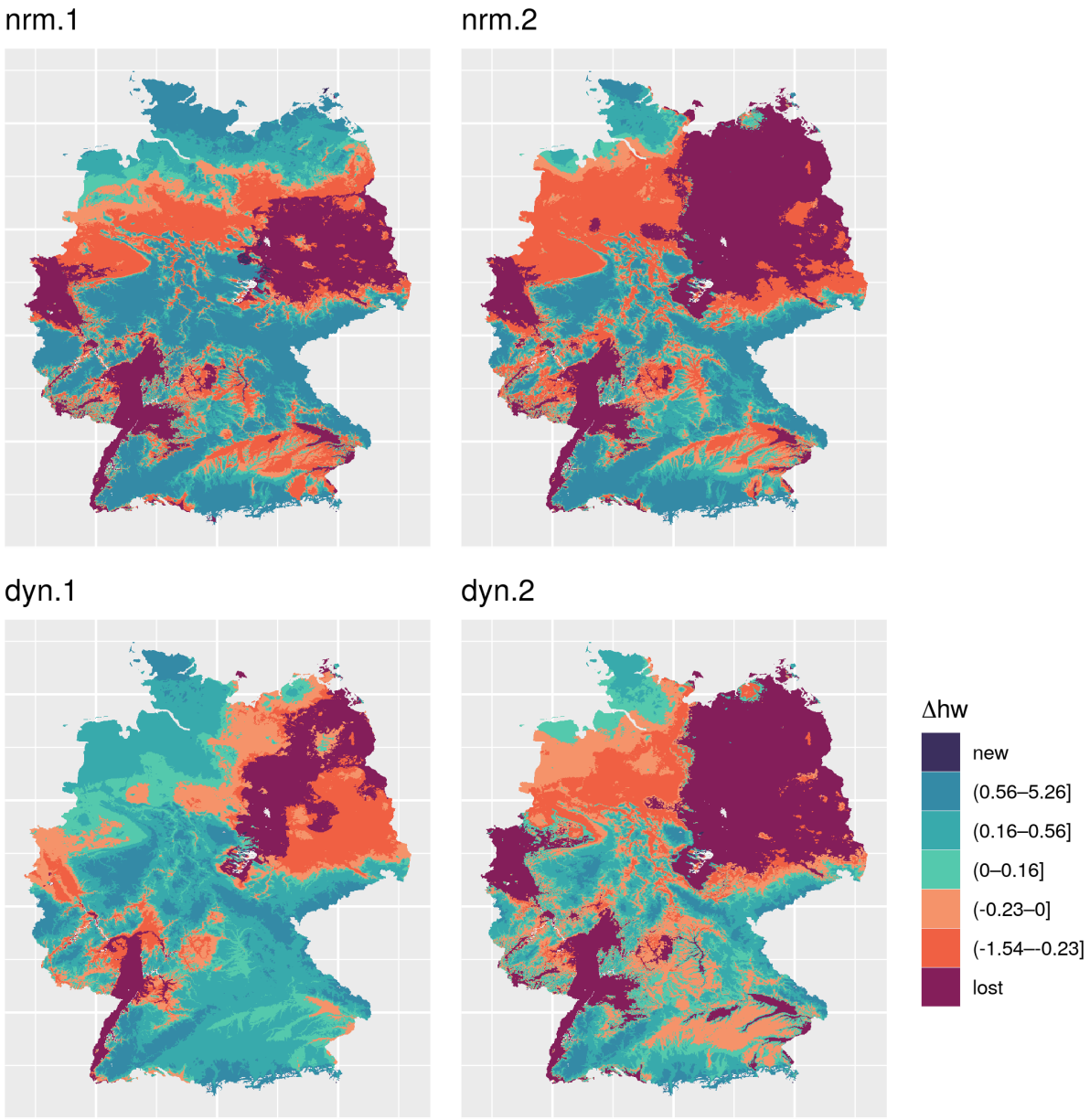


Figure 3.13: Differences in beech `hw` [m] at age 100 between 2012 and 2050 for the four projection scenarios. *Lost* areas were part of the climate range of the final model fit of step I.3 in 2012 but not in 2050. *New* areas were not part of the climate range in 2012 but entered it in 2050.

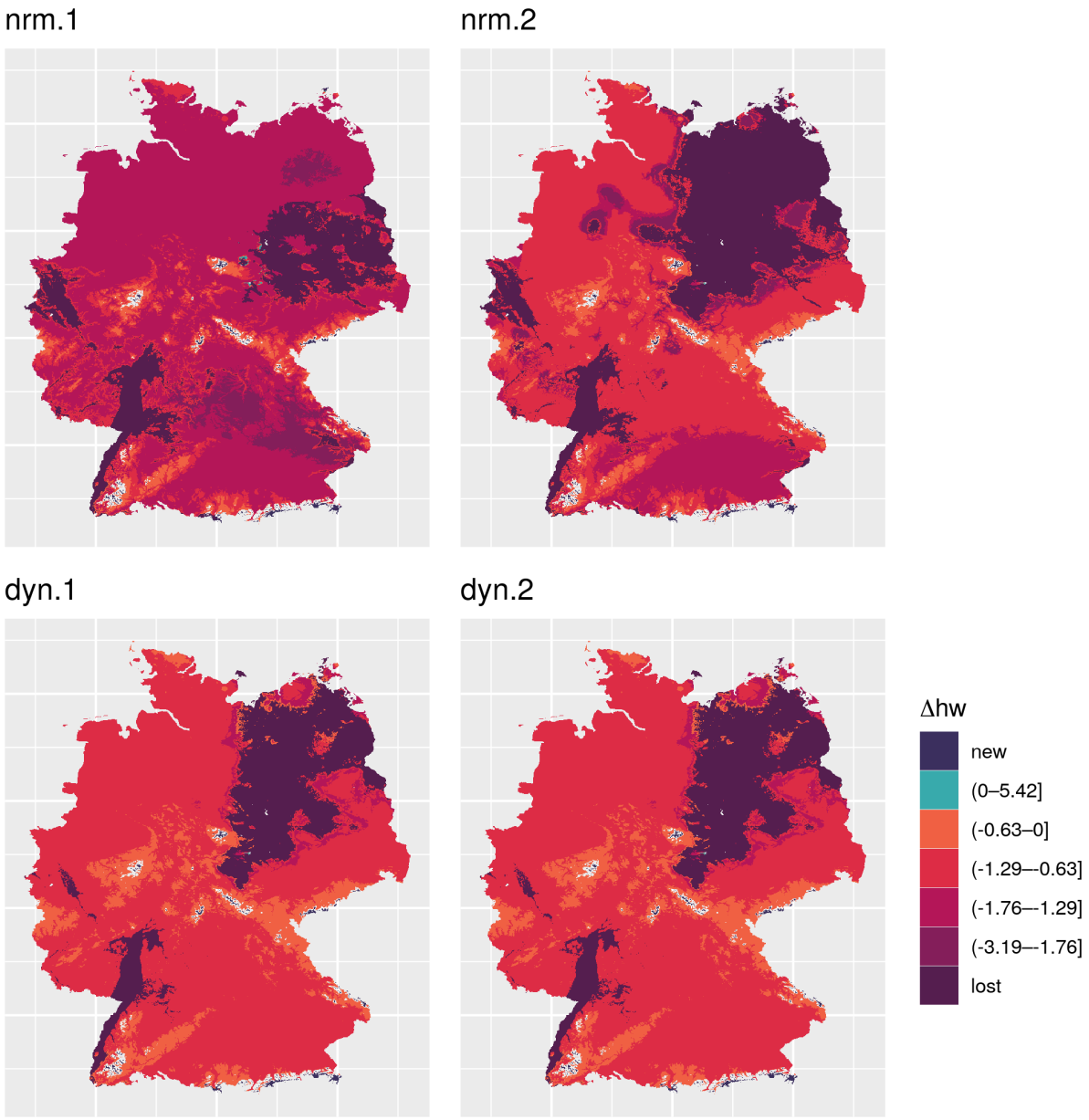


Figure 3.14: Differences in pine `hw` [m] at age 100 between 2012 and 2050 for the four projection scenarios. *Lost* areas were part of the climate range of the final model fit of step I.3 in 2012 but not in 2050. *New* areas were not part of the climate range in 2012 but entered it in 2050.

dynamic yield tables (Nuske et al. 2022). They display larger ranges, especially into the upper margins of tabulated productivity, given the modeled `aw` ranges.

`dw` values were also higher compared to dominant diameter values modeled in other studies (Sharma et al. 2012; Díaz-Yáñez et al. 2017; Stankova et al. 2022). Few studies give information on the corresponding age (Palahí et al. 2003; Castaño-Santamaría et al. 2023), which could be due to the fact that a dominant diameter is often used as substitute for stand age.

The larger `dw`s compared to other studies and yield tables may express a factually larger range in productivity, which is not within the range of current yield tables. They might also have been caused by the differing modeling approaches. Indeed, it is likely that the selected bandwidth of 3 for the KDEs yielded systematically higher values compared to the yield table values. In order to get more realistic `dw` and consequently `hw` and SI, the bandwidth could be derived from experimental plots yielding full diameter samples. They could also simply be calibrated using yield tables by trying different bandwidths and comparing them to `dw` and age pairs in the tables.

Although modeling precise dominant heights was not within the scope of the study, the `hw` ranges turned out to be realistic. `hw` was similar to dominant heights found in the literature (Sharma et al. 2016; Brandl et al. 2018), with a tendency to larger values given the respective age (Díaz-Yáñez et al. 2017; González-Rodríguez and Diéguez-Aranda 2021; Castaño-Santamaría et al. 2023). The larger values might be caused by the bandwidth, but also by the method of using the 20% largest trees, while dominant heights in the referenced studies were mostly based on the 100 thickest trees (`h100`).

The calculated SIs were also much larger than those calculated in similar studies, especially the minimum values (Palahí et al. 2003; Albert and Schmidt 2010; Nothdurft et al. 2012; Brandl et al. 2018; Schmidt 2020, ; González-Rodríguez and Diéguez-Aranda 2021). Only one study showed similarly high, and even larger minimum SI values for pine (Nothdurft et al. 2012). Compared to the aforementioned yield tables, SI values covered a larger range than the tabulated ones, reaching into the extrapolation space for higher and especially lower yield classes.

The values for the original parameters λ and c were higher than in most other studies (Lappi 1997; Mehtätalo 2004). This might have been caused by the flexible iterative grid optimization that was used in this study, with grid ranges being adjusted if the search reached the margins of the grid. Another study which also used an iterative grid search showed higher values for λ and c compared to the observed ones (Schmidt et al. 2018).

The aggregated climate periods for the model fit showed an expected structure, with `nrm.1` located in a colder climate, while `nrm.2` covered a warmer climate after the turn of the millenium. The dynamic aggregation over `aw` (`dyn`) covered a medium climate between

nrm.1 and nrm.2 for both T and P sum. This was to be expected, since aw covers climate values until 2012 but often goes far back beyond 30 years of age. The higher P of nrm.2, compared to dyn and nrm.1, corresponded to a slightly higher mean precipitation of 64.9 mm for 1991-2020 compared to 60.8 mm for 1961-1990 in Germany (Deutscher Wetterdienst 2023). Higher P values might also have been caused by an increased length of the vegetation period, over which climate parameters were aggregated.

The values for the original parameter λ for the h-a model were, as expected, larger than for the original h-d model, since transformed aw was passed instead of dbh. The values can be interpreted as a difference in hw between a 50 and a 100 year old tree. However, the optimized values for λ were not realistic in that sense, with 36 m for pine and especially 293 m for beech. In comparison, hw differences between 50 and 100 year old stands from yield tables range between 11.6 and 15.3 m for beech and 6.8 to 9.6 m for pine. It must be stressed, that the grid optimization based on statistical measures can potentially be repeated to infinity. It was, however, cut off at a point where further iterations led to AIC changes at $\leq 10^{-1}$ magnitude. Accordingly, the values for λ became increasingly implausible in an ecological sense, and these values were selected for the best model fit rather than ecological plausibility.

The different aggregation periods led to different levels of climate effects for both species but very similar shapes. Visible differences in effect shapes could be expected if aggregation periods yielded different spatial configurations of climate, so that a warmer climate assigned to one stand would be relatively colder in another period. This was not the case and it can be concluded that the model fit itself is not very sensitive to different forms of aggregation. However, this might not be the case if model fits are based on other climate projections or aggregation periods.

T displayed distinct parabolic covariate effects for both species. Specific effect shapes are often assumed but not backed by data and might be biased due to the applied statistical method (Loehle 2000; Oksanen and Minchin 2002). Parabolic effects are often suggested to be close to ecological reality. They have also been identified in previous studies using GAMs or related spline-based approaches (Nothdurft et al. 2012; Schmidt 2020). Other similar studies identified asymptotic functions, without an identifiable optimum at higher temperatures (Albert and Schmidt 2010; Burggraef et al. 2016; Pya and Schmidt 2016; Brandl et al. 2019). It is likely that asymptotic effects are identified if temperature data doesn't range far enough into hot areas.

P effects showed asymptotic shapes for both species, with the effect for pine having to be fixed using a segmented covariate in order to maintain significance. Asymptotic precipitation effects have often been observed (Burggraef et al. 2016; Schmidt 2020), while other studies showed parabolic effects for precipitation (Nothdurft et al. 2012; Brandl et al. 2019). Those were probably caused by sample plots located in mountainous areas, with high precipitation

correlating with cold temperatures. The relatively smaller trees at those sites led to a declining effect of precipitation on productivity in the model. The data available on dry sites cannot balance out this effect. This was also the problem with the `P` effect for pine in the models presented here, which is why a segmented linear effect had to be used to keep this important component of climate sensitivity in the model.

Adding soil parameters to the model turned out to be difficult, as covariate effects often lacked significance or showed implausible shapes. The reason may have been that regionalized soil parameters don't represent the site conditions exactly enough. Implausible effect shapes may often be caused by correlations. For instance, high base saturation may be found on shallow, calcareous sites, correlating small `hw` values to high base saturation.

For the final model, `ph` was selected as a soil covariate. It displayed significance and plausibility for both species. Beech displayed a plausible optimum at a `ph` value of 4.5. A recent study on tree growth depending on soil properties determined, on average, similar values, locating optimal growth for beech on soil types *Haplic Luvisols* and *Eutric Planosols*, with `ph` values of 3.5-4.2 and 5-6.5 respectively (Bončina et al. 2023). However, the optimal growth might have been affected more by the good drainage of the aforementioned soil types than `ph` value. The values mentioned in that study match the observed optimum for pine, which reached its maximum growth at *Dystric Leptosols* with `ph` values of 4-5. A study on pine monitoring plots found similar `ph` values of 4-4.7 for 20-30 cm soil depth (Prietzel et al. 2020), while another found much higher values of 4.98-7.87 for 60 cm soil depth (Aguirre et al. 2022). However, traditional indicator values characterize both beech and pine as indifferent to `ph`, so it is possible that the effects were correlated to other soil characteristics (Ellenberg 1991).

In the final models, z_{ki} displayed the highest relative amplitudes, which was to be expected due to most of the `hw` signal being explained by `aw`. The order of the relative amplitudes of climate and soil covariates was interesting, in that the soil parameter `ph` displayed similar or more explanatory power compared to `P`. This was in spite of the difficulty in the selection of soil covariates, and of soil covariates mostly being inferior to climate covariates in previous model fitting attempts (Burggraef et al. 2016). However, soil parameters displaying higher explanatory power than precipitation has occasionally been observed in site-productivity modeling (Brandl et al. 2019; Schmidt 2020). Temperature is known to display the highest impact on site productivity measures after age or diameter (Albert and Schmidt 2010; Nothdurft et al. 2012; Pya and Schmidt 2016; Rita and Borghetti 2019; Brandl et al. 2019; Schmidt 2020; Cheng et al. 2023). It should also be added, that statistics on the model fit don't always represent the actual impact of a covariate effect on model predictions. A difference between extreme sites can still lead to a difference of several meters in `hw` in the presented models. Thus, even a spline of low significance can be helpful in model application.

This is why the relative amplitude might be a better measure to evaluate covariate effects. It also needs to be stressed that the low significance and plausibility of soil covariates in the presented models do not mean a generally low effect of soil relative to climate in real life. Rather, it might indicate a problem with using regionalized parameters to predict productivity measured on a forest inventory grid, which differs from the soil inventory grid or the grid being too coarse.

As opposed to model fit, large differences were observed between the projections for different models. Differences could be observed both in the projections for `hw` as well as the underlying climate data. Differences in climate were caused both by the location of the aggregation period on the timeline and by the length of the aggregation period for each scenario.

This was especially apparent for `T` and the dynamic aggregation over 100 years for `dyn.1`. Since it covered more cold years, the resulting `T` was far lower compared to scenarios `nrm.1`, `nrm.2` and `dyn.2`, which closely overlapped in terms of `T` values. The popular 30-year baseline periods, which are often used for climate aggregation, may thus overestimate the climatic effect on older stands. Static climate aggregations for different periods are often used, depending on the available data (Nothdurft et al. 2012; Pya and Schmidt 2016; Brandl et al. 2019). Especially 30 year means from the WorldClim data set are very popular (Hijmans et al. 2005; Aguirre et al. 2022; Cheng et al. 2023).

An interesting observation was the increase in `P` in the projection for `nrm.1`, while `P` decreased for all other scenarios. This was due to the projection scenario in the early 2000s lying in the abovementioned period where precipitation increased slightly, overlapping with the aggregation period of `nrm.2` for the fit. However, `nrm.1` also showed the highest increase in `T` compared to the other three scenarios.

With projected `T` and `P` moving in different directions along the covariate effects of the models, changes in `hw` from 2012 to 2050 also showed high variations between projection scenarios.

For beech, `nrm.1` showed the largest increases in `hw` compared to the other two scenarios due to the increase in `P`. This shows that `P`, which is notoriously hard to predict, is one of the key factors determining the direction of productivity trends, in spite of the low amplitudes shown by the respective covariate effects in the models. `nrm.1` and `nrm.2` turned out to be the most positive scenarios. The `dyn.1` scenario was the most moderate one – it didn't show high decreases or increases, which were closer to zero, supporting the hypothesis that 30-year means might overestimate predictions for older stands. The most negative scenario was `dyn.2` which showed mostly decreases and fewer increases in `hw`.

For pine, all scenarios predicted high `hw` losses. Only `dyn.1` and `nrm.2` showed increases in `hw` in some areas while `nrm.2` and `dyn.2` showed the highest losses. These patterns were

likely rooted in the small effect of P in the pine models, which means the species couldn't profit from increases in P, especially in scenario nrm.1.

The maps of the four scenarios showed SI losses for beech especially in northeast, west and southwest Germany and increases in the colder north and in mountainous regions in central and south Germany. This is in accordance with predictions of other studies (Albert and Schmidt 2010; Brandl et al. 2018). The regions in northwest Germany with the federal states of Brandenburg, Berlin and Mecklenburg-West Pomerania are notoriously dry, while the Rhine lowlands in the southwest are known to be hot and dry. The magnitudes of increase and decrease of beech SI fit to other projections, given longer projection scenarios until 2070 or 2100 (Brandl et al. 2018; Maleki et al. 2022).

The same patterns were visible for pine, even though hw losses prevailed here. Losses were higher in the regions where beech also experienced the highest losses while productivity gains were restricted to the mountainous regions. The ranges of decreases and the few increases of hw also showed similarities to other projections (Nothdurft et al. 2012; Maleki et al. 2022)

Based on these results, it is obvious that different aggregation periods heavily affect model predictions. It has become apparent that not only the intensity of increase or decrease in productivity is affected but also the direction of effects. The same areas might be predicted to increase in productivity for one aggregation period and decrease for another.

It is highly recommended to use a dynamic aggregation method if possible, since those are closer to ecological reality. In any case, if predictions are to be used to inform practitioners and political actors, ensemble studies with different modeling approaches and assumptions have to be considered. Also, syntheses of multiple modeling studies within the scope of meta-studies are imperative, since model predictions may contradict each other for certain areas based on aggregation of climate data alone.

An important caveat that has to be mentioned about the approach presented here is that differences in the effect of climatic covariates on stands of different ages were not modeled. The effect might have been represented in the two dynamic periods to a certain degree: while old stands (dyn.1) might not be heavily influenced by a changing climate, younger stands (dyn.2) might suffer more dramatically from drought and increased T. This was to some degree represented in the shorter aggregation period of dyn.2, which was located in the warmer periods with stands not spending the major part of their life in the colder 20th century. It was also represented in the larger aggregation period of dyn.1 which showed changes in both directions (hw decrease and increase), but with less extreme ranges.

3.5 Conclusion

The goal of the study presented here was to analyze how static and dynamic climate aggregation methods, as a basis for site-productivity models, influence covariate selection and predictions. In conclusion, covariate selection was only marginally affected by climate aggregation periods, while predictions were influenced dramatically. Projections based on different aggregation periods may lead to different magnitudes of increase and decrease in site productivity as well as to contradictory results concerning the direction of productivity changes. It is recommended to use dynamic aggregation methods across tree or stand age if possible, and to keep in mind possible over- or underestimations of productivity changes if static aggregation periods are used.

4 Synthesis

The thesis presented here was comprised of two studies. The first one tested the potential of a continental scale set of inventory data for the development of site productivity models. The second study analyzed the sensitivity of site productivity models to different aggregation periods of the underlying climate data.

In the development of models for both studies, soil parameters were also evaluated for plausibility and significance as potential covariates.

In both studies, a tendency towards parabolic covariate effects was observed for temperature in all species. Comparing temperature effects for beech in parts 1 and 2, it was clear that the range of the effect which was covered by the bulk of the data (5%-95% inter quantile range) was located below the optimum and the decrease of the effect was mainly observed where data was sparse. For pine, a different pattern was observed: while the optimum effect for temperature was far more pronounced in part 1, the biggest part of the data was still located below the optimum. The opposite was observed in part 2. This was likely caused by the different distribution of the data within the temperature-precipitation climate envelope (Figure 4.1). While the pan-European dataset from the first study showed a distribution from high precipitation-cold temperatures to low precipitation-warm temperatures, data was distributed across a wide temperature range, but it was clustered in a narrow range of precipitation in part 2. With pine already being located at its precipitation limit, increasing temperatures only lead to a negative effect on height.

Precipitation effects showed asymptotic shapes for beech in both studies, with the asymptote being more pronounced in the second study. For pine, the same pattern was observed, with the aforementioned narrow range of precipitation in the second study, which led to its inclusion as a segmented linear effect. In the second study, the precipitation effect was also heavily influenced by low quality, cold sites in high elevation areas. Here, a definite advantage of the pan-European data set became visible, with the large scale breaking up the correlation between cold temperature and high precipitation.

However, plausible effects depend not only on the range, but also on the quantity of the data across the environmental gradient. Plausible precipitation effects for pine have been identified based on NFI data when it was pooled with data from regional inventories (Schmidt 2020), sometimes including the negative effect due to the temperature correlation (Nothdurft et al. 2012). Some studies used data from two to three repeated inventories, adding a small real time series which might also influence effect shapes (Brandl et al. 2019; Schmidt 2020).

Similar problems were observed in both studies when trying to include soil parameters in the models. Both the pan-European and the German data sets showed problems with clustered data and implausible shapes, likely due to correlations with plots on shallow soils.

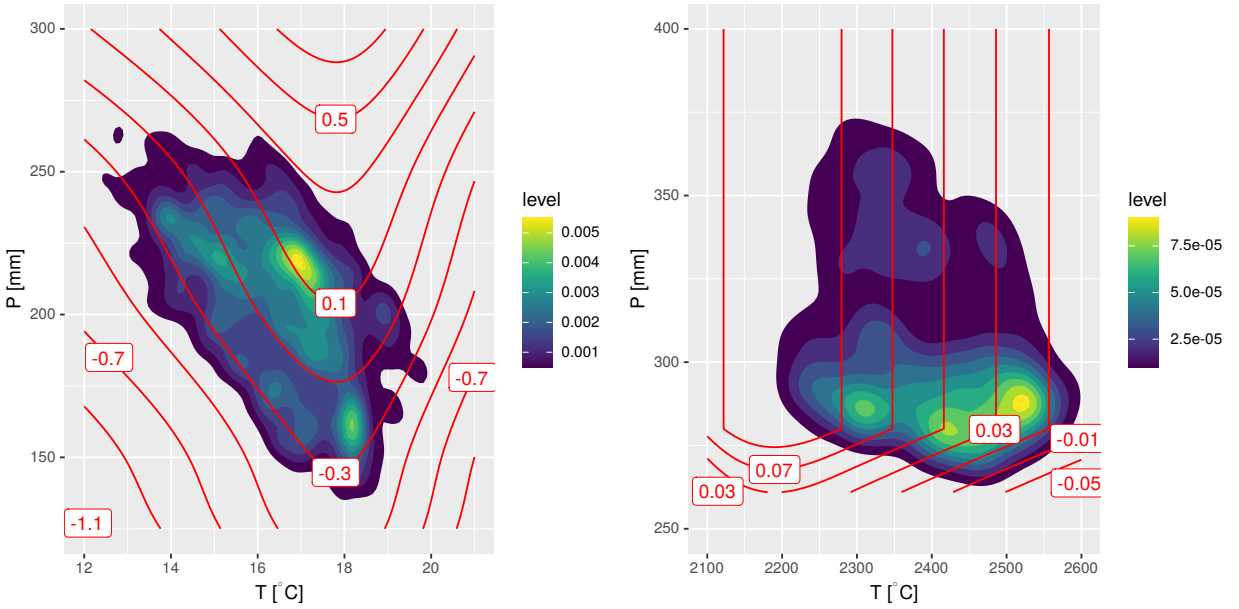


Figure 4.1: Grids showing the density of pine data within a temperature-precipitation envelope for the model from the first study (left) and the second study (right). Note that the first study used temperature mean for the warmest quarter, while the second study used temperature sum for the vegetation period averaged over stand age. Red isolines denote equal effect on the single tree height-diameter relation (first study) and dominant height-age relation (second study) for each model in m.

Yet at least one parameter could be identified in each study, the C:P-ratio in the first one and the pH-value in the second. Multiple candidate paramaters were available for each model, such as pH value, potassium stock and base saturation in the first study and base saturation or available water capacity in the second study. However, few parameters were plausible and significant across all species and scenarios.

In the second study, problems might have been caused by the regionalization of soil data. The same does not hold for the first study though, where samples were taken on site and parameters were obtained from laboratory tests (Höhle et al. 2018).

Another problem might be the modeling approach. GAMs, or regression approaches in general, might not be the best basis to identify soil parameter effects, while the literature shows multiple examples of successful soil-data based site productivity modeling using machine learning algorithms, such as regression trees (Aertsen et al. 2012; Mellert and Ewald 2014; Jiang et al. 2015).

Projected changes in productivity for the two studies were similar for beech, in spite of the different modeling approaches. The results from the first study contain more areas of decreasing productivity in south and central Germany if compared to projection scenarios `nrm.1`, `nrm.2` and `dyn.1` from the second study.

Since temperature effects were similarly shaped and precipitation splines were steeper in the second study, this was likely due to different projection times. The basic projections for the first study were further in the past (1960-1990 as compared to 2012) and future projections further in the future (2050-2080 as compared to 2050), representing higher temperatures and less precipitation. Also, the static aggregations of the WorldClim dataset which were used in the first study were most similar to `nrm.1` in the second study, meaning projections would likely have been be more negative for young stands and more positive for older stands if dynamic aggregations had been used.

Projected productivity differences for pine were far less dramatic in the first study compared to the second one, which showed mostly decreases. Here, the reason was to be found in the covariate effects, especially the precipitation covariate. In the first study, precipitation could be included as a spline with a relatively modereate decrease in the direction of decreasing precipitaton, while the decrease for the segmented linear effect in the second study was very strong. The same observation could be made for the temperature effects, which showed a more moderate slope before and after the optimum in the first study, in comparison to the effect in the second study.

In conclusion, the pine model from the second study can barely be considered realistic due to the restricted data basis. If very steep or insignificant climate effects are observed, the data basis should be improved by adding data from regional inventories or increasing the spatial extent.

Important caveats concerting both studies, which have to be kept in mind, are the assumptions of the SFT approach and the missing differentiation in the effects on different development phases of a stand.

The problems with the SFT approach have, in part, already been discussed in the discussion of the first study. The SFT approach assumes *time* to be a proxy (or surrogate) parameter for those effects which can be completely explained by varying site conditions across space. However, histories are expected to differ across sampling plots. Observed sites may come from different species compositions in the past and are likely to have experienced different calamities and management practices. This may not be a problem and it may be assumed that the major effects are still visible and that historical differences are averaged out across the data set. This assumption, however, usually remains untested. An early review about SFT-studies on successional stages of plant communities concluded that the SFT-approach is helpful if general trends are to be analyzed or for hypothesis building, while problems arise when historical effects are of a large magnitude (Pickett 1989). A similar warning has been issued for predictions based on modeling studies in general, pointing out that not only the history of ecological events but also the chronology might lead to different results (Pilkey and Pilkey-Jarvis 2007).

The problem of the unknown difference in each environmental effect depending on the age of the stand has also been mentioned. Temperature and precipitation as well as soil covariates are likely to have a different effect on a young stand compared to older stands which have aged past the culmination point of their growth. A possible solution would be to weigh the effects using growth curves obtained from studies based on real time series, an approach which is currently being tested (Schick et al. 2023).

4.1 Conclusion

This thesis investigated the development of site productivity models within the scope of two studies. The first study evaluated the potential of a pan-European dataset for model development. The second study analyzed the sensitivity of site productivity models to different aggregation periods of climate data, including dynamic aggregation over stand age. With regard to the research questions posed, it became apparent that forest inventories on a continental scale may be superior to national scale inventories with respect to the identification of plausible effects, especially at the edges of the data. However, results may be similar if national scale inventories are supplemented with fine scale regional inventory data. Advantages depend on the geographic area covered by the species in question. Dynamic aggregation of climate data over the stand age is clearly advisable, in order to avoid over- or underestimations or even contrasting directions of productivity changes when projecting into the future. The identification of plausible soil covariate effects proved difficult in both studies, with most parameters returning implausible or insignificant effects. However, reasonable effects were identified for C:P ratio and pH value. Other modeling approaches, such as regression trees, might be better suited to identify effects of soil parameters on site productivity.

References

Aertsen W, Kint V, De Vos B, et al (2012) Predicting forest site productivity in temperate lowland from forest floor, soil and litterfall characteristics using boosted regression trees. Plant and Soil 354:157–172

Aguirre A, Moreno-Fernández D, Alberdi I, et al (2022) Mapping forest site quality at national level. Forest Ecology and Management 508:120043. https://doi.org/10.1016/j.foreco.2022.120043

Akaike H (1974) A new look at the statistical model identification. IEEE Transactions on Automatic Control 19:716–723. https://doi.org/10.1109/TAC.1974.1100705

Albert M, Schmidt M (2010) Climate-sensitive modelling of site-productivity relationships for Norway spruce (Picea abies (L.) Karst.) and common beech (Fagus sylvatica L.). Forest Ecology and Management 259:739–749

Albrecht A, Hanewinkel M, Bauhus J, Kohnle U (2012) How does silviculture affect storm damage in forests of South-Western Germany? Results from empirical modeling based on long-term observations. European Journal of Forest Research 131:229–247. https://doi.org/10.1007/s10342-010-0432-x

Apaydin H, Sonmez FK, Yildirim YE (2004) Spatial interpolation techniques for climate data in the GAP region in Turkey. Climate Research 28:31–40

Bastrup-Birk A, Neville P, Chirici G, Houston T (2007) The BioSoil Forest Biodiversity Field Manual. Working Group on Forest Biodiversity

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48. https://doi.org/10.18637/jss.v067.i01

Battaglia M, Sands P (1997) Modelling site productivity of Eucalyptus globulus in response to climatic and site factors. Functional Plant Biology 24:831–850

Battaglia M, Sands PJ (1998) Process-based forest productivity models and their application in forest management. Forest Ecology and Management 102:13–32

Bergès L, Chevalier R, Dumas Y, et al (2005) Sessile oak (Quercus petraea Liebl.) Site index variations in relation to climate, topography and soil in even-aged high-forest stands in northern France. Annals of Forest Science 62:391–402

Bioclim (2021) Bioclim | WorldClim - Global Climate Data. http://www.worldclim.org/data/bioclim.html. Accessed 1 Oct 2021

Bitterlich W (1947) Die Winkelzählmessung. Allgemeine Forst- und Holzwirtschaftliche Zeitung 58:94ff

Böckmann T, Saborowski J, Dahm S, et al (1998) Die Weiterentwicklung der Betriebsinventur in Niedersachsen. Forst und Holz 53:219–226

Bolte A, Ammer C, Löf M, et al (2010) Adaptive forest management: A prerequisite for sustainable forestry in the face of climate change. In: Sustainable Forest Management in a Changing World. Springer, pp 115–139

Bončina A, Klopčič M, Trifković V, et al (2023) Tree and stand growth differ among soil classes in semi-natural forests in central Europe. Catena 222:106854

Bontemps J-D, Bouriaud O (2013) Predictive approaches to forest site productivity: Recent trends, challenges and future perspectives. Forestry 87:109–128

Bouwman M, Forrester D, Ouden J, et al (2020) Species interactions under climate change in mixed stands of Scots pine and pedunculate oak. Forest Ecology and Management 481:118615. https://doi.org/10.1016/j.foreco.2020.118615

Brandl S, Falk W, Klemmt H-J, et al (2014) Possibilities and limitations of spatially explicit site index modelling for spruce based on national forest inventory data and digital maps of soil and climate in Bavaria (SE Germany). Forests 5:2626–2646

Brandl S, Falk W, Rötzer T, Pretzsch H (2019) Assessing site productivity based on national forest inventory data and its dependence on site conditions for spruce dominated forests in Germany. Forest Systems 28:e007. https://doi.org/10.5424/fs/2019282-14423

Brandl S, Mette T, Falk W, et al (2018) Static site indices from different national forest

inventories: Harmonization and prediction from site conditions. Annals of Forest Science 75:56

Burggraef L, Schmidt-Walter P, Hilbrig L, Schmidt M (2016) Standort-Leistungsmodelle als Grundlage für realistische Waldentwicklungsszenarien unter Klimawandel. Tagungsband der Jahrestagung der Sektion Ertragskunde im DVFFA 8–19

Campbell EM, Magnussen S, Antos JA, Parish R (2021) Size-, species-, and site-specific tree growth responses to climate variability in old-growth subalpine forests. Ecosphere 12:e03529

Castaño-Santamaría J, López-Sánchez CA, Obeso JR, Barrio-Anta M (2023) Development of a site form equation for predicting and mapping site quality. A case study of unmanaged beech forests in the Cantabrian range (NW Spain). Forest Ecology and Management 529:120711

Cheng R, Zhang J, Wang X, et al (2023) Predicting the growth suitability of Larix principis-rupprechtii Mayr based on site index under different climatic scenarios. Frontiers in Plant Science 14:1097688–1097688

Ciais Ph, Reichstein M, Viovy N, et al (2005) Europe-wide reduction in primary productivity caused by the heat and drought in 2003. Nature 437:529–533. https://doi.org/10.1038/nature03972

Coops N, Waring R, Landsberg J (1998) Assessing forest productivity in Australia and New Zealand using a physiologically-based model driven with averaged monthly weather data and satellite-derived estimates of canopy photosynthetic capacity. Forest Ecology and Management 104:113–127

Daly C, Halbleib M, Smith JI, et al (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology: a Journal of the Royal Meteorological Society 28:2031–2064

Dănescu A, Albrecht AT, Bauhus J, Kohnle U (2017) Geocentric alternatives to site index for modeling tree increment in uneven-aged mixed stands. Forest Ecology and Management 392:1–12

De Vos B, Cools N (2011) Second European Forest Soil Condition Report. Volume I: Results of the BioSoil Soil Survey. Research Institute for Nature and Forest, Brussel

Deutscher Wetterdienst (2023) Zeitreihen und Trends. In: Wetter und Klima – Leistungen – Zeitreihen. https://www.dwd.de/DE/leistungen/zeitreihen/zeitreihen.html-#buehneTop, Accessed: 2023-05-17

Díaz-Yáñez O, Mola-Yudego B, Olabarria J, Pukkala T (2017) How does forest composition and structure affect the stability against wind and snow? Forest Ecology and Management 401:215–222. https://doi.org/10.1016/j.foreco.2017.06.054

Dietrich H, Wolf T, Kawohl T, et al (2019) Temporal and spatial high-resolution climate data from 1961 to 2100 for the German National Forest Inventory (NFI). Annals of Forest Science 76:1–14

Eggers J, Lindner M, Zudin S, et al (2008) Impact of changing wood demand, climate and land use on European forest resources and carbon stocks during the 21st century. Global Change Biology 14:2288–2303. https://doi.org/10.1111/j.1365-2486.2008.01653.x

Ellenberg H (1991) Zeigerwerte von Pflanzen in Mitteleuropa. Scripta geobotanica 18:1–248

Fisher JB, Huntzinger DN, Schwalm CR, Sitch S (2014) Modeling the terrestrial biosphere. Annual Review of Environment and Resources 39:91–123

Friedrichs DA, Trouet V, Büntgen U, et al (2009) Species-specific climate sensitivity of tree growth in Central-West Germany. Trees 23:729–739

Fries A, Lindgren D, Ying CC, et al (2000) The effect of temperature on site index in western Canada and Scandinavia estimated from IUFRO Pinus contorta provenance experiments. Canadian Journal of Forest Research 30:921–929

Fukami T, Wardle DA (2005) Long-term ecological dynamics: Reciprocal insights from natural and anthropogenic gradients. Proceedings of the Royal Society B: Biological Sciences 272:2105–2115

Gadow K von (2003) Waldstruktur und Wachstum. Universitätsverlag Göttingen

Galbraith D, Levy PE, Sitch S, et al (2010) Multiple mechanisms of Amazonian forest biomass losses in three dynamic global vegetation models under climate change. New Phytologist 187:647–665

Gillis MD, Omule A, Brierley T (2005) Monitoring Canada's forests: The national forest inventory. The Forestry Chronicle 81:214–221

González-Rodríguez MÁ, Diéguez-Aranda U (2021) Rule-based vs parametric approaches for developing climate-sensitive site index models: A case study for Scots pine stands in northwestern Spain. Annals of Forest Science 78:1–14

Hastie T, Tibshirani R (1993) Varying-coefficient models. Journal of the Royal Statistical Society: Series B (Methodological) 55:757–779

Hijmans RJ, Cameron SE, Parra JL, et al (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978

Hilbrig L, Wellbrock N, Bielefeldt J, et al (2014) Harmonisierte Bestandesinventur: Zweite Bundesweite Bodenzustandserhebung BZE II. Methode. Johann Heinrich von Thünen Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries

Hoffman CM, Sieg CH, Linn RR, et al (2018) Advancing the science of wildland fire dynamics using process-based models. Fire 1:32

Höhle J, Bielefeldt J, Dühnelt P, et al (2018) Bodenzustandserhebung im Wald-Dokumentation und Harmonisierung der Methoden. Thünen Working Paper

Hutchinson MF (1989) A new objective method for spatial interpolation of meteorological variables from irregular networks applied to the estimation of monthly mean solar radiation, temperature, precipitation and windrun. Need for climatic and hydrologic data in agriculture in SE Asia Proc UN University workshop, Canberra, 1983 89/5:95–104

IPCC (2007) Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland

Jacob D, Bülow K, Kotova L, et al (2012) Regionale Klimaprojektionen für Europa und

Deutschland: Ensemble Simulationen für die Klimafolgenforschung. Helmholtz-Zentrum Geesthacht, Climate Service Center (CSC)

Jaiswal RK, Mukherjee S, Raju KD, Saxena R (2002) Forest fire risk zone mapping from satellite imagery and GIS. International Journal of Applied Earth Observation and Geoinformation 4:1–10

Jiang H, Radtke PJ, Weiskittel AR, et al (2015) Climate-and soil-based models of site productivity in eastern US tree species. Canadian Journal of Forest Research 45:325–342

Kahn M (1995) Quasikausale Modellierung des Standort-Leistung-Bezuges als Voraussetzung zum Aufbau flexibler MischbestandsmodelleModelling. Forstwissenschaftliches Centralblatt vereinigt mit Tharandter forstliches Jahrbuch 114:175–187

Kändler G (2006) The Design of the Second German National Forest Inventory. In: Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium

Kautz M, Anthoni P, Meddens AJ, et al (2018) Simulating the recent impacts of multiple biotic disturbances on forest carbon cycling across the United States. Global Change Biology 24:2079–2092

Kitahara F, Mizoue N, Yoshida S (2009) Evaluation of data quality in Japanese national forest inventory. Environmental Monitoring and Assessment 159:331–340

Kramer H, Akça A (2008) Leitfaden zur Waldmesslehre. Sauerländer, Frankfurt a.M.

Lappi J (1997) A longitudinal analysis of height/diameter curves. Forest Science 43:555–570

Lester RE, Close PG, Barton JL, et al (2014) Predicting the likely response of data-poor ecosystems to climate change using space-for-time substitution across domains. Global Change Biology 20:3471–3481

Loehle C (2000) Forest ecotone response to climate change: Sensitivity to temperature response functional forms. Canadian Journal of Forest Research 30:1632–1645

Löf M, Madsen P, Metslaid M, et al (2019) Restoring forests: Regeneration and ecosystem function for the future. New Forests 50:139–151

Maleki K, Astrup R, Kuehne C, et al (2022) Stand-level growth models for long-term projections of the main species groups in Norway. Scandinavian Journal of Forest Research 37:130–143

Malhi Y, Doughty C, Galbraith D (2011) The allocation of ecosystem net primary productivity in tropical forests. Philosophical Transactions of the Royal Society B: Biological Sciences 366:3225–3245

Mann ME, Bradley RS, Hughes MK (1999) Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. Geophysical Research Letters 26:759–762

Mauri A, Strona G, San-Miguel-Ayanz J (2017) EU-forest, a high-resolution tree occurrence dataset for Europe. Scientific data 4:1–8

McDowell NG, Allen CD, Anderson-Teixeira K, et al (2020) Pervasive shifts in forest dynamics in a changing world. Science 368:eaaz9463

Mehtätalo L (2004) A longitudinal height–diameter model for Norway spruce in Finland. Canadian Journal of Forest Research 34:131–140

Mellert K, Ewald J (2014) Nutrient limitation and site-related growth potential of Norway spruce (Picea abies [L.] Karst) in the Bavarian Alps. European journal of forest research 133:433–451

Mellert KH, Lenoir J, Winter S, et al (2018) Soil water storage appears to compensate for climatic aridity at the xeric margin of European tree species distribution. European Journal of Forest Research 137:79–92. https://doi.org/10.1007/s10342-017-1092-x

Mette T, Kölling C (2015) Wald, Wachstum, Umwelt. Großes Gemeinschaftsprojekt WP-KS-KW verschneidet die Bundeswaldinventur mit Boden- und Umweltdaten. LWF aktuell 107:46–49

Morin X, Augspurger C, Chuine I (2007) Process-based modeling of species' distributions: What limits temperate tree species' range boundaries? Ecology 88:2280–2291

Nothdurft A, Wolf T, Ringeler A, et al (2012) Spatio-temporal prediction of site index based

on forest inventories and climate change scenarios. Forest Ecology and Management 279:97–111. https://doi.org/10.1016/j.foreco.2012.05.018

Nuske R (2017) vegperiod: Determine thermal vegetation periods [software]. Zenodo. https://doi.org/https://doi.org/10.5281/zenodo.7272438

Nuske R, Staupendahl K, Albert M (2022) et.nwfva: Forest yield tables for Northwest Germany and their application. R package version 011

Oksanen J, Minchin PR (2002) Continuum theory revisited: What shape are species responses along ecological gradients? Ecological Modelling 157:119–129. https://doi.org/10.1016/S0304-3800(02)00190-4

Overbeck M, Schmidt M (2012) Modelling infestation risk of Norway spruce by Ips typographus (L.) In the Lower Saxon Harz mountains (Germany). Forest Ecology and Management 266:115–125

Palahí M, Pukkala T, Miina J, Montero G (2003) Individual-tree growth and mortality models for Scots pine (Pinus sylvestris L.) In north-east Spain. Annals of Forest Science 60:1–10

Palahí M, Tomé M, Pukkala T, et al (2004) Site index model for Pinus sylvestris in north-east Spain. Forest Ecology and Management 187:35–47

Pecchi M, Marchi M, Burton V, et al (2019) Species distribution modelling to support forest management. A literature review. Ecological Modelling 411:108817

Pickett STA (1989) Space-for-time substitution as an alternative to long-term studies. In: Likens GE (ed) Long-Term Studies in Ecology. Springer, pp 110–135

Pilkey OH, Pilkey-Jarvis L (2007) Useless arithmetic: Why environmental scientists can't predict the future. Columbia University Press

Polley H (2011) Aufnahmeanweisung für die dritte Bundeswaldinventur (BWI^3) (2011-2012. TI: Johann Heinrich von Thünen-Institut; Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz (BMELV), Bonn, Germany

Pretzsch H (2009) Forest Dynamics, Growth, and Yield. Springer, Berlin, Heidelberg

Pretzsch H, Kahn M (1995) Modelling growth of Bavarian mixed stands in a changing environment. In: Caring for the forest: Research in a changing world. Congress report. pp 234–248

Prietzel J, Falk W, Reger B, et al (2020) Half a century of Scots pine forest ecosystem monitoring reveals long-term effects of atmospheric deposition and climate change. Global Change Biology 26:5796–5815

Pya N (2019) Scam: Shape constrained additive models

Pya N, Schmidt M (2016) Incorporating shape constraints in generalized additive modelling of the height-diameter relationship for Norway spruce. Forest Ecosystems 3:1–14. https://doi.org/10.1186/s40663-016-0061-z

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Rammig A, Heinke J, Hofhansl F, et al (2018) A generic pixel-to-point comparison for simulated large-scale ecosystem properties and ground-based observations: An example from the Amazon region. Geoscientific Model Development 11:5203–5215. https://doi.org/10.5194/gmd-11-5203-2018

Riedel T, Hennig P, Kroiher F, et al (2017) Die dritte Bundeswaldinventur (BWI 2012). Inventur-und Auswertungsmethoden [The third National Forest Inventroy (BWI 2012) Inventory and analysis methods] Johann Heinrich von Thünen-Institut Braunschweig

Rita A, Borghetti M (2019) Linkage of forest productivity to tree diversity under two different bioclimatic regimes in Italy. Science of the total environment 687:1065–1072

Romijn E, Lantican CB, Herold M, et al (2015) Assessing change in national forest monitoring capacities of 99 tropical countries. Forest Ecology and Management 352:109–123

Rustad LE (2008) The response of terrestrial ecosystems to global climate change: Towards an integrated approach. Science of the total environment 404:222–235

Schelhaas M-J, Nabuurs G-J, Schuck A (2003) Natural disturbances in the European forests in the 19th and 20th centuries. Global Change Biology 9:1620–1633. https://doi.org/10.1046/j.1365-2486.2003.00684.x

Schick J, Albert M, Schmidt M (2023) A new approach for modeling stand height development of German forests under climate change. In review

Schiermeier Q (2018) Droughts, heatwaves and floods: How to tell when climate change is to blame. Nature 560:20–22. https://doi.org/10.1038/d41586-018-05849-9

Schmidt M (2020) Standortsensitive und kalibrierbare Bonitätsfächer: Wachstumspotenziale wichtiger Baumarten unter Klimawandel. Allgemeine Forst- und Jagdzeitung 190:136–160

Schmidt M, Breidenbach J, Astrup R (2018) Longitudinal height-diameter curves for Norway spruce, Scots pine and silver birch in Norway based on shape constraint additive regression models. Forest Ecosystems 5: https://doi.org/10.1186/s40663-017-0125-8

Schmidt M, Hanewinkel M, Kändler G, et al (2010) An inventory-based approach for modeling single-tree storm damage–experiences with the winter storm of 1999 in southwestern Germany. Canadian Journal of Forest Research 40:1636–1652

Schober R (1975) Ertragstafeln wichtiger Baumarten. J.D. Sauerländer Verlag, Frankfurt a. M.

Schulla J, Jasper K (2007) Model description WaSiM-ETH. Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich, Switzerland

Seynave I, Gégout J-C, Hervé J-C, et al (2005) Picea abies site index prediction by environmental factors and understorey vegetation: A two-scale approach based on survey databases. Canadian Journal of Forest Research 35:1669–1678

Seynave I, Gégout J-C, Hervé J-C, Dhôte J-F (2008) Is the spatial distribution of European beech (Fagus sylvatica L.) limited by its potential height growth? Journal of Biogeography 35:1851–1862

Sharma RP, Brunner A, Eid T (2012) Site index prediction from site and climate variables

for Norway spruce and Scots pine in Norway. Scandinavian Journal of Forest Research 27:619–636. https://doi.org/10.1080/02827581.2012.685749

Sharma RP, Vacek Z, Vacek S (2016) Modeling individual tree height to diameter ratio for Norway spruce and European beech in Czech Republic. Trees 30:1969–1982

Skovsgaard JP, Vanclay JK (2008) Forest site productivity: A review of the evolution of dendrometric concepts for even-aged stands. Forestry 81:13–31. https://doi.org/10.1093/forestry/cpm041

Stankova TV, Ferezliev A, Dimitrov DN, et al (2022) A parsimonious generalised height-diameter model for Scots pine plantations in Bulgaria: A pragmatic approach. South-east European forestry: SEEFOR 13:37–51

Staupendahl K, Schmidt M (2016) Kalibrierung der Ertragstafeln für die Hauptbaumarten in Hessen mithilfe der Daten der Bundeswaldinventur. Allgemeine Forst und Jagdzeitung 187:197–216

Talkner U, Meiwes KJ, Potočić N, et al (2015) Phosphorus nutrition of beech (Fagus sylvatica L.) is decreasing in Europe. Annals of Forest Science 72:919–928

Tinkham WT, Mahoney PR, Hudak AT, et al (2018) Applications of the united states forest inventory and analysis dataset: A review and future directions. Canadian Journal of Forest Research 48:1251–1268

Tveito O, Førland E, Alexandersson H, et al (2001) DNMI report: Nordic climate maps. Norwegian Meteorological Institute, Oslo, Norway

Udali A, Andrighetto N, Grigolato S, Gatto P (2021) Economic impacts of forest storms—taking stock of after-Vaia situation of local roundwood markets in northeastern Italy. Forests 12:414

Uhl E, Hilmers T, Pretzsch H (2021) From acid rain to low precipitation: The role reversal of Norway spruce, silver fir, and European beech in a selection mountain forest and its implications for forest management. Forests 12: https://doi.org/10.3390/f12070894

Ung C-H, Bernier PY, Raulier F, et al (2001) Biophysical site indices for shade tolerant and intolerant boreal species. Forest Science 47:83–95

Van Oldenborgh GJ, Philip S, Kew S, et al (2019) Human contribution to the record-breaking June 2019 heat wave in France. World Weather Attribution (WWA): Amsterdam, Netherlands 32

Van Sundert K, Radujković D, Cools N, et al (2019) Towards comparable assessment of the soil nutrient status across scales—review and development of nutrient metrics. Global Change Biology

Vauhkonen J, Berger A, Gschwantner T, et al (2019) Harmonised projections of future forest resources in Europe. Annals of Forest Science 76:1–12

Vidal C, Bélouard T, Hervé J-C, et al (2007) A new flexible forest inventory in France. In: McRoberts RE, Reams GA, Van Deusen PC, McWilliams WH (eds) Proceedings of the seventh annual forest inventory and analysis symposium; 3-6. US Department of Agriculture, Forest Service: 67-73. Portland, ME. Gen. Tech. Rep. WO-77., Washington DC

Vogel MM, Zscheischler J, Wartenburger R, et al (2019) Concurrent 2018 hot extremes across northern hemisphere due to human-induced climate change. Earth's future 7:692–703

Wang T, Hamann A, Spittlehouse D, Carroll C (2016) Locally downscaled and spatially customizable climate data for historical and future periods for North America. PLOS ONE 11:e0156720. https://doi.org/10.1371/journal.pone.0156720

Watanabe M, Suzuki T, O'ishi R, et al (2010) Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. Journal of Climate 23:6312–6335

Wessolek G, Kaupenjohann M, Renger M (2009) Bodenphysikalische Kennwerte und Berechnungsverfahren für die Praxis. In: Rote Reihe. Institut für Ökologie, Fachgebiet Bodenkunde, Standortkunde und Bodenschutz, Berlin

Wilpert K von, Puhlmann H, Zirlewagen D, et al (2017) Schätzung von Bodendaten an den Messorten der Bundeswaldinventur aus Umweltmonitoring- und Standortskundedaten im Vergleich. In: Spellmann H, Kehr I, Gaertner U (eds) Waldböden: Nutzung und Schutz –

Beiträge aus der Nordwestdeutschen Forstlichen Versuchsanstalt, 17th edn. Göttingen University Press, pp 137–150

Wilpert Kv, Zirlewagen D, Puhlmann H (2016) Regionalisierung von Bodendaten für Deutschland–Datenbasis, Zielgrößen und Modellgüte am Beispiel zweier Testgebiete. Waldökologie, Landschaftsforschung und Naturschutz 16:109–120

Wood S (2019) Summary.gam manual page from mgcv package

Wood SN (2017) Generalized Additive Models: An introduction with R, 2nd edn. Chapman; Hall/CRC

Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73:3–36

Wood S, Scheipl F (2020) gamm4: Generalized additive mixed models using 'mgcv' and 'lme4'

World Weather Attribution (2018) Heatwave in Northern Europe, summer 2018. https://www.worldweatherattribution.org/attribution-of-the-2018-heat-in-northern-europe/. Accessed 7 Apr 2020

Yener I, Altun L, et al (2018) Predicting site index for oriental spruce (Picea orientalis L.(Link)) using ecological factors in the eastern Black Sea, Turkey. Fresenius Environmental Bulletin 27:3107–3116

Yiou P, Cattiaux J, Faranda D, et al (2020) Analyses of the Northern European summer heatwave of 2018. Bulletin of the American Meteorological Society 101:S35–S40

Yue C, Kahle H-P, von Wilpert K, Kohnle U (2016) A dynamic environment-sensitive site index model for the prediction of site productivity potential under climate change. Ecological Modelling 337:48–62

Zeng W, Tomppo E, Healey SP, Gadow KV (2015) The national forest inventory in China: History-results-international context. Forest Ecosystems 2:1–16

Zhang K, Almeida Castanho AD, Galbraith DR, et al (2015) The fate of Amazonian ecosystems over the coming century arising from changes in climate, atmospheric CO2, and land use. Global Change Biology 21:2569–2587

Zimmermann L, Rötzer T, Hera U, et al (2007) Konzept für die Erstellung neuer hochaufgelöster Klimakarten für die Wälder Bayerns als Bestandteil eines forstlichen Standortinformationssystems. In: Matzarakis A, Mayer H (eds) Proceedings zur 6. Fachtagung BIOMET des Fachausschusses Biometeorologie der Deutschen Meteorologischen Gesellschaft e.V. Meteorologisches Institut der Universität Freiburg, Freiburg, Germany, pp 152–159

Zirlewagen D, Wilpert K von (2010) Upscaling of environmental information: Support of land-use management decisions by spatio-temporal regionalization approaches. Environmental management 46:878–893

Table of abbreviations

`aw`	dominant age (age corresponding to Weise's dominant height (`hw`)
`awc`	available water capacity
`CP`	carbon-phosphorous ratio
`dbh`	diameter at breast height (1.3 m)
`dw`	mean diameter of the 20% strongest trees on a plot, corresponding to Weise's dominant height (`hw`)
h-a	the height-age relation
h-d	the height-diameter relation
`hw`	Weise's dominant height, height corresponding to the 20% strongest trees on a plot
`ph`	pH value
`P`	precipitation covariate; in the models in section 2 this denotes the precipitation sum for the warmest quarter, in the models in section 3 it denotes the precipitation sum for the vegetation period
`qmD`	the quadratic mean diameter
SI	site index (in projections of section 2 and 3 defined as the dominant height at age 100)
SFT	Space for Time
`T`	temperature covariate; in the models in section 2 this is the mean temperature for the warmest quarter, in section 3 it denotes the temperature sum for the vegetation period

A Appendix

A.1 Parameters for the h-d model fit

Quantiles (5% and 95%) for all parameters used for model selection in study I, by species. Note that climate parameters (`tmn_*`, `tmax_`, `tmin_*` and `psum_*`) and available water capacity (`awc`) are based on the complete dataset including the French NFI, while all other parameters are based on BioSoil and NFSI. See Table 2.1 for further information on the included surveys and Table 2.3 for units of the respective parameters.

Species	Quantile	tmax_wm	tmin_cm	tmn_wq	psum
Beech	5	19.3	-6.1	13.2	677
Beech	95	24.9	0.5	18.2	1259
Oak	5	21.3	-3.1	15.8	644
Oak	95	26.1	2.2	19.2	1049
Pine	5	20.1	-12.4	13.8	548
Pine	95	25.4	0.2	18.5	1029
Spruce	5	18	-11.6	12.1	619
Spruce	95	24.1	-1.4	17.3	1401
Species	Quantile	psum_wq	t01	t07	t_5to9
Beech	5	164	-2.8	14.1	11.84
Beech	95	320	3.7	19	16.78
Oak	5	151	-0.2	16.4	14.56
Oak	95	262	5.3	19.9	17.98
Pine	5	157	-8.7	14.9	11.68
Pine	95	254	3.5	19.4	17.08
Spruce	5	196	-7.8	13.1	10.74
Spruce	95	368	1.7	18.2	15.96
Species	Quantile	t_678	tyr	p_5to9	p_678
Beech	5	13.233	5.733	281	160
Beech	95	18.167	10.825	525	320
Oak	5	15.733	8.408	266	151
Oak	95	19.267	12.333	435	261
Pine	5	13.8	3.125	274	155

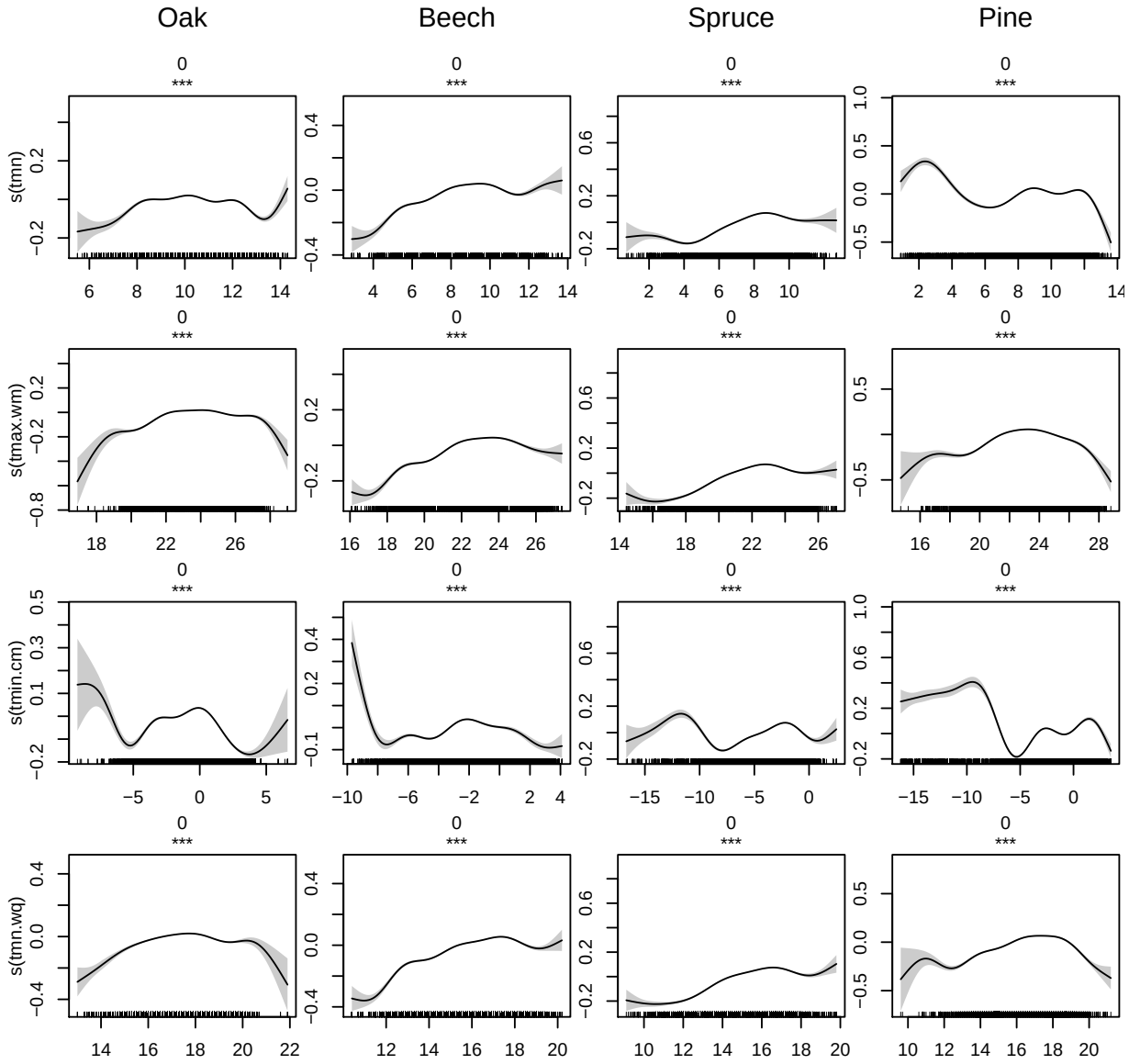
Pine	95	18.5	11.1	436	256
Spruce	5	12.167	3.183	307	196
Spruce	95	17.333	9.708	582	368
Species	Quantile	pyr	awc	aircap_min	fbv_org
Beech	5	677	20	4	4.785
Beech	95	1259	222	24	127.457
Oak	5	644	26.4	2	3.5
Oak	95	1049	231.5	25	132.8
Pine	5	548	24	7	16.7
Pine	95	1029	192	30	155.072
Spruce	5	619	26.414	3	10.22
Spruce	95	1401	216	25	170.7
Species	Quantile	oc_org	oc_minst_010	oc_minst_040	oc_minst_soildepth
Beech	5	0	16.484	34.025	44.561
Beech	95	34.703	72.787	168.25	229.61
Oak	5	1.046	16.685	33.701	39.69
Oak	95	35.991	56.09	131.439	222.737
Pine	5	4.944	6.388	15.986	20.574
Pine	95	50.3	59.504	136.775	182.85
Spruce	5	0	7.692	21.976	27.972
Spruce	95	53.835	63.578	147.461	213.649
Species	Quantile	ton_org	ton_minst_010	ton_minst_040	ton_minst_soildepth
Beech	5	0	0.896	2.043	2.753
Beech	95	1.523	4.94	13.106	17.363
Oak	5	0.014	0.944	1.546	2.301
Oak	95	1.558	3.717	8.317	14.741
Pine	5	0.163	0.285	0.828	1.192
Pine	95	1.813	2.62	6.327	9.673
Spruce	5	0	0.342	1.574	2.472
Spruce	95	2.238	3.577	8.884	13.14
Species	Quantile	p_org	p_minst_010	p_minst_040	p_minst_soildepth
Beech	5	2.592	104.575	353.565	620.885

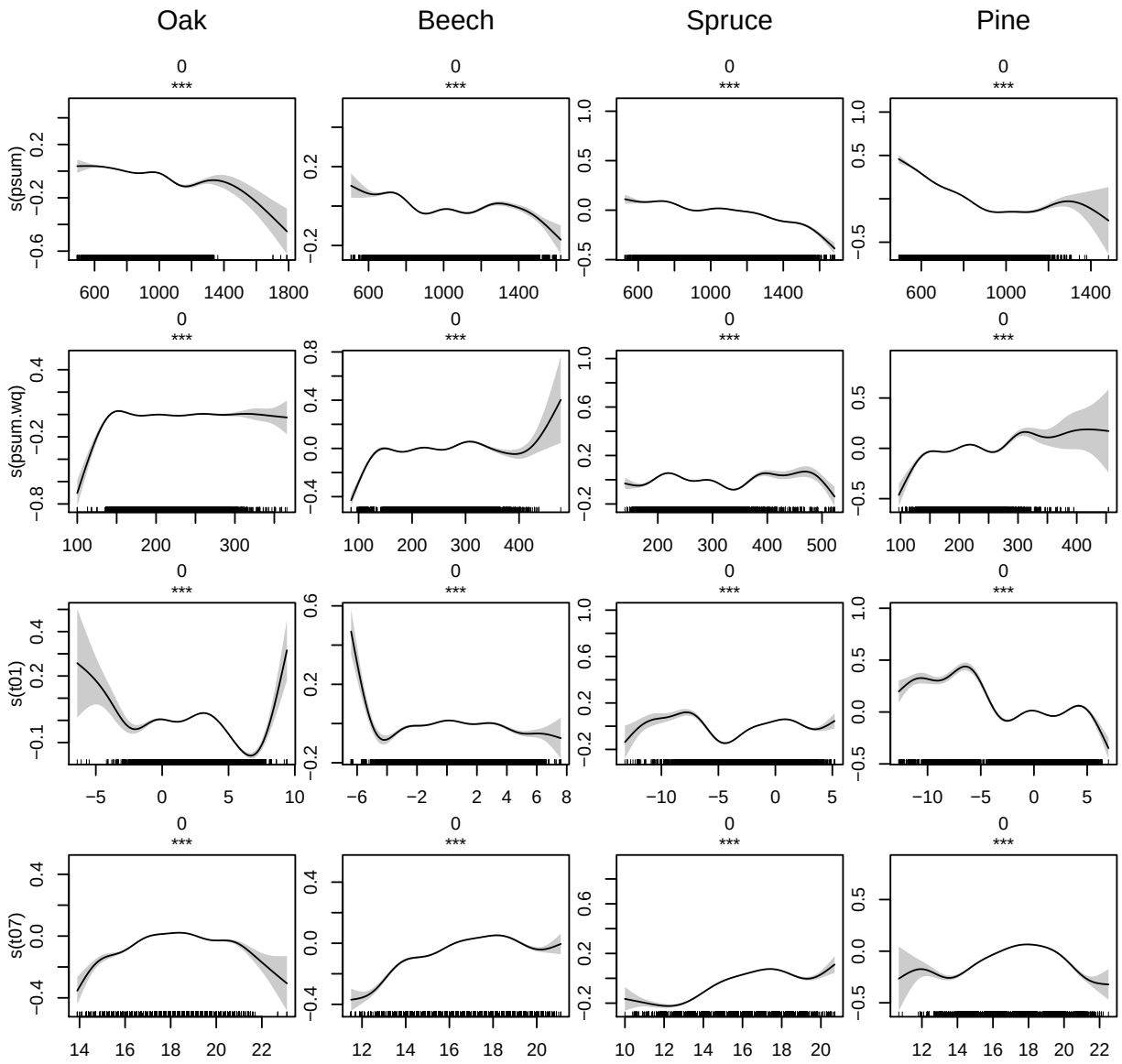
Beech	95	87.286	834.126	3391.583	7243.175
Oak	5	2.349	108.405	335.553	678.577
Oak	95	108.963	826.884	3720.07	8971.4
Pine	5	8.641	48	261.586	492.44
Pine	95	102.295	462.331	2156.949	4282.119
Spruce	5	0	54.789	408.232	693.341
Spruce	95	133.813	689.287	3134.091	8323.349
Species	Quantile	cn_org	cn010	cn020	cp_org
Beech	5	18.49	12.04	10.81	249.045
Beech	95	33.05	26.205	26.86	765.029
Oak	5	16.763	11.83	10.842	191.413
Oak	95	32.42	26.055	27.192	711.914
Pine	5	21.57	14.275	13	296.503
Pine	95	38.65	38.57	38.05	860.303
Spruce	5	19.21	12.975	11.482	201.883
Spruce	95	33.041	28.58	28.177	683.757
Species	Quantile	cp010	cp020	ph_org	ph010
Beech	5	43.48	33.289	3.25	3.3
Beech	95	277.86	227.13	6.094	7
Oak	5	29.882	21.742	3.255	3.217
Oak	95	267.673	191.33	5.64	6.46
Pine	5	33.792	27.795	2.85	3.07
Pine	95	455.387	331.09	5.104	5.64
Spruce	5	36.102	27.198	2.94	3.08
Spruce	95	347.663	295.631	5.2	6.3
Species	Quantile	ph020	ph040	ph4080	basesat_org
Beech	5	3.487	3.676	3.795	39.96
Beech	95	7.135	7.3	7.565	99.11
Oak	5	3.41	3.57	3.74	41.331
Oak	95	6.683	6.884	7.325	98.04
Pine	5	3.285	3.501	3.9	29.98
Pine	95	5.73	6.152	6.57	93.102

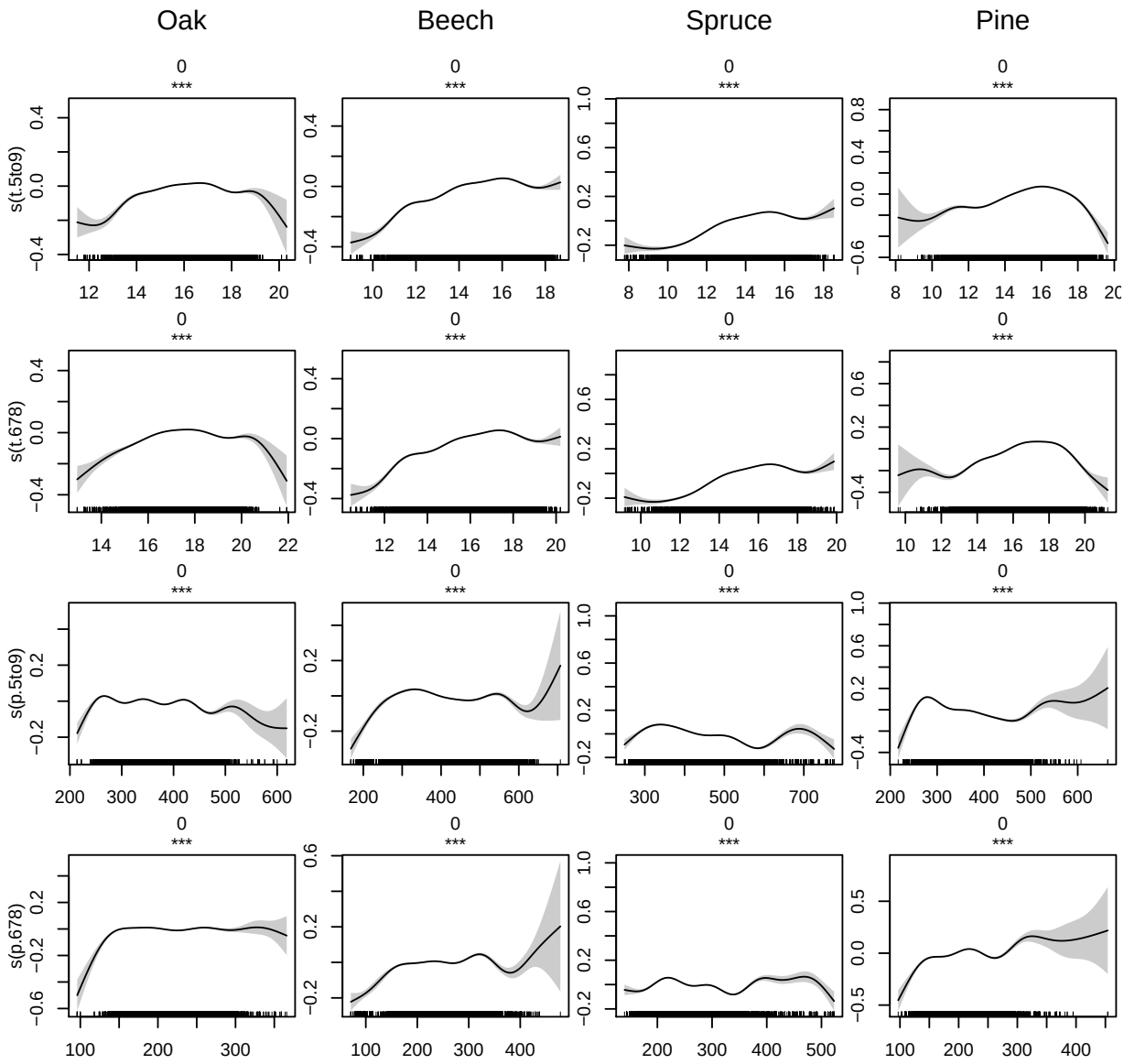
Spruce	5	3.343	3.525	3.835	26.68
Spruce	95	6.685	6.857	7.375	97.23

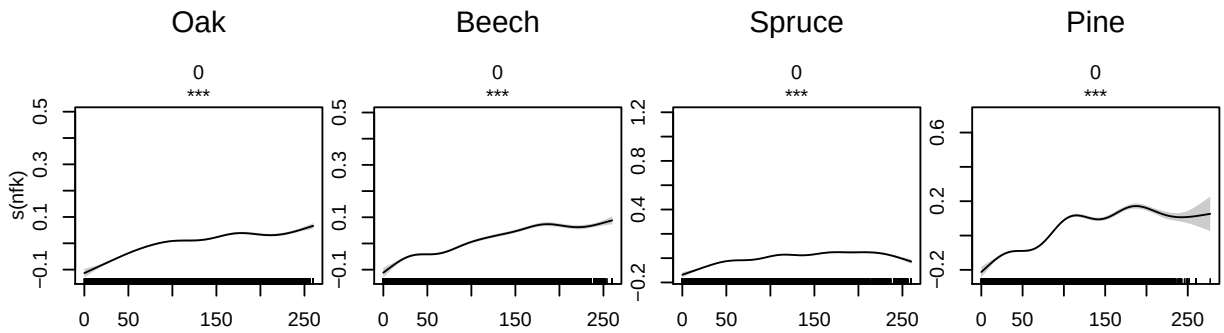
A.2 Splines for the h-d model climate covariates

Splines for climatic covariates for oak, beech, spruce and pine in part I, included in the first model step with p-values and significance codes as given in R (R Core Team 2020). The codes symbolize the following significance levels: "`***`" for $p < 0.001$, "`**`" for $p < 0.\ 01$, "`*`" for $p < 0.05$, "`.`" for $p < 0.1$ and "`n.sig`" for $p > 0.1$. All models have the form of the minimal model Equation 2.3 with the climatic parameter added as a single, unconstrained spline.



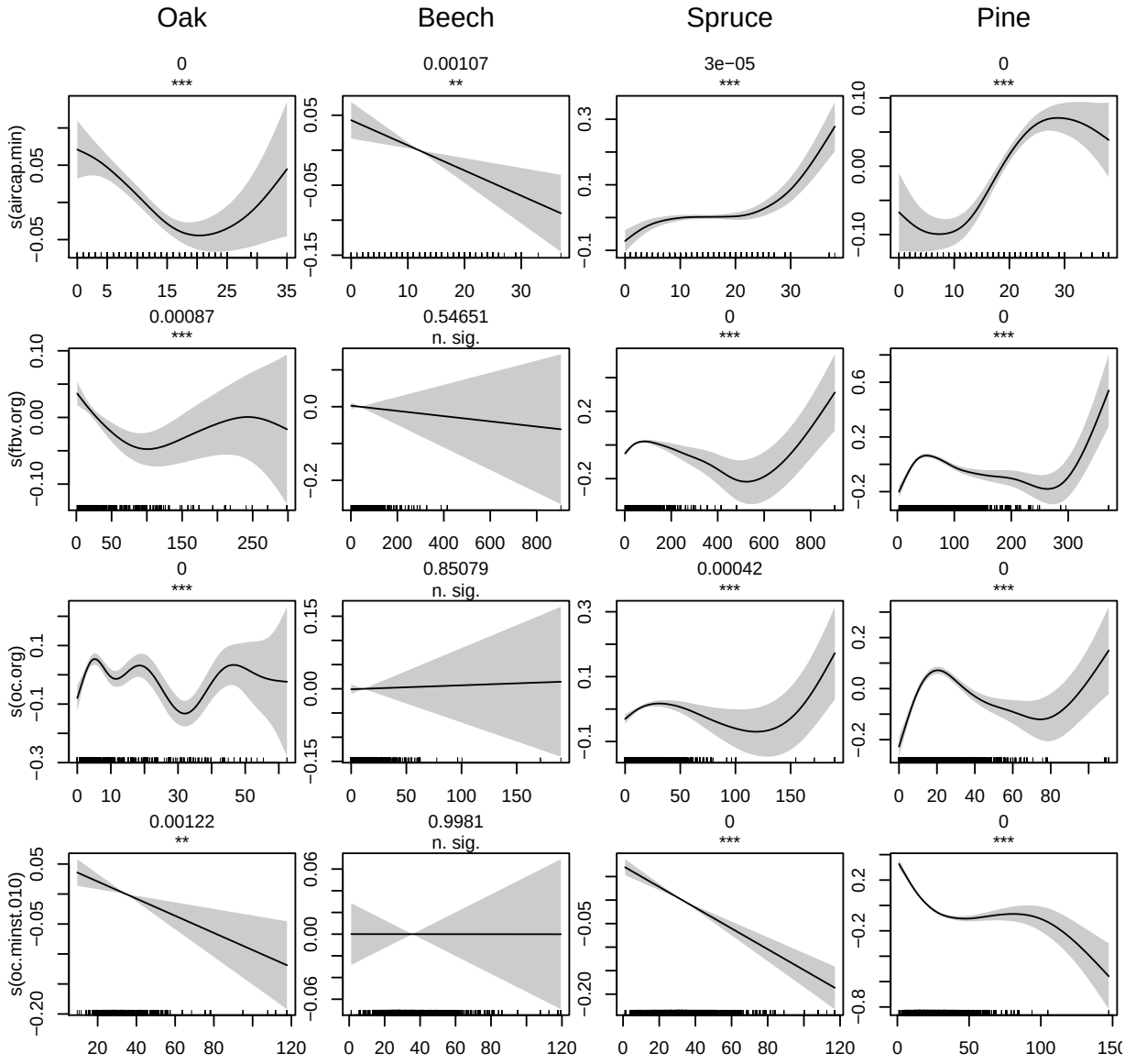

Oak
Beech
Spruce
Pine
s(psum)
s(psum.wq)
s(t01)
s(t07)
0
***

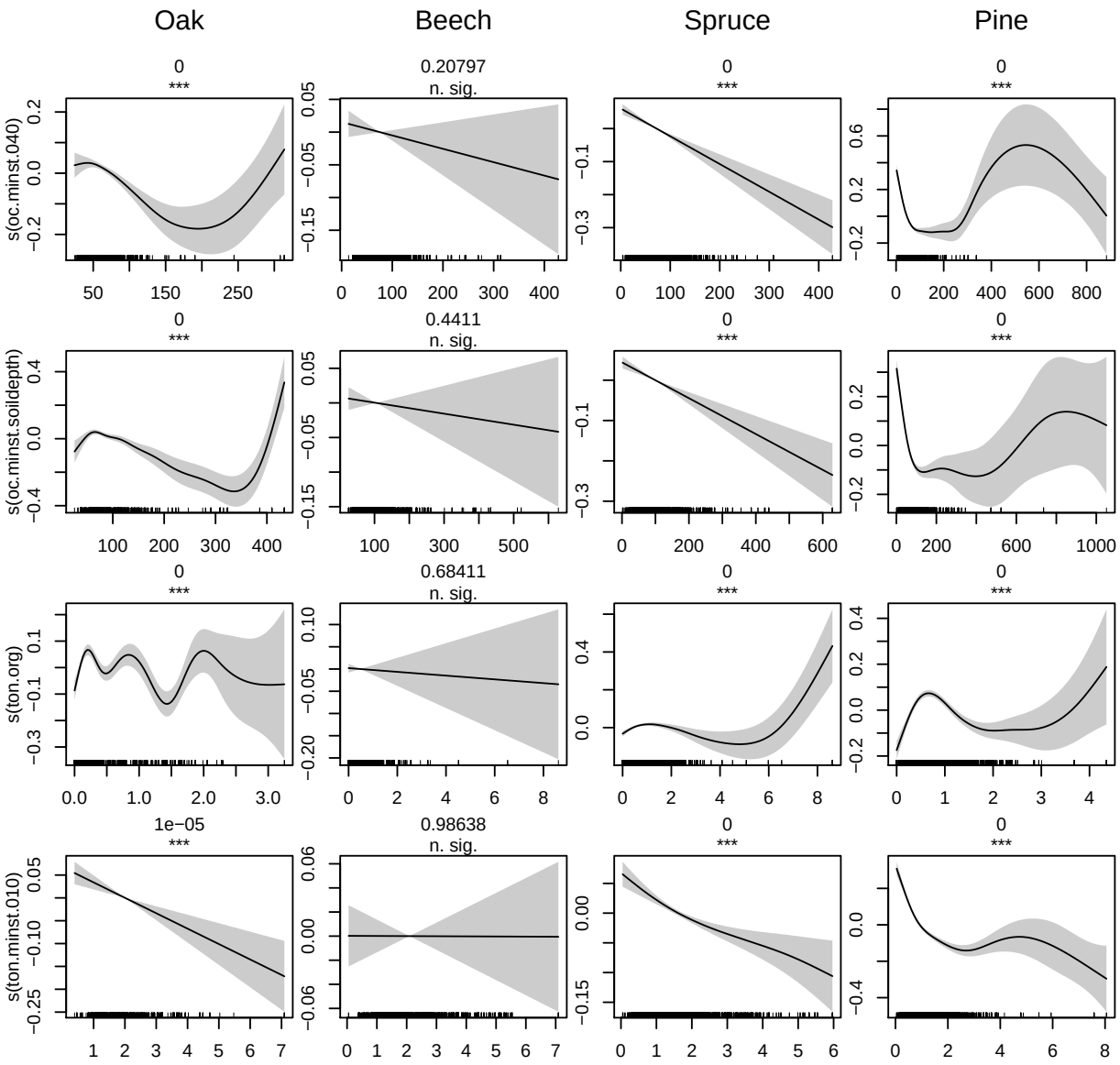

Oak
Beech
Spruce
Pine
s(t.5to9)
s(t.678)
s(p.5to9)
s(p.678)
0
***

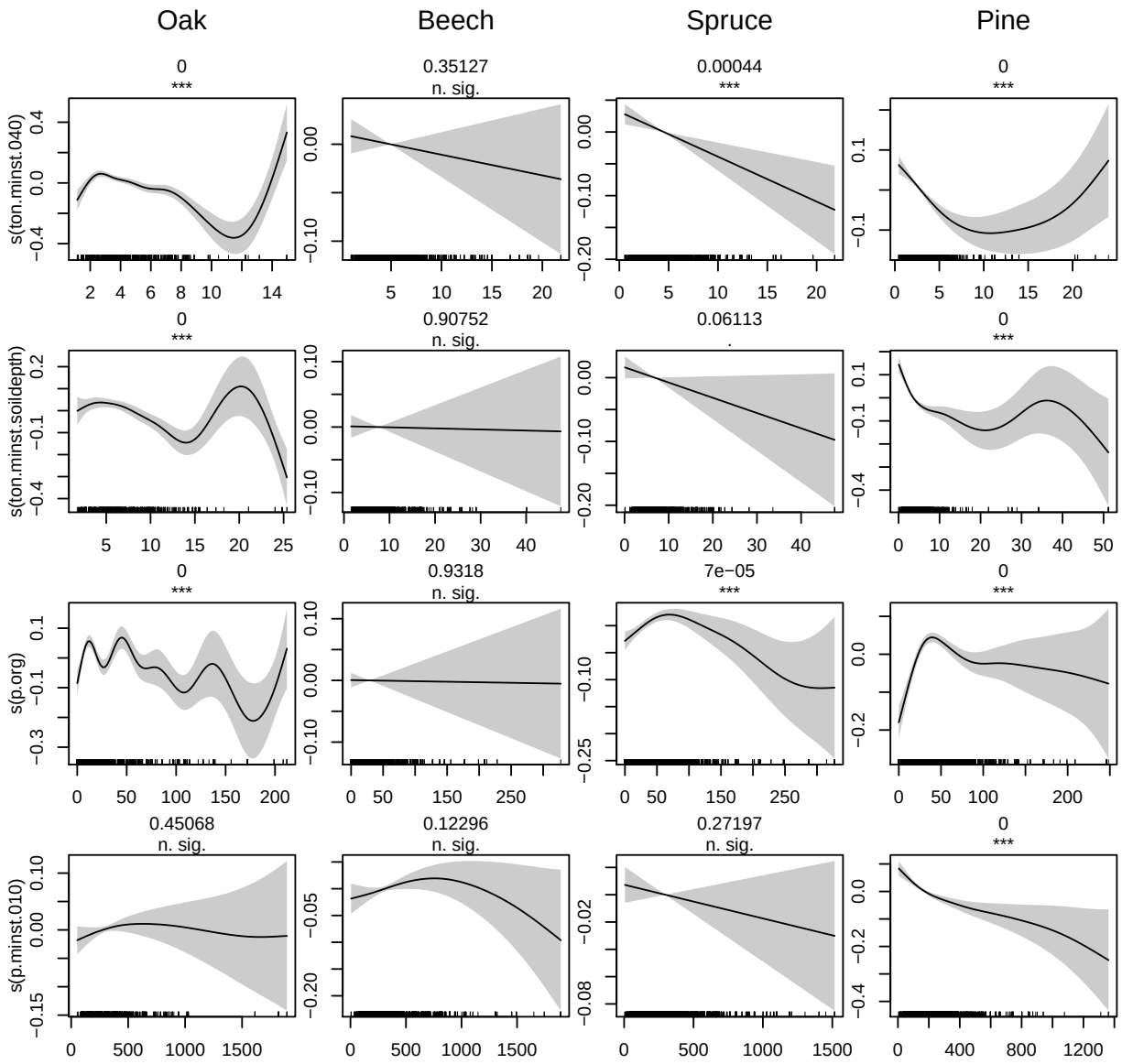

Oak
Beech
Spruce
Pine
0
***
s(nfk)
−0.1
0.1
0.3
0.5
−0.2
0.4
0.8
1.2
0.2
0.6
0
50
150
250

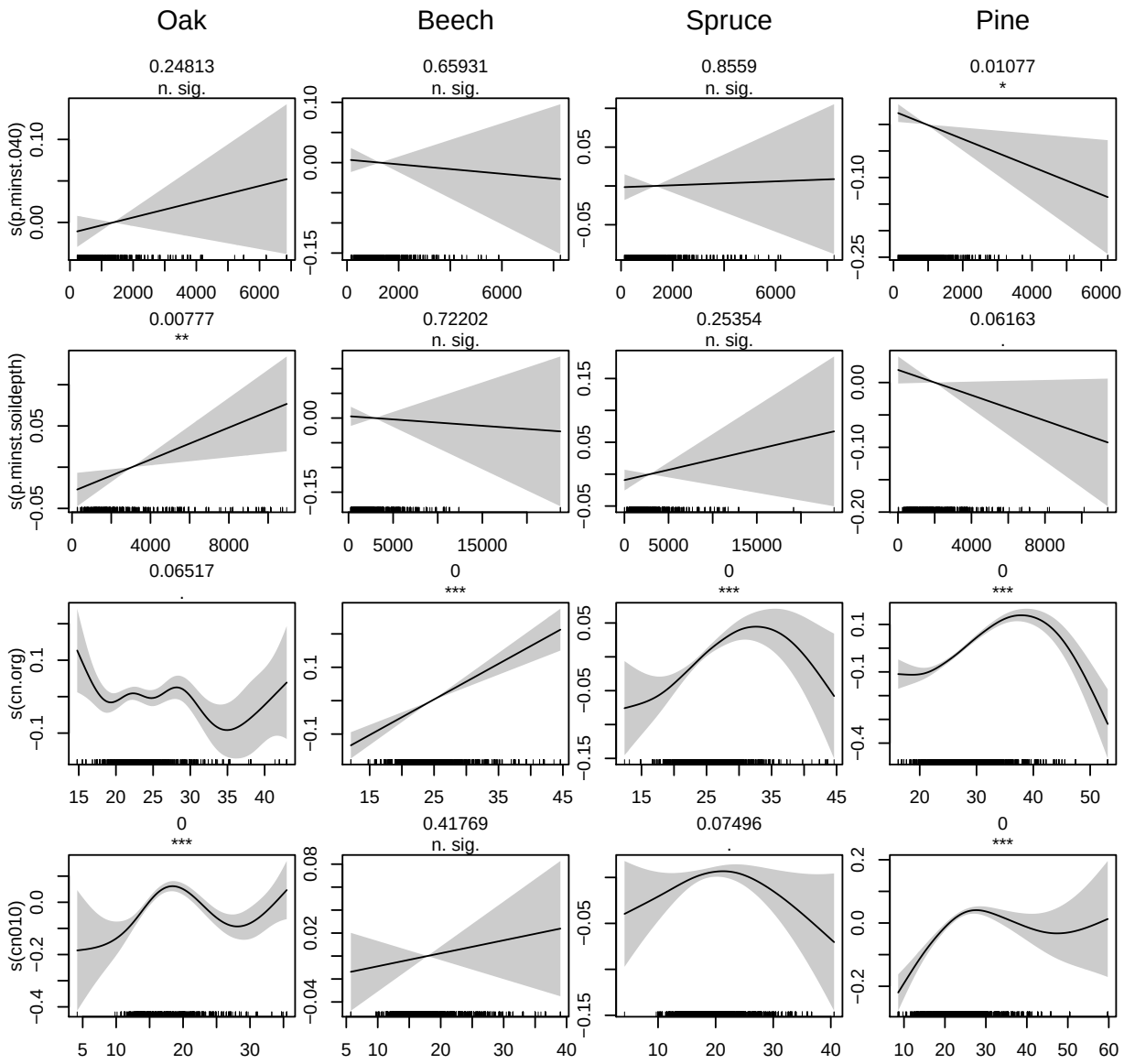
A.3 Splines for the h-d model soil covariates

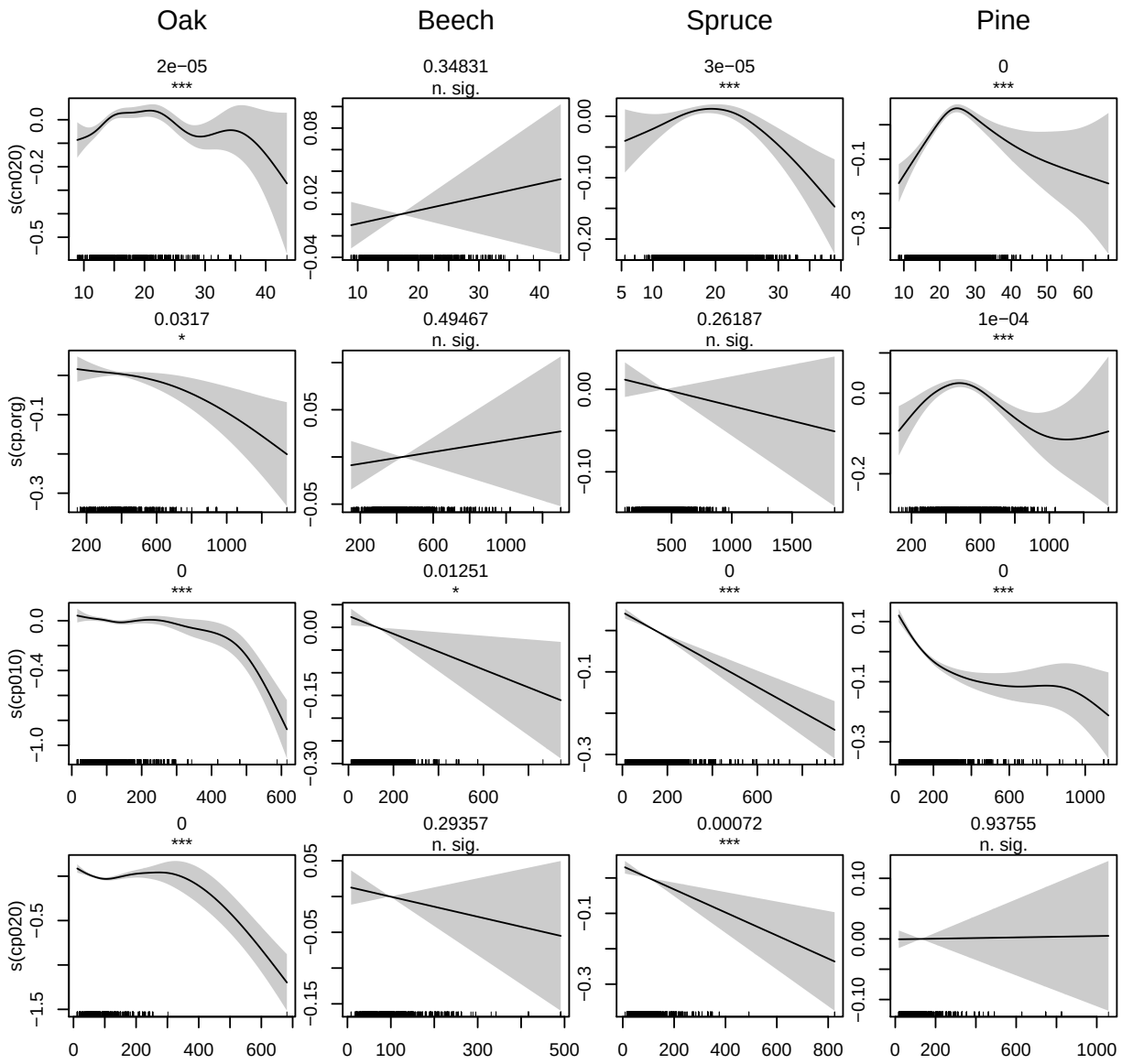
Splines for BioSoil parameters for oak, beech, spruce and pine included in the second model step of study I with p-values and significance codes as given in R (R Core Team 2020). The codes symbolize the following significance levels: "`***`" for $p < 0.001$, "`**`" for $p < 0.\ 01$, "`*`" for $p < 0.05$, "`.`" for $p < 0.1$ and "`n.sig`" for $p > 0.1$. All models have the form of the minimal model of the second model step Equation 2.6 with each soil parameter added as a single, unconstrained spline.

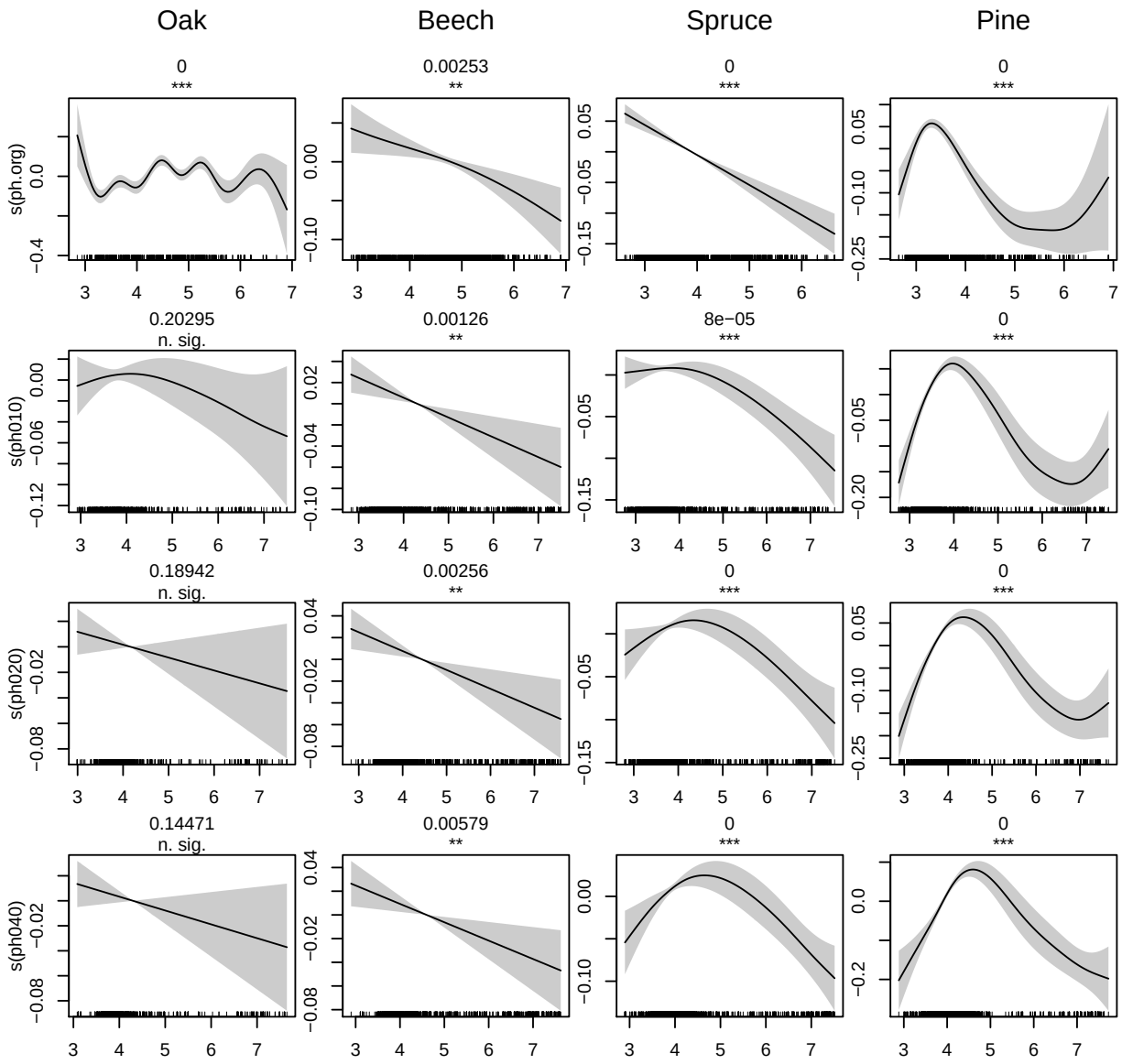


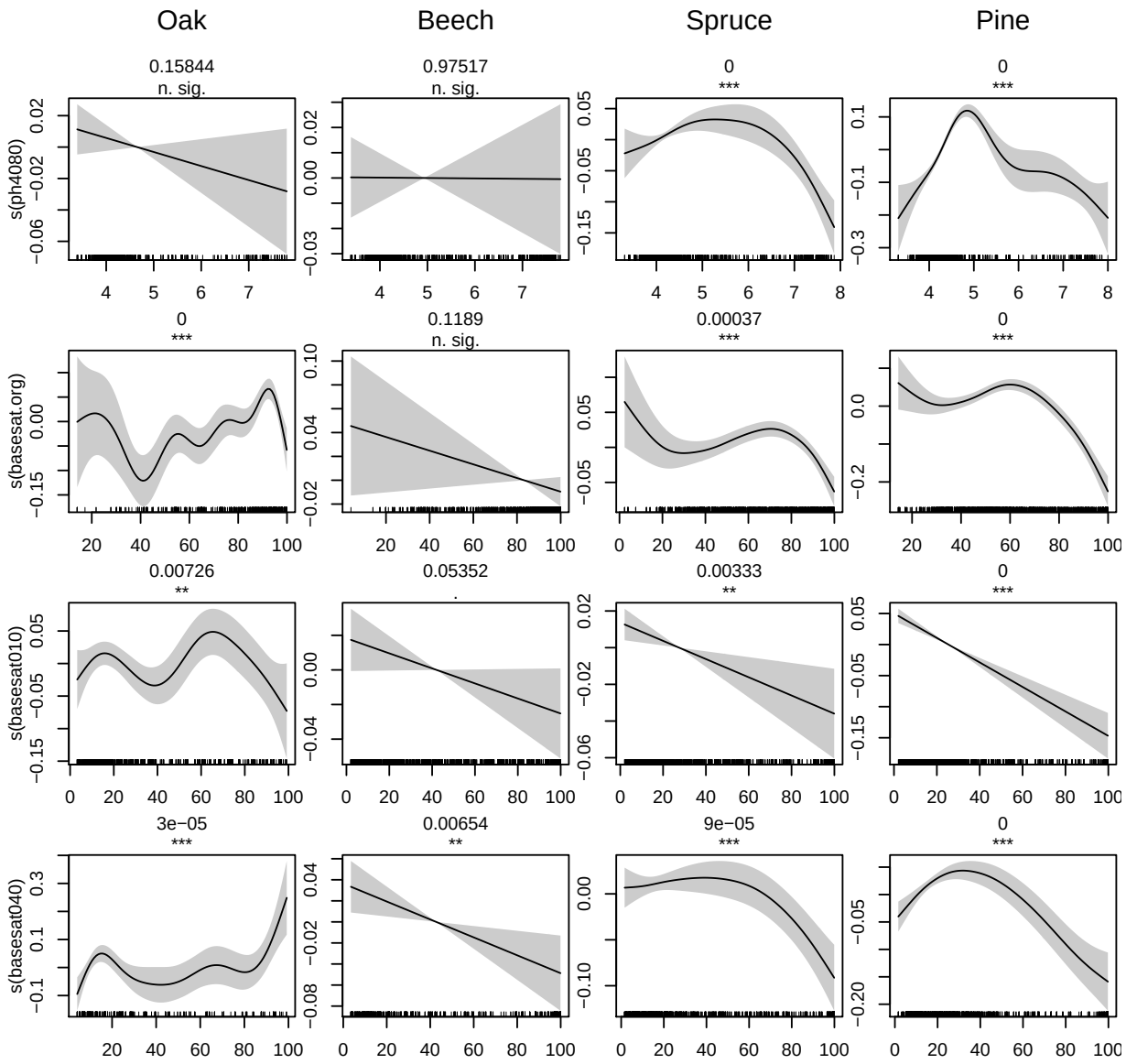

Oak
Beech
Spruce
Pine
0
***
0.20797
n. sig.
0
***
0
***
s(oc.minst.040)
0
***
0.4411
n. sig.
0
***
0
***
s(oc.minst.soildepth)
0
***
0.68411
n. sig.
0
***
0
***
s(ton.org)
1e−05
***
0.98638
n. sig.
0
***
0
***
s(ton.minst.010)

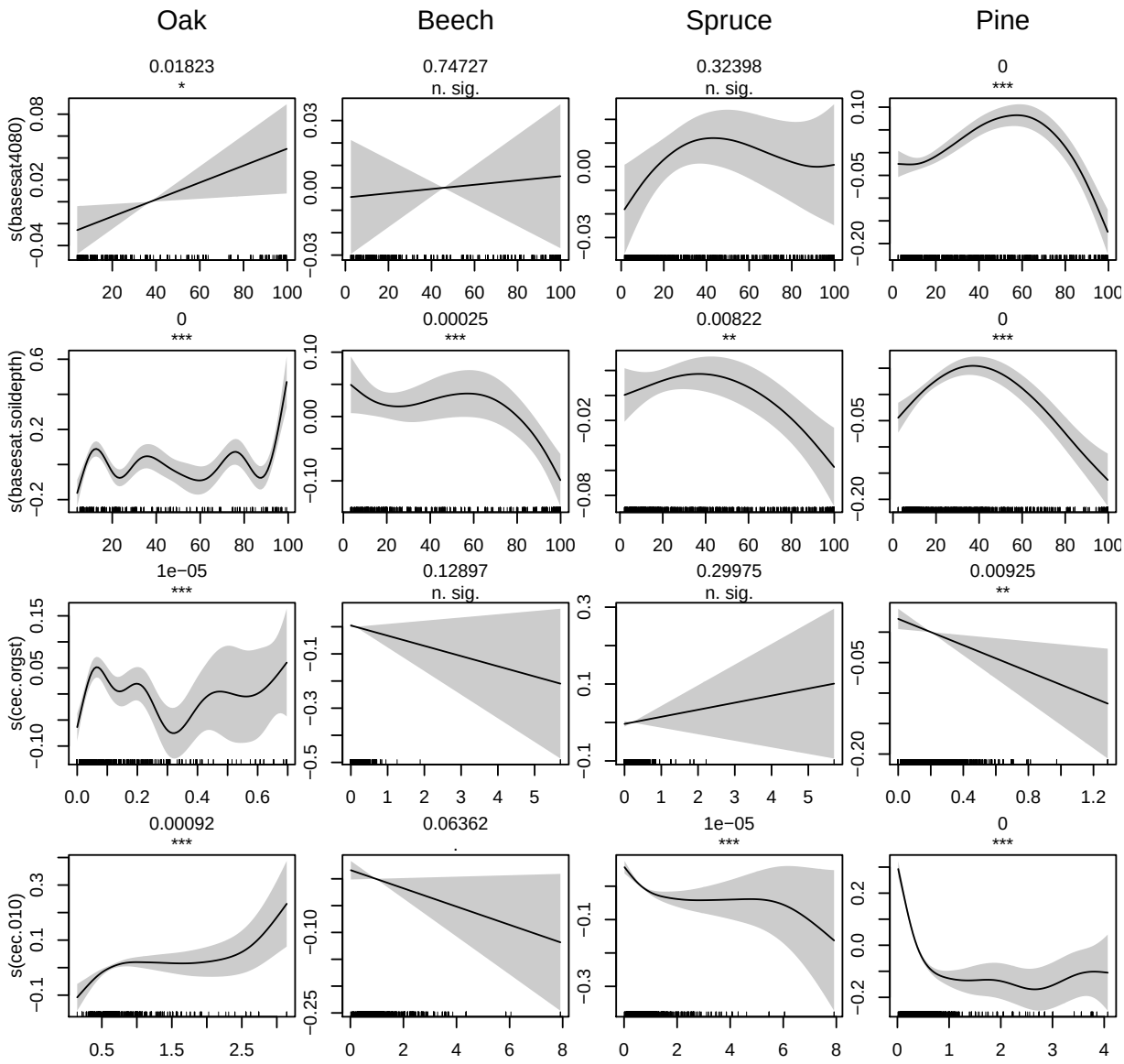

Oak
Beech
Spruce
Pine
0
***
0.35127
n. sig.
0.00044
***
0
***
s(ton.minst.040)
0
***
0.90752
n. sig.
0.06113
.
0
***
s(ton.minst.soildepth)
0
***
0.9318
n. sig.
7e−05
***
0
***
s(p.org)
0.45068
n. sig.
0.12296
n. sig.
0.27197
n. sig.
0
***
s(p.minst.010)

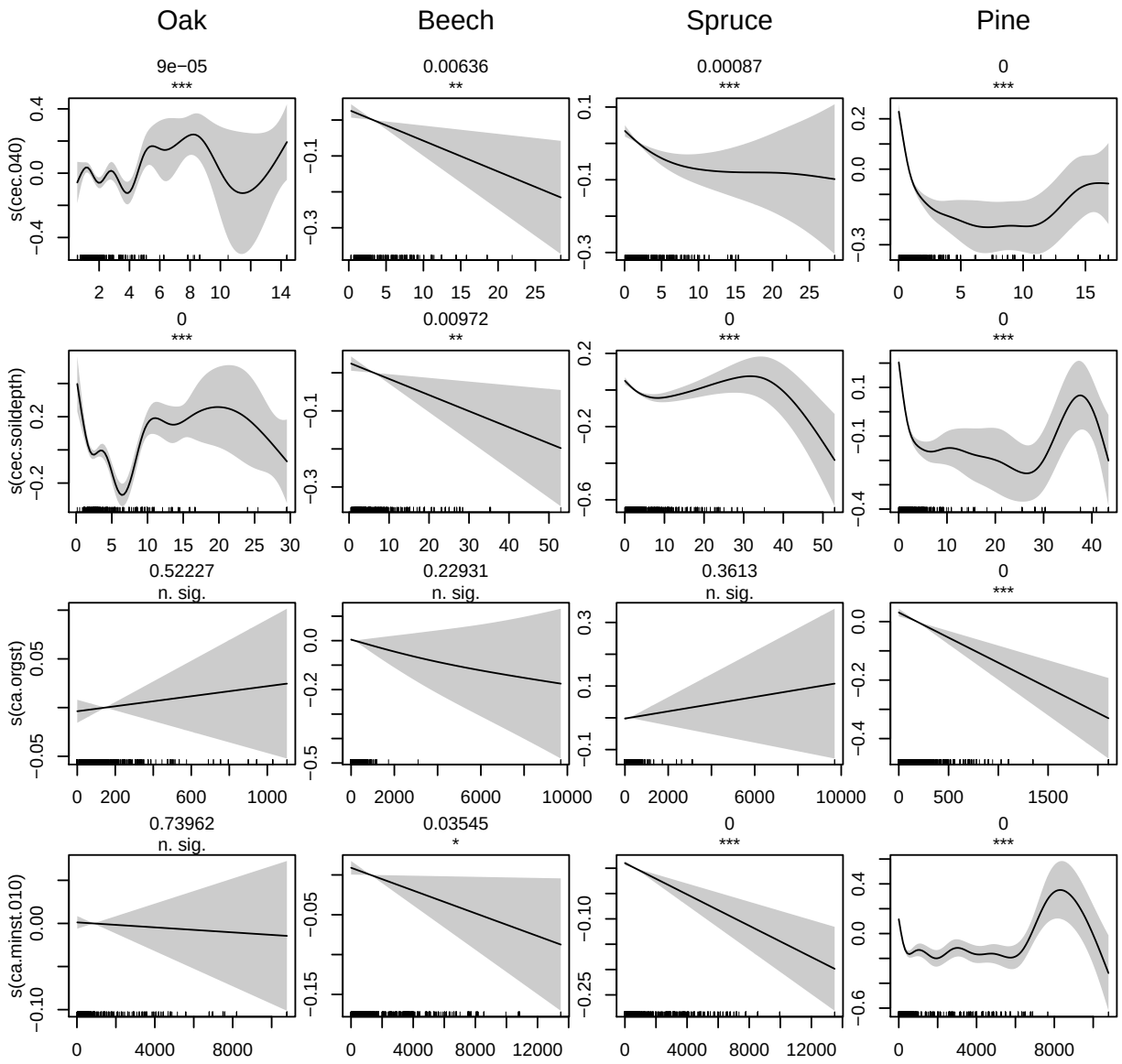

Oak
Beech
Spruce
Pine
0.24813
n. sig.
0.65931
n. sig.
0.8559
n. sig.
0.01077
*
s(p.minst.040)
0.00777
**
0.72202
n. sig.
0.25354
n. sig.
0.06163
.
s(p.minst.soildepth)
0.06517
.
0
***
0
***
0
***
s(cn.org)
0
***
0.41769
n. sig.
0.07496
.
0
***
s(cn010)

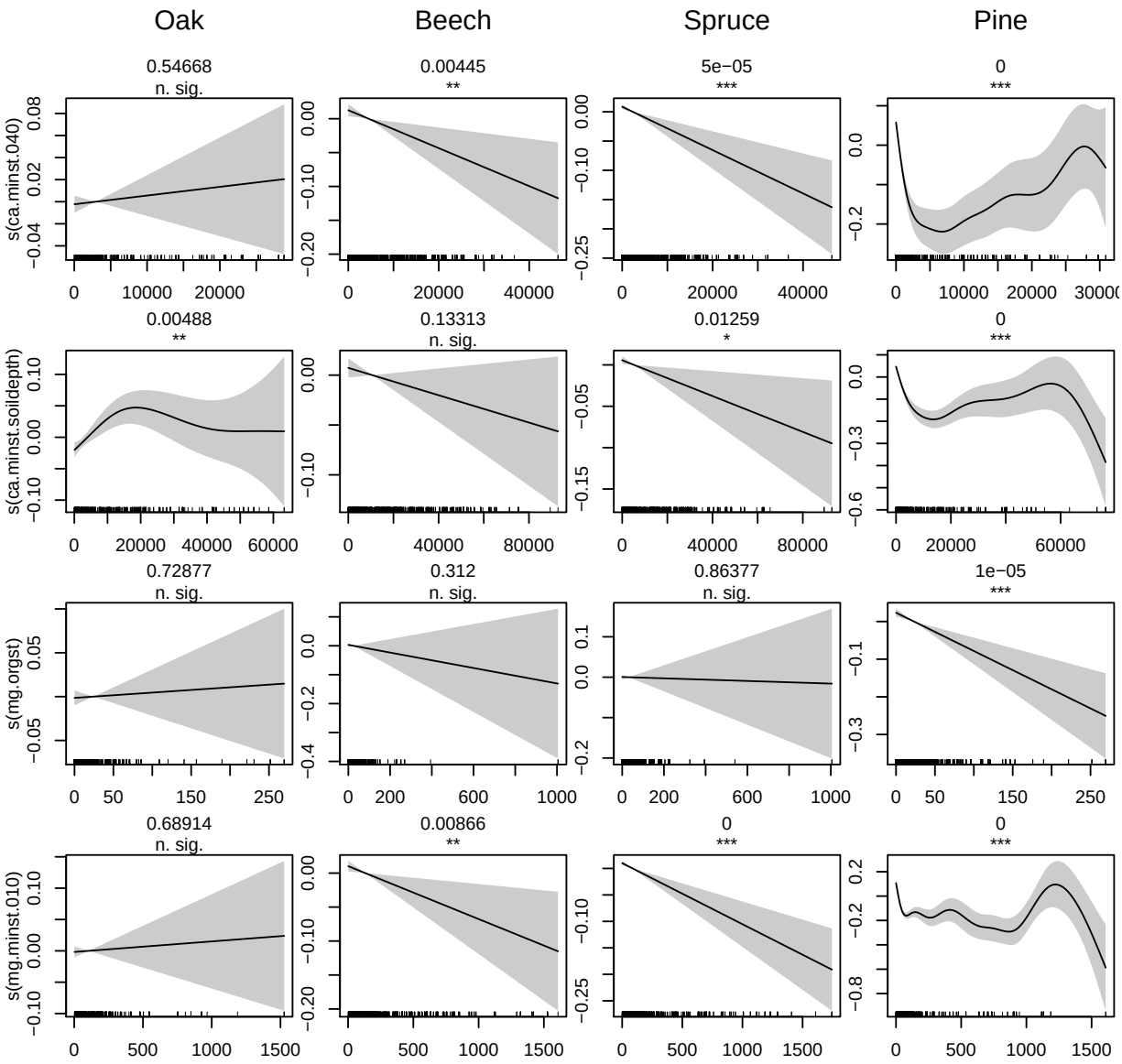

Oak
Beech
Spruce
Pine
2e−05
***
0.34831
n. sig.
3e−05
***
0
***
s(cn020)
0.0317
*
0.49467
n. sig.
0.26187
n. sig.
1e−04
***
s(cp.org)
0
***
0.01251
*
0
***
0
***
s(cp010)
0
***
0.29357
n. sig.
0.00072
***
0.93755
n. sig.
s(cp020)

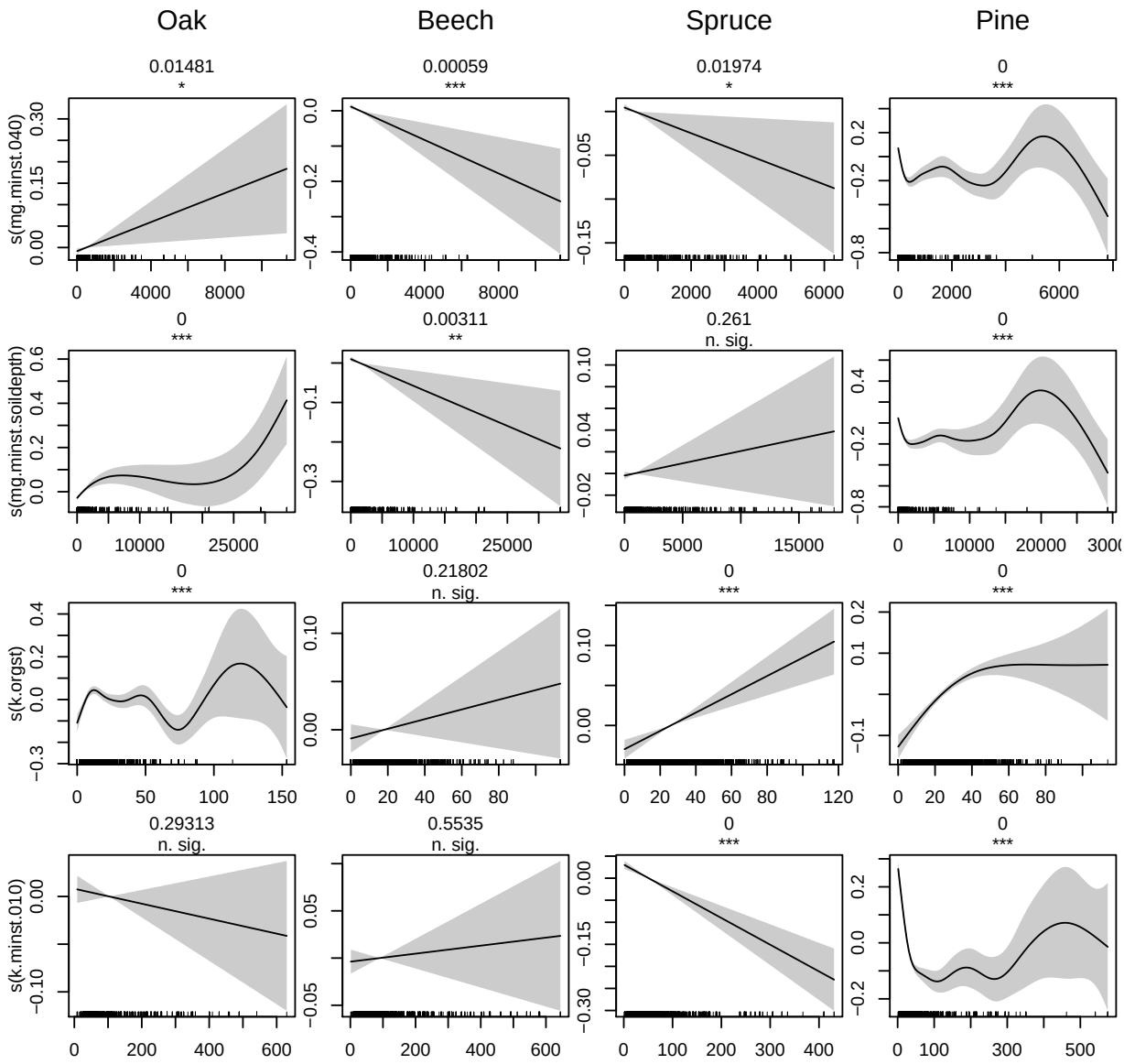

Oak
Beech
Spruce
Pine
0
***
0.00253
**
0
***
0
***
s(ph.org)
0.20295
n. sig.
0.00126
**
8e−05
***
0
***
s(ph010)
0.18942
n. sig.
0.00256
**
0
***
0
***
s(ph020)
0.14471
n. sig.
0.00579
**
0
***
0
***
s(ph040)

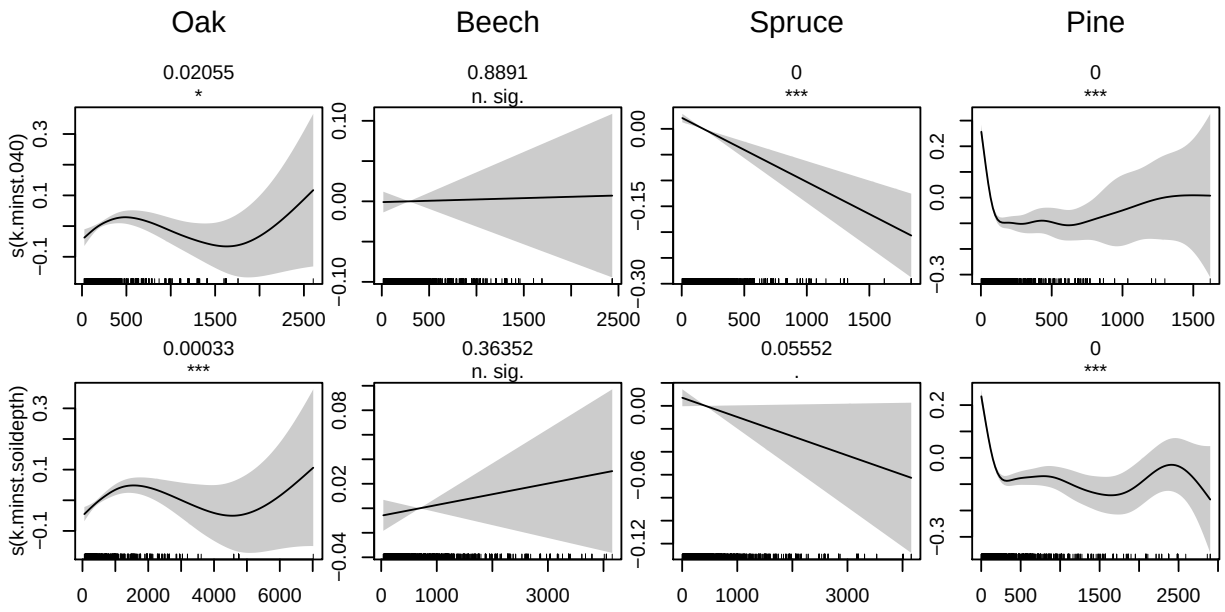

Oak
Beech
Spruce
Pine
0.15844
n. sig.
0.97517
n. sig.
0
***
0
***
s(ph4080)
0
***
0.1189
n. sig.
0.00037
***
0
***
s(basesat.org)
0.00726
**
0.05352
.
0.00333
**
0
***
s(basesat010)
3e−05
***
0.00654
**
9e−05
***
0
***
s(basesat040)


Oak
Beech
Spruce
Pine
0.01823
*
0.74727
n. sig.
0.32398
n. sig.
0
***
s(basesat4080)
0
***
0.00025
***
0.00822
**
0
***
s(basesat.soildepth)
1e−05
***
0.12897
n. sig.
0.29975
n. sig.
0.00925
**
s(cec.orgst)
0.00092
***
0.06362
.
1e−05
***
0
***
s(cec.010)


Oak
Beech
Spruce
Pine
9e−05
***
0.00636
**
0.00087
***
0
***
s(cec.040)
0.00972
0.52227
n. sig.
0.22931
0.3613
s(cec.soildepth)
s(ca.orgst)
0.73962
0.03545
*
s(ca.minst.010)


Oak
Beech
Spruce
Pine
0.54668
n. sig.
0.00445
**
5e−05
***
0
***
0.00488
**
0.13313
n. sig.
0.01259
*
0
***
0.72877
n. sig.
0.312
n. sig.
0.86377
n. sig.
1e−05
***
0.68914
n. sig.
0.00866
**
0
***
0
***
s(ca.minst.040)
s(ca.minst.soildepth)
s(mg.orgst)
s(mg.minst.010)

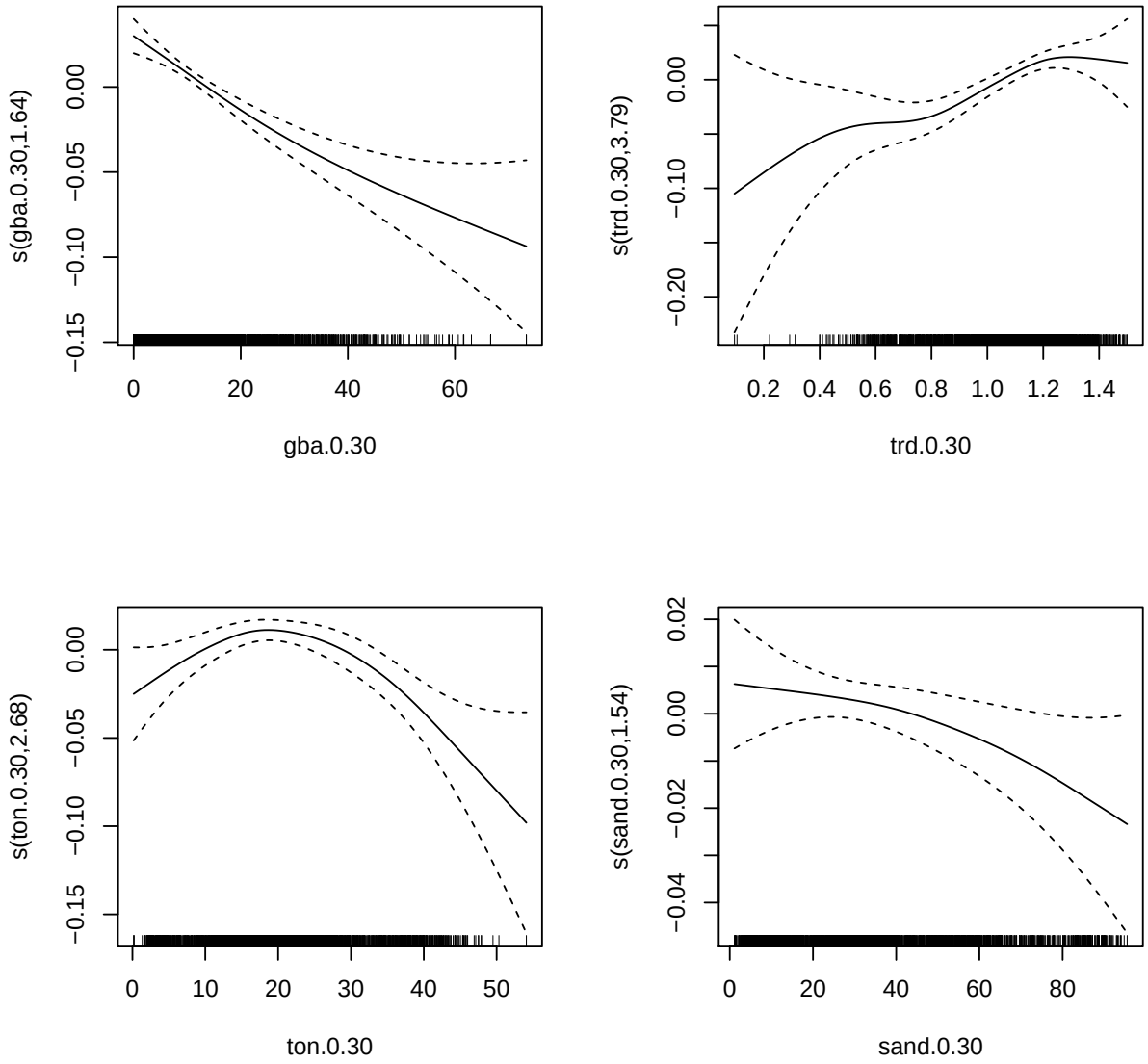

Oak
Beech
Spruce
Pine
0.01481
*
0.00059
***
0.01974
*
0
***
s(mg.minst.040)
0
***
0.00311
**
0.261
n. sig.
0
***
s(mg.minst.soildepth)
0
***
0.21802
n. sig.
0
***
0
***
s(k.orgst)
0.29313
n. sig.
0.5535
n. sig.
0
***
0
***
s(k.minst.010)

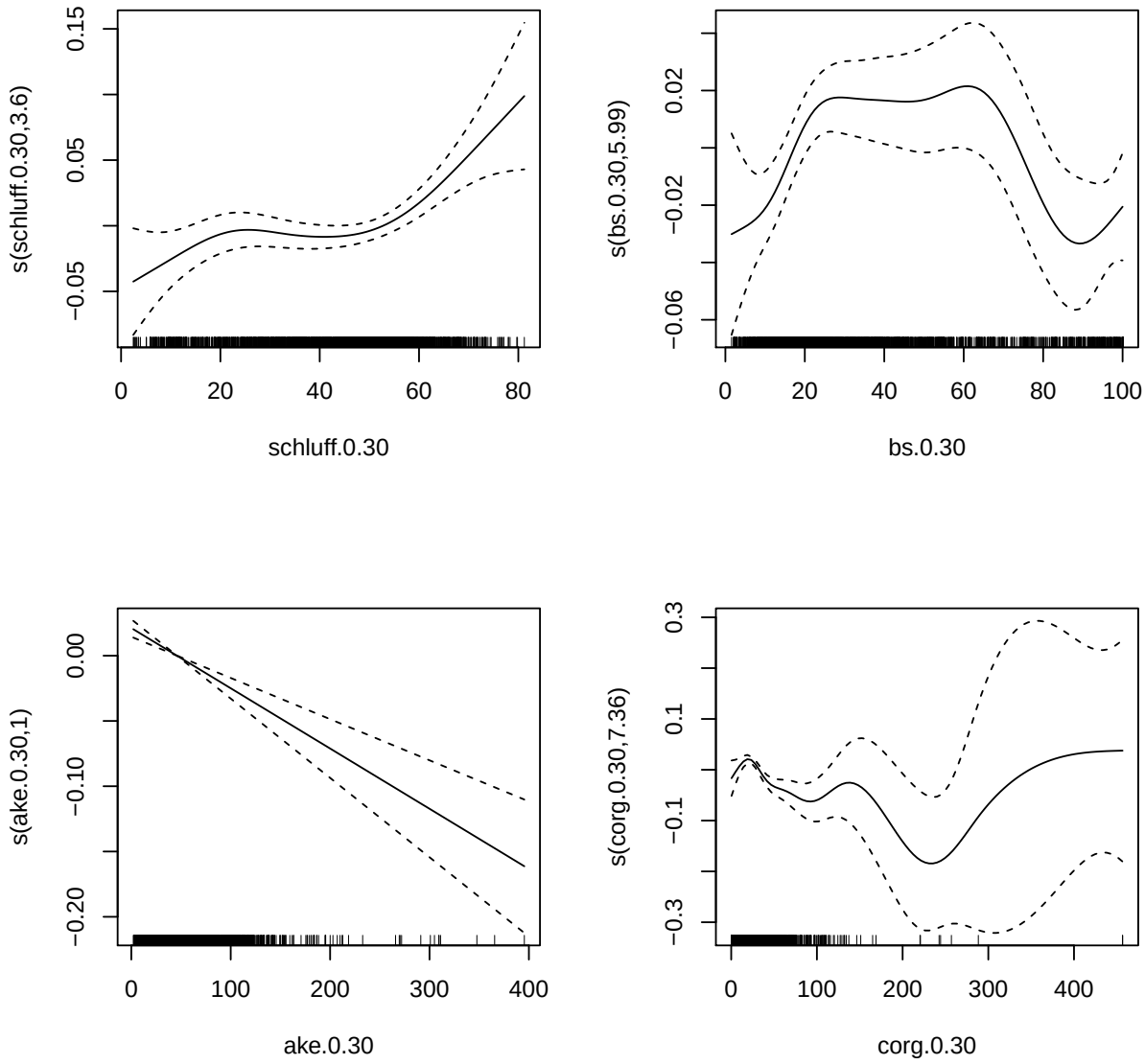

Oak
Beech
Spruce
Pine
0.02055
*
0.8891
n. sig.
0
***
0
***
0.00033
***
0.36352
n. sig.
0.05552
.
0
***
s(k.minst.040)
s(k.minst.soildepth)

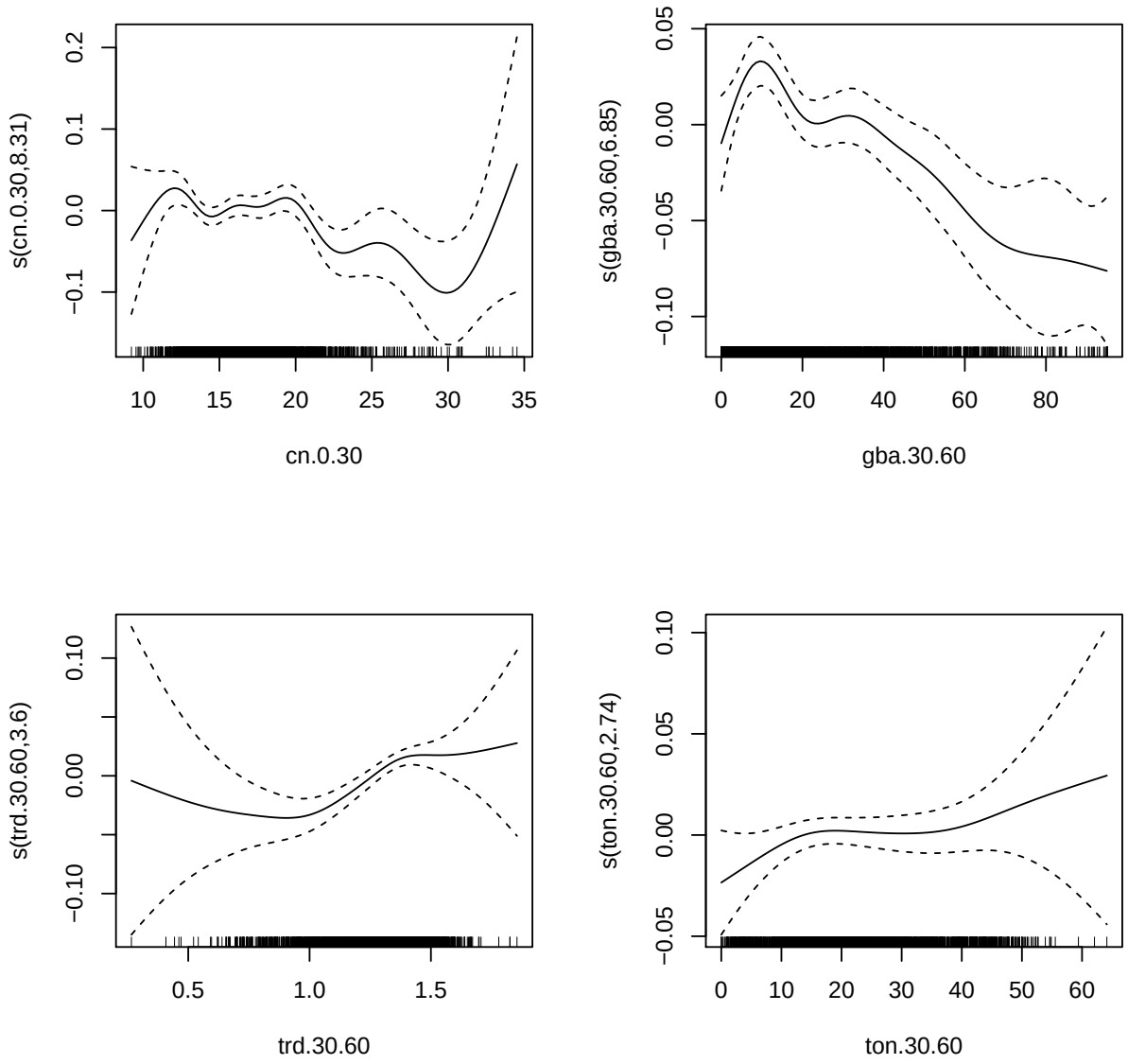
A.4 Splines for the h-a model soil covariates

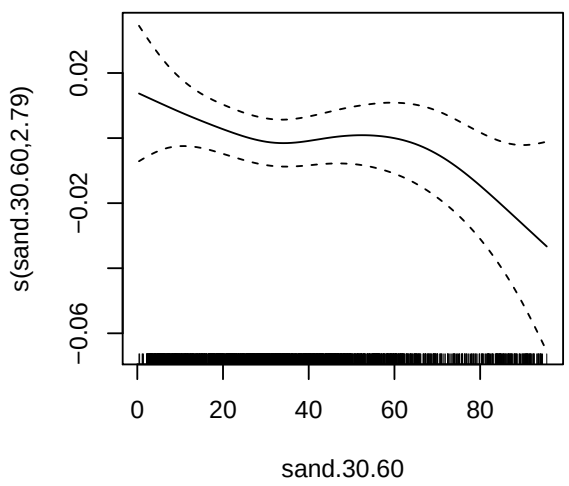
Splines for soil parameters for beech and pine evaluated as covariates for the second model step of study II.

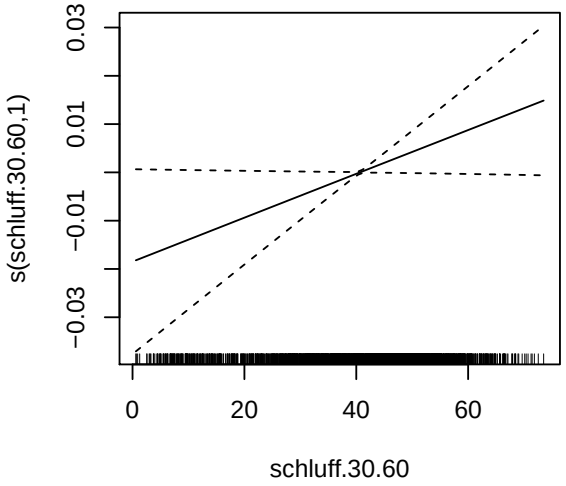
A.4.1 Beech

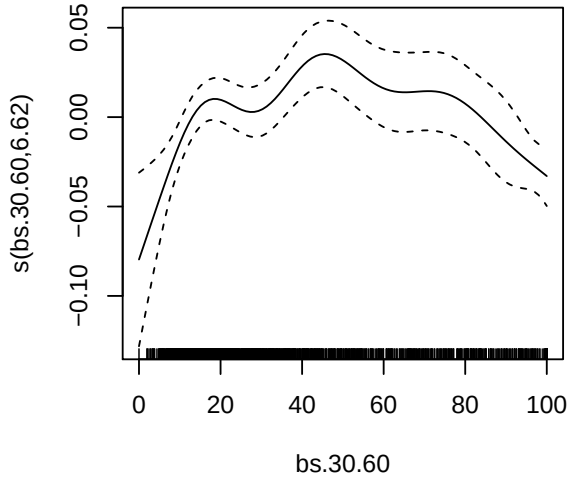


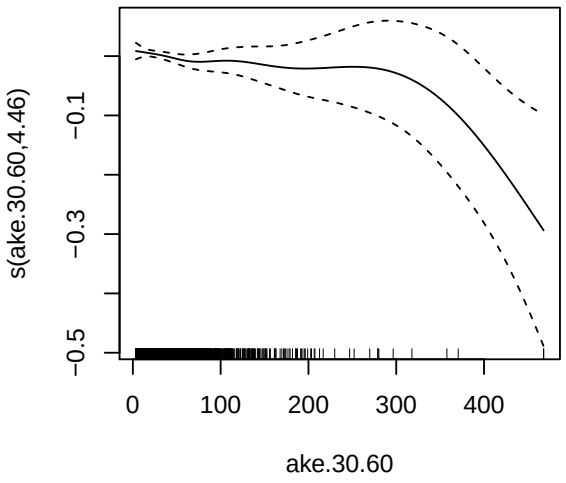

s(schluff.0.30,3.6)
0.15
0.05
−0.05
0
20
40
60
80
schluff.0.30
s(bs.0.30,5.99)
0.02
−0.02
−0.06
0
20
40
60
80
100
bs.0.30
s(ake.0.30,1)
0.00
−0.10
−0.20
0
100
200
300
400
ake.0.30
s(corg.0.30,7.36)
0.3
0.1
−0.1
−0.3
0
100
200
300
400
corg.0.30

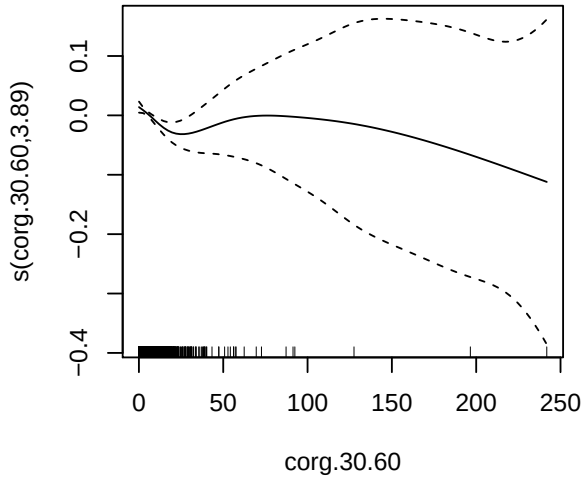

s(cn.0.30,8.31)
−0.1
0.0
0.1
0.2
10
15
20
25
30
35
cn.0.30
s(gba.30.60,6.85)
−0.10
−0.05
0.00
0.05
0
20
40
60
80
gba.30.60
s(trd.30.60,3.6)
−0.10
0.00
0.10
0.5
1.0
1.5
trd.30.60
s(ton.30.60,2.74)
−0.05
0.00
0.05
0.10
0
10
20
30
40
50
60
ton.30.60

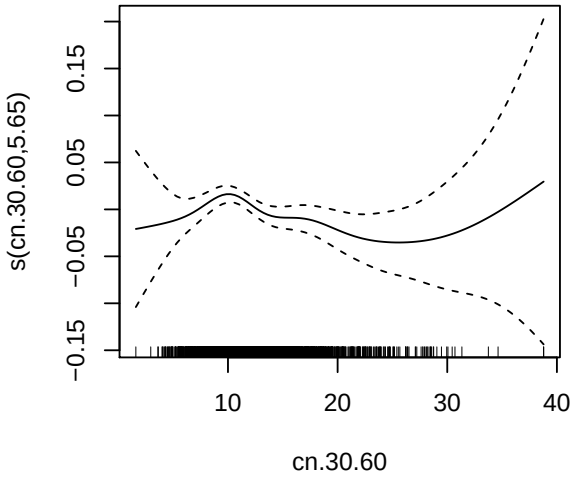

s(sand.30.60,2.79)
0.02
−0.02
−0.06
0
20
40
60
80
sand.30.60

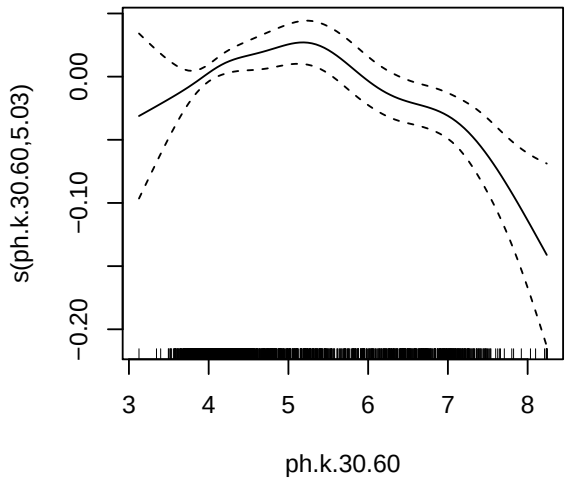

s(schluff.30.60,1)
0.03
0.01
−0.01
−0.03
0
20
40
60
schluff.30.60

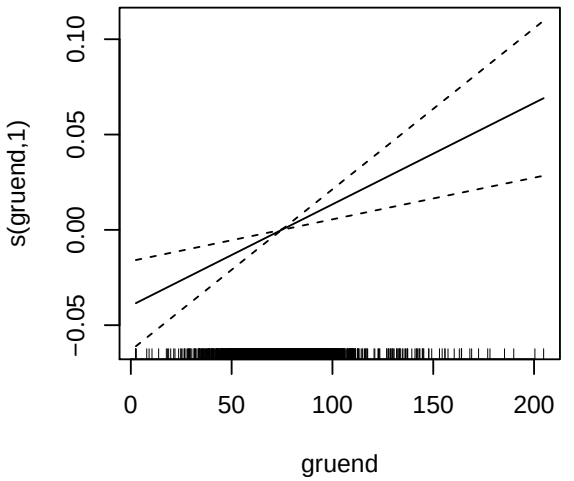

s(bs.30.60,6.62)
0.05
0.00
−0.05
−0.10
0
20
40
60
80
100
bs.30.60

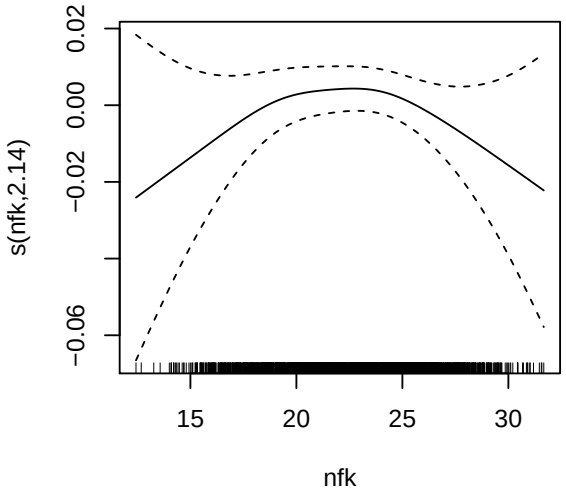

s(ake.30.60,4.46)
−0.1
−0.3
−0.5
0
100
200
300
400
ake.30.60


s(corg.30.60,3.89)
0.1
0.0
−0.2
−0.4
0
50
100
150
200
250
corg.30.60

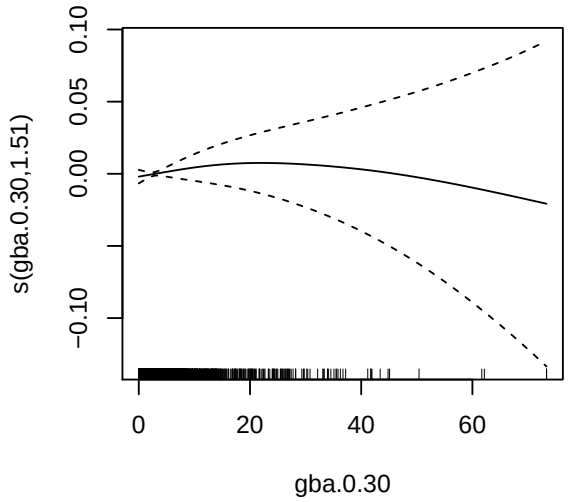

s(cn.30.60,5.65)
0.15
0.05
−0.05
−0.15
10
20
30
40
cn.30.60

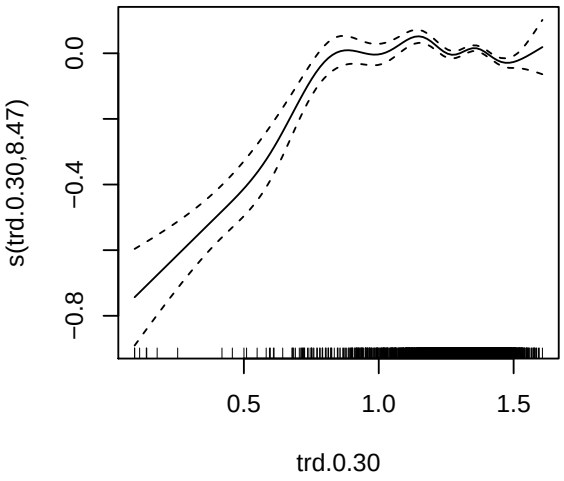

s(ph.k.30.60,5.03)
0.00
−0.10
−0.20
3
4
5
6
7
8
ph.k.30.60

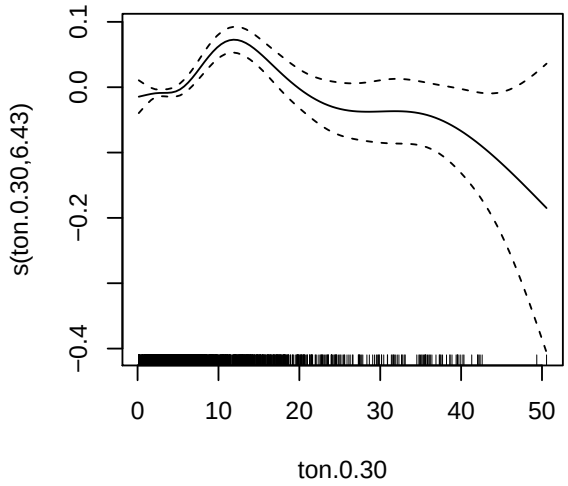

s(gruend,1)
0.10
0.05
0.00
−0.05
0
50
100
150
200
gruend

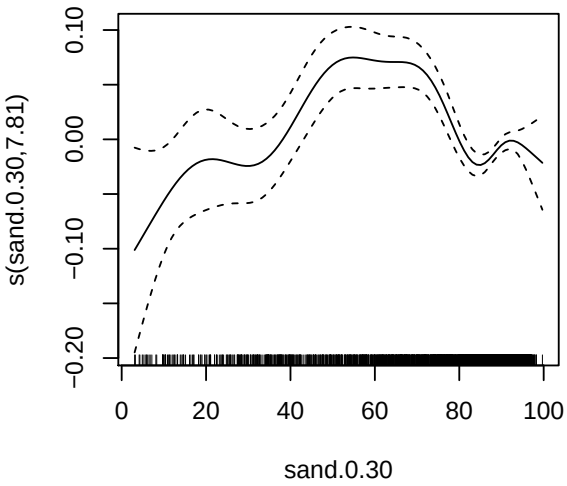

s(nfk,2.14)
0.02
0.00
−0.02
−0.06
15
20
25
30
nfk

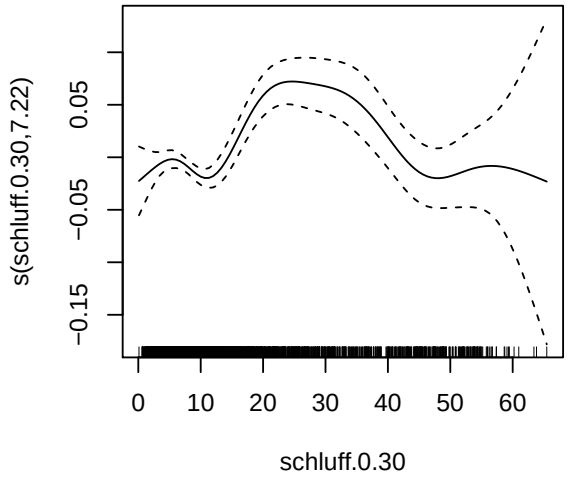
A.4.2 Pine

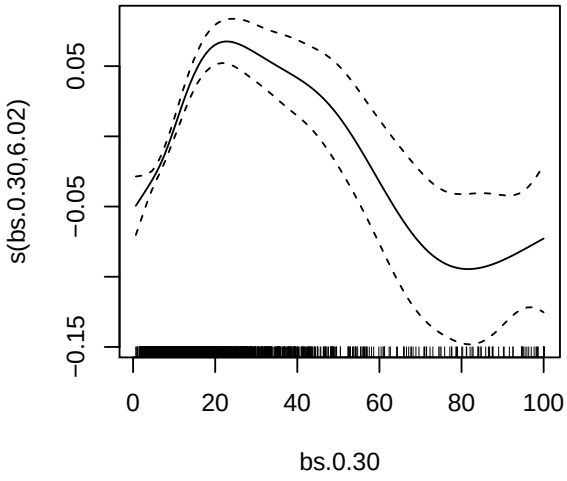

s(gba.0.30,1.51)
−0.10
0.00
0.05
0.10
0
20
40
60
gba.0.30

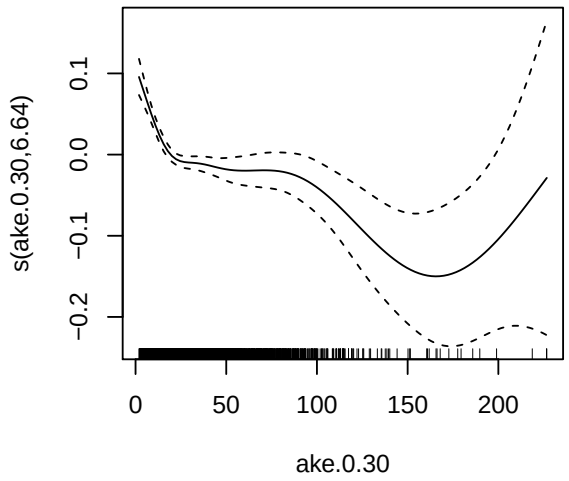

s(trd.0.30,8.47)
−0.8
−0.4
0.0
0.5
1.0
1.5
trd.0.30

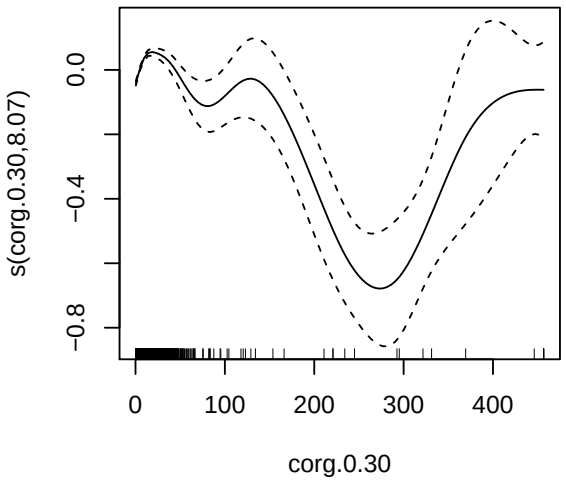

s(ton.0.30,6.43)
−0.4
−0.2
0.0
0.1
0
10
20
30
40
50
ton.0.30

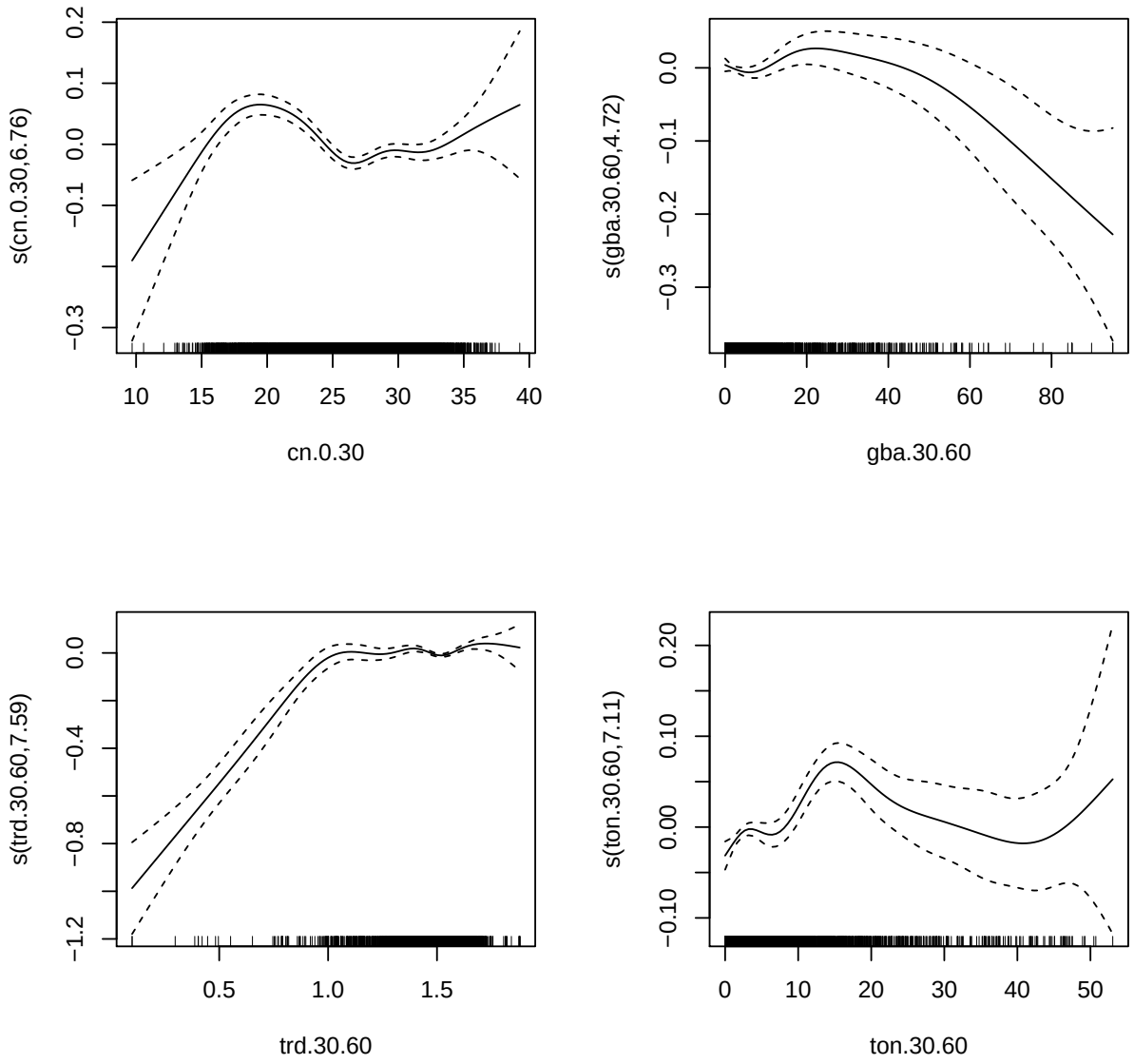

s(sand.0.30,7.81)
−0.20
−0.10
0.00
0.10
0
20
40
60
80
100
sand.0.30

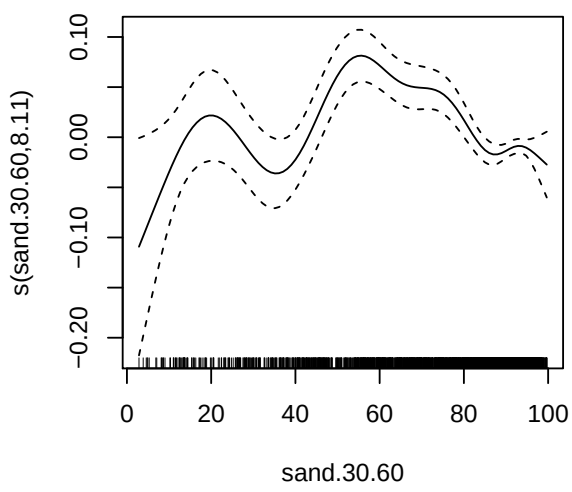

s(schluff.0.30,7.22)
0.05
−0.05
−0.15
0
10
20
30
40
50
60
schluff.0.30

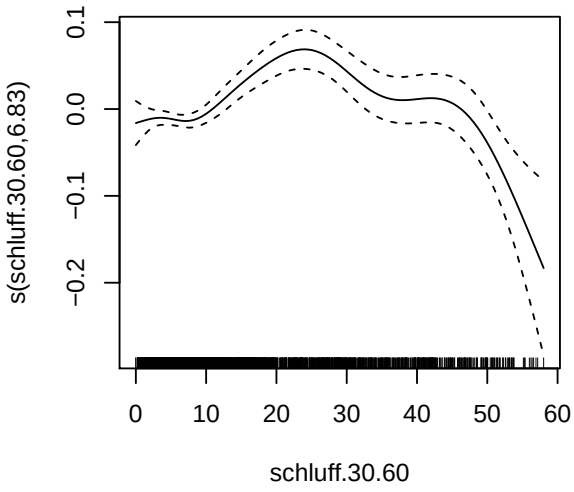

s(bs.0.30,6.02)
0.05
−0.05
−0.15
0
20
40
60
80
100
bs.0.30

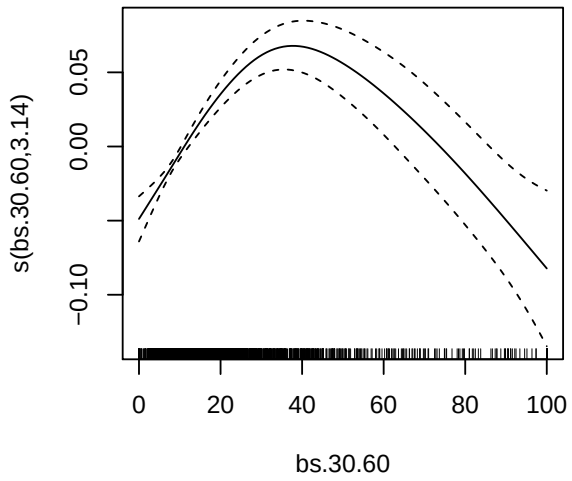

s(ake.0.30,6.64)
0.1
0.0
−0.1
−0.2
0
50
100
150
200
ake.0.30

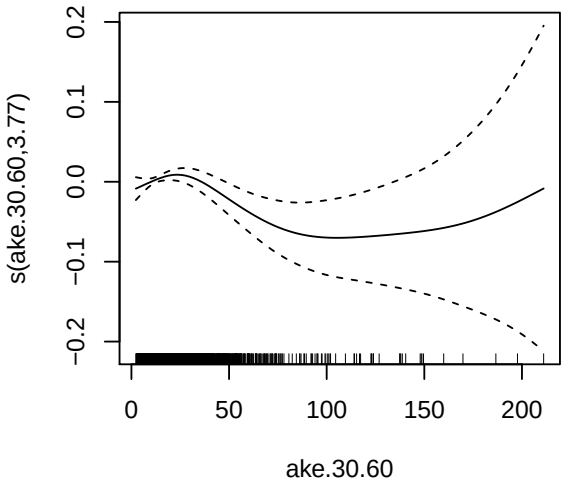

s(corg.0.30,8.07)
0.0
−0.4
−0.8
0
100
200
300
400
corg.0.30

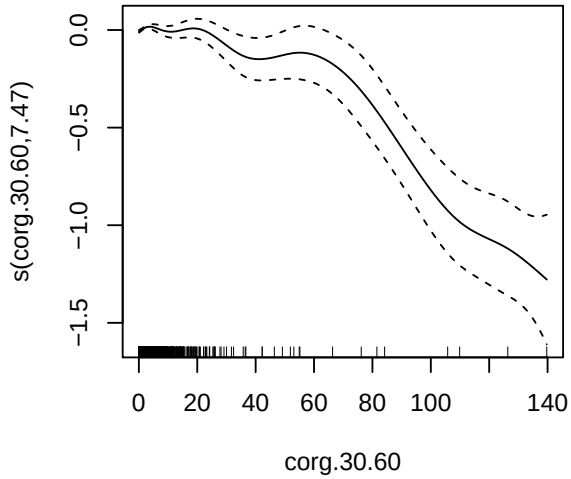

s(cn.0.30,6.76)
cn.0.30
s(gba.30.60,4.72)
gba.30.60
s(trd.30.60,7.59)
trd.30.60
s(ton.30.60,7.11)
ton.30.60

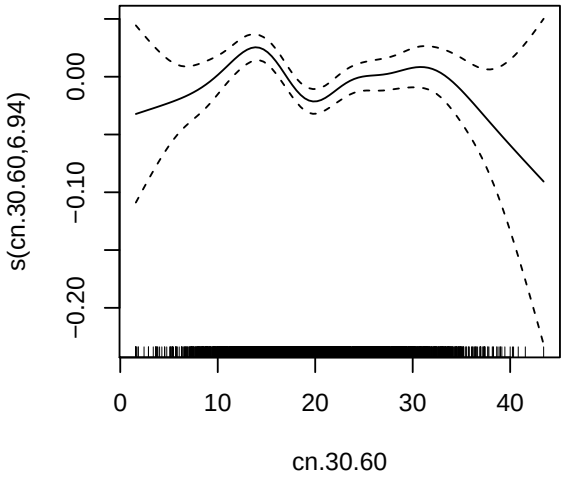

s(sand.30.60,8.11)
0.10
0.00
−0.10
−0.20
0
20
40
60
80
100
sand.30.60

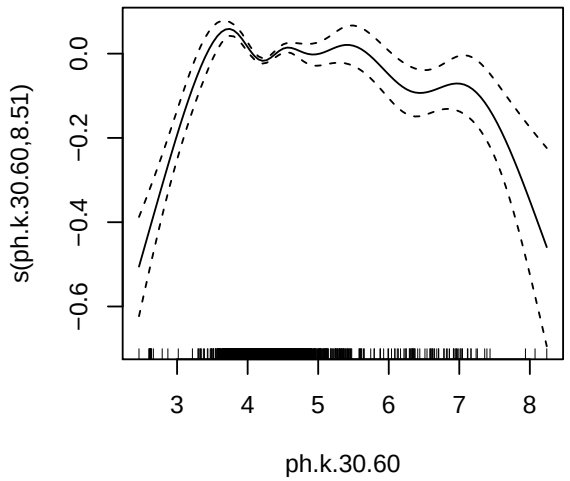

s(schluff.30.60,6.83)
0.1
0.0
−0.1
−0.2
0
10
20
30
40
50
60
schluff.30.60

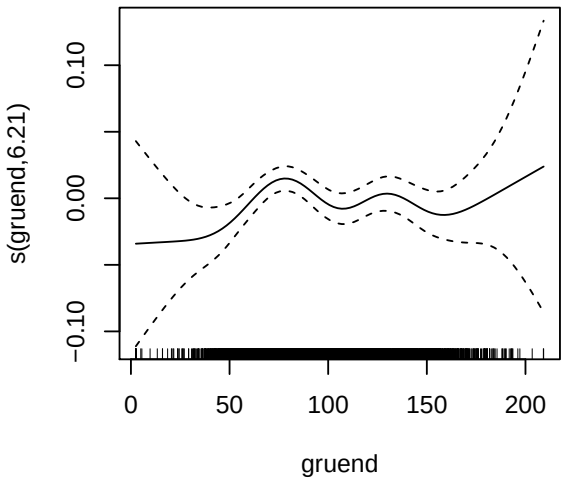

s(bs.30.60,3.14)
0.05
0.00
−0.10
0
20
40
60
80
100
bs.30.60

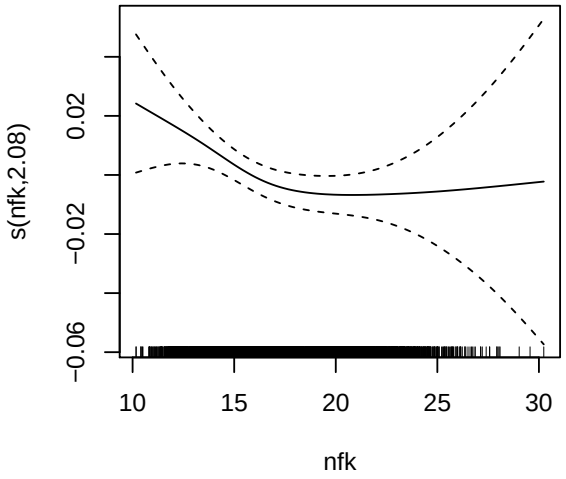

s(ake.30.60,3.77)
0.2
0.1
0.0
−0.1
−0.2
0
50
100
150
200
ake.30.60


s(corg.30.60,7.47)
0.0
−0.5
−1.0
−1.5
0
20
40
60
80
100
140
corg.30.60


s(cn.30.60,6.94)
0.00
−0.10
−0.20
0
10
20
30
40
cn.30.60


s(ph.k.30.60,8.51)
0.0
−0.2
−0.4
−0.6
3
4
5
6
7
8
ph.k.30.60


s(gruend,6.21)
0.10
0.00
−0.10
0
50
100
150
200
gruend


s(nfk,2.08)
0.02
−0.02
−0.06
10
15
20
25
30
nfk