Article (refereed) - postprint

Reyer, Christopher P.O.; Flechsig, Michael; Lasch-Born, Petra; Van Oijen, Marcel. 2016. **Integrating parameter uncertainty of a process-based model in assessments of climate change effects on forest productivity**. *Climatic Change*, 137 (3). 395-409. 10.1007/s10584-016-1694-1

Integrating parameter uncertainty of a process-based model in assessments of climate change effects on forest productivity

Christopher P.O. REYER[1,2]*, Michael FLECHSIG[1], Petra LASCH-BORN[1], Marcel VAN OIJEN[3]

[1]Potsdam Institute for Climate Impact Research

[2]Department of Geography, Humboldt University Berlin, Berlin, Germany

[3]Centre for Ecology and Hydrology (CEH-Edinburgh)

*Corresponding Author:

Potsdam Institute for Climate Impact Research

Telegrafenberg, P.O. Box 601203

14412 Potsdam, Germany

reyer@pik-potsdam.de

Abstract

The parameter uncertainty of process-based models has received little attention in climate change impact studies. This paper aims to integrate parameter uncertainty into simulations of climate change impacts on forest net primary productivity (NPP). We used either prior (uncalibrated) or posterior (calibrated using Bayesian calibration) parameter variations to express parameter uncertainty. We assessed the effect of parameter uncertainty on projections of the process-based model 4C in Scots pine (*Pinus sylvestris*) stands under climate change. We compared the uncertainty induced by differences between climate models with the uncertainty induced by parameter variability and climate models together. This paper shows that the uncertainty of simulated changes in NPP induced by climate model and parameter uncertainty is substantially higher than the uncertainty of NPP changes induced by climate model uncertainty alone. It however also highlights that the direction of NPP change is mostly consistent between the simulations using the standard parameter setting of 4C and the majority of the simulations including parameter uncertainty. Climate change impact studies that do not consider parameter uncertainty may be appropriate for projecting directions of change but not for quantifying the exact degree of change. Moreover, models that were calibrated to data may not much show reduced output uncertainty under climate change if parameter combinations are selected that are particularly climate sensitive. Our findings are highly relevant because most climate change impact studies do not integrate parameter uncertainty

and may thus be over- or underestimating climate change impacts on forest ecosystems.

Keywords: 4C; Bayesian calibration; climate models; Europe; Monte Carlo analysis; National Forest Inventory data

1. Introduction

Process-based models are widely used to assess the impacts of climate change on forest ecosystems because they are constructed to represent forest processes under non-analogues conditions such as the ones expected under future climate change (Fontes et al. 2010; Reyer 2015). However, their results depend on the reliability of the input data (*input uncertainty*), the representation or the lacking of processes (*structural uncertainty*) and the uncertainty about model parameter values (*parameter uncertainty*). All these uncertainties need to be accounted for when interpreting the results of model simulations (Lindner et al. 2014).

In many cases parameter values of process-based models are uncertain since they are derived from few and very specific ecophysiological measurements and observations (Mäkelä et al. 2000). This leads to considerable parameter uncertainty especially if a model is applied to sites across the distribution range of a tree species in which phenotypic and genotypic variation prevail. For example, carbon balance models from stand-scale forest growth models (e.g. Mäkelä 1986) to dynamic global vegetation models (e.g. Sitch et al. 2003) often include the pipe model (Shinozaki et al. 1964). These models assume that the leaf to sapwood area ratio is constant for a particular species or plant functional type. However, empirical studies show that this ratio varies with climate (Mencuccini and Grace 1995), stand density and site fertility (Berninger et al. 2005; Espinosa-Bancalari et al. 1987; Long and Smith 1988; Pothier and Margolis 1991). If this variation is included in a model, it influences the model results by altering the allocation of net primary productivity to the stem (Berninger and Nikinmaa 1997). While the effects of input uncertainty and of structural uncertainty have been partly addressed elsewhere (e.g. Medlyn et al. 2011; Reyer et al. 2014) and although there are methods that use widely available data sources to address uncertain parameter values (Hartig et al. 2012; van Oijen et al. 2005; van Oijen & Thompson 2010; van Oijen et al. 2013), so far parameter uncertainty has received less attention in climate change impact studies.

Therefore, the objectives of this paper are (1) to combine an analysis of parameter uncertainty with simulations of climate change impacts on forest productivity and (2) to compare the effects of input uncertainty arising from several climate models with the combined effects of both climate model input uncertainty and parameter uncertainty. We used Bayesian calibration with a Markov Chain Monte Carlo algorithm to assess the effects of parameter uncertainty on the projections of the process-based forest model 4C in Scots pine (*Pinus sylvestris*) stands under climate change in Austria, Belgium, Estonia and Finland. More specifically, we calibrated the model parameters of 4C in two different ways: for each country separately and for all countries simultaneously. Thereby two types of parameter distribution were derived: country-specific (calibrated on the stands available in the country) and generic (calibrated on the stands available from all four countries). These distributions were used to test whether calibration improved the model predictions in comparison to the standard, uncalibrated parameter set. We assessed the prior (before calibration) and posterior (after calibration) model output uncertainty for past conditions. Finally, we compared the uncertainty of net primary productivity (NPP), height and diameter at breast height (DBH) projections induced by using climate data from several climate models including the uncertainty induced by parameter variations with the uncertainty of NPP projections under climate change excluding parameter variations.

2. Material and Methods

a. Overview of methodology

This study builds upon a model comparison study where national forest inventory (NFI) data were used to calibrate forest models of different complexity (van Oijen et al. 2013). Van Oijen et al. (2013) calibrated parameter distributions of six models with Bayesian calibration techniques. They used either country-specific data from two NFI plots in each country (thus generating country-specific posterior parameter distributions) or a generic dataset consisting of the data of all the available NFI plots for that study (i.e. eight plots from four countries, leading to a generic posterior parameter distribution). Including also uncalibrated (i.e. prior) parameter distributions, they aimed to determine whether the models predicted the data of a third plot (a permanent sampling plot, PSP) in each country better without calibration or with the country-specific or the generic calibration. For more details on and formal descriptions of Bayesian calibration and applications with forest process-based models see van Oijen et al. (2005; 2013).

Here, we first compared the simulation results of the prior, the country-specific posterior and the generic posterior parameter distributions of the 4C model with the PSP data of van Oijen et al. (2013) to assess the influence of the country-specific and generic calibration datasets in more detail. Secondly, we ran the 4C model with its standard parameters and with climate change pathways from three regional climate models to assess the uncertainty of NPP projections induced by different climate models. Thirdly, we compared this climate model-induced uncertainty in NPP projections with the uncertainty induced by climate model and parameter uncertainties together. For the climate change simulations, we also studied the influence of continuous CO_2-fertilization as (i.e. increasing CO2-concentraions according to the A1B emission pathway) opposed to an acclimation of photosynthesis to 20th century CO_2-levels (i.e. fixed at 350ppm). Fig. 1 provides a schematic overview of the methodology.

b. Data

We used data from four European countries where Scots pine is part of commercial forestry, namely Austria (A), Belgium (B), Estonia (E) and Finland (F) (Table 1). In each country, we used two plots from national forest inventories (NFI, e.g. referred to as A1 and A2) and one PSP (e.g. referred to as A3) (Table 1). NFIs are usually carried out to assess forest resources over large spatial scales by systematic sampling and only measuring a few key variables while PSP are typically established in a few typical forests only but therefore monitored with much greater effort. In Estonia, no NFI plots but three PSPs were available. Hence for the first two of them the data were prepared as if originating from NFI to assure consistency with the other countries. For each stand, we initialized the forest model 4C (see below) with the stand data of the first available observation. The management of all stands was mimicked in 4C by removing trees following a thinning-from-above management strategy until the measured tree number was reached. Further descriptions of the stand, climate and soil data we used for the validation and calibration runs can be found in van Oijen et al. (2013).

For the climate change simulations we used the same soil and stand data but also modeled past climate data to ensure compatibility between past and future model simulations. We prepared data from three Regional Climate Models (RCMs) driven by three different General Circulation Models (GCMs) using the A1B emission scenario (Nakicenovic et al. 2000). The RCM/GCM combinations were CCLM/ECHAM5, HadRM3/HadCM3 and HIRHAM3/Arpège. The data of the latter two RCM/GCM combinations originated from the ENSEMBLES project (van der Linden and Mitchell 2009) while the CCLM/ECHAM5 data were from Lautenschlager et al. (2009) (henceforth we refer to the RCM/GCM combinations simply as RCMs). We bias-corrected and interpolated the simulated climate data to the sites by calculating a monthly mean model bias (absolute difference for temperature and relative for precipitation), adding (for temperature) or multiplying (for precipitation) this bias to/with daily simulated climate of past and future and interpolating the climate to the plots accounting for altitudinal dependencies of the climatic variables using a digital elevation model and external-drift-Kriging as described in Reyer et al. (2014). Table 2 shows the changes in temperature and precipitation featured in each climate model and at each plot.

c. The model 4C

The model 4C (‘FORESEE’ - Forest Ecosystems in a Changing Environment; http://www.pik-potsdam.de/4c/) describes forest development under changing environmental conditions (Bugmann et al. 1997; Lasch et al. 2005). Processes are modeled on the tree- and stand-level describing ecosystem carbon and water balances, leaf area index and forest structure. Establishment, growth, competition for light, water and nutrients and mortality of tree cohorts are modeled spatially implicit on a patch on which horizontal homogeneity is assumed. The soil sub-model describes temperature and water, carbon and nitrogen dynamics in different soil layers.

Photosynthesis is modelled as a function of environmental influences (temperature, water and nitrogen availability, radiation and CO2) modified from Haxeltine and Prentice (1996). Elevated CO2 increases the internal partial pressure of CO2 which increases light-use efficiency and gross assimilation and reduces stomatal conductance and the potential transpiration water demand thus increasing water-use efficiency. Water stress (described in Reyer et al. 2010) and nutrient limitations reduce assimilation. Respiration is a constant fraction of annual GPP (Landsberg and Waring 1997). The resulting NPP is allocated to different tree organs according to the pipe model (Shinozaki et al. 1964), the functional balance (Davidson 1969), height growth depending on foliage mass and light availability and a rise in bole height if the photosynthetic production of the lowermost branches drops below compensation of the sum their respiratory losses and senescence fluxes.

Temperature affects photosynthesis, growing season length, evapotranspiration which determines water demand and thus drought stress, and mineralization/decomposition and hence nutrient availability. Precipitation determines the soil water content and hence the water availability for uptake by trees.

The water balance is calculated from potential evapotranspiration depending on temperature, relative humidity, solar radiation according to Turc/Ivanov (Dyck and Peschke 1995), interception and percolation transport of water in the multi-layered soil is calculated (Grote and Suckow 1998). Root uptake is determined by the transpiration demand of all trees and the plant available water.

4C requires meteorological driving forces at daily resolution as well as a soil and a forest stand description for the model initialization. During initialization, the observed basal area and age of the stand are matched. Each of its currently 13 tree species, is represented by a set of 45 species-specific parameter values. These parameter values originate from literature, aggregated datasets and expert assessment and are henceforth referred to as the 'standard parameter' values (Table ESM1). A more detailed description of 4C, recent model applications as well as a model validation can be found in Reyer et al. (2010; 2014).

For all the Bayesian calibration and Monte Carlo simulation experiments, we interfaced 4C to the generic and model-independent simulation environment SimEnv (Flechsig et al. 2013).

d. Evaluation and comparison of calibration datasets

We constructed two different prior (i.e. uncalibrated) parameter distributions from independent marginal distributions for the individual model parameters. In the first one, each parameter was assumed to be uniformly distributed between 50% and 150% of its standard value in 4C (Table ESM1). This ±50% range of parameter values reflects a large uncertainty about parameter values across the broad variety of geographic distribution, stands, sites and climates considered in this study. In the second one, each parameter was assumed to be normally distributed with the distributions being truncated based on the literature and data that was used to define the standard parameters (Table ESM2). This second prior parameter distribution reflects the 'most plausible prior' and was introduced to test the influence of the more arbitrary ±50% range of parameter values of a uniform prior on model output uncertainty.

Using Monte Carlo simulations with Latin hypercube sampling, we then sampled 1000 parameter vectors from the prior parameter distributions and ran 4C for each parameter vector with the measured soil, stand, management and climate data for each PSP-site (codes A3, B3, E3, F3 in Table 1). The simulations were run for the time period between the first and the last available data point. This yielded 1000 simulation results that express the prior model output uncertainty under current climate.

The prior parameter distributions were then updated during the country-specific and generic calibrations using NFI data (codes A1, A2, B1, B2, E1, E2, F1, F2 in Table 1) and Bayesian calibration using a Markov-Chain Monte Carlo algorithm (see ESM2 for details). This resulted in four country-specific and one generic posterior parameter distribution. For the ‘most plausible prior’ we only performed the generic calibration and the country-specific calibration for F3 (referred to as F3*) but not for the other sites because the F3 site has the longest record of test data. From each of the posterior parameter distributions we sampled another 1000 parameter vectors and ran 4C with each parameter vector with the measured soil, stand, management and climate data of each PSP (codes A3, B3, E3, F3 in Table 1) which had not been used for calibration for a period from the first to the last available data point.

The results of these 1000 simulations express the country-specific and generic posterior model output uncertainty respectively under current climate. From the country-specific and generic posterior parameter distribution, we also derived the maximum a posteriori estimate (MAP), which is the most probable parameter vector (van Oijen et al. 2005).

To assess how the simulations fitted the observed stand data and which calibration dataset improved the predictions the most, we compared observed and simulated mean tree height and DBH for each plot. DBH and mean height were chosen since these are commonly reported variables in forest science. We calculated the Normalized Root Mean Square Error (NMRSE, see ESM1), based on the whole distribution (i.e. calculated as an average across the samples from the probability distributions) (van Oijen et al. 2013).

e. Influence of climate model and parameter uncertainty

For the climate change simulations, we ran 4C with the 1000 prior, country-specific posterior and generic posterior parameter vectors as well as with the standard parameter values (in case of the prior) and the MAPs (in case of the posterior simulations) at each of the four PSPs in the four countries using the measured stand, management and soil data for 30 years of climatic data from the three climate models for the periods 1971-2000 and 2061-2090. We calculated the change in the mean NPP and the height and DBH of the last simulation year for the period 2061-2090 compared to the period 1971-2000. To test the sensitivity of our results to the choice of the parameter uncertainty range of ±50%, we also repeated the prior simulations assuming a smaller uncertainty of initial parameter values of ±25% variation.

Although the changes in climate are driven by an increase in atmospheric CO_2 according to the A1B storyline (see section 'data'), the long-term effect of increasing CO_2 concentrations on forests is unclear (Körner 2006; Reyer et al. 2015). Therefore, in our simulations we made two assumptions about CO_2 concentrations and the persistence of its effects on photosynthesis: Firstly, we ran all simulations with increasing CO_2 concentrations according to the A1B emission scenario (i.e. persisting stimulation of photosynthesis by CO_2, hence the upper margin of CO_2-effects) and secondly we kept CO_2 concentration constant at 350ppm (i.e. an acclimation of photosynthesis to CO_2 at 350ppm, hence the lower margin of CO_2-effects) (see Reyer et al. (2014) or Medlyn et al. 2011 for a more thorough discussion of CO_2-effects in forest models).

Our simulation design resulted in a total of 192 192 simulation runs (three RCMs x two time periods x four stands x two assumptions about CO_2 x four parameter distributions based on two priors and two posteriors x 1001 parameter vectors). To assess the uncertainties induced by the ensemble of climate models and by parameter uncertainty, we considered the results of the simulations with standard parameter values, the MAPs and of the full range of simulations with prior, country-specific posterior and generic posterior parameter distributions.

3. Results

a. Bayesian calibration

Table 3 shows that even without calibration, 4C simulates height and DBH with reasonably low NRSME except for site F3. As expected, the calibration improves the model results as expressed by a lower NRMSE at all sites and for both diameter and height. The results of the generic calibration fit the data best (with the exception of height at E3) but generally the NRMSE for both calibration datasets are similar. The Bayesian calibration also reduced output uncertainty for both the country-specific and generic calibration. In most cases both the posterior mean as well as the MAP provide better fit to the data than the standard parameter run and the output range is much smaller than for prior simulations (see Fig. 2 and Fig ESM1 for an example for F3 and F3* respectively). Interestingly, the output uncertainty for height is smaller when considering F3 compared to F3* while the opposite is true for DBH. For F3*, the maximum values for height and DBH are also further reduced in comparison to F3 but some of the parameter combinations found for F3* lead to a die-off of trees while this is not the case under F3. For most marginal parameter distributions the posterior standard deviation was 1-2% less than the prior standard deviation. Parameter correlations were small and exceeded correlations of 0.4 in only one case. A full list of all prior and posterior parameter estimates is available in Table ESM2-3.

b. Influence of climate change on NPP projections

Across the four plots used in this study and across the three climate models, climate change leads to NPP changes ranging from -9 to 29% during the period 2061-2090 relative to 1971-2000 under an acclimation of CO_2-effects (Fig. ESM2). In the two Central European locations (Austria and Belgium) the responses are mostly small but negative, while in the two Northern European locations (Estonia and Finland) the responses are positive. Under persistent CO_2-effects, climate change always leads to positive NPP changes ranging from 11 to 78% across the four plots (Fig. ESM2).

c. Influence of climate change and parameter uncertainty on NPP projections

When parameter uncertainty is included in the climate change simulations under an acclimation of CO_2-effects, the range of possible NPP changes increases across all sites, varying from -21 to 62% for the prior assuming ±25% uncertainty ranges, from -48 to 136% for the prior assuming ±50% uncertainty ranges and from -46 to 141% and -45 to 231% for the posterior generic and the posterior country-specific model output distribution respectively, but the median changes remain comparable (Fig. ESM2). The F3* simulations show very similar ranges of results but slightly less negative NPP changes. The two different assumptions about parameter uncertainty, namely ±50% and ±25%, do not lead to large differences in median and the lower and the upper quartiles of NPP change. However, fewer extreme NPP changes are found under a parameter uncertainty of ±25%. There is no large difference between calibrated and uncalibrated (assuming ±50% parameter uncertainty) model output distributions but overall, the posterior model output uncertainty is slightly larger than the prior model output uncertainty.

Under persistent CO_2-effects, the range of possible NPP changes is much larger and mostly positive, varying from 0 to 147% for the prior assuming ±25% uncertainty ranges, from -35 to 478% for the prior assuming ±50% uncertainty ranges and from -36 to 489% and -15 to 539% for the posterior generic and the posterior country-specific model output distribution respectively, but again the median changes and the lower and upper quartiles remain comparable (Fig. ESM2). The F3* simulations show very similar ranges of results but slightly less negative NPP changes. Under persistent CO_2-effects, also the difference between ±50% and ±25% prior parameter uncertainty is less pronounced for E3 and F3.

Fig. 3 and 4 show the relative NPP changes at each of the four plots used in this study split up per regional climate model. In most cases, the NPP change induced by the standard parameter vector is close to the median and the MAP of the distribution of NPP change induced by parameter uncertainty. The largest deviations of the medians and MAPs of NPP change compared to the NPP change of the standard parameter simulations occur under persistent CO_2-effects at the E3 site. The medians, lower and upper quartiles and interquartile ranges of the prior assuming 50% uncertainty ranges and the posterior model output distributions are similar for the same RCM. They differ however between the different RCMs. While the median of the prior assuming 25% uncertainty ranges is similar to the medians of the other output distributions, its lower and upper quartiles and interquartile ranges are, with the exception of E3, much smaller than for the other output distributions. These general patterns are consistent between the simulations featuring different assumptions about CO_2 although persistent CO_2-effects lead to much larger values and ranges.

The results for height and DBH mainly mirror the NPP results but are characterized by slightly lower negative relative changes for the CCLM RCM (Fig. ESM3-8).

4. Discussion

a. Evaluation and comparison of calibration datasets

This paper shows that calibration of model parameters with even small amounts of NFI data helped to reduce the NRMSE of height and diameter predictions of a parameter-rich, process-based forest model driven with observed climate (Table 3). In a recent model comparison study using the same data, 4C was identified as the most plausible model for simulating height and DBH after calibration (van Oijen et al. 2013). Despite the low number of data points used for calibration and our assumptions about the prior parameter distribution (see discussion below), our findings supports evidence from other studies that Bayesian methods combined with NFI data improve model parameterizations (Mäkelä et al. 2012; van Oijen et al. 2013). Although the generic posterior parameter distribution yielded mostly lower NRMSE values than the country-specific posterior parameter distribution, there were no large differences between the two methods. This is noteworthy since the country-specific posterior parameter distribution included fewer data points. Thus, the advantage of having more data points in the generic calibration was partly compensated for by having only country-specific data points in the country-specific calibration. This shows that process-based models can actually be calibrated to represent local conditions but as well for larger regions if enough calibration data is available. Given that process-based models are increasingly designed for the latter and that more and more data for model calibration is becoming available, we see good prospect for further improving our understanding of parameter uncertainty at larger scales. Further studies are needed to determine at which level of data availability a generic calibration would perform better than a country-specific calibration and should consider testing the difference of using regional prior parameter information as opposed to generic priors used here.

b. Influence of climate model and parameter uncertainty

This paper highlights that the uncertainty about changes in NPP induced by climate model and parameter uncertainty can be substantially higher than the uncertainty about NPP changes induced by climate models alone. While this is a trivial statement as such, it means that model-based projections of climate change-induced changes in NPP and their implications for carbon cycling and forest growth may be more uncertain than previously thought. Our findings partly rely on the assumption that the climate change uncertainty induced by the three climate models and the prior parameter uncertainty are realistic and hence can be compared.

It is also important to note that some parameter values may in reality be more, others less variable than the parameter variations we assumed here. Especially, the truncation of the normally distributed parameters (for the F3*) simulations seems to be too wide given that certain parameter combinations lead to stand decline (Figure ESM1). Similarly, the higher NRSME values for height in the F3* simulations as opposed to the F3 simulations (Table 3) are possibly related to the larger parameter ranges assumed for key parameters governing carbon allocation to height growth and light extinction in the F3* simulations (pfext and pnus in Table ESM2). Also, the distribution of the prior may differ from a uniform or normal distribution. While using another distribution may decrease uncertainty (Wramneby et al. 2008), here we took examples of assuming 1) a simple uniform distribution and the same relative uncertainty for each parameter and 2) a normal distribution with parameter mean and truncation derived from the original literature and data used to parameterize 4C as a first attempt to account for parameter uncertainty. The variation around the standard parameter as well as the shape of the prior parameter distribution could be further refined in future studies by gathering information of possible parameter values from traits-databases (e.g. Kattge et al. 2011).

The large prior uncertainties however also mean that, if several species would be considered, as is usually done in climate change impact studies (e.g. Reyer et al. 2014), species-specific parameter uncertainty ranges may overlap. This may complicate risk assessments for individual tree species or for the competition of tree species (Wramneby et al. 2008) and highlights the need for the use of existing data assimilation techniques such as in this study or in van Oijen et al. (2013) with more data (i.e. longer time series, more sites) to improve species-specific parameterizations of process-based models, handle more complex forest structures and/or even derive regional, sub-species level parameterizations. Especially, data from a wider array of sources could help to directly constrain the wider range of processes encapsulated in process-based models.

To test how sensitive our prior model output uncertainties are to the assumption of ±50% parameter variation, we included results from the Monte Carlo simulations without calibration assuming only ±25% variation around the standard value and the calibrations including 'the most plausible prior' using normally distributed parameter ranges taken from the literature and data available for model parametrization. In the former case, the uncertainties about the NPP changes due to the choice of climate model and parameter uncertainty were reduced (Fig. ESM2). However, they were still considerably larger than the variability in NPP changes induced by the climate models alone. When considering the simulations with the 'most plausible prior', model output uncertainty was not much different from the ±50% parameter uncertainty runs. This result is not surprising given that the 'most plausible prior' contains ranges larger than ±50% for some parameters or information availability was low so that values had to be kept at a ±50% range (Table ESM2). Thus, our results are qualitatively robust across a large range of assumed parameter uncertainties.

A restriction of our study is that we build the prior from independent marginal distributions. While this is a natural starting point when information is scarce it is likely that some of the parameter combinations which lead to very extreme results in our simulations may not be realistic (c.f. Wramneby et al. 2008), but without additional data no parameter combinations could be excluded at this stage. Moreover, our study was not very dependent on the prior since we analyzed output uncertainty by posterior distributions. Even our simulations using the posterior parameter distributions (hence after including data) show a wide range of possible productivity changes despite very unrealistic parameter combinations having been eliminated by the calibration procedure. It is however important to note that the calibration was done for past climates measured at the specific study sites and that the climate model data differ from measured data even for the past and after a bias correction and interpolation (c.f. Reyer et al. 2014). Thus, calibrated parameters are not necessarily fully realistic under climate change.

Another important assumption of our study is that the climate models we have chosen adequately represent uncertainty about possible climate change. The projections of the RCMs used here range from 1.5 to 4.5°C warming and from -16 to 15% changes in precipitation between the different stands (Table 2) which is well in line with the range of projections by the IPCC for Europe for a similar period (IPCC 2007). Even though using a wider range of climate scenarios would certainly encapsulate stronger climate changes and hence lead to stronger NPP changes, a recent study with the 4C model found that these are rarely lager than 45% or smaller than -15% in the countries covered here (Reyer et al. 2014). Even though these results were found for a different set of forest in the respective countries, the changes seem substantially smaller than the changes induced by parameter uncertainty and climate change in our study. Thus uncertainty in climate input introduced by the three RCMs seems wide enough to be compared with the uncertainty induced by the variation of parameter values. It is noteworthy that the input uncertainty induced by the different climate models alone already leads to a variation in NPP changes from 3 to 29% in the most extreme case of E3.

The influence of model structural uncertainty can also increase the range of climate model-induced uncertainty (but also in the simulations including parameter uncertainty) as exemplified by our two different but very influential assumptions about the persistence of CO_2-effects on photosynthesis and water use. While this is an attempt to assess model structural uncertainty regarding the influence of CO_2, it does not fully account for the true range of structural uncertainty since the actual model formulation of how CO_2 affects photosynthesis in 4C remains unchanged. This can be better tested by driving structurally diverse models with the same data (e.g. Warszwaski et al. 2013).

Our results reveal one more interesting particularity: Figures 3 and 4 show that the posterior model output uncertainty (of both the generic and country-specific posterior parameter distributions) is sometimes larger than the prior model output uncertainty. This is counterintuitive since for the simulations using measured climate in the first part of our analysis, the posterior model output NRMSE was reduced in comparison to the prior model output NRMSE values. The posterior parameter uncertainty was slightly reduced as well. This means that forward propagation of posterior parameter uncertainty to model output uncertainty (of NPP change) leads to increased uncertainty when comparing the effects of multiple climate models. This could be because our comparably small calibration dataset might have led to parameter combinations that were coming from inappropriate regions of the parameter space. While we cannot fully rule out this possibility, we think that the reduction in posterior output uncertainty for past conditions, even though not a substantial one, rather points towards another explanation: the posterior parameter distribution assigns higher probability to a subregion of parameter space where climate sensitivity is high and varies much. This is possible because in 4C, NPP is nonlinearly related to the model parameters and therefore parameter combinations that may not seem to have much effect under current climatic conditions, may lead to larger output variation under different climates. We speculate that especially those parameters related to the photosynthesis model would be particularly sensitive to such effects, because in 4C NPP is strongly linked to photosynthesis which is itself sensitive to tmeperatures. This also means that when calibration reduces a model's output uncertainty for present-day conditions, it does not guarantee that the model's output uncertainty for future, climatically changed conditions is reduced too.

c. Implications for climate change impact studies

This paper shows that – while the absolute magnitude of climate change-induced NPP changes is highly uncertain if considering parameter uncertainties – the direction of NPP change is mostly consistent between the simulations using the standard parameter setting of 4C and the majority of the simulations using the parameter variation induced by prior or posterior parameter uncertainties (as expressed by the boxes in Figs 3 and 4 which include 50% of the values). Figs 3 and 4 show that typically the median of the NPP change due to climate change and parameter uncertainty mirrors the NPP change induced by climate change alone. Although projections using the standard parameters of 4C do not take into account parameter uncertainty, the direction and quality of change (i.e. small or large) are met quite well. Thus, the standard parameters may be appropriate for projecting directions of climate change impacts, especially if including some information on input uncertainty, but not their exact magnitude. This increases the confidence in the overall pattern of NPP change under climate change found in recent applications of 4C at the European scale (Reyer et al. 2014). However, it is important that for quantitative assessments of climate change impacts on forests using complex process-based models, parameter uncertainty is considered more thoroughly as it adds significantly to input uncertainty induced by climate models. Our study also shows that this can be done either using country-level calibrations or more generic calibrations as the climate sensitivity of NPP is rather similar for these two different calibrations in our study. Given that process-based modelling is often focused on finding general parameter values that are applicable across the range of a species or plant functional type , generic calibrations may be favored but further research is needed to determine when a more localized calibration is to be preferred to a more generic one. Finally, our findings are highly relevant for climate change impact assessment because most such studies do not yet integrate parameter uncertainty and may thus be over- or underestimating impacts on forest ecosystems and may not provide the full range of uncertainties to decision makers. Integrating more thorough assessments of different kinds of uncertainties would allow increasing the robustness of climate change impact studies.

5. Acknowledgements

This work would not have been possible without the data and support provided by Werner Rammer (Austria), Gaby Deckmyn (Belgium), Andres Kiviste (Estonia), Annikki Mäkelä and Sanna Härkönen (Finland). We are grateful to the support of COST Actions FP0603 and FP1304 as well as the FP7 project MOTIVE (grant agreement no. 226544). We further acknowledge the help and support of our colleagues Felicitas Suckow, Tobias Pilz, Martin Gutsch, Sebastian Ostberg and Stefan Lange.

6. References

Berninger F, Coll L, Vanninen P, Mäkelä A, Sari Palmroth, Nikinmaa E (2005) Effects of tree size and position on pipe model ratios in Scots pine. Can J For Res 35:1294–1304

Berninger F, Nikinmaa E (1997) Implications of varying pipe model relationships on Scots Pine growth in different climates. Funct Ecol 11 (2):146-156. doi:10.1046/j.1365-2435.1997.00067.x

Bugmann H, Grote R, Lasch P, Lindner M, Suckow F (1997) A new forest gap model to study the effects of environmental change on forest structure and functioning. In: Mohren GMJ, Kramer K, Sabate S (eds) Impacts of Global Change of Tree Physiology and Forest Ecosystem. Proceedings of the International Conference on Impacts of Global Change on Tree Physiology and Forest Ecosystems, held 26-29 November 1996, Wageningen. Forestry Science. Kluwer Academic Publisher, Dordrecht, pp 255-261

Buisson L, Thuiller W, Casajus N, Lek S, Grenouillet G (2010) Uncertainty in ensemble forecasting of species distribution. Glob Chang Biol 16 (4):1145-1157

Cipra B (2000) Revealing Uncertainties in Computer Models. Sci 287 (5455):960-961. doi:10.1126/science.287.5455.960

Cramer W, Bondeau A, Woodward FI, Prentice C, Betts R, Brovkin V, Cox P, Fisher V, Foley J, Friend A, Kucharik C, Lomas MR, Ramakutty N, Sitch S, Smith B, White A, Young-Molling C (2001) Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models. Glob Chang Biol 7:357-373

Espinosa-Bancalari MA, Perry DA, Marshall JD (1987) Leaf area-sapwood area relationships in adjacent young Douglas-fir stands with different early growth rates. Can J For Res 17 (2):174-180. doi:10.1139/x87-030

Flechsig M, Böhm U, Nocke T, Rachimow C (2005) Techniques for quality assurance of models in a multi-run simulation environment. In: Hanson KM, Hemez FM (eds) Sensitivity Analysis of Model Output - Proceedings of the 4th International Conference on Sensitivity Analysis of Model Output (SAMO 2004). Los Alamos National Laboratory, Los Alamos, USA, pp 297-306

Flechsig M, Böhm U, Nocke T, Rachimow C (2013) The multi-run simulation environment SimEnv: User Guide for Version 3.1. Potsdam Institute for Climate Impact Research, Potsdam, Germany

Fontes L, Bontemps JD, Bugmann H, van Oijen M, Gracia CA, Kramer K, Lindner M, Rötzer T, Skovsgaard JP (2010) Models supporting forest management in a changing environment. For Sys 19(SI):8-29

Hartig F, Dyke J, Hickler T, Higgins SI, O'Hara RB, Scheiter S, Huth A (2012) Connecting dynamic vegetation models to data – an inverse perspective. J Biogeogr. doi:10.1111/j.1365-2699.2012.02745.x

IPCC (2007) Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on

Climate Change. In: Solomon S, Qin D, Manning M et al. (eds). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,, p 996

Kattge, J., et al. (2011). TRY – a global database of plant traits. Global Change Biology 17(9): 2905-2935

Landsberg J (2003) Modelling forest ecosystems: state of the art, challenges, and future directions. Can J For Res 33 (3):385-397

Lasch P, Badeck FW, Suckow F, Lindner M, Mohr P (2005) Model-based analysis of management alternatives at stand and regional level in Brandenburg (Germany). For Eco Manage 207 (1-2):59-74

Lautenschlager M, Keuler K, Wunram C, Keup-Thiel E, Schubert M, Will A, Rockel B, Boehm U (2009a) Climate Simulation with CLM, Climate of the 20th Century run no.1, Data Stream 3: European region MPI-M/MaD. World Data Center for Climate. doi:10.1594/WDCC/CLM_C20_1_D3

Lautenschlager M, Keuler K, Wunram C, Keup-Thiel E, Schubert M, Will A, Rockel B, Boehm U (2009c) Climate Simulation with CLM, Climate of the 20th Century run no.2, Data Stream 3: European region MPI-M/MaD. World Data Center for Climate. doi:10.1594/WDCC/CLM_C20_2_D3

Lautenschlager M, Keuler K, Wunram C, Keup-Thiel E, Schubert M, Will A, Rockel B, Boehm U (2009b) Climate Simulation with CLM, Scenario A1B run no.1, Data Stream 3: European region MPI-M/MaD. World Data Center for Climate. doi:10.1594/WDCC/CLM_A1B_1_D3

Lautenschlager M, Keuler K, Wunram C, Keup-Thiel E, Schubert M, Will A, Rockel B, Boehm U (2009d) Climate Simulation with CLM, Scenario A1B run no.2, Data Stream 3: European region MPI-M/MaD. World Data Center for Climate. doi:10.1594/WDCC/CLM_A1B_2_D3

Lindner M, JB Fitzgerald, NE Zimmermann, C Reyer, S Delzon, E van der Maaten, MJ Schelhaas, P Lasch-Born, J Eggers, M van der Maaten-Theunissen, F Suckow, A Psomas, B Poulter, M Hanewinkel (2014). Climate Change and European Forests: What do we know, what are the uncertainties, and what are the implications for forest management? Journal of Environmental Management 146:69-83 DOI: 10.1016/j.jenvman.2014.07.030

Long JN, Smith FW (1988) Leaf area - sapwood area relations of lodgepole pine as influenced by stand density and site index. Can J For Res 18 (2):247-250. doi:10.1139/x88-036

Mäkelä A (1986) Implications of the Pipe Model Theory on Dry Matter Partitioning and Height Growth in Trees. J Theor Biol 123:103-120

Mäkelä A, Landsberg J, Ek AR, Burk TE, Ter-Mikaelian M, Agren GI, Oliver CD, Puttonen P (2000) Process-based models for forest ecosystem management: current state of the art and challenges for practical implementation. Tree Physiol 20 (5-6):289-298. doi:10.1093/treephys/20.5-6.289

Mäkelä A, Río Md, Hynynen J, Hawkins MJ, Reyer C, Soares P, van Oijen M, Tomé M (2012) Using stand-scale forest models for estimating indicators of sustainable forest management. For Ecol Manage 285:164-178

Medlyn BE, Duursma RA, Zeppel MJB (2011) Forest productivity under climate change: a checklist for evaluating model studies. Wiley Interdiscip Rev Clim Chang 2 (3):332-355. doi:10.1002/wcc.108

Mencuccini M, Grace J (1995) Climate influences the leaf area/sapwood area ratio in Scots pine. Tree Physiol 15 (1):1-10. doi:10.1093/treephys/15.1.1

Nakicenovic N, Alcamo J, Davis G, Vries Bd, Fenhann J, Gaffin S, Gregory K, Grübler A, Jung T, Kram T, Rovere EL, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Riahi K, Roehrl A, Rogner H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, Rooijen Sv, Victor N, Dadi Z (2000) IPCC Special Report Emission Scenarios. Cambridge University Press, Cambridge, United Kingdom

Pothier D, Margolis A (1991) Analysis of growth and light interception of balsam fir and white birch saplings following precommercial thinning. Ann For Sci 48 (2):123-132

Purves D, Pacala S (2008) Predictive Models of Forest Dynamics. Sci 320:1452-1453

Reyer C, Lasch P, Mohren GMJ, Sterck FJ (2010) Inter-specific competition in mixed forests of Douglas-fir (*Pseudotsuga menziesii*) and common beech (*Fagus sylvatica*) under climate change—a model-based analysis. Ann For Sci 67:805

Reyer C, Lasch-Born P, Suckow F, Gutsch M, Murawski A, Pilz T (2014) Projecting regional changes in forest net primary productivity in Europe driven by climate change and carbon dioxide concentration. Ann For Sci 71:211–225 DOI 10.1007/s13595-013-0306-8

Reyer C (2015). Projections of changes in forest productivity and carbon pools under environmental change – A review of stand scale modeling studies. Current Forestry Reports 1:53-68. DOI: 10.1007/s40725-015-0009-5

Reyer C, N Brouwers, A Rammig, B Brook, J Epila, RF Grant, M Holmgren, F Langerwisch, S Leuzinger, W Lucht, B Medlyn, M Pfeiffer, J Steinkamp, M Vanderwel, H Verbeeck, D Villela (2015). Forest resilience and tipping points at different spatio-temporal scales: approaches and challenges. Journal of Ecology 103:5-15 DOI: 10.1111/1365-2745.12337

Shinozaki K, Yoda K, Hozumi K, Kira T (1964a) A quantitative analysis of plant form - the pipe model theory I. Basic analysis. Jap J Ecol14 (3):97-105

Shinozaki K, Yoda K, Hozumi K, Kira T (1964b) A quantitative analysis of plant form - the pipe model theory II. Further evidence of a theory and its application in forest ecology. Jap J Ecol 14 (4):133-139

Sitch S, Smith B, Prentice IC, Arneth A, Bondeau A, Cramer W, Kaplan JO, Levis S, Lucht W, Sykes MT, Thonicke K, Venevsky S (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Glob Chang Biol 9 (2):161-185

Valentine HT (1985) Tree-growth models: Derivations employing the pipe-model theory. J Theor Biol 117 (4):579-585

van der Linden P, Mitchell J (eds) (2009) ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, Exeter, UK van Oijen M, Cameron DR, Butterbach-Bahl K, Farahbakhshazad N, Jansson PE, Kiese R, Rahn KH, Werner C, Yeluripati JB (2011) A Bayesian framework for model calibration, comparison and analysis: Application to four models for the biogeochemistry of a Norway spruce forest. Agric For Meteorol 151 (12):1609-1621

van Oijen M, Reyer C, Bohn FJ, Cameron DR, Deckmyn G, Flechsig M, Härkönen S, Hartig F, Huth A, Kiviste A, Lasch P, Mäkelä A, Mette T, Minunno F, Rammer W (2013) Bayesian calibration, comparison and averaging of six forest models, using data from Scots pine stands across Europe. For Ecol Manage 289:255-268

Van Oijen M, Rougier J, Smith R (2005) Bayesian calibration of process-based forest models: bridging the gap between models and data. Tree Physiology 25 (7):915-927

Van Oijen M, Thompson A (2010) Toward Bayesian uncertainty quantification for forestry models used in the United Kingdom Greenhouse Gas Inventory for land use, land use change, and forestry. Climatic Change 103:55–67

Warszawski, L., et al. (2013). A multi-model analysis of risk of ecosystem shifts under climate change. Environmental Research Letters 8(4): 044018

Wramneby A, Smith, B, Zaehle S, Sykes MT (2008) Parameter uncertainties in the modeling of vegetation dynamics – Effects on tree community structure and ecosystem functioning in European forest biomes. Ecol Mod 216:277-290

7. Tables

Table 1 Forest stands used in this study. The data refer to the last measurement at each plot. For more information see van Oijen et al. (2013). NFI = National Forest Inventory; PSP = Permanent Sampling Plot; DBH = Diameter at Breast Height. "N observations" indicates how many data points for both height and diameter combined were available from each site and in brackets the years of the first and last measurement. The first data point was always used for model initialization.

Site code	Data type	Lat.	Long.	Age (y)	Stem number (ha^{-1})	Height (m)	DBH (cm)	N observations
A1	NFI	48.31°	14.79°	~64	526	18.5	32.4	4 (1987-2000)
A2	NFI	48.51°	15.70°	~66	1363	17.7	20.7	4 (1989-2002)
A3	PSP	48.51°	15.70°	59	690	18.1	23.9	4 (1980-1995)
B1	NFI	51.28°	5.52°	67	380	18.4	27.1	4 (2000-2004)
B2	NFI	51.28°	5.52°	66	393	23.2	29.3	4 (2000-2008)
B3	PSP	51.3°	4.52°	79	362	21.3	31.9	6 (1994-2007)
E1	PSP*	57.85°	25.92°	70	402	25.0	27.4	6 (2000-2010)
E2	PSP*	57.98°	25.63°	67	692	24.9	23.7	6 (2000-2010)
E3	PSP	57.58°	25.28°	73	667	25.6	24.5	6 (2000-2010)
F1	NFI	61.97°	27.67°	75	899	17.8	19.1	4 (1985-1995)
F2	NFI	63.83°	24.65°	55	1067	10.1	14.6	4 (1985-1995)
F3	PSP	61.33°	25.03°	79	1710	21.8	17.0	14 (1948-1997)

*PSP-data but presented in the format of and used as if originating from NFI data

Table 2 Mean annual temperature (T; degree Celsius) and mean annual precipitation sum (P; mm) of the periods 1971-2000 and 2061-2090 for climate models considered in this study. They result from three RCMs forced with the A1B emission scenario at the four permanent sampling plots (A3, B3, E3, F3) used in this study.

RCM	**Period**	**T [°C]**	**P [mm]**	**T [°C]**	**P [mm]**	**T [°C]**	**P [mm]**	**T [°C]**	**P [mm]**
		A3		*B3*		*E3*		*F3*	
CCLM	1971-2000	10.0	607	10.5	806	6.1	684	4.4	638
HadRM3	1971-2000	10.0	643	10.3	873	5.9	729	4.0	689
HIRHAM3	1971-2000	10.2	584	10.4	832	6.2	713	4.5	675
CCLM	2061-2090	12.9	605	13.0	852	9.2	787	7.8	718
HadRM3	2061-2090	14.0	635	13.6	809	10.4	734	8.4	739
HIRHAM3	2061-2090	11.7	647	12.1	700	8.9	642	8.1	670

Table 3 Normalized Root Mean Square Error (NRMSE, c.f. ESM1) from simulations compared to measured heights and DBHs (Diameter at Breast Height) at four permanent sampling plots in four European countries without calibration and with country-specific and generic calibration.

Site	Uncalibrated	Country-specific calibration	Generic calibration
		Height	
A3	0.29	0.15	0.12
B3	0.23	0.15	0.09
E3	0.13	0.12	0.14
F3	0.52	0.28	0.27
F3*	0.47	0.43	0.38
		DBH	
A3	0.23	0.16	0.13
B3	0.14	0.13	0.08
E3	0.06	0.06	0.05
F3	1.00	0.68	0.52
F3*	0.76	0.57	0.46

8. Figures

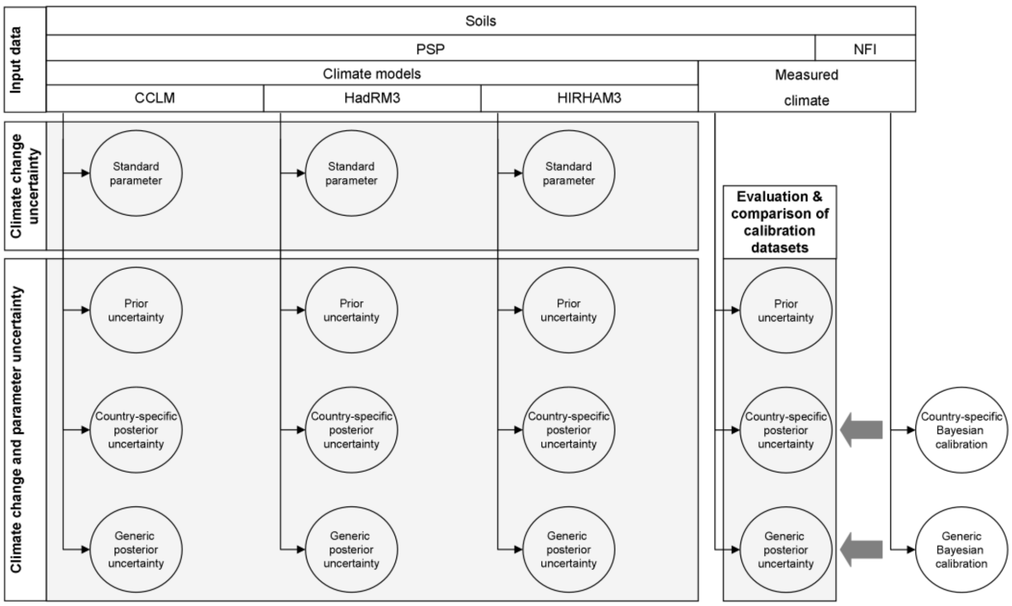

Fig. 1 Schematic overview of the methodology and the steps of the analysis (PSP = Permanent sampling plot; NFI = National Forest Inventory). The grey shaded areas represent aspects analyzed in this paper

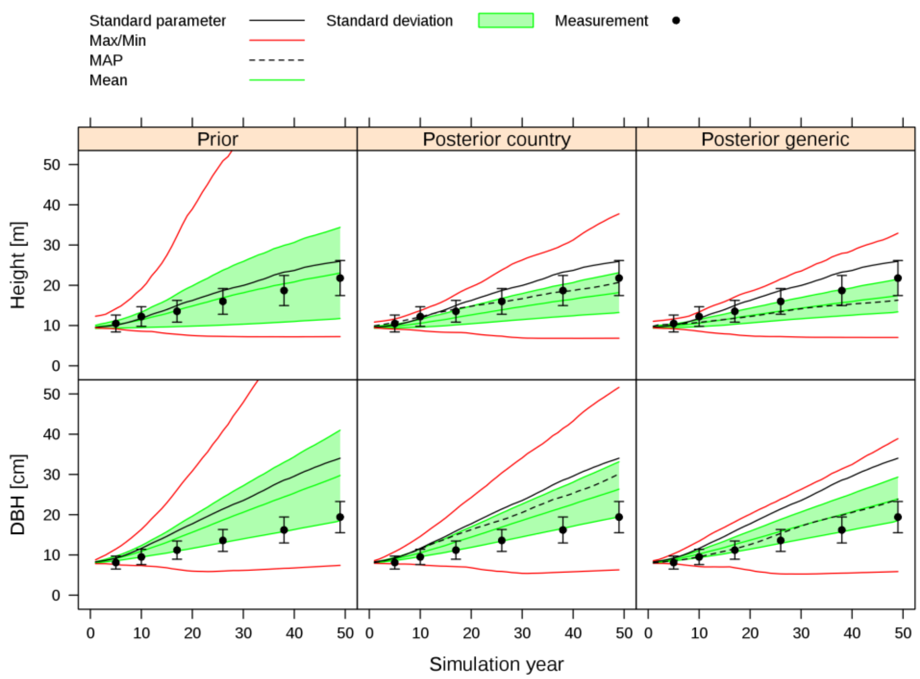

Fig. 2 Prior and posterior output uncertainty for height and DBH of the F3 plot. Posterior output uncertainty is depicted once for the country-specific ("posterior country") and generic ("posterior generic") calibration.

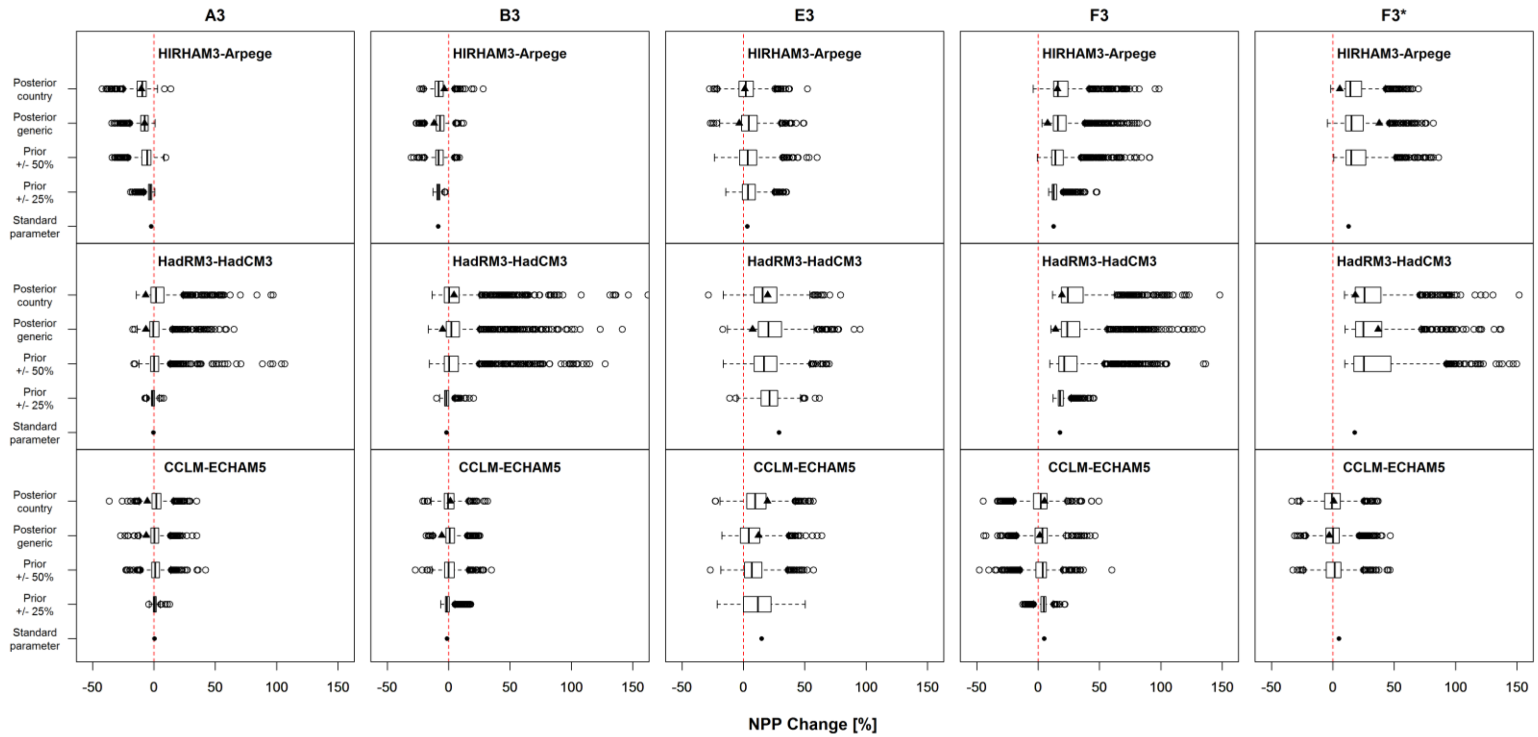

Fig. 3 Change in net primary productivity (NPP) under climate change for individual climate models under the assumption of an acclimation of photosynthesis to CO_2-effects for four plots in Austria, Belgium, Estonia and Finland (A3-F3, see Table 1). F3* denotes the simulations assuming the most plausible prior parameter distribution. The data are sorted according to climate model uncertainty alone (Label 'Standard parameter' (i.e. using 4C's standard parameter set)) and due to climate model and parameter uncertainty of uncalibrated (two degrees of prior parameter uncertainty, 'Prior ±50%' or

‘Prior ±25%’, respectively) or calibrated (‘Posterior generic’ or ‘Posterior country’) parameter distributions. Please note that for F3*, ‘Prior ±50%’ actually designates the simulations with the updated prior parameter ranges as described in Table ESM2. The responses are split up for each climate model. The triangles represent the simulations using the MAP. See the text for further explanation. The x-axis is cut at 150% for better legibility. The boxplots show the following information: thick line= median, bottom and top of the box = lower and upper quartiles, whiskers = maximum value or 1.5 times the interquartile range of the data depending on which is smaller. Points = outliers larger than 1.5 times interquartile range. The dotted line indicates no change

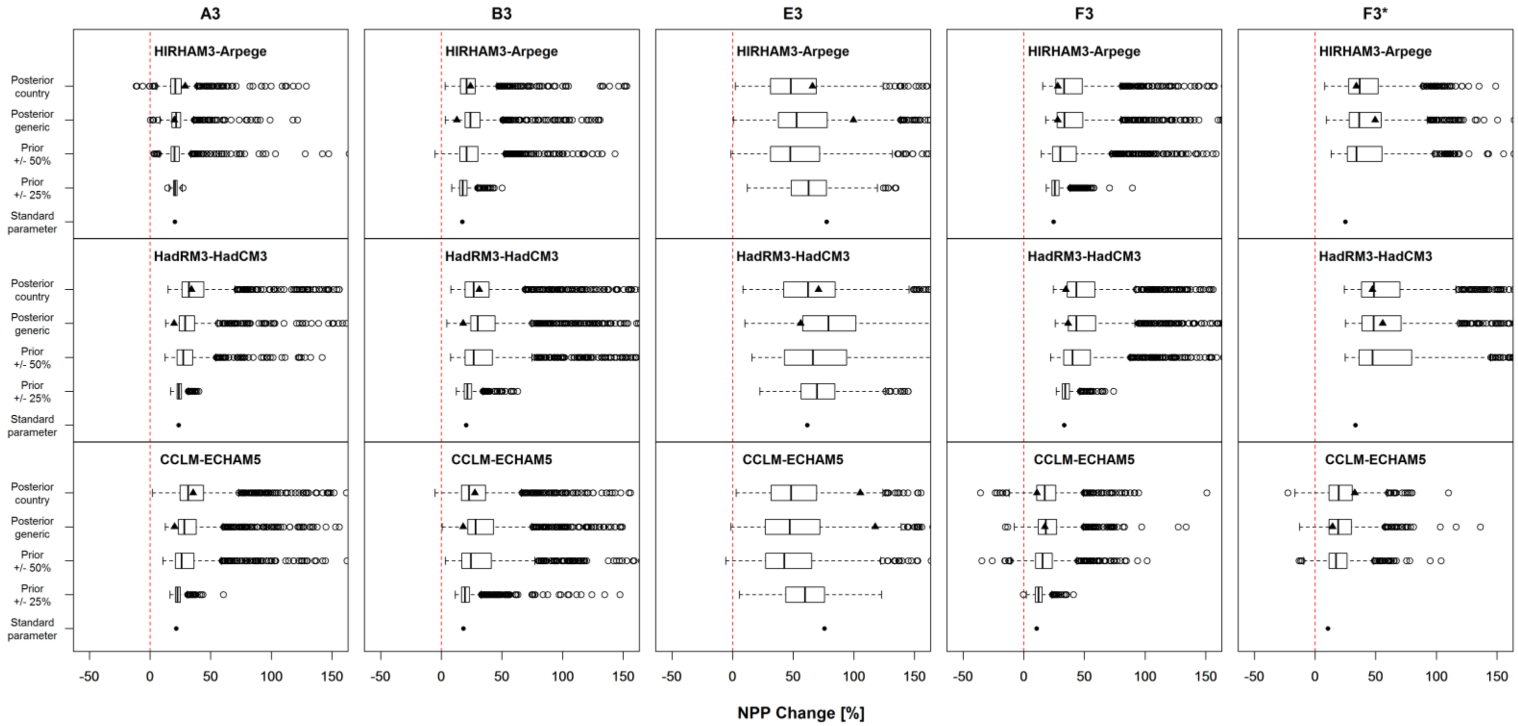

Fig. 4 Same as Fig. 3 but under the assumption of persistent CO_2-effects