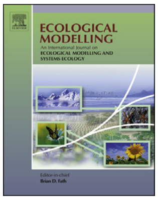
Ecological Modelling 341 (2016) 37–52

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



Calibration and validation of a semi-empirical flux ecosystem model for coniferous forests in the Boreal region

F. Minunno[a,*], M. Peltoniemi[b], S. Launiainen[b], M. Aurela[c], A. Lindroth[d], A. Lohila[c], I. Mammarella[e], K. Minkkinen[a], A. Mäkelä[a]

[a] *Department of Forest Sciences, University of Helsinki, PO Box 27, Helsinki 00014, Finland*
[b] *Natural Resources Institute Finland (Luke), Jokiniemenkuja 1, 01301 Vantaa, Finland*
[c] *Finnish Meteorological Institute, FI-00560 Helsinki, Finland*
[d] *Lund University, Department of Physical Geography and Ecosystem Science, Lund, Sweden*
[e] *Department of Physics, P.O. Box 48, FIN-00014, University of Helsinki, Finland*

ARTICLE INFO

Article history:
Received 16 December 2015
Received in revised form
20 September 2016
Accepted 22 September 2016

Keywords:
Forest modelling
PRELES
Eddy-fluxes
Boreal forest
Bayesian statistics

ABSTRACT

Simple models are less input demanding and their calibration involves a lower number of parameters, however their general applicability to vast areas must be tested. We analysed if a simple ecosystem model (PRELES) can be applied to estimate carbon and water fluxes of Boreal forests at regional scale.

Multi-site (M-S) and site-specific (S-S) calibrations were compared using evapotranspiration (ET) and gross primary production (GPP) measurements from 10 sites. The performances of M-S were similar to S-Ss except for a site with agricultural history. Although PRELES predicted GPP better than ET, we concluded that the model can be reliably used at regional scale to simulate carbon and water fluxes of Boreal forests.

We further found that, in the calibration, the use of a long and carefully collected flux dataset from one site that covers a wide range of climate variability leads to better model performance in other sites as well.

1. Introduction

Biogeochemical flux models quantify the mass and energy exchanges between the atmosphere, biosphere and soil as a function of soil and vegetation characteristics and climate forcing (Meyers and Baldocchi, 1988). Flux models are focal components of forest growth models and dynamic vegetation models (Friend et al., 2014) that describe the interactions and long-term feedbacks between the vegetation cover, soils and the atmosphere. Information about flux rates is also useful for monitoring the current carbon and water balances, such as in national greenhouse gas inventories (Peltoniemi et al., 2015a). Although the physical and physiological processes related to biogeochemical fluxes are theoretically fairly well understood (Farquhar et al., 1980; Monteith, 1981), their reliable quantification in the large geographical scale still remains a challenge. This has been demonstrated by several model comparison studies providing vastly variable predictions (e.g. Medlyn et al., 2011a). For example, a recent comparison of seven dynamic vegetation models concluded that although the net primary productivity (NPP) predictions were very similar, the related vegetation biomass predictions varied vastly, implying that the models also differed in their descriptions of photosynthesis and/or respiration rates for a given vegetation type and biomass (Friend et al., 2014).

The models of ecosystem carbon and water exchange range from complex descriptions of canopy structure accompanied with short sub-daily time steps (Juang et al., 2008; Launiainen et al., 2011; Leuning et al., 1995; Meyers and Baldocchi, 1988; Ogée et al., 2003; Olchev et al., 2008), to big-leaf models often also operating at lower temporal resolution (Kimball et al., 1997; Liu et al., 1997). On one hand the more complex mechanistic models reproduce in detail the processes of ecosystems, potentially covering a variety of responses and interactions, but also dependent on a large number of inputs with relatively high uncertainty (Van Oijen et al., 2013). On the other hand, the more simple summary type models are less input demanding, involve a lower number of parameters, and could more easily be incorporated in larger-scale vegetation models and other applications. However, because of the simplifications, some of the mechanistic interactions generating site-specific differences may have been excluded, establishing a need for site-specific calibration.

The light-use-efficiency (LUE) approach provides a simple model for describing vegetation carbon fluxes and has already been applied to the regional scale in the MODIS algorithm, where the

---

[*] Corresponding author.
E-mail address: francesco.minunno@helsinki.fi (F. Minunno).

http://dx.doi.org/10.1016/j.ecolmodel.2016.09.020

gross-primary productivity (GPP) and NPP are estimated from daily weather data and leaf area index retrieved from remote sensing images (Heinsch et al., 2006). The LUE approach was further developed by Mäkelä et al. (2008) to be suited for boreal and temperate conifers. The resulting model was tested at different sites and it was found to describe daily GPP rather generally (Mäkelä et al., 2008; Peltoniemi et al., 2012). In a recent study, Peltoniemi et al. (2015b) extended this approach to include evapotranspiration (ET) through its coupling to photosynthesis by assuming that GPP is a good proxy of transpiration of coniferous forests that are aerodynamically well-coupled to the atmosphere (Brümmer et al., 2012). They calibrated the resulting model, PRELES, by means of Bayesian analysis applied to eddy-covariance (EC) flux and soil moisture data at two Scots pine–dominated boreal sites. In a separate study Peltoniemi et al. (2015a) also demonstrated that the GPP predicted by PRELES across Finland, using field-based leaf area measurements as structural input, was similar to predictions by the JSBACH dynamic vegetation model (Raddatz et al., 2007) calibrated for Finland. Both predicted lower GPP values than the standard MODIS algorithm, possibly due to leaf area index input data differences.

In model development, model calibration represents a crucial step that strongly affects the reliability of predictions. Process-based models need parameters that are directly related to physiological, functional and structural properties of the system. While detailed process-based ecosystem models that upscale processes from the canopy level to a stand scale, can mostly be calibrated based on scale-appropriate measurements or literature values (i.e. leaf gas-exchange data, soil properties etc.), simpler semi-empirical models often require calibration against ecosystem level data. The calibration is required especially for the parameters where direct measurements are difficult or impossible and must thus be estimated inversely, comparing model outputs with observed data (Hartig et al., 2012; Van Oijen et al., 2005). In environmental sciences large amounts of data (e.g., EC-fluxes, national forest inventory data, remote-sensing data, and physiological measurements) are becoming available for model calibration and validation purposes. At the same time, developments in computational techniques allow to quantify model uncertainties efficiently, analyse model structure and evaluate prediction accuracy and reliability (Minunno et al., 2013a,b; Van Oijen et al., 2011). The EC flux-tower network (Baldocchi, 2008), which already provides more than a decade of continuous measurements, offers a good opportunity to calibrate and test models of carbon and water fluxes by providing model input variables as well as stand and site characteristics.

For the development of a generally applicable, calibrated model with explicitly expressed uncertainty bounds, systematic methods of parameter estimation from data are useful. In ecological models the parameters can usually be assigned a plausible range of variability that should be taken into account in the calibration, rather than finding the over-all best statistical fit of the model to data. Bayesian calibration offers a good method for taking into account such prior distributions which can be modified so as to reduce the uncertainty by systematic comparisons of model predictions with available data (Green et al., 2000; Van Oijen et al., 2005). Recently, calculation methods have been developed to the use of Bayesian methods in combination with sensitivity analysis, error propagation and uncertainty estimates (Minunno et al., 2013b; Van Oijen et al., 2011). Even if, in the carbon cycle field, many model calibrations have been carried out in the last decade, multi-site calibrations are quite rare, especially those that take into account data and parameter uncertainties.

The objective of this study was to assess if PRELES can be used as a tool to estimate the carbon and water fluxes of boreal coniferous forests in Fennoscandia. Firstly, we prepared a comprehensive sensitivity analysis of PRELES and then used the Bayesian framework to calibrate and evaluate the model against data from multiple boreal coniferous sites in Fennoscandia. Using these analyses as basis, we sought answers to three questions: (1) Can we find a generic set of model parameters that adequately performs at all sites? (2) Under what conditions – if any – should the multi-site calibration be used in favour of the site-specific calibration, if both exist for a site? (3) How should data be selected for model calibration to extend model predictions of GPP and ET to a site with no prior data?

2. Materials and methods

2.1. PRELES model

PRELES (PREdict with LESs – or – PREdict Light-use efficiency, Evapotranspiration and Soil water) is a semi-empirical ecosystem model of intermediate complexity developed by Peltoniemi et al. (2015b), in which the dependent variables, GPP (P, gC m^{-2} day^{-1}), ET (E, mm) and soil water (θ, mm), are interlinked by simplified processes, so that GPP influences ET, ET decreases soil water, and soil water restricts GPP and ET during drought. The model operates with one leaf, for which GPP is predicted using a reformulation of LUE-based model of Mäkelä et al. (2008). ET is predicted using an empirical equation utilizing GPP, vapour pressure deficit and radiation. Water balance is depicted with a one-pool model for soil, one pool snow pack model, and a pool of surficial (intercepted) water in the ecosystem. The model works at daily time-step and requires minimal input data. The climatic driving variables are daily mean temperature (T, °C), vapour pressure deficit (D, kPA), precipitation (R, mm) and photosynthetic photon flux density (PPFD, Φ, μmol m^{-2} d^{-1}). The only stand structural information is the fraction of absorbed PPFD (f_{aPPFD}), estimated using the Beer-Lambert law as f_{aPPFD}

$$f_{aPPFD} = 1 - exp^{-kL} \tag{1}$$

where L is the leaf area index (m^2 m^{-2}) and k the extinction coefficient.

A detailed description of PRELES can be found in Peltoniemi et al., 2015b; in Appendix A we briefly outline model structure and provide all the equations.

2.2. Carbon and water flux data

Stand-scale net ecosystem exchange (NEE) of CO$_2$, evapotranspiration and meteorological data from ten boreal coniferous forest sites located in Finland and Sweden were used in this study (Table 1). The sites cover a latitudinal band from 60°N to 67°N with annual mean temperatures ranging from 0.8 to 7.1 °C, and precipitation from ~550 mm to ~850 mm. Leaf area index (LAI) at each site was treated as one lumped LAI, i.e. all the canopy layers were included in one unique layer. The total (all-sided) LAI varies between ~3.8 and ~12 m^2 m^{-2} offering a good possibility to address both climatic and LAI controls on forest GPP and ET. A brief summary of the sites is provided in Table 1, and complete descriptions can be found in the respective References

The NEE and ET were measured above the forest canopies by the eddy-covariance method and the ½ h fluxes computed according to common practices (Aubinet et al., 2012). Gaps in data caused by instrumental failures or methodological issues, such as insufficient turbulent mixing, were gap-filled, and NEE was partitioned into component fluxes before the ½ h data was aggregated into daily averages or sums. The gap-filling of NEE was done using a combination of look-up tables and mean diurnal variability according to Reichstein et al. (2005). The gaps in meteorological data were filled either by linear interpolation or by the mean diurnal variability determined in a 14-day moving window.

Table 1
Site characteristics.

	Lat (deg)	Long (deg)	Elev (m)	Site type	Dominant species	all-sided LAI including understory ($m^2\,m^{-2}$)	Age (yrs)
Hyytiälä	61.51	24.17	180	haplic podzol, mean depth 0.6 m	Scots pine	7.9	40–49
Sodankylä	67.22	26.38	179	haplic podzol, mean depth 1.5 m	Scots pine	3.8	50–160
Flakaliden	64.07	19.27	300	Sandy podzolic till	Norway spruce	9.5	43
Norunda	60.1	17.5	45	Sandy podzolic till	Scots pine, Norway spruce	12.7	ca. 100
Kalevansuo	60.39	24.22	123	Originally ombotrophic dwarf-shrub pine bog, drained in 1969. Fertilized with P and K.	Scots pine	5.7	<40
Knottåsen	61	16.13	320	Sandy podzolic till	Norway spruce	7.0	39
Alkkia	62.11	22.47	153	Former Sphagnum bog drained for agriculture in 1936-38, amended with mineral soil. Regular agricultural fertilization. Afforested in 1971 with Scots pine	Scots pine, very dense understory reflecting high nutrient content of the soil	9.0	32
Skyttorp	60.07	17.5	40	Sandy podzolic till	Scots pine	8.0	60
CAge 12yr	61.51	24.17	170	haplic podzol	Scots pine	7.0	12
CAge 75yr	61.51	24.17	170	haplic podzol	Scots pine	7.9	75

	Annual P (mm)	Annual T (°C)	Years of flux measurements	Ndata GPP	Ndata E	Reference
Hyytiälä	709	2.9	2000–2010	3391	3601	Hari and Kulmala (2005); Kolari et al. (2009)
Sodankylä	527	−0.4	2001–2009	2698	2878	Thum et al. (2008)
Flakaliden	600	2.3	1997, 1998, 2001, 2002, 2007–2009	1414	1653	Kleja et al. (2008)
Norunda	527	5.5	1996–1999, 2003	1476	1499	Lundin et al. (1999); Lindroth et al. (2008)
Kalevansuo	606	4.3	2004–2009	1144	1154	Pihlatie et al. (2010); Lohila et al. (2011); Ojanen et al. (2012)
Knottåsen	613	3.4	2007, 2009	680	699	Berggren et al. (2008)
Alkkia	681	4.1	2002–2004	357	404	Lohila et al. (2007)
Skyttorp	830	7.1	2005	267	282	–
CAge 12yr	709	2.9	2002	235	237	Kolari et al. (2004)
CAge 75yr	709	2.9	2002	204	198	Kolari et al. (2004)

The GPP was separated from the measured NEE as GPP = −NEE + R_e, where the ecosystem respiration R_e is (Kolari et al., 2009; Reichstein et al., 2005)

$$R_e = R_{10} Q_{10}^{[(T-10)/10]} \tag{2}$$

The R_{10} (μmol $m^{-2}\,s^{-1}$) represents the temporally varying base respiration rate at 10 °C temperature (T) and Q_{10} (unitless) represents the short-term temperature sensitivity, which is assumed constant in time but can vary among the sites. The R_e model parameters were determined for each site by a non-linear least squares fit of Eq. (2) to nighttime NEE measured in turbulent conditions (friction velocity u^* exceeding an empirically defined site-specific threshold) using measured soil or air temperature as an independent variable. The Q_{10} was first computed by pooling all available growing season data, defined here as May–Sept (Jun-Aug in Sodankylä due to northern location). Secondly, Q_{10} was fixed and the temporal variability of R_{10} was determined by fitting Eq. (2) to data in four-day non-overlapping windows, and linearly interpolating between the window centres. Finally, R_e was computed by extrapolating the obtained regression model to daytime temperatures, allowing the GPP to be approximated for each ½ h period.

After gap-filling, the fluxes and meteorological variables were aggregated at daily time step. A quality flag (F) varying between 0 and 1 was assigned to each day to represent the fraction of gap-filled data used to compute the daily value, and used in later analysis to weight the observations error (see "*Model calibration and model comparison*" section).

2.3. *Overview of the model analyses*

Bayesian calibration (BC) and Bayesian model comparison (BMC) were used to quantify the uncertainty in model parameters and model structure. For a comprehensive understanding of model behaviour, the Bayesian analyses were combined with a model-data mismatch analysis and a global sensitivity analysis following the framework proposed by Van Oijen et al. (2011) and improved by Minunno et al. (2013b).

The work consisted of three analyses, where we compared multi-site (M-S) and site-specific (S-S) calibrations. M-S has the advantage that the data involved in the calibration covers a wider variability in terms of climate and forest structure since they come from different sites, including measurement and other errors which may or may not partially cancel out when all data are used in parameter inference. In contrast, S-S could provide good correspondence to local data, but may not be spatially generalizable, firstly because the processes may not be generic, and secondly because the risk of bias increases with less measurements. More specifically, we conducted the following comparative analyses:

Analysis #1–'Global or local' – *Is PRELES generic enough to be applied at regional scale using one generic calibration?* We compared M-S and S-S calibrations, in order to test if PRELES is a model of general applicability, and to test how well one calibration can predict ecosystem fluxes. In total, 11 BCEs were performed; the model was independently calibrated for each of the ten sites (S-S calibrations) and a multi-site calibration was achieved using data from all the sites in one BC. Parameter estimates and model outputs from the M-S and S-Ss were compared in order to detect any significant differences between the calibrations.

Analysis #2–'Forward prediction' – *Is a multi-site calibration better than site-specific in predicting fluxes for a site for which data are already available?* The aim of this exercise was to compare M-S and S-S calibrations in predicting future carbon and water fluxes of a site for which data are already available. For this analysis the datasets of each site were split in two parts, the first half was used for model calibrations (calibration dataset) and the second half for the comparison (comparison dataset). Similarly to Analysis #1, but using just the shorter calibration datasets, 11 BCEs were performed. In addition, 10 model comparisons were carried out, one for each site, using the comparison dataset and outputs from the M-S and the S-S versions of PRELES.

Analysis #3–'New site' – *To predict GPP and ET for a new site, should a single site calibration or a multi-site calibration be used?* In this analysis we compared the M-S and S-S to test which calibration is more suitable for predicting ET and GPP for a site where the model has not been calibrated before. In this case we used 10 single site and 10 multi-site calibrations of the model. First, PRELES was calibrated for each site and then used to predict the fluxes of the other sites (site-specific calibrations). Second, 10 M-Ss were carried out excluding each site in turn from the calibration process, and the M-S model versions were run for the site excluded from the calibration; so for each site we had model predictions from a multi-site version independently calibrated. The M-S predictions were combined to be compared with the predictions of the site specific calibrations. Finally we carried out 10 comparisons between the multi-site and the single site versions of the model. Data used to assess the performance of these calibrations was the full data set of each site, which was always excluded from the calibrations.

2.4. Sensitivity analyses

Peltoniemi et al. (2015b) found that variation of output sensitivity to certain model parameters is strongly regulated by soil moisture status. At other sites, other parameters could become sensitive due to varying site conditions and weather inputs. Here we calculated sensitivities with a global sensitivity analysis method (Morris, 1991) at all sites in order to resolve what parameters are globally the most sensitive. Morris method (Appendix B) allows us to determine which parameters have linear or additive effects on model output and which ones have non-linear effects and interact with other parameters. The analyses were carried out for the parameter space defined by the minimum and maximum values (Table 2), corresponding to the parameter space of the prior distribution of the Bayesian calibrations. The prior parameter space was chosen because it helps us to understand model behaviour in relation to the dataset used in the calibration process, thus helping us to interpret the results of the Bayesian calibrations.

In addition to the Morris method, we analysed the response of PRELES outputs to LAI that is the stand structural variable used to drive the model. The Morris method and an analysis of sensitivity to LAI were used at each site in order to test if the results change with the climatic conditions in Fennoscandia. The details of the methods and results are provided in Appendix B.

2.5. Model calibration and comparison

Bayesian calibration provides an updated joint probability distribution of the parameters (*posterior* distribution) combining existing parameter knowledge (*prior* distribution) and new information enclosed in the data (*likelihood*).

The 12 most influential parameters on PRELES outputs (Peltoniemi et al., 2015b) were included in the BC (Table 2); the priors were uniformly distributed between the minimum and maximum values reported (Table 2).

A variable number of GPP and ET data points were available for the BC at each site (Table 1). The data, that are daily sums of half hour measurements, were considered to be normally distributed so the likelihood was a Gaussian distribution. The error of flux measurements varies with their magnitude (Richardson et al., 2008). The data were grouped in two bins and the standard deviatons of the two groups were included in the calibrations as well as the threshold that separated the two bins.

The daily fluxes used in the calibration included some gap-filled data. In order to evaluate the impact of gap-filled data on parameter estimates, we carried out different calibrations including data with a flag F of 0, 0.3, 0.5, 0.7 and 0.9, where F is the fraction of gapfilled data in a day. Parameter estimates and model predictions were really similar independently of the data used to calibrate the model (data not shown); i.e., gap filled data seem to be consistent with measurements. For the analyses conducted in this work we used data with a quality flag lower than 0.7.

The posterior distributions achieved by the Bayesian calibrations were sampled by means of Markov chain Monte Carlo simulations (MCMC). We used the differential evolution Markov Chain with fewer chains and snooker updater algorithm (DE-MCzs) (Ter Braak and Vrugt, 2008).

In the Analysis #2 and #3 we compared the performances of different model calibrations. For the evaluations, beside more classical approaches (see next Section), we applied the Bayesian model comparison. BMC is an extension of Bayesian calibration that evaluates different model structures in the light of their relative likelihoods, taking into account parameter uncertainty and provides the probability of each model of being correct (Kass and Raftery, 1995). Applications of the BMC to forest models can be found in Van Oijen et al. (2013, 2011) and in Minunno et al. (2013b). Note that in this work we were comparing model calibrations instead of different model structures.

2.6. Model-data mismatch

The Bayesian approach jointly uses the *prior* and all the data to calculate the posterior distribution, but it provides little information about the strengths and weaknesses of a model. On the contrary, the more classical analyses of the mismatch between the simulated and the observed data give useful insights into model behaviour. In particular the decomposition of the mean squared error (*MSE*) provides indications about the accuracy and the precision of the predictions (Minunno et al., 2013b; Van Oijen et al., 2011).

MSE can be decomposed into three components: the bias error, the variance error and the correlation error (Kobayashi and Salam, 2000) (Eq. (3)).

$$MSE = \overline{(S-O)^2} = \left(\bar{S}-\bar{O}\right)^2 + (\sigma_S - \sigma_O)^2 + 2(\sigma_S\sigma_O)(1-r) \qquad (3)$$

where O are the observations, S the model predictions, σ_O and σ_S are the standard deviation of the observed and simulated data respectively and r is the correlation between the O and S.

The bias error quantifies the distance of the predictions from the data; the variance error expresses if the model is able to catch data variability; the correlation error indicates if the model is able to reproduce the pattern of data fluctuations. The latter component expresses the lack of positive correlation between the observed and simulated data and is weighted with standard deviations, there-

Table 2
List of parameters used in the calibration.

Name	Symbol	Units	Minimum	Maximum
Potential light use efficiency	β	gC molPPFD^{-1}	0.2	2.5
Delay parameter for ambient temperature response	τ	–	1	25
Threshold for state of acclimation change	X_0	°C	−20	20
Acclimation state maximum	S_{max}	°C	2.3	30
Sensitivity parameter for VPD response	κ	kPA^{-1}	−1	−0.001
Light modifier parameter	γ	molPPFD^{-1} m^{-2}	1.03e-4	5.03e-1
Threshold for linear decrease of &z.hfl;$_{W,P}$	ρ_P	–	0	0.999
Transpiration parameters	α	mm (gC m^{-2} kPa$^{1-\lambda}$)$^{-1}$	1e-6	10
Parameter adjusting water use efficiency with VPD	λ	–	1e-4	1.2
Evaporation parameter	χ	mm molPPFD^{-1}	0	2.5
Threshold for linear decrease of &z.hfl;$_{W,E}$	ρ_E	–	0	0.999
Parameter adjusting water use efficiency if soil water is limiting GPP (&z.hfl;$_{W,P}$)	υ	–	1e-4	5

fore there is an overlap between the variance and the correlation error (Kobayashi and Salam, 2000). In practice this *MSE* component seems to capture all types of error, such as random errors, left after accounting for the bias and the differences in the variances. *MSEs* were calculated for both GPP and ET and for each site.

The predictive power of the model was also assessed through the Nash Suttclife coefficient (*ME*, Eq. (4)), also known and the modelling efficiency (Janssen and Heuberger, 1995).

$$ME = 1 - \frac{\sum_{i=1}^{N}(S_i - O_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \tag{4}$$

where S_i and O_i denote the predicted and observed value i; N is the number of observed values; $\bar{O}$ is the mean of the observed values.

ME can vary between $-\infty$ and 1; *ME* = 1 indicates a perfect match of model predictions to the data. If *ME* = 0 model predictions are as accurate as the mean of the data. *ME* assumes negative values when the observed mean is more accurate than the model predictions.

3. Results

3.1. Analysis #1 – 'global or local' –

The data were highly informative in determining the values of the parameters that were assessed highly influential in the Morris sensitivity analysis (Appendix B), as demonstrated by the constrained posterior distributions (Fig. 1a) of the parameters compared with their priors (Table 2). In the multi-site calibration even the less important parameters were well constrained in the posterior distributions, however, a lot of uncertainty remained in some of these parameters in the site-specific calibrations (Fig. 1b). These include υ for Alkkia, Kalevansuo, CAge12yr, CAge75yr and Skyttorp site-specific calibrations; ρ_E for Alkkia, CAge12yr, CAge75yr, Knottasen, Skyttorp and Flakaliden; ρ_P for CAge75yr and Skyttorp (Fig. 1b). Parameter estimates across the different calibrations (i.e., S-S and M-S) were consistent for the most influential parameters, in particular for α, γ and χ (Fig. 1a), whereas differences occurred in estimates of the parameters at which model outputs are less sensitive (Fig. 1b). In the CAge12yr site-specific calibration γ, X_0 and β marginal posterior distributions were quite different from the rest of the calibrations.

Bayesian calibration provides a joint posterior distribution of model parameters, i.e. BC also considers the interactions between parameters. This kind of information is not derivable from the marginal posterior distributions (Fig. 1), but it can be expressed through the correlations (r) between parameters calculated from the posterior sampled by the MCMC. We report the correlation matrices of the 11 calibrations of Analysis #1 in Appendix C, while here we briefly summarize the results showing the correlations between the parameter correlations of M-S and the S-S calibrations (Fig. 2). The full posterior distributions can be downloaded at the link provided in the APPENDIX D. Parameter correlations of M-S were really similar with the Hyytiälä, Flakaliden, Kalevansuo, Norunda and Knottasen site specific calibrations. CAge 12yr and CAge 75yr were the site specific calibrations for which parameter calibrations differed most from the M-S (Fig. 2).

We evaluated model performances in terms of R^2, slopes (with intercept forced to 0) and modelling efficiency (*ME*) of the simulated *vs.* observed data, calculated for each calibration and each model output (i.e., GPP and ET) at daily time step (Table 3). The predictions were generated using the maximum *a posteriori* (MAP, i.e. the parameter vector that had the highest posterior distribution) parameter vectors of M-S and S-S. The variance explained by the model was higher for GPP than for ET, both being in most of the cases higher than 70% (R^2 of Table 3); however the model tended to underestimate carbon and water fluxes (slopes lower than 1) (Table 3). *ME* was always higher than 0.6 and in many cases close to 1 (Table 3). Model fit to the Flakaliden data was generally rather poor. Furthermore, the multi-site calibration significantly underestimated evapotranspiration at Alkkia site (slope = 0.62). In general, after BC, model outputs were characterized by low uncertainty (not shown in the plots).

In model predictions of GPP and ET for each site, the differences between the multi-site and site-specific calibrations were small in most of the cases (Fig. 3a and b). The exception is the evapotranspiration at the Alkkia site, where ET by the multi-site calibration was clearly different from the site-specific calibration prediction (Fig. 3a).

Mean squared errors were calculated for GPP and ET and decomposed to the bias, variance and correlation errors for the multi-site and site-specific calibrations (Fig. 4a and b). For the site-specific calibrations the main component of model error was the correlation error, while the other two components were negligible (Fig. 4a and b). Also the MSEs of the multi-site calibration predictions were mainly constituted from the correlation error (Fig. 4a and b); however the other two error components were significant at some sites, varying between 10 and 30% of MSE. ET predictions at the Alkkia site for the multi-site calibration were characterized by the highest bias error (i.e. 40% of MSE, Fig. 4b).

Both M-S and S-S calibrations showed robust performances in predicting the photosynthetic activity of boreal forests also at annual time step (Figs. 5a and 7a); while the model was less accurate in reproducing the annual evapotranspiration (Figs. 5b and 7b). Note that to compute the 'annual' fluxes, the daily fluxes were summed only if the quality flag was lower than 0.7 (i.e. at maximum 70% of 1/2 h fluxes were missing and gap-filled for that particular day), and the numbers in Figs. 5 & 6 are not representative for true annual balances. PRELES was able to catch the pattern of GPP interannual fluctuations for the sites with the long-term datasets (i.e., Hyytiälä, Sodankylä, Flakaliden, Norunda and Kalevansuo) (Fig. 6a), but at Flakaliden for some years (i.e., 1997, 2002) the relative difference between the observed and modelled GPP was around 50%.

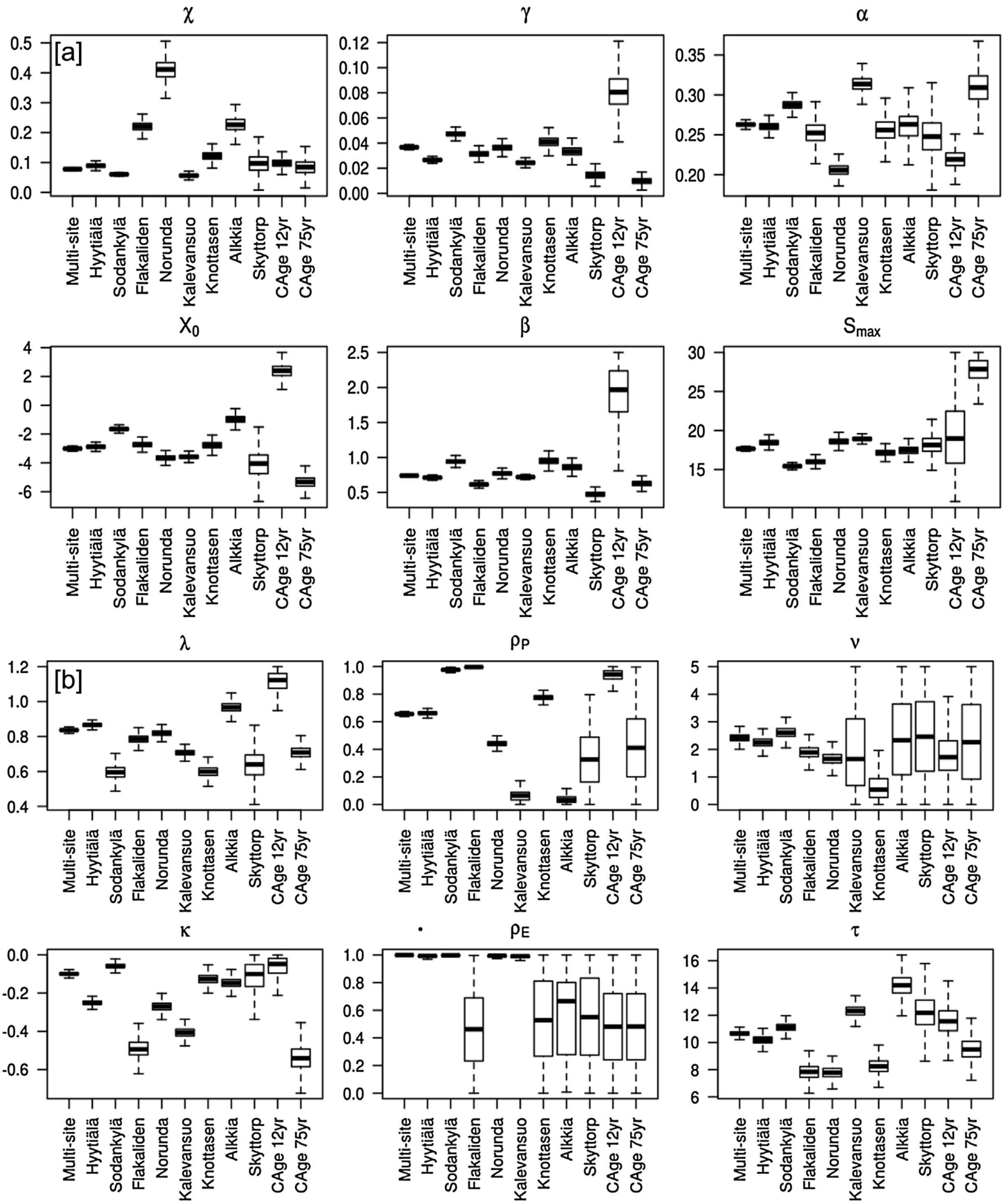


Fig. 1. (a) Marginal posterior distributions of high sensitivity parameters obtained through the multi-site calibration and the site-specific calibrations. The Morris method was used to rank the parameters according to their impact on model outputs. (b) Marginal posterior distributions of medium and low sensitivity parameters obtained through the multi-site calibration and the site-specific calibrations.

The M-S and S-S annual prediction were really similar (Figs. 5a,b, 6a and b), apart from the GPP at Flakaliden (Fig. 6a).

3.2. *Analysis #2 – 'forward predictions' –*

M-S and S-S calibrations were evaluated at each site using the validation dataset and considering both output fluxes (i.e., ET and GPP) at the same time. In seven sites S-S had 100% probability of being the best model version, while in the other sites BMCs supported the M-S calibration (Table 4). At four of the five sites with long-term datasets (i.e. Hyytiälä, Sodankylä, Flakaliden and Norunda) the S-S calibration performed better than the M-S. In general the NRMSEs calculated with the two types of calibration did not differ substantially (Fig. 7a and b). At Skyttorp the GPP NRMSE of S-S was 30% higher than the GPP NRMSE of M-S, while at Flakaliden the M-S calibration error was significantly higher (Fig. 7a). The NRMSEs of the evapotranspiration M-S were always higher than those of the S-S calibration, except for Skyttorp (Fig. 7b) and Cage

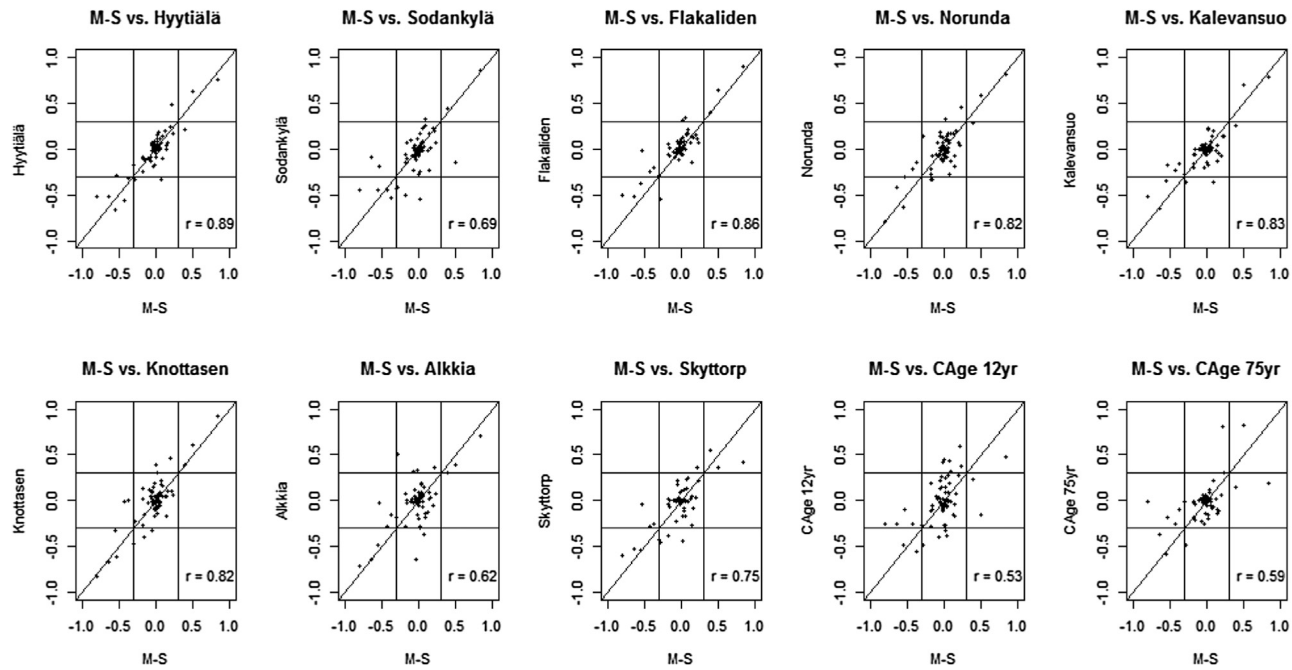


Fig. 2. The parameter correlations of M-S (x-axis) were plotted against the parameter correlations found in the S-Ss (y-axes). In the bottom right corner the correlation (r) between the M-S and S-S correlation.

Table 3
R^2, slopes (with intercept forced to 0) and modelling efficiency (ME) calculated for the multi-site and site-specific calibration of Analysis #1–“Global or local”.

	GPP						E					
	mutli-site			site-specific			mutli-site			site-specific		
	R^2	slope	ME	R^2	slope	ME	R^2	slope	ME	R^2	slope	ME
Hyytiälä	0.96	0.98	0.96	0.96	0.97	0.96	0.89	0.90	0.88	0.88	1.02	0.86
Sodankylä	0.89	0.85	0.88	0.89	0.87	0.88	0.75	0.79	0.73	0.75	0.94	0.73
Flakaliden	0.79	1.15	0.63	0.75	0.77	0.73	0.67	0.95	0.61	0.69	1.04	0.58
Norunda	0.89	0.98	0.88	0.87	0.92	0.87	0.82	0.89	0.81	0.82	1.02	0.80
Kalevansuo	0.93	0.96	0.93	0.94	0.98	0.94	0.86	0.86	0.86	0.88	1.00	0.87
Knottåsen	0.91	0.80	0.85	0.89	0.88	0.87	0.89	0.79	0.85	0.86	1.07	0.82
Alkkia	0.89	0.86	0.87	0.80	0.87	0.78	0.82	0.66	0.63	0.71	0.99	0.66
Skyttorp	0.80	0.98	0.79	0.81	0.94	0.81	0.72	0.98	0.69	0.72	0.95	0.69
CAge 12yr	0.74	0.93	0.73	0.84	0.90	0.83	0.71	1.05	0.63	0.75	0.98	0.72
CAge 75yr	0.93	1.14	0.85	0.93	0.95	0.93	0.88	0.93	0.88	0.91	1.01	0.91

75yr. At Alkkia the M-S had a NRMSE of about 60%, 20% higher than the error of the S-S calibration; on the contrary, at Skyttorp the NRMSE of S-S was 20% higher than the M-S NRMSE. For the rest of the sites the ET NRMSEs of the two calibrations differed less than 5%.

Table 4
Results of the BMC between multi-site and site-specific calibrations for the Analysis #2–“Forward prediction”.

	multi-site	site-specific
Hyytiälä	0	100
Sodankylä	0	100
Flakaliden	0	100
Norunda	0	100
Kalevansuo	100	0
Knottåsen	100	0
Alkkia	0	100
Skyttorp	0.49	0.51
CAge 12yr	100	0
CAge 75yr	0	100

Table 5
Results of the BMC between multi-site and single site calibrations for the Analysis #3–“New site”.

	multi-site	single site
Hyytiälä	0	100
Sodankylä	100	0
Flakaliden	100	0
Norunda	100	0
Kalevansuo	100	0
Knottåsen	100	0
Alkkia	100	0
Skyttorp	100	0
CAge 12yr	100	0
CAge 75yr	100	0

3.3. Analysis #3 – ‘new site’ –

According to the BMC, the Hyytiälä S-S calibration had 100% probability of being better than the M-S calibration on every site; on the contrary, the M-S calibrations were always better than the other S-S calibrations (Table 5). The normalized root mean squared errors calculated for Analysis #3 are consistent with the BMC probabili-

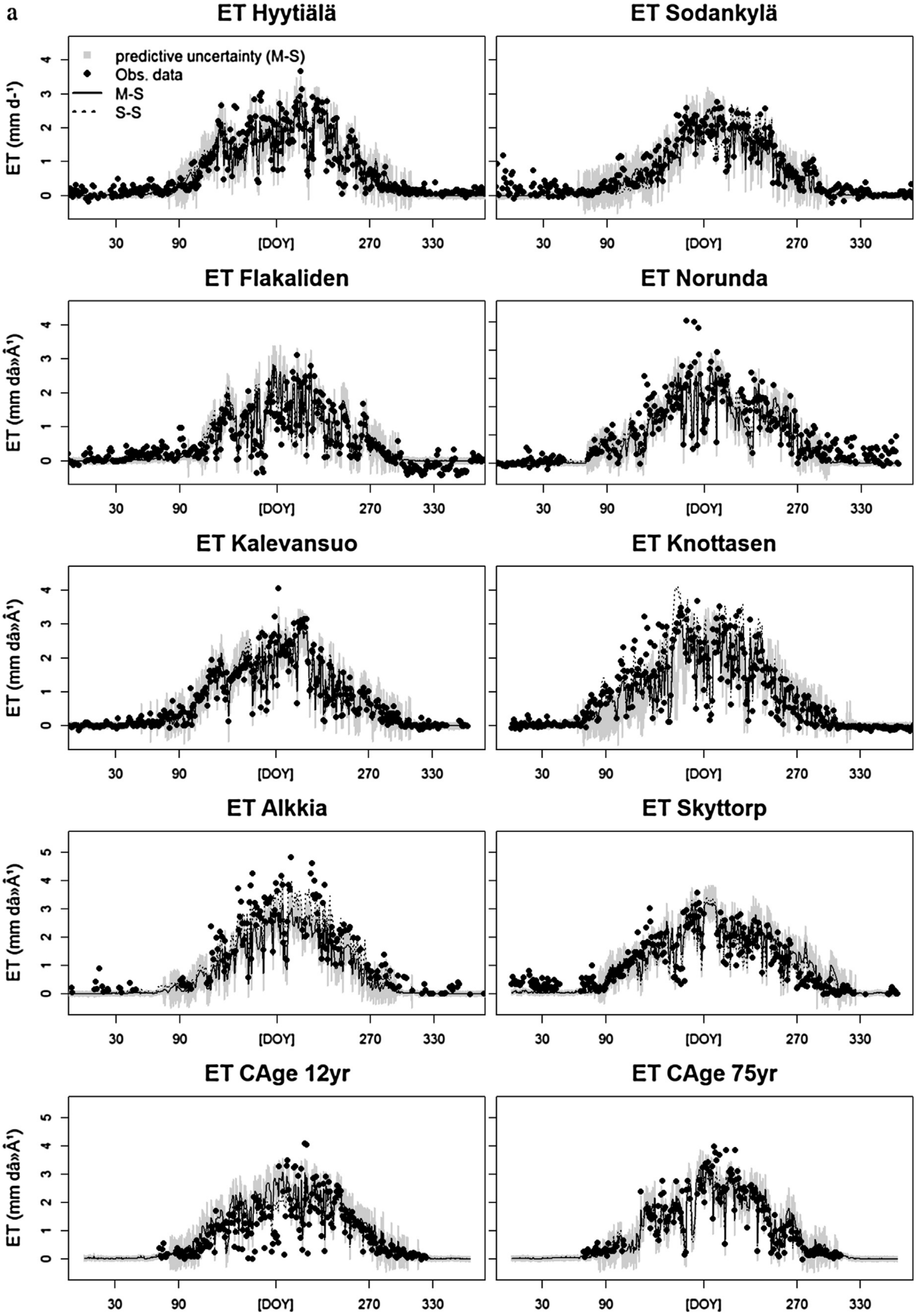


Fig. 3. (a) Daily evapotranspiration at each experimental site for a year randomly selected from the dataset. Sites are ordered according to the number of data points available for model calibration. Dots represent the observations. The lines are PRELES predictions generated by MAP (the maximum a *posteriori* parameter vector); the dashed line is the output from the site-specific calibrations, while the continuous lines represent the multi-site calibration. The grey area represents the M-S predictive uncertainty given by parameter uncertainty and the observation error. (b) Daily gross primary production at each experimental site for a year randomly selected from the dataset. Sites are ordered according to the number of data points available for model calibration. Dots represent the observations. The lines are PRELES predictions generated by MAP (the maximum a *posteriori* parameter vector); the dashed line is the output from the site-specific calibrations, while the continuous lines represent the multi-site calibration. The grey area represents the M-S predictive uncertainty given by parameter uncertainty and the observation error.

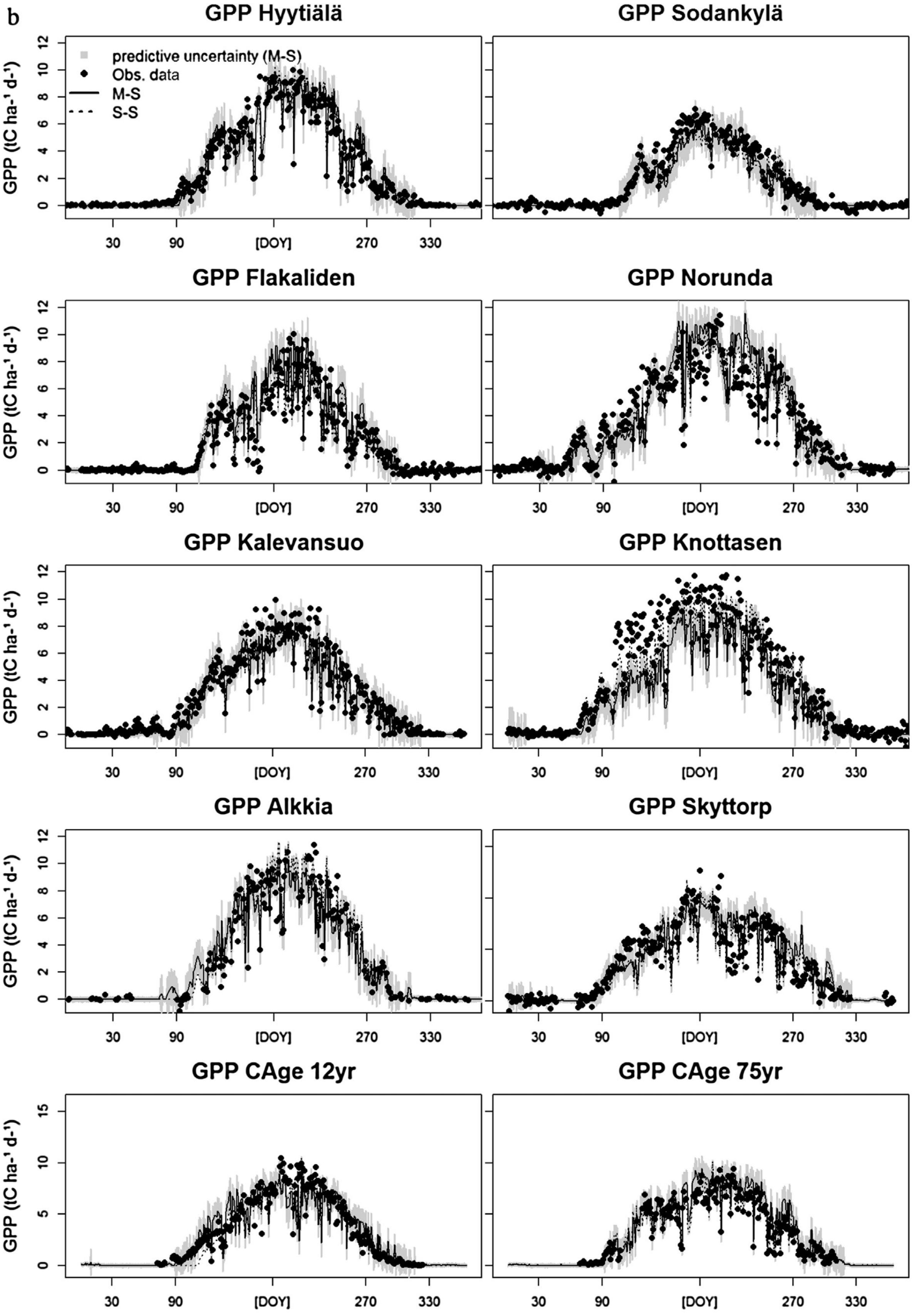


Fig. 3. (*Continued*)

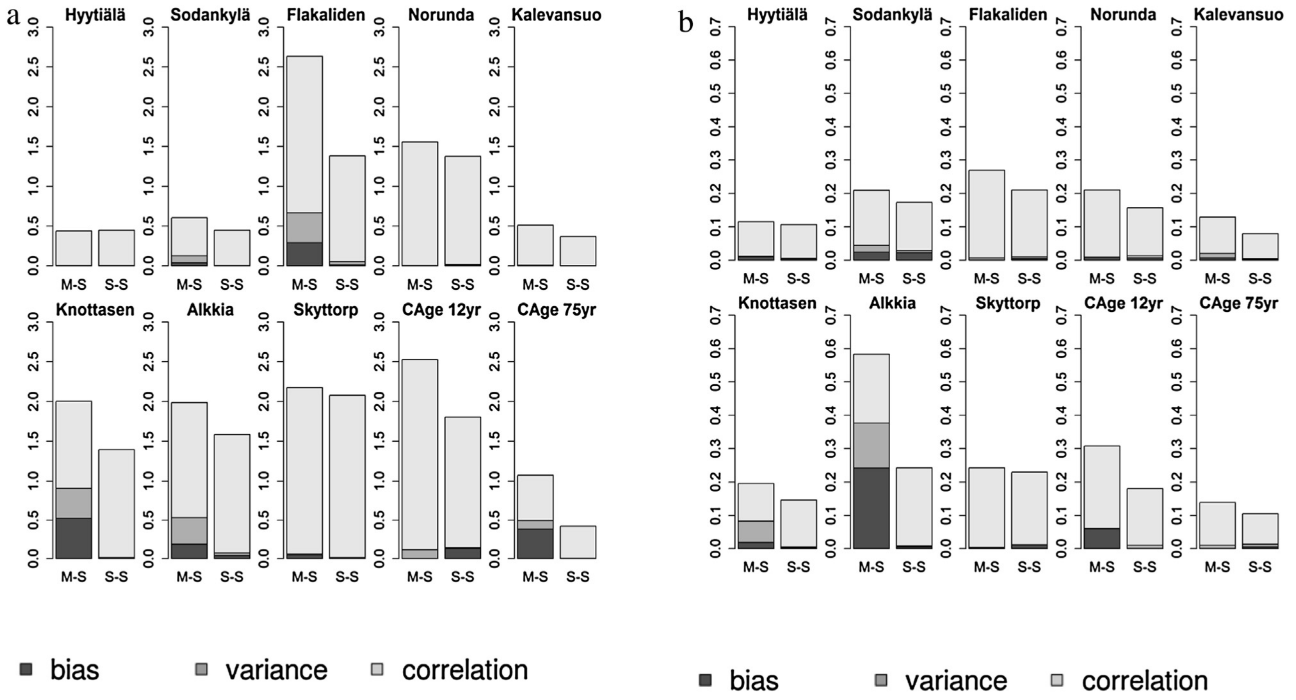


Fig. 4. (a) Mean squared error decomposition for GPP. The first bar in the plots is the *MSE* calculated with PRELES outputs generated from the multi-site calibration (M-S), while the second bar is the error of the site-specific calibration (S-S). The maximum values of the y-axes were set to the square of the mean of the GPP observed values of all the sites. (b) Mean squared error decomposition for ET. The first bar in the plots is the *MSE* calculated with PRELES outputs generated from the multi-site calibration (M-S), while the second bar is the error of the site-specific calibration (S-S). The maximum values of the y-axes were set to the square of the mean of the ET observed values of all the sites.

ties (Fig. 8, Table 4). The ET and GPP NRMSEs of the M-S calibration were slightly higher than those of Hyytiälä S-S calibration, but the differences are negligible (Fig. 8). The NRMSEs of the other S-S calibrations were always higher than the errors generated from the M-S calibration (Fig. 8).

4. Discussion

Evaluating model performances in the light of site-specific calibrations and a multi-site calibration gives useful information about the general applicability of the model. Many models, such as

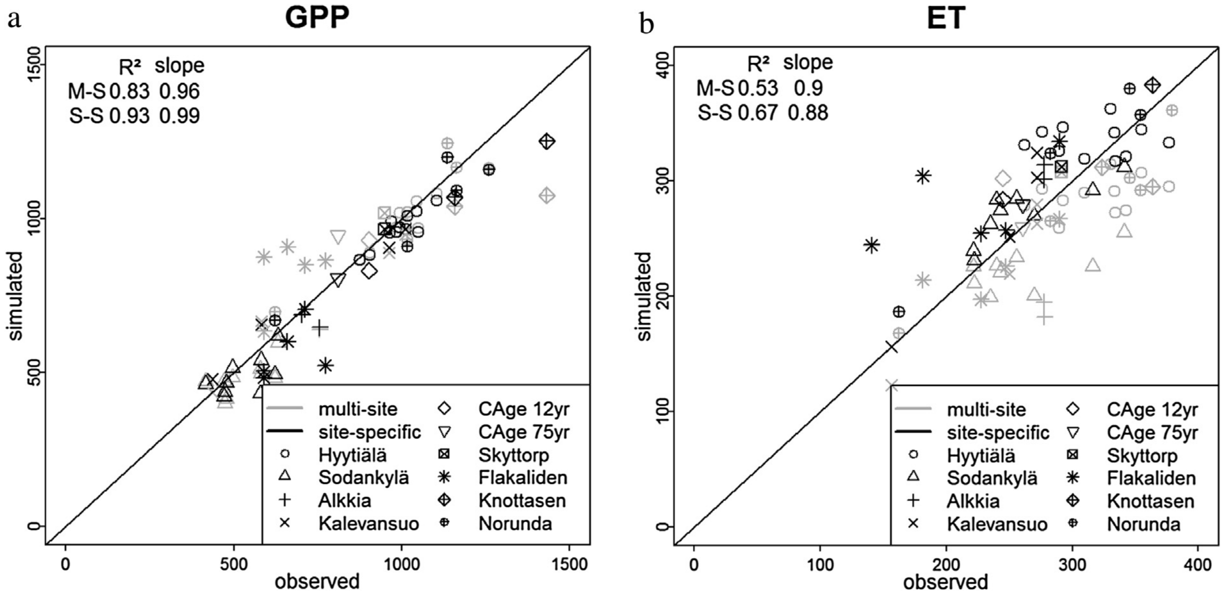


Fig. 5. (a) Observed versus simulated annual gross primary production. Each symbol corresponds to different site; while the colours, grey and black, refer to the multi-site (M-S) and site-specific (S-S) calibration, respectively. The daily observed and simulated data were summed to obtain the annual GPP values in the figure only if the quality flags of the daily GPP measurements were lower than 0.7. In the left up corner of the plotting area a table with the R^2 and slopes are reported for M-S and S-S. (b) Observed versus simulated annual evapotranspiration. Each symbol corresponds to different site; while the colours, grey and black, refer to the multi-site (M-S) and site-specific (S-S) calibration, respectively. The daily observed and simulated data were summed to obtain the annual ET values in the figure only if the quality flags of the daily ET measurements were lower than 0.7. In the left up corner of the plotting area a table with the R^2 and slopes are reported for M-S and S-S.

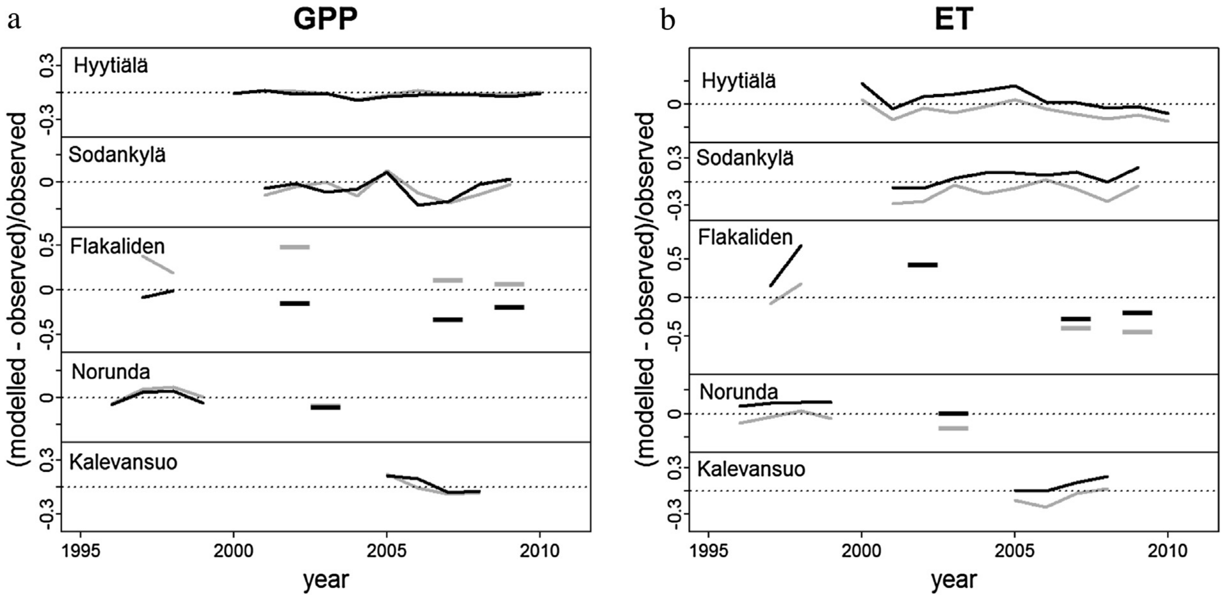


Fig. 6. (a) Relative differences (i.e., y = (observed – modelled)/observed) between the modelled and observed annual sums of gross primary production; note that only days with flag lower than 0.7 were considered for annual sum calculations. The grey line refers to the predictions generated with the multi-site version, while the black line refers to the site-specific calibration. (b) Relative differences (i.e., y = (observed – modelled)/observed) between the modelled and observed annual sums of evapotranspiration; note that only days with flag lower than 0.7 were considered for annual sum calculations. The grey line refers to the predictions generated with the multi-site version, while the black line refers to the site-specific calibration.

Dynamic Global Vegetation Models, assume that the concept of plant functional type is valid and can be used for parameterising not only submodels of GPP and ET like here, but in fact the entire dynamics of the PFT and the related vegetation type. However, very few studies tested model performances against data to compare site-specific and multi-site parameterisations (Kuppel et al., 2012). In contrast, there has been increasing concern in the literature about the justification of this assumption (e.g. Reichstein et al., 2014). For example, Reichstein et al. (2014) suggested that a set of individual plant traits should be taken as a basis, instead of plant functional types that may consist of species with some shared and some separate traits. In fact, they used a simple version of the LUE model as an example of observed wide variability of parameters within a plant functional type.

With the analyses carried out in this study we comprehensively tested PRELES analysing its strenghs and weaknesses. In the following subsections we discuss the possible applications of the model answering the three questions posed in the introduction. In the first two Subsections (4.1 and 4.2) we analyse the behaviour of PRELES; in 4.3 we discuss the model applicability to a site when data are already available; finally in 4.4 we draw conclusions on the regional applicability of the model.

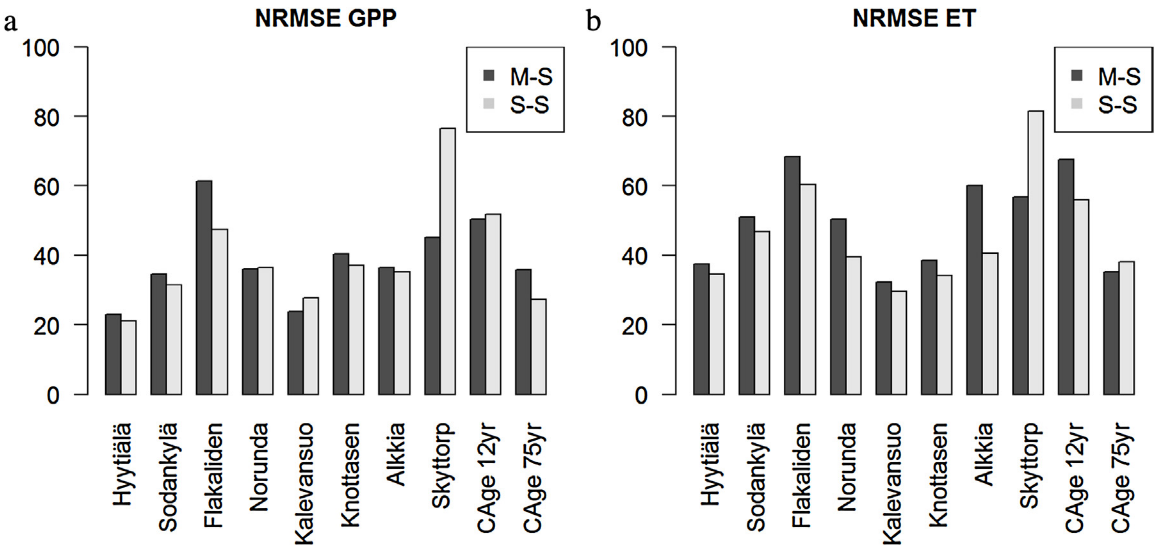


Fig. 7. (a) Normalized root mean squared errors, for GPP. MSEs were normalized using the standard deviations of the observations. Sites are ordered from left to right according to the number of data points available for model calibration and evaluation. M-S and S-S refer to the multi-site and the site-specific calibration, respectively. (b) Normalized root mean squared errors, for ET. MSEs were normalized using the standard deviations of the observations. Sites are ordered from left to right according to the number of data points available for model calibration and evaluation. M-S and S-S refer to the multi-site and the site-specific calibration, respectively.

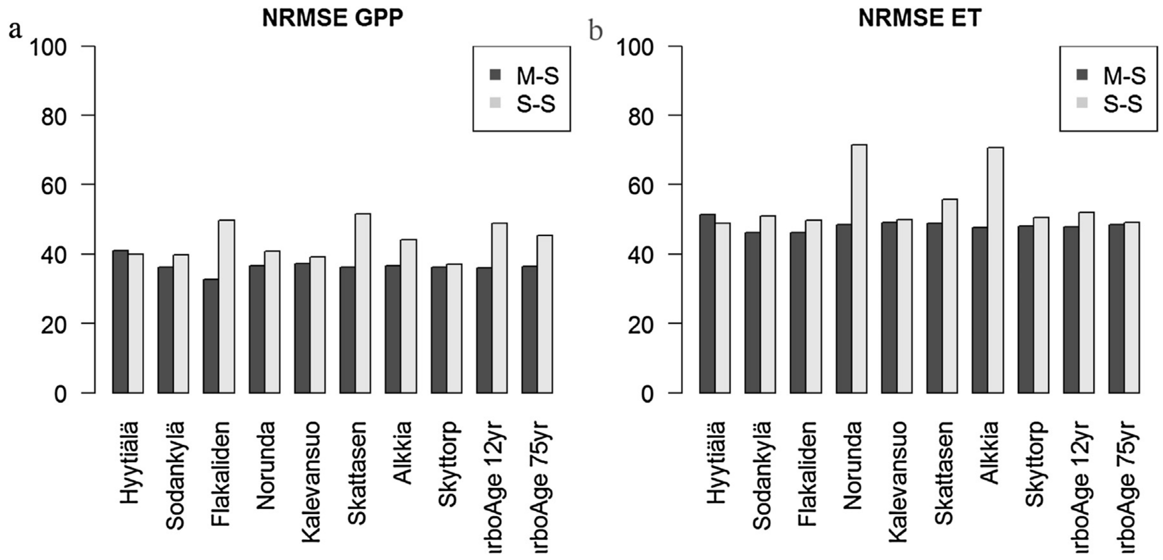


Fig. 8. (a) Normalized root mean squared errors, for GPP. MSEs were normalized using the standard deviations of the observations. Sites are ordered from left to right according to the number of data points available for model calibration and evaluation. (b) Normalized root mean squared errors, for ET. MSEs were normalized using the standard deviations of the observations. Sites are ordered from left to right according to the number of data points available for model calibration and evaluation.

4.1. On the general applicability of the model

PRELES is a simple model, with a strong empirical component, however our results showed that a generic calibration can be used to estimate the gross primary production and the evapotranspiration of all the sites considered in this study. The eddy-covariance method measurement has a tendency to average over sites, because the footprint area is larger than the actual central site used for model parameterisation. The forest landscape in Finland and Sweden is quite heterogeneous, such that the footprint area could cover different soil types and species. We believe that this could explain some but certainly not all of the similarity of the sites included here. Particularly, the similarity between the upland and wetland sites does not seem expected. A more thorough analysis of the footprint area and the contribution of the surroundings would be required to asses this quantitatively, which we believe is beyond the scope of this study.

It is also possible that the environmental responses of GPP and ET are genuinely relatively independent of site and species, once the f_{APAR} of the stand is known. The latter carries information about stand structure which is known to be strongly dependent on both site and species (Medlyn et al., 2005; Duursma et al., 2009) and may also develop dynamically in a site and species specific manner, which was not considered here. This independence has also been found in some previous studies, especially in the case of GPP (Duursma et al., 2009; Medlyn et al., 2005), although other studies have found the opposite (Kaminski et al., 2012; Chen et al., 2014; Reichstein et al., 2014). A lack of independence of site could also be caused by model inadequacy (rather than simplicity), where parameters are needed to compensate for shortcomings in model structure. A typical cause for this in many GPP studies could be a sub-optimal representation of the seasonal cycle (Reichstein et al., 2014; Chen et al., 2014; Richardson et al., 2012). Increasing model complexity would also probably increase the need to make the description more species and site specific (such as in Kasurinen et al., 2014). The question is then, how much of the variability could be explained with a generic model, and how large a gain could be obtained by increasing complexity. We have found in previous studies that even if parameters vary between sites as a result of site-specific fitting, the processes can be (almost) equally well represented by a generic parameterisation, because there are mutual correlations between the parameters (Mäkelä et al., 2008). This could of course be interpreted as opposite errors canceling out each other. Importantly, it could also signify a tendency of ecosystems to evolve towards essentially similar environmental responses, albeit through different structural and functional adaptations (e.g. Valladares et al., 2002). In this study, we conclude that a large percentage of the variability in GPP and ET can be explained using a generic parameterisation. Part of the remaining variability could likely be explained by increasing the complexity of the model. One should bear in mind, however, that the role of the complexities and differences between sites and species may become more pronounced if the environment changes. The need of complexity in such cases could be studied, for example, by investigating how far a complex model can be reduced in such a situation.

Model performances obtained using the multi-site calibration were similar to those achieved by the site-specific versions at both daily and annual time steps (Table 3, Figs. 6–8), with exception of a site with agricultural history (Alkkia). Although errors in the data cannot be excluded as potential reasons, the most likely explanation is that Alkkia forest is located on a peatland drained for agriculture in 1930s and subsequently afforested about 35 years ago. The agricultural history of the site is seen as high nutrient contents in the soil (due to use of fertilizers) that are reflected in the amount and species composition of the understory vegetation (Lohila et al., 2007). The vigorously growing understory at Alkkia is composed of deciduous species that have less conservative water use strategies than Scots pine and Norway spruce that dominate the LAI at other sites. Kalevansuo is also a drained peatland forest, but for this site M-S and S-S ET predictions were similar. Contrary to Alkkia, Kalevansuo has no agricultural history and its understory consists of dwarf shrubs and mosses similar to the mineral soil sites in this study (Table 1). The failure of the M-S calibration to predict ET at Alkkia could also be partially related to an improper representation of LAI. We used a lumped LAI for trees and understory, but they have different seasonal dynamics and different physiology. The problem can also be due to uncertainty in soil hydraulic characteristics (e.g., field capacity) as well as to simplistic representation of the soil water balance in PRELES. The water storages are described by small superficial water storage and a simple bucket model with a pre-defined fixed drainage coefficient and has no explicit description of lateral flows such as drainage to ditches.

In the S-S calibration at Alkkia, the model structural deficiency may have been compensated by different parameter estimates (see parameter χ of Fig. 1a).

PRELES predicts GPP better (Table 3) than evapotranspiration, the total water flux from several sources. While transpiration is highly correlated with GPP through linkage between stomatal conductance and assimilation rate (Katul et al., 2010; Medlyn et al., 2011b), other water sources are more constrained by stand characteristics, microclimate and soil properties. Modelling highly dynamic processes such as interception and evaporation at daily time step and neglecting the layered structure of forest ecosystems could be one reason for poorer ET predictions. Also, EC-based evapotranspiration estimates have in most cases higher uncertainty than carbon fluxes due to unclosed energy balance (Foken, 2008) and technical problems measuring water vapour at high air humidity (Mammarella et al., 2009).

The analysis on the MSE decomposition (Fig. 4) allowed us to understand model behaviour better. The main component of the MSE was usually the correlation error, probably due to the summer peaks that occur in particular environmental conditions and that the model is not able to reproduce. Furthermore, in this study we were using fixed annual values of LAI; a better representation of seasonal LAI dynamics might help improve the predictions of rainfall interception and thereby the water flux partitioning between transpiration and evaporation. However, the seasonal cycle modifier f_S, (Eq. A.12–A.14) partially accounts for both the annual cycle of photosynthetic capacity and LAI but it only influences GPP and transpiration predictions and has an upper limit of 1 that is reached typically before midsummer. Earth observed data would allow us to integrate the intra annual LAI variability in the model. In spite of these reservations, we found that the mean squared error was low for most of the sites and the deviation from the annual aggregated data was lower than 10% in most of the cases, confirming that the model can be considered a reliable tool to predict the carbon and water fluxes of boreal forests.

4.2. Sensitivity and uncertainty analysis

The sensitivity analysis carried out through the prior parameter space allowed us to identify the parameters that *a priori* were the most influential on the outputs. The differences between the sensitivity results at Norunda and the rest of the sites were mainly due to the differences in leaf area index, LAI being much higher at Norunda. LAI has a strong impact on PRELES outputs, especially on GPP. The sensitivity of photosynthesis to LAI follows the exponential curve of equation 1, since GPP is linearly related to the fraction of absorbed PPFD (Eq. A.9). It is important to have accurate estimates of LAI, especially for stands with low foliage biomass (e.g, young stands, low productive sites), because small errors in LAI strongly affect GPP calculations. Nowadays, thanks to remote sensing techniques, it is possible to obtain inputs for ecosystem models at high spatial and temporal scale, making possible the application of process-based models in practice (Härkönen et al., 2011). The relatively weak response of forest ET to LAI (Fig. 4) is in line with EC-based observations in boreal Canada (Amiro et al., 2006).

Combining the sensitivity of model output to the parameters and the uncertainty analysis it was possible to extract useful information about PRELES behaviour and its general applicability. Eddy-covariance data can typically constrain 4–5 parameters (Keenan et al., 2013); however in the M-S calibration the uncertainty of all the 12 parameters was significantly reduced, probably because we were assimilating data from 10 sites.

The assumptions made on data uncertainties can greatly affect the posterior distribution (Keenan et al., 2011). A proper representation of data uncertainty is often difficult to achieve, because of lack of knowledge and/or computational limitations. We carried out a few calibrations using different likelihoods and indeed the posterior distribution slightly changed accordingly to the different choices made; however the S-S and M-S posteriors were still similar, in support of the fact that the model can be calibrated and applied for regional simulations.

The posterior distribution should always be provided in order to have a proper rappresentation of model uncertainty (Van Mourik et al., 2014). The uncertainty of the most influential parameters was strongly reduced by the data in all the calibrations and the parameter estimates were quite similar for the different versions of the model (Fig. 1a). Peltoniemi et al. (2015b), performing a local sensitivity analysis, found that PRELES outputs were sensitive to soil-water related parameters (ρ_P, ρ_E, υ) under water limited years, but less sensitive in a wet year (Peltoniemi et al., 2015b). These parameters did not contribute to the output uncertainty in this study and their posterior distribution remained quite uncertain. The reason for this is that the boreal forests are not often water stressed so there is little information to estimate these parameters. However, those parameters could become crucial if the model is applied to more xeric sites. In the CAge12yr site-specific calibration, some of the most important parameters (i.e., γ, X_0 and β) were quite different from the rest of the calibrations. This could be due to the understory that accounted for almost 50% of the LAI of this site. Also, in young regenerating stands the contribution of deciduous tree species is more abundant than at older sites. The physiological differences of forest plant species might influence the stand level carbon and water fluxes. In the future it must be investigated if modelling stand layers separately as well as describing the soil water balance in more detail will improve model performance.

No significant differences were encountered in the parameter estimates and model outputs of the Scots pine and Norway spruce dominated stands. The delay parameter for ambient temperature response (τ, Fig. 1b) was the only parameter for which the Norway spruce dominated sites (i.e., Flakaliden, Norunda and Knottåsen) had similar marginal posterior distributions, while τ estimates for the Scots pine dominated sites were different. Consistently with recent results by Linkosalo et al. (2014), the photosynthetic activity of Norway spruce starts earlier than Scots pine, explaining the lower values of τ. Flakaliden was the site were the model showed the worst performance in predicting GPP. Nevertheless, since the model performed well at the other two spruce sites, we believe that there is no need for a species specific calibration of PRELES, which also speaks for the generality of model calibration, at least given the uncertainties involved.

Parameter uncertainty was strongly reduced by the BC since thousands of data points were involved in the calibrations; for this reason also the uncertainties of model predictions were low. As expected M-S resulted in more accurate parameter estimates; however, PRELES parameter uncertainty was low also for S-S of the sites with the long-term datasets (i.e., Hyytiälä, Sodankylä, Flakaliden, Norunda and Kalevansuo). Furthermore the sites with the long-term datasets were the most influential on M-S, having more weight on the likelihood. In particular Hyytiälä data had a strong influence on the M-S calibration. Instead of giving a weight to each site dataset according to the number of data points we preferred to use all the available information to calibrate and test the model.

4.3. Predicting fluxes for a site for which data are already available

The analysis #2 and #3 tested the reliability of the model. Results from BMC can look quite severe (Tables 4 and 5) but in reality, while BMC tells which model is more likely to be the best, this does not mean that the worst model gives completely wrong predictions. Combining BMC with more classical model error quantifications provided a more complete picture about the models under evaluation. In our analyses BMC was assigning in most of the cases near to 100% probability of being correct to one version of the model because thousands of data points were used in the comparisons. However the NRMSEs (Figs. 7 and 8) showed quantitatively that the differences between M-S and S-S calibrations in predicting carbon and water fluxes are quite low for most of the sites.

The eddy-covariance network is expanding and flux data is currently available for hundreds of sites (Baldocchi, 2008). At some sites measurements have already been collected for more than a decade and it is likely that the inter-annual variability is well represented by the measurements. On the contrary, for other sites the fluxes have been measured just for a few years. Analysis #2 gave insights into which model version (i.e. M-S or S-S) is more appropriate to predict carbon and water fluxes for a site for which data are available. The multi-site calibration showed robust performances in predicting carbon and water fluxes when compared to site-specific versions. M-S was the best calibration for 4 sites over 10 (Table 4) and the NRMSE between M-S and S-Ss were not significantly different in most of the cases.

Except for Skyttorp, the evapotranspiration NRMSEs of the multi-site version were always higher than S-S NRMSE (Fig. 7b). As suggested by Analysis #1, the ET module in PRELES seems to be too simplistic, rendering the S-S calibration with better performance. Likely there are site-specific differences in flux-environmental driver relationships that could have been compensated by site-specific parameter estimates. Comparison results are more significant when a high number of data are involved and the measurements cover different years. We conclude that both versions of the model (the S-S and the M-S) can be used to predict GPP and ET of sites for which flux data are available but the S-S should be preffered when long-time series are available (Table 4) at a site, while the M-S is more appropriate for sites with short-term data series.

4.4. Analysis #3. The regional applicability of PRELES

Results from Analysis #3 provided key indications about the regional applicability of PRELES. The performances of the M-S calibration and the Hyytiälä site-specific calibration were more reliable in predicting ET and GPP compared to the other site-specific versions. Hyytiälä S-S provided the most robust performances and was slightly better than the M-S. This is probably due to the fact that Hyytiälä was the most comprehensive dataset in terms of site years and data quality. This underlines the importance of long and carefully collected flux datasets, even a single site can provide model calibrations that can be applied at a wider spatial scale, since it covers a wide range of variability in climatic conditions. However, the good fit to Hyytiälä data can also stem from the development history and structure of PRELES (Mäkelä et al., 2008; Peltoniemi et al., 2015b) that was partly based on the understanding gained from the Hyytiälä measurements. Therefore, using a multi-site calibrated model for regional analysis represents a more conservative choice in terms of spatial representativeness, and the fact that not all required site conditions can be extrapolated by the modeller. From carbon modelling perspective, the use of a few aerially representative sites with long and high quality records would likely improve the reliability of model predictions.

5. Conclusions

Some ecosystem processes, such as photosynthesis, have been found to be sufficiently well understood and generalizable. Our analyses involved just one model and these results could surely not to be expected with all models unless models are generic and calibrated with high quality data. PRELES has a simplified structure but as a computationally efficient model that requires easily available input data it is suitable for applications at regional scale. A useful application of PRELES could be in estimating the impact of climate change on Boreal forests. On one hand, we tested the model with data that covered a wide range of climatic conditions (i.e., D, PPFD, R, T); but, on the other hand, we used fixed CO_2 atmosphere concentrations. Nevertheless, according to future climate scenario simulations (IPCC, 2007), CO_2 concentration is expected to increase dramatically in the future, having a strong impact on Boreal forests (Kalliokoski et al., in prep.). The CO_2 effects on forest photosynthetic activity have already been estimated by other studies (Kolari et al., 2009); implementing this information in PRELES structure (Kalliokoski et al., in prep.) will allow us to use PRELES to make estimates under future climatic scenarios.

Code availability

PRELES is available in C language and in an R package. Model code and the posterior parameter distributions of this study can be obtained on request from Francesco Minunno (francesco.minunno@helsinki.fi) or can be downloaded at https://github.com/MikkoPeltoniemi/Rpreles/. The APPENDIX D describes how to install and run PRELES using R programming language.

Acknowledgements

We thank Pasi Kolari for his general advice in acquiring the data.

The study was supported by the European project “Enabling Intelligent GMES Services for Carbon and Water Balance Modeling of Northern Forest Ecosystems” and the LIFE+ financial instrument of the European Union (LIFE12 ENV/FI/000409 Monimet, LIFE09 ENV/FI/000571 Climforisk).

The authors thank their colleagues for continuing support and discussion around the coffee breaks. The editor thanks X. Y. Furore and another referee for assisting in evaluating this paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2016.09.020.

References

Amiro, B.D., Orchansky, A.L., Barr, A.G., Black, T.A., Chambers, S.D., Chapin III, F.S., Goulden, M.L., Litvak, M., Liu, H.P., McCaughey, J.H., McMillan, A., Randerson, J.T., 2006. The effect of post-fire stand age on the boreal forest energy balance. Agric. For. Meteorol. 140, 41–50.

Aubinet, M., Vesala, T., Papale, D., 2012. Eddy Covariance: A Practical Guide to Measurement and Data Analysis. Springer Science & Business Media, pp. 460pp.

Baldocchi, D., 2008. TURNER REVIEW No. 15.'Breathing'of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust. J. Bot. 56, 1–26.

Brümmer, C., Black, T.A., Jassal, R.S., Grant, N.J., Spittlehouse, D.L., Chen, B., Nesic, Z., Amiro, B.D., Arain, M.A., Barr, A.G., Bourque, C.P.-A., Coursolle, C., Dunn, A.L., Flanagan, L.B., Humphreys, E.R., Lafleur, P.M., Margolis, H.A., McCaughey, J.H., Wofsy, S.C., 2012. How climate and vegetation type influence evapotranspiration and water use efficiency in Canadian forest, peatland and grassland ecosystems. Agric. For. Meteorol. 153, 14–30.

Chen, J., Zhang, H., Liu, Z., Che, M., Chen, B., 2014. Evaluating parameter adjustment in the MODIS gross primary production algorithm based on eddy covariance tower measurements. Remote Sens. 6, 3321–3348.

Duursma, R.A., Kolari, P., Perämäki, M., Pulkkinen, M., Mäkelä, A., Nikinmaa, E., Hari, P., Aurela, M., Berbigier, P., Bernhofer Ch Grunwald, T., Loustau, D., Molder, M., Verbeeck, H., Vesala, T., 2009. Contributions of climate, leaf area index and leaf physiology to variation in gross primary production of six coniferous forests across Europe: a model-based analysis. Tree Phys. 29, 621–639.

Farquhar, G.D., von Caemmerer, S., von Berry, J.A., 1980. A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta 149, 78–90.

Foken, T., 2008. The energy balance closure problem: an overview. Ecol. Appl. 18, 1351–1367.

Friend, A.D., Lucht, W., Rademacher, T.T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D.B., Dankers, R., Falloon, P.D., Ito, A., Kahana, R., Kleidon, A., Lomas, M.R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N., Warszawski, L., Wiltshire, A., Woodward, F.I., 2014. Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2. Proc. Natl. Acad. Sci. 111, 3280–3285, http://dx.doi.org/10.1073/pnas.1222477110.

Green, E.J., MacFarlane, D.W., Valentine, H.T., 2000. Bayesian synthesis for quantifying uncertainty in predictions from process models. Tree Physiol. 20, 415–419, http://dx.doi.org/10.1093/treephys/20.5-6.415.

Härkönen, S., Lehtonen, A., Eerikäinen, K., Peltoniemi, M., Mäkelä, A., 2011. Estimating forest carbon fluxes for large regions based on process-based modelling, NFI data and Landsat satellite images. For. Ecol. Manag. 262, 2364–2377, http://dx.doi.org/10.1016/j.foreco.2011.08.035.

Hari, P., Kulmala, M., 2005. Station for measuring ecosystem-atmosphere relations. Boreal Environ. Res. 10, 315–322.

Hartig, F., Dyke, J., Hickler, T., Higgins, S.I., O'Hara, R.B., Scheiter, S., Huth, A., 2012. Connecting dynamic vegetation models to data – an inverse perspective. J. Biogeogr. 39, 2240–2252, http://dx.doi.org/10.1111/j.1365-2699.2012.02745.x.

Heinsch, F.A., Zhao, M., Running, S.W., Kimball, J.S., Nemani, R.R., Davis, K.J., Bolstad, P.V., Cook, B.D., Desai, A.R., Ricciuto, D.M., Law, B.E., Oechel, W.C., Kwon, H., Luo, H., Wofsy, S.C., Dunn, A.L., Munger, J.W., Baldocchi, D.D., Xu, L., Hollinger, D.Y., Richardson, A.D., Stoy, P.C., Siqueira, M.B.S., Monson, R.K., Burns, S.P., Flanagan, L.B., 2006. Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. Geosci. Remote Sens. IEEE Trans. 44, 1908–1925.

IPCC, 2007. In: Solomon, S., D. Qin, M., Manning, Z., Chen, M., Marquis, K.B., Averyt, M. Tignor (Eds.), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University PressCambridge, United Kingdom and New York, NY, USA.

Janssen, P.H.M., Heuberger, P.S.C., 1995. Calibration of process oriented models. Ecol. Model. 83 (1–2), 55–66.

Juang, J.-Y., Katul, G.G., Siqueira, M.B., Stoy, P.C., McCarthy, H.R., 2008. Investigating a hierarchy of Eulerian closure models for scalar transfer inside forested canopies. Bound. – Layer Meteorol. 128, 1–32.

Kaminski, T., Rayner, P.J., Voßbeck, M., Scholze, M., Koffi, E., 2012. Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design. Atmos. Chem. Phys. 12, 7867–7879.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 430 (90), 773e795, American Statistical Association. Retrieved from /2291091 http://www.jstor.org/pss.

Katul, G., Manzoni, S., Palmroth, S., Oren, R., 2010. A stomatal optimization theory to describe the effects of atmospheric CO2 on leaf photosynthesis and transpiration. Ann. Bot. 105, 431–442, http://dx.doi.org/10.1093/aob/mcp292.

Keenan, T.F., Carbone, M.S., Reichstein, M., Richardson, A.D., 2011. The model–data fusion pitfall: assuming certainty in an uncertain world. Oecologia 167, 587–597.

Keenan, Trevor F., Davidson, Eric A., Munger, J. William, Richardson, Andrew D., 2013. Rate my data: quantifying the value of ecological data for the development of models of the terrestrial carbon cycle. Ecol. Appl. 23 (1), 273–286.

Kimball, J.S., Thornton, P.E., White, M.A., Running, S.W., 1997. Simulating forest productivity and surface–atmosphere carbon exchange in the BOREAS study region. Tree Physiol. 17, 589–599, http://dx.doi.org/10.1093/treephys/17.8-9.589.

Kleja, D.B., Svensson, M., Majdi, H., Jansson, P.-E., Langvall, O., Bergkvist, B., Johansson, M.-B., Weslien, P., Truusb, L., Lindroth, A., Ågren, G.I., 2008. Pools and fluxes of carbon in three Norway spruce ecosystems along a climatic gradient in Sweden. Biogeochemistry 89, 7–25.

Kobayashi, K., Salam, M.U., 2000. Comparing simulated and measured values using mean squared deviation and its components. Agron. J. 92, 345–352.

Kolari, P., Pumpanen, J., Rannik, Ü., Ilvesniemi, H., Hari, P., Berninger, F., 2004. Carbon balance of different aged Scots pine forests in southern Finland. Global Change Biol. 10, 1106–1119.

Kolari, P., Kulmala, L., Pumpanen, J., Launiainen, S., Ilvesniemi, H., Hari, P., Nikinmaa, E., 2009. CO2 exchange and component CO2 fluxes of a boreal Scots pine forest. Boreal Environ. Res. 14, 761–783.

Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., Richardson, a. D., 2012. Constraining a global ecosystem model with multi-site eddy-covariance data. Biogeosci. Discuss. 9 (3), 3317–3380, http://dx.doi.org/10.5194/bgd-9-3317-2012.

Launiainen, S., Katul, G.G., Kolari, P., Vesala, T., Hari, P., 2011. Empirical and optimal stomatal controls on leaf and ecosystem level CO 2 and H 2 O exchange rates. Agric. For. Meteorol. 151, 1672–1689.

Leuning, R., Kelliher, F.M., Pury, D., de Schulze, E.-D., 1995. Leaf nitrogen, photosynthesis, conductance and transpiration: scaling from leaves to canopies. Plant Cell Environ. 18, 1183–1200.

Lindroth, A., Klemedtsson, L., Grelle, A., Weslien, P., Langvall, O., 2008. Measurement of net ecosystem exchange: productivity and respiration in three spruce forests in Sweden shows unexpectedly large soil carbon losses. Biogeochemistry 89, 43–60.

Linkosalo, T., Heikkinen, J., Pulkkinen, P., Mäkipää, R., 2014. Fluorescence measurements show stronger cold inhibition of photosynthetic light reactions in Scots pine compared to Norway spruce as well as during spring compared to autumn. Front. Plant Sci. 5, 264.

Liu, J., Chen, J.M., Cihlar, J., Park, W.M., 1997. A process-based boreal ecosystem productivity simulator using remote sensing inputs. Remote Sens. Environ. 62, 158–175.

Lohila, A., Laurila, T., Aro, L., Aurela, M., Tuovinen, J.-P., Laine, J., Kolari, P., Minkkinen, K., 2007. Carbon dioxide exchange above a 30-year-old Scots pine plantation established on organic-soil cropland. Boreal Environ. Res. 12, 141–157.

Lohila, A., Minkkinen, K., Aurela, M., Tuovinen, J.-P., Penttilä, T., Laurila, T., 2011. Greenhouse gas flux measurements in a forestry-drained peatland indicate a large carbon sink. Biogeosciences 8, 3203–3218, http://dx.doi.org/10.5194/bg-8-3203-2011.

Lundin, L.-C., Halldin, S., Lindroth, A., Cienciala, E., Grelle, A., Hjelm, P., Kellner, E., Lundberg, A., Mölder, M., Morén, A.-S., Nord, T., Seibert, J., Stähli, M., 1999. Continuous long-term measurements of soil-plant-atmosphere variables at a forest site. Agric. For. Meteorol. 98, 53–73.

Mäkelä, A., Pulkkinen, M., Kolari, P., Lagergren, F., Berbigier, P., Lindroth, A., Loustau, D., Nikinmaa, E., Vesala, T., Hari, P., 2008. Developing an empirical model of stand GPP with the LUE approach: analysis of eddy covariance data at five contrasting conifer sites in Europe. Glob. Change Biol. 14, 92–108, http://dx.doi.org/10.1111/j.1365-2486.2007.01463.x.

Mammarella, I., Launiainen, S., Gronholm, T., Keronen, P., Pumpanen, J., Rannik, Ü., Vesala, T., 2009. Relative humidity effect on the high-frequency attenuation of water vapor flux measured by a closed-path eddy covariance system. J. Atmos. Ocean Technol. 26, 1856–1866, http://dx.doi.org/10.1175/2009JTECHA1179.1.

Medlyn, B.E., Berbigier, P., Clement, R., Grelle, A., Loustau, D., Linder, S., Wingate, L., Jarvis, P.G., Sigurdsson, B.D., McMurtrie, R.E., 2005. Carbon balance of coniferous forests growing in contrasting climates: model-based analysis. Agric. For. Meteorol. 131, 97–124.

Medlyn, B.E., Duursma, R.A., Zeppel, M.J., 2011a. Forest productivity under climate change: a checklist for evaluating model studies. Wiley Interdiscip. Rev. Clim. Change 2, 332–355.

Medlyn, B.E., Duursma, R.A., Eamus, D., Ellsworth, D.S., Prentice, I.C., Barton, C.V., Crous, K.Y., de Angelis, P., Freeman, M., Wingate, L., 2011b. Reconciling the optimal and empirical approaches to modelling stomatal conductance. Glob. Change Biol. 17, 2134–2144.

Meyers, T.P., Baldocchi, D.D., 1988. A comparison of models for deriving dry deposition fluxes of O3 and SO2 to a forest canopy. Tellus B 40B, 270–284, http://dx.doi.org/10.1111/j.1600-0889.1988.tb00297.x.

Minunno, F., van Oijen, M., Cameron, D.R., Pereira, J.S., 2013a. Selecting parameters for bayesian calibration of a process-Based model: a methodology based on canonical correlation analysis. SIAMASA J. Uncert. Quantif. 1, 370–385, http://dx.doi.org/10.1137/120891344.

Minunno, F., van Oijen, M., Cameron, D.R., Cerasoli, S., Pereira, J.S., Tomé, M., 2013b. Using a Bayesian framework and global sensitivity analysis to identify strengths and weaknesses of two process-based models differing in representation of autotrophic respiration. Environ. Model. Softw. 42, 99–115, http://dx.doi.org/10.1016/j.envsoft.2012.12.010.

Monteith, J.L., 1981. Evaporation and environment. In: Fogg, C.E. (Ed.), The State and Movement of Water in Living Organisms. Cambridge University Press, Cambridge, UK, pp. 205–234.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33, 161–174.

Ogée, J., Brunet, Y., Loustau, D., Berbigier, P., Delzon, S., 2003. MuSICA, a CO2, water and energy multilayer: multileaf pine forest model: evaluation from hourly to yearly time scales and sensitivity analysis. Glob. Change Biol. 9, 697–717.

Ojanen, P., Minkkinen, K., Lohila, A., Badorek, T., Penttilä, T., 2012. Chamber measured soil respiration: a useful tool for estimating the carbon balance of peatland forest soils? For. Ecol. Manage. 277, 132–140, http://dx.doi.org/10.1016/j.foreco.2012.04.027.

Olchev, A., Ibrom, A., Priess, J., Erasmi, S., Leemhuis, C., Twele, A., Radler, K., Kreilein, H., Panferov, O., Gravenhorst, G., 2008. Effects of land-use changes on evapotranspiration of tropical rain forest margin area in Central Sulawesi (Indonesia): modelling study with a regional SVAT model. Ecol. Model. 212, 131–137.

Peltoniemi, M., Pulkkinen, M., Kolari, P., Duursma, R.A., Montagnani, L., Wharton, S., Lagergren, F., Takagi, K., Verbeeck, H., Christensen, T., Vesala, T., Falk, M., Loustau, D., Mäkelä, A., 2012. Does canopy mean nitrogen concentration explain variation in canopy light use efficiency across 14 contrasting forest sites? Tree Physiol. 32, 200–218, http://dx.doi.org/10.1093/treephys/tpr140.

Peltoniemi, M., Markkanen, T., Härkönen, S., Muukkonen, P.T., Aalto, T., Mäkelä, A., 2015a. Consistent estimates of gross primary production of Finnish Forests – comparison of estimates of two process models. Boreal Environ. Res. 20, 196–212.

Peltoniemi, M., Pulkkinen, M., Aurela, M., Pumpanen, J., Kolari, P., Mäkelä, A., 2015b. A semi-empirical model of boreal forest gross primary production, evapotranspiration, and soil water – calibration and sensitivity analysis. Boreal Environ. Res. 20, 151–171.

Pihlatie, M.K., Kiese, R., Brüggemann, N., Butterbach-Bahl, K., Kieloaho, A.-J., Laurila, T., Lohila, A., Mammarella, I., Minkkinen, K., Penttilä, T., et al., 2010. Greenhouse gas fluxes in a drained peatland forest during spring frost-thaw event. Biogeosciences 7, 1715–1727.

Raddatz, T.J., Reick, C.H., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., Jungclaus, J., 2007. Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century? Clim. Dyn. 29, 565–574, http://dx.doi.org/10.1007/s00382-007-0247-8.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J.-M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., Valentini, R., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Glob. Change Biol. 11, 1424–1439, http://dx.doi.org/10.1111/j.1365-2486.2005.001002.x.

Reichstein, M., Bahn, M., Mahecha, M.D., Kattge, J., Baldocchi, D.B., 2014. Linking plant and ecosystem functional biogeography. PNAS 111, 13697–13702.

Richardson, Andrew D., Mahecha, Miguel D., Falge, Eva, Kattge, Jens, Moffat, Antje M., Papale, Dario, Reichstein, Markus, Stauch, Vanessa J., Braswell, Bobby H., Churkina, Galina, Kruijt, Bart, Hollinger, David Y., 2008. Statistical properties of random CO_2 flux measurement uncertainty inferred from model residuals. Agricult. For. Meteorol. 148 (1), 38–50.

Richardson, A.D., Anderson, R.S., Arain, M.A., Barr, A.G., Bohrer, G., Chen, G., Chen, J.M., Ciais, P., Davis, K.J., Desai, A.R., Dietze, M.C., Dragoni, D., Garrity, S.R., Gough, C.M., Grant, R., Hollinger, D.Y., Margolis, H.A., McCaughey, H., Migliavacca, M., Monson, R.K., Munger, J.W., Poulter, B., Raczka, B.M., Ricciuto, D.M., Sahoo, A.K., Schaefer, K., Tian, H., Vargas, R., Verbeeck, H., Xiao, J., Xue, Y., 2012. Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site synthesis. Global Change Biol. 18, 566–584, http://dx.doi.org/10.1111/j.1365-2486.2011.02562.x.

Ter Braak, C.J., Vrugt, J.A., 2008. Differential evolution Markov chain with snooker updater and fewer chains. Stat. Comput. 18, 435–446.

Thum, T., Aalto, T., Laurila, T., Aurela, M., Lindroth, A., Vesala, T., 2008. Assessing seasonality of biochemical CO 2 exchange model parameters from micrometeorological flux observations at boreal coniferous forest. Biogeosciences 5, 1625–1639, et al., 2008.

Valladares, F., Skillman, J.B., Pearcy, R.W., 2002. Convergence in light capture efficiencies among tropical forest understory plants with contrasting crown architectures: a case of morphological compensation. Am. J. Bot. 89 (8), 1275–1284.

Van Mourik, S., ter Braak, C., Stigter, H., Molenaar, J., 2014. Prediction uncertainty assessment of a systems biology model requires a sample of the full probability distribution of its parameters. Peer J., e433, http://dx.doi.org/10.7717/peerj.433.

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

Van Oijen, M., Cameron, D.R., Butterbach-Bahl, K., Farahbakhshazad, N., Jansson, P.-E., Kiese, R., Rahn, K.-H., Werner, C., Yeluripati, J.B., 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, 1609–1621, http://dx.doi.org/10.1016/j.agrformet.2011.06.017.

Van Oijen, M., Reyer, C., Bohn, F.J., Cameron, D.R., 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. Manag. 289, 255–268, http://dx.doi.org/10.1016/j.foreco.2012.09.043.

Further reading

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Model. Softw. 22, 1509–1518. http://dx.doi.org/10.1016/j.envsoft.2006.10.004.

Kuusisto, E., 1984. Snow Accumulation and Snowmelt in Finland. No. 55, Publications of the Water Research Institute, National Board of Waters Finland, pp. 149.

Landsberg, J.J., Waring, R.H., 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. For. Ecol. Manag. 95, 209–228. http://dx.doi.org/10.1016/S0378-1127(97)00026-1.

Mäkelä, A., Hari, P., Berninger, F., Hänninen, H., Nikinmaa, E., 2004. Acclimation of photosynthetic capacity in Scots pine to the annual cycle of temperature. Tree Physiol. 24, 369–376. http://dx.doi.org/10.1093/treephys/24.4.369.

Song, X., Bryan, B.A., Paul, K.I., Zhao, G., 2012. Variance-based sensitivity analysis of a forest growth model. Ecol. Model. 247, 135–143. http://dx.doi.org/10.1016/j.ecolmodel.2012.08.005.

Song, X., Bryan, B.A., Almeida, A.C., Paul, K.I., Zhao, G., Ren, Y., 2013. Time-dependent sensitivity of a process-based ecological model. Ecol. Model. 265, 114–123. http://dx.doi.org/10.1016/j.ecolmodel.2013.06.013.